N. d'ordre : ...

Ecole Doctorale SPI de Lille
Université d'Artois

Méthaheuristiques pour le Problème de Sélection d'Attributs
Metaheuristcs For The Feature Selection Problem:
Memetic, Adaptive And Swarm Approaches

THÈSE

présentée et soutenue publiquement le .. septembre 2011
en vue de l'obtention du

Doctorat de l'Université d'Artois

(Spécialité : Génie Informatique et Automatique)

par

Mohamed Amir Esseghir

Composition du jury

Rapporteurs :	M. Patrick Siarry	Professeur, Université Paris-Est Créteil
	M. Jean-Charles Creput	MCF HDR, Université de Technologie de Belfort-Montbéliard (UTBM)
Examinateurs :	M. Rémy Dupas	Professeur, Université Bordeaux I
	Mme. Laetitia Jourdan	Professeur, INRIA / USTL
	M. Daniel Jolly	Professeur, Université d'Artois
Directeur de thèse :	M. Gilles Goncalves	Professeur, Université d'Artois

Laboratoire de Génie Informatique et d'Automatique de l'Artois, (E.A. 3926)

Contents

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

1.1 Context

We experienced fast data evolution, in which high dimensional data become increasingly common [90]. Although the expansion of storage technologies, networking systems, and information system methodologies, the capabilities of conventional data processing techniques remain limited.

The need to knowledge extraction, compact representation and data analysis are highly motivated by data expansion. Nevertheless, learning from data might be a complex task, particularly when it includes noisy, redundant and information-less attributes.

Such challenges, stimulate the development of feature selection research [53, 54, 87, 90], since it is considered as a valuable way to provide compact representations of a given context and to prepare data for other learning methods. In fact, in the last decade we witnessed the expansion of the Feature Selection research in multiple disciplines and application fields (data mining, machine learning, combinatorial optimization, bio-informatics, statistics, fraud and spam detection, bankruptcy prediction, *etc*).

In this thesis we investigate different optimization paradigms as well as its adaptation to the requirements of the feature selection challenges.

1.2 Problem and Motivations

Feature Selection (FS) tries to select the most relevant attributes from raw data, and hence guides the construction of final classification models or decision support systems. Selected features should be representative of the underlying data and provide effective usefulness to the targeted learning paradigm (*i.e.* classification technique, knowledge representation technique).

In addition to the compact representation of data, feature selection allows:

- the selection of useful features: appropriate features for a particular context
- discarding irrelevant attributes: noisy or information-less.
- removing redundancy: attributes that are correlated to the selected ones and providing the same informational content.
- reducing models complexity: enhancing comprehensibility of the devised models

The selection of a subset of attributes, according to a given criterion, involves the selection of a search space more larger than the number of attributes, since it should take into account possible combinations of attributes. The combinatorial nature of the problem [54, 90] should be tackled by adapted methods reducing search complexity and keeping an acceptable effectiveness level.

The above challenging problems will be detailed through the review of the existing approaches as well as the recent advances in the field. Various aspects of the challenging points will be studied throughout different contributions.

1.3 Contributions

According to *H. liu* and *H. Motoda* [90], the expansion of feature selection research could be characterized in two ways: (i) The first is to develop and adapt existing approaches to meet the arising challenges (ii) while the second trend targeted the development of new algorithms tackling the above mentioned challenges. Both aspects are taken into consideration, throughout this thesis. In fact, contributions cover enhancement of existing approaches as well as the investigation of new FS modeling schema.

1.3.1 Goals

Among goals that will be studied throughout the thesis,

- The enhancement of Genetic Algorithms (GA) search capabilities by endowing it with a variety of local search operators [34, 58]. The design of the LS operators should take into consideration the requirement of FS modeling, in addition to the complexity to the problem at hand.
- Looking for an optimization paradigm which is able to combine the advantages of both filters and wrappers in a natural way.
- the investigation of the swarm modeling as a framework for the combination of knowledges provided by different filters as well as the hybridization of filter-wrapper capabilities.

1.3.2 Empirical validation

The empirical study of the different experimented approaches were done according to validation protocol which take into consideration the requirement and the specificities of the FS problem. The validation protocol involves metrics for the assessment of the results of both stages: search and validation. The search stage requires a procedure for fitness evaluation which reflects the accuracy of the retained feature subset, while the validation stage assesses the reliability of returned solution regardless of the search context (*i.e.* different classifiers, different validation data sets). The validation protocol used for the assessment of the different approaches will be detailed in Annex I (*see p.* 165).

1.4 Thesis outline

Thesis chapters are organized as follows:

- Chapter II reviews feature selection modeling, details main problem aspects, and surveys featured, recent and reference approaches.
- Chapter III is devoted to the enhancement of evolutionary algorithms with local search (LS) capabilities. We focus on global local-search hybridization as well as the adaptation of the LS operators to the requirement of high dimensional FS problems.
- Chapter IV investigates the Greedy Randomized Adaptive Search Procedures (GRASP) [38] metaheuristic as a scheme for FS modeling through the effective combination wrapper-filter approaches [32]. We extend an existent recent proposed GRASP to a more accurate approach which is able to handle high dimensional data sets.
- Chapter V explores the swarm FS modeling abilities. A continuous PSO (Particle Swarm Optimization [31]) is devised and compared to a reference approach based on a binary PSO [21]. We study, also, the wrapper filter hybridization issues as well as the adaptation and the control of swarm parameters and velocities.

Finally, we conclude the thesis by summarizing, main design issues, and potential results. We provide some perspectives in relation with the investigated search methodologies and the new trends and application in the context of feature selection and classification optimization paradigms.

Chapter 2

Feature Selection Problem: Rewiew and Recent Advances

2.1 Introduction

The identification of useful and informative attributes for given data set, broadly referred to as Feature Selection (FS), is an attractive and challenging research topic for several domains including predictive Data Mining, Pattern recognition, Machine Learning and information retrieval [54, 96, 100].

Recently, the interest in feature selection has been on the increase for several reasons including the expansion of dataset dimensions, the need to analyze with the same efficiency as with small datasets and the development of new and urgent needs of robust and reliable techniques for fraud detection, multimedia information retrieval, and predictive modeling for medical data [103].

One of the fundamental motivations for feature selection is the curse of dimensionality [90]. In fact, the presence of useless features may not only deteriorate the performance of learning algorithms but also obscure information behind data [89]. Considered as a fundamental problem in machine learning [122], the role of FS is critical, especially, in a context deemed with irrelevant features (*i.e.* redundant and noisy features).

In addition to the curse of dimensionality, effective FS has the ability to reduce measurement and storage costs, alleviate computational complexity for classification and data analysis processes, improve classification accuracy, reduce over-fitting in learning and enhance model comprehensibility [53, 54].

Besides, the multi-disciplinary nature of the FS problem has resulted in a plethora of approaches and techniques that have attempted to tackle the problem from different purposes consideration and targets and using various methodologies as well as optimization paradigms [54]. Most of the existing approaches rely on heuristic search, and thus, cannot provide any guarantee of optimality.

The aim of this chapter is to introduce basic problem materials, provides recent survey of existing FS approaches and to enumerate challenging issues for FS modeling and applications. The reminder of the chapter is organized as follows:

Section 2 details different formulations and concepts associated to FS. Section 3 reviews basic apporaches and new advances in FS. Section 4 discusses both challenges and future research direction in FS modeling. Section 5 concludes the chapter.

2.2 Problem formulation and main concepts

This section starts by reviewing different feature selection problem formulation alternatives. Next, we detail, some definitions and basic concepts in relation with feature selection modeling. After that, an overview of the feature selection process will be detailed.

2.2.1 Problem formulation

The FS problem was defined as an optimization problem [87, 90] and existing formulations could be grouped into tree families. The common one, which is being widely used, suggests the selection of the subset that could achieve the highest classification accuracy (*i.e.* The ability to represent and predicit a given target value) [1] [54, 90]. The problem is formlulated as follows: Let D be a data set where each instance is described by n attributes [2] N ($\| N \|= n$), and let X ($X \subseteq N$) be a subset of N.

---

[1] Classification and associated concepts and paradigms will be subsequently detailed

[2] Attributes and features would be used interchangeably

Let $J(X)$ be the function capable of assessing the relevance of the subset X. The problem of feature selection states the selection of a subset Z such that:

$$J(Z) = max_{X \subseteq N} J(X) \tag{2.1}$$

The selection of the best subset requires the exploration of 2^n solutions, which would be prohibitive, even, for moderate values of n. The typical structure of the data set D includes a set of attributes (features) describing the context and a set of classes representing the target to predict or to represent by a compact set of features. Table 2.1 illustrates such a structure.

	$\mathbf{f_1}$	$\mathbf{f_2}$	. . .	$\mathbf{f_n}$	***class***
instance 1	val_{11}	val_{12}	. . .	val_{1n}	c_1
instance 2	val_{21}	val_{22}	. . .	val_{2n}	c_2
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
instance m	val_{m1}	val_{m2}	. . .	val_{mn}	c_j

Table 2.1: Typical data set structure

In other words, the retained feature subset should be compact and representative of the dataset instances or the underlying context. This can be done by both ignoring redundant and/or irrelevant attributes and keeping the minimal information loss.

Some alternative formulations, tried to alleviate search exploration complexity by considering a penalty term within the objective function (*i.e.* multi-objective formulation [35, 42, 51]). In fact, an additional member of the objective function computes the number of selected attributes and associated cost. By doing so, search is focused on subsets of features well performing with a reduced cardinality.

The third formulation targets a subspace of original search space, by limiting the search space exploration to solutions with a given cardinality ($d << n$). Therefore, the search space is reduced to the combinations $\binom{n}{d}$ of d-subsets. Indeed, all solutions representing larger subset or different cardinalities are discarded. Such formulation could be useful for high dimensional datasets. However, we should define the appropriate d value, because each d provides a new search space different from the others. Featured apporaches adopting such formulation alternative include [27, 101, 120, 133, 134].

The common argument in favor of such formulation is that the target of the FS process, is to find a subset of attributes that can be used to carry out the classification task in an optimal way.

In addition to the three above formulations, another one, which is rarely used, consists in searching for compact subsets reaching a given level of accurracy [79].

Throughout this thesis we opt for the first formulation not only because it is commonly and widely used, but also because we don't have to cope with the definition neither the approximation of the appropriate d value, or the weighting coefficients to scale between the solution quality and subset cardinality. Further studies could focus and study in depth multi-objective formulation issues or both theoretical and empirical approximation of the optimal d value during the search.

2.2.2 Concepts and Definitions

In this section, we review some featured definitions and concepts frequently used in feature selection resolution modeling.

One important question that could be discussed here, is how to categorize features and to consider them as: *relevant*, *irrelevant*, *noisy*, or/and *redundant*?

Relevance

To reply to the above question, we need to recall definitions of relevance property as they were introduced by Kohavi [68, 76].

Definition 1.

Strong Relevance: An attribute f_i is strongly relevant if its removal yields a deterioration of the performance criterion (*i.e.* classification rate)

Definition 2.

Weak Relevance: An attribute f_i is weakly relevant if it is not strongly relevant and there exists a subset of features X such that the performance on $X \bigcup \{f_i\}$ is better than the performance on X.

Therefore, features that are neither strongly relevant or weakly relevant are *irrelevant*. The presence of many irrelevant features, particularly in high dimensional contexts, could lead classifiers to *overfit training data*: learning from irrelevant feature made the classifier unable to generalize and predict correct outcome on new data [27, 53, 54, 135]. Some related concepts add another degree of complexity, especially when we consider overlapping boundaries between relevant, redundant and interacting features (*see Figure* 2.2).

We think that attributes relevance could not be defined in an absolute manner (*i.e.* considering only attribute-class dependency level). In fact, the attribute add/removal impact on performance contribution/deterioration of a given feature might vary not only, from one feature to another but also from one selected subset to another.

Redundancy

The notion of feature redundancy, is intuitively expressed in terms of correlation [90]. Indeed, two features highly correlated might be seen as redundant. Nevertheless, the attribute redundancy depends on the metric used for correlation assessment (*i.e.* linear, non-liear correlation) and also, on their respective dependency to the context (*i.e.* class, target outcome). Several studies consider the correlation and attributes distribution similarity as a sufficient metric for redundancy detection [54], while recent ones focus on both attributes correlation and attribute-class dependency [27, 35, 54, 90]. This notion was, usually, associated to the concept of Markov Blankets [54].

Definition 3.

Redundant feature [90]:

for a given subset X, a feature is redundant and hence should be removed, if it has a Markov blanket M_i within X.

Definition 4.

Markov blanket: given a feature $f_i \in F$, let $M_i \subset F$, such that $f_i \notin M_i$, M_i is said to be a markov blanket for f_i iif

$$P(F \setminus M_i \setminus \{f_i\}, (C|f_i, M_i)) = P(F \setminus M_i \setminus \{f_i\}, C|M_i) \tag{2.2}$$

where $P(.)$ denotes a probalility distribution. In other words, a Markov blanket of a given variable $f_i \notin X$ is defined as the minimal union of all variables M_i that makes X independent from M_i [77].

Several works have attempted to tackle the FS problem with two stages search. Indeed, rather than selecting representative features, they try to, first, identify relevant ones, then the redundent attributes are discarded. This two stages scheme, is illustrated by Figure 2.1. Featured approaches, which have adopted such a scheme include [15, 35, 88, 106, 130, 135].

Since Markov Blanket process requires exponential time for a full dependency network exploration, a set of alternatives attempted to overcome this limitation [90, 131].

Figure 2.1: Relevance-redundancy serach process [131]

A *Blanket Markov* approximation was proposed in [77], then reformulated in [130]. The approximation is based on information theory measures. Torkkola reports in [54], that the returned subset is not optimal and might contain unnecessary variables. He also, suggested that redundancy must be dealt by other methods. Nevertheless, the yielding dimensionality reduction was interesting in a context of large number of irrelevant variables (*i.e.* biomarker identification [130, 139]). A figure which could, in a part, summarizes the different concepts previously introduced was proposed in [131] and [90] (*see Figure* 2.2). the feature set includes *irrelevant* (I), *strongly relevant*

(IV), and *weakly relevant*. The optimal feature subset according to [131] must cover strongly relevant features. However, these features might be redundant and optimal subset obtained after an exhaustive search space exploration might include feature not strongly relevant but performing better when they interact with the selected attributes.

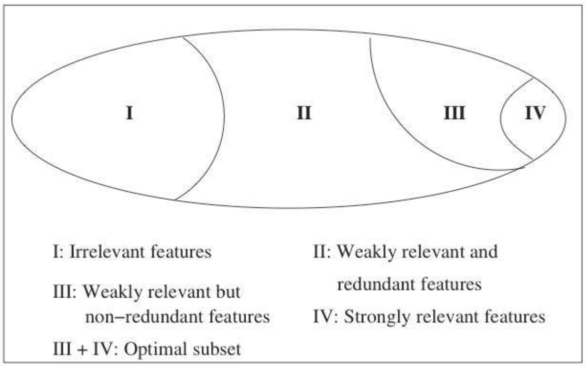


Figure 2.2: Feature categorization [90]

Attributes interaction

Since feature selection is considered as one of the effective means to remove irrelevant features, particularly, those for which the removal did not deteriorate the performance, feature removal is a very challenging task. In fact, attributes individually irrelevant (*i.e.* according to a given attribute), might become useful when combined with other features [54, 135]. In addition to the risk of useful information loss, the selection of the suitable features, which are not necessarily relevant to the problem context, but contribute, with some of the selected attributes, to enhance performance accuracy (*i.e.* classification rate), would generate another combinatorial problem. In [66], Jakulin and Bratko investigated the interactions among subsets of retained features. Indeed, a feature might loose its relevance due to the absence of interacting feature [67]. Some heuristics and criteria localizing dependency and assessing interaction levels among subsets of features, were devised. These works include the *interaction gain* [66] measure, and the Interact algorithm [135] which will be detailed in the next section.

Researchers often resort to various approximation to find relevant features [135]. It has been shown that estimating the relevance of individual features may not be difficult. However, the real challenge is to estimate the relevance of a subset of features [104]. The following section details various approaches accessing individual feature relevance as well as attribute subsets search strategies.

2.2.3 Feature selection process

Figure 2.3 illustrates the generic process that could summarizes the different steps by any feature selection apporach. As input the feature selection process requires the dataset for which the relevant features will be identified. The outcome should include the retained features as well as their performance level. Generally, such process consists of two stages: *search* and *validation*. Regardless of the nature of the feature selection approach, it should provide a search mechanism (heuristic) and an evaluation procedure. In other words, within the first stage the subset generation produces candidate feature subsets based on a certain search strategy. Next, each candidate subset is evaluted according to certain criterion and compared to the best solution found. generation and evaluation is repeated until a given stopping criterion is satisfied. The best subset resulting from the first stage is provided as input for the second stage where it is usually, validated on a different data set and event on a different evaluation criterion. Generally, the validation stage involves the use of a classifier to assess the performance of the selected features. The second stage is usually separated from the search process, to reduce the risk of a non biased selection procedure, nevertheless some approaches limit its process to the first stage [54]. It is clear that the evaluation procedure is different from the validation stage because the second stage asseses the robustness and the stability of the search stage. Besides, it is not recommanded to consider validation metrics as criteria to optimize within the first stage (*i.e.* assessment procedure) [54].

A validation protocol (*see Annex I p.* 165) derivated from the feature selection process, was developed and adopted for the empirical assessment and comparison of

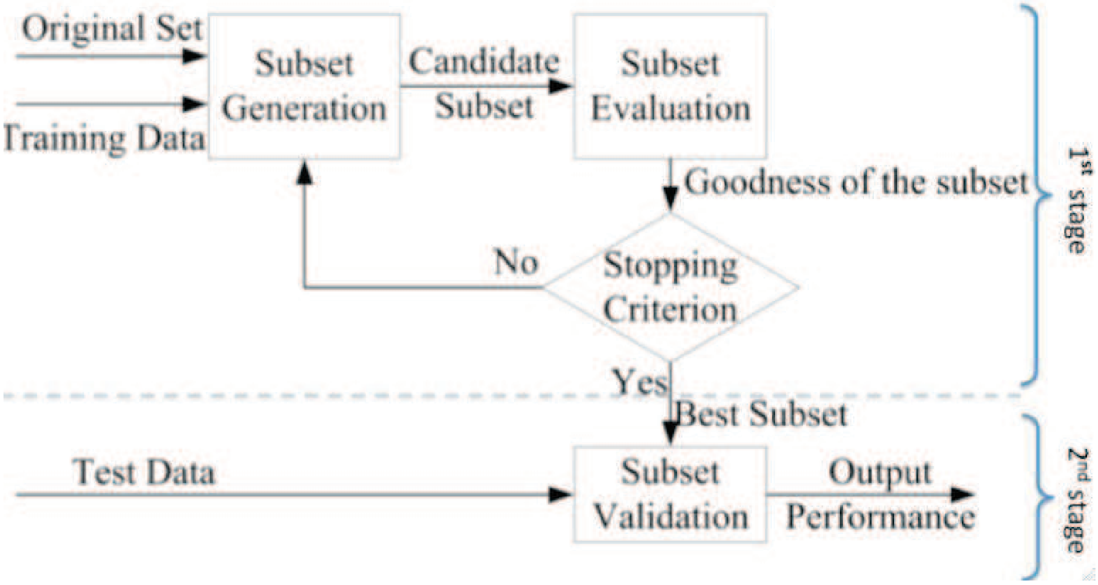


Figure 2.3: Feature selection Process [131]

the devised approaches as well as reference ones.

2.3 Feature selection modeling: state of the art

There are various ways in which the FS methods could be categorized. Aiming at the coverage of the recent advances in FS modeling, we extend the classical wrapper-filter classification, by devising a classification relying on five groups of methods: *filters*, *wrapper*, *embdded*, *hybrid*, and distributed alternatives.

2.3.1 Filters

Considered as the earliest approach to feature selection, filter methods discard irrelevant features, without any reference to a data mining technique for subsets evaluation. It applies a search based on intrinsic attribute properties and mainly its relation with the data set class [54]. Several measures were proposed to assess attribute relevance and redundancy (*i.e.* distance measures, correlation, information theory, consistency, causality, *etc*). Filter approaches apply the evaluation criterion to each attribute and a score is assigned to each of them. Typically filters sort attribute scores and, usually, return *top-K* attributes with best scores. Filters are also known as *scoring methods*.

The main advantage of the filter methods is its reduced computational complexity which is due to the independent criterion used for feature evaluation (often, evaluation complexity is in the range of $O(N * M)$ [3] and $O(M^2)$) and, the relatively simple heuristic search which is comparable to a sort algorithm complexity ranges between $O(NlogN)$ and $O(N^2)$. In most of the cases, filters rank attributes according to a predefined criterion. Nevertheless, considering one feature at a time cripple the filter to handle with either redundant or interacting features. Such limitations have paved the way to the multivariate approaches (*i.e.* wrappers, embedded alternatives [54], multivariate filters *etc*) which take into consideration subsets of features in both search and evaluation.

Distance measure filters

Filters based on distance measure use the class separability as discrimination criterion. In fact, greater difference between the two class conditional probabilities, for a given distribution, is preferable to distinguish an attribute from another. Difference between attribute distributions could be evaluated by χ^2 statistics [92]. Similarly, distribution difference could be assessed by Kolmogorov probabilities measure [54].

As a reference approach, the *Relief* remains one of the more representative algorithm, based on distance measure, used for effective feature selection. *Relief* adopts an iterative search procedure based on neighborhood and assesses features according to their discriminative power. A weight $W[i]$ is assigned to each feature. The weights should reflect the respective features ability to distinguish between the classes by comparing attribute values for similar instances belonging to different classes. In fact, the iterative process (*see Algo.* 1 Lines 3-8) involves the selection of a random instance x_k. The nearest instances from each class are selected as neighborhood (same class:

---

[3] M: number of dataset instances; N: number of attributes

nearest hit x_H and opposite class: nearest miss x_M). The weights are updated according to the merit of each feature in the differentiation of the appropriate instance against neighborhood ones (*see Algo.* 1 *Line 8* and *eq.* 2.3).

$$diff(i, x_j, x_k) = \begin{cases} \dfrac{|x_{j,i} - x_{k,i}|}{max(f_i) - min(f_i)}, & attribute \quad f_i\, is\, numerical \\ 0, \quad x_{j,i} = x_{k,i}, \quad f_i\, is\, nominal \\ 1, \quad x_{j,i} \neq x_{k,i}, \quad f_i\, is\, nominal \end{cases} \tag{2.3}$$

Features are ranked according to final weights, and negative ones could be considered as irrelevant. The more the score is high, the more the attribute is relevant. The threshold of zero can be used as a cut-off to return the proportion of relevant features. This method is not well sweeten for datasets with redundant or highly correlated features [76]. *Relief* was initially designed for binary classification problems (data sets with two classes). Some extensions have aimed to provide generalized alternative and overcome basic *Relief* limitations. *ReliefF* [78], is able to deal with both missing values and noisy data. Besides, weights ($W[i]$) update relies on a new generalized version of the $diff(.)$ function and it is able to handle multi-class problems in an effective manner.

An extension of the *ReliefF* was proposed to cope with regression problems in [112]. Another formulation of the basic *Relief* which takes into consideration context sensitivity was proposed by Hong in [59]. A detailed review of *Relief* algorithms family as well as empirical studies could be found in [113] and [90].

Relevance and Correlation measures

Correlation or dependence measures evaluate the ability to predict the value of one variable from the value of another variable [24]. Feature relevance could be measured in terms of level of the correlation between the class and a given feature.

A popular correlation coefficient in statistics, is the Pearson's coefficient (*see eq.*

Algorithm 1: Basic *Relief*

```
Input:
$F$: Initial Feature set
$C$: Target class Attribute
$m$: sampling parameter (max. iterations)
Output: $W$ : features weights
begin
    $\forall \ W[i] \in W, \quad W[i] \leftarrow 0$
    for $(l = 1 \ to \ m)$ do
        $x_k \leftarrow getRandomInstance()$
        for $(y = 1 \ to \ |C|)$ do
            $x_H \leftarrow getNearestHit(x_k)$; $x_M \leftarrow getNearestMiss(x_k)$;
            for $(i = 1 \ to \ |F|)$ do
                $W[i] \leftarrow W[i] - \frac{diff(i, x_k, x_H)}{m} + \frac{diff(i, x_k, x_M)}{m}$
    Return $(W)$
```

2.4):

$$\varrho(x, y) = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_i (x_i - \overline{x})^2 \sum_i (y_i - \overline{y})^2}} \tag{2.4}$$

where x_i and y_i respectively denote x and y values, $\overline{x}$ and $\overline{y}$ for average features values. Features might be either positively or negatively dependent. But, for coefficient around zero features are considered uncorrelated. Such a measure is usualy used with continuous features, and its generally considered as a linear dependence measure.

A Non-linear correlation measure was also proposed. They were in most of the cases based on information theory measure as *mutual information* and *symmetrical uncertainty* [57]. Symmetrical uncertainty [4] was usually used with categorical variables [96] and considered more reliable than mutual information measure [56, 57, 90, 96] and low biased for multivalued features. Detailed survey on correlation, statistical or information based relevance scoring methods could be found, in [29].

In [96], authors discussed issues in relation with significance of relevance based

---

[4] Usage, formulation and applications will be discussed on the two following paragraphs

correlation criteria in practice and argued that in statistics literature the strength (respectively weakness) of correlation could be due purely to chance. In other words, the attribute-class correlation coefficient could be interpreted in different ways.

In most of the cases, the design of filter techniques only considers, feature-class relation to assess attribute relevance or features rank, but ignores the inter-features relations. Feature independence assumptions could not be validated in practice. To overcome this problem, the same relevance criterion could be used to evaluate correlation between features to identify redundent ones.

Consistency as relevance criterion

Consistency criterion assesses how well the instances could be distinguished and not in contradiction according to a subset of feature values. The result for an instance with a given features subset is either consistent or inconsistent.

Example: Let us illustrate it by a simple example. The two following instances belong to the same dataset and the last value represents the class label. $X1 : (1, 2, 1, C2)$ and $X2 : (1, 2, 1, C1)$. $X1$ and $X2$ uses three identical feature values but they do not belong to the same class. Consequently this two instances are considered *inconsistent*.

The score associated to a given subset S by such a measure is the rate of inconsistency $IC_{rate}(S)$ among the whole dataset instances. The Subset S is said to be consistent, if the rate is below a given threshold α ($iff\ IC_{rate}(S) \leq \alpha$). The current measure differs from the other selection criteria by the following properties [24]. Contrarily to almost FS evaluation measures, this one is monotone. Consequently, the consistency rate might decrease (and never increase) by the addition of new feature to a given subset. This property, reduces considerably the FS problem search space. Besides, the complexity of subset evaluation is in the order of $\Theta(N)$ [24]. The level of noise could be assessed by the inconsistency rate. We should also note that such a measure could be only used in discrete or binary context. Furthermore, consistency criterion heavily relies on the data provided as input training data. The structure of a new set of instances (*i.e.* test and validation Datasets) belonging to the same context might generate a new inconsistency forms that were not detected with feature selection process. A set of algorithms were developed to explore search space with this criterion (*i.e.* Focus [9], ABB [91], SetCover [23]).

Focus [9] is one of the earliest algorithms, that was initially designed for binary features. It looks for the minimal set of features able to predict pure classes. FocusM [24] extends original Focus to discrete contexts. Both Focus and FocusM apply exhaustive search and guarantee optimal solutions (according to the consistency criterion used). The search strategy starts with an empty set of features and incrementally enumerates all subsets of a given size until stopping condition is met. In [24], *FocusM* algorithm was considered as only efficient with a small number of relevant features.

ABB algorithm was proposed as an application of the Branch and Bound technique to the FS problem [91]. In contrast to Focus, it starts with a full set of features, and removes one feature at a time. The early pruning of inconsistent subsets, makes the ABB more efficient with large number of relevant subset sizes.

A detailed survey of consistency based approaches could be found in [24, 83].

Information theory measures

Information theory provides a plethora of measures that characterize the variable relations. Typically, an information measure quantifies the information gain from a feature. A feature is preferred to another one if the information gain is greater. In this section, we overview basic concepts, next we present effective measures that are being used in FS.

Starting from the begining, Shanon introduced the concept of *entropy* as a measure that quantifies the amount of information in a transmitted message or the amount of information missing before message reception. $H(X)$ assesses uncertainty among X.

The definition is expressed in terms of discrete probabilities $p(.)$ (*see eq.* 2.5).

$$H(X) = -\sum_{x_i \in X} p(x_i) \log p(x_i). \tag{2.5}$$

The dependency level of two random variables could be measured by *Mutual information* (MI). MI could be expressed in terms of variable entropy (*see eq.* 2.6), where

$H(X|Y)$ is the conditional entropy which represents the uncertainty in X after knowing Y. Indeed, the amount by which the entropy of X decreases reflects addition information about X provided by Y [131]. The *Information Gain* criterion used by Quinlan, in decison trees construction [110], is equivalent to the concept of MI (*see eq.* 2.6). *Information Gain* measure is used in FS by assigning, to each attribute, a score evaluating the *attribute-class* relation.

$$I(X;Y) = H(X) - H(X|Y). \tag{2.6}$$

MI could be formulated as well for discrete contexts as for continuous ones. Mutual information for discrete attributes (*eq.* 2.7):

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \tag{2.7}$$

MI for continuous attributes (*eq.* 2.8):

$$I(X;Y) = \int_Y \int_X p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \tag{2.8}$$

where $p(x,y)$ denotes join probability (density function for continuous variables and mass function with discrete variables) [35]. In comparison with existing dependency measures, MI has the capacity of measuring any kind of relationship between variables. Another advantage, is its invariance under space transformation [35]. *Information Gain Ratio* (*see eq.* 2.9), also used as a criterion for attribute relevance evaluation. It is a variant of *information Gain* metric which is considered less biased for attribute assessment [13] and decision tree construction [110].

$$Gain_{Ratio}(X;Y) = \frac{H(X) - H(X|Y)}{H(X)}. \tag{2.9}$$

Gain Ratio normalizes the MI gain by dividing it by the entropy of X. Another well known measure normalizes the MI gain by using both entropies in the denominator: *Symmetrical Uncertainty* SU. As a result the SU coefficient is in the range of [0..1].

$$SU(X;Y) = 2\frac{H(X) - H(X|Y)}{H(X) + H(Y)}. \tag{2.10}$$

The measure is widely used and considered as a robust measure for attribute ranking [54, 57, 96, 130, 131]. Besides, SU criterion was used by *Huan Liu* for the *Blanket Markov* approximation and redundancy removal [130, 131].

Multivariate filters

Almost all criteria, described above, assume the independence of the features with the respect of the class, simply because they evaluate the quality of the feature independently of the context of the other features that could be selected. The major problem is feature redundancy. Two features could be highly relevant, and useful for the class but, also, highly correlated, redundant for the retained subset of features or/and share the same informational content. On the other hand, features individually irrelevant could behave differently, and contribute to improve considerably classification accuracy, when they are together. Such features would not have a great chance to be selected with one of the above criteria, simply, because only individual intrinsic attribute properties are considered. Multivariate filters, tries to overcome this limitation by endowing FS process with an additional evaluation criterion. In other words, multivariate approaches compete with wrappers (*see section* 2.3.2) by enhancing their abilities in subset selection and evaluation.

$$CFS_s = \frac{|S|\overline{r_{cf}}}{\sqrt{|S| + |S|(|S|-1)\overline{r_s}}}. \tag{2.11}$$

Mark Hall [57] was among the first who introduced multivariate filters. In fact, he formulated a selection assessment criterion implementing a trade-off between redundancy and relevance. Consequently, the CFS criterion has the ability to evaluate subsets of features as in the wrapper approaches. In the numerator of the equation 2.11 computes the mean correlation between individual selected feature and the class ($\overline{r_{cf}}$). It gives an estimation of the relevance of the selected attributes, whereas the denominator returns an expression based on mean features pairs correlation[5]. The SU measure was used to assess feature dependency and relevance [57]. The fact that the criterion is based on average correlation could penalize features that are individually low correlated with class and relevant together (*i.e.* interacting features).

Recently, similar filter measures that aimed at subset evaluation were proposed. For example, Ooi *et al.* [104] introduced the concept of *differential vaporization* which is illustrated by equation 2.12:

$$DP_s = (\overline{r_{cf}})^{\alpha}.(U_s)^{1-\alpha} \tag{2.12}$$

The first member measures the average of feature-class pairs correlations, while the second corresponds to the *anti-redundancy measure* (*see eq.* 2.13).

$$U_s = \frac{1}{|S|^2} \sum_{i,j \in S, i \neq j} 1 - |r_{i,j}| \tag{2.13}$$

Battiti in [15], tried to select k relevant features from an initial set of N features and proposed a greedy search (*see Algo.* 2). MIFS algorithm starts with an empty set, and selects the best attributes according to the MI criterion. Next, attributes are selected with two criteria:

- relevance: MI_i
- redundancy: $\sum_{f_s \in S} MI_{i,s}$

The attribute f_i^* that maximizes $(MI_i - \beta \sum_{f_s \in S} MI_{i,s})$ is added to S. In fact, the selection process reflects a relative trade-off between relevance and redundancy. The user defined parameter β, regulates the impact of redundancy on the selection criteria. MIFS-U applies the same incremental selection procedure using an enhanced selection criterion (*see eq.* 2.14). In other words, the same process is applied with a different selection rule (9^{th} instruction).

$$\forall f_i \in F : maximize[I(f_i;C) - \beta \sum_{f_s \in S} \frac{I(f_s;C)}{H(f_s)} I(f_s;f_i)] \tag{2.14}$$

---

[5] all subset pairs are considered. Besides $|S|$ refers to the number of selected attributes

Algorithm 2: MIFS: Mutual Information based Feature Selection

```
Input:
$F$: Initial Feature set
$C$:target class Attribute
$\beta$: regulation parameter
$K$ : number of attribute to select
Output:
$S$ : Selected Features
begin
    $S=\emptyset$
    $\forall f_i \in F$, $MI_i \longleftarrow I(C; f_i)$ /* Computing mutual information */
    $i^* \longleftarrow argmax_i(MI_i)$
    $S \longleftarrow S \cup \{f_i^*\}$
    $F \longleftarrow F \setminus \{f_i^*\}$
    while ($|S| < K$) do
        $\forall f_s \in S, \forall f_i \in F$ , $MI_{i,s} \longleftarrow I(f_i, f_s)$
        $i^* \longleftarrow argmax_i(MI_i - \beta \sum_{f_s \in S} MI_{i,s})$
        $S \longleftarrow S \cup \{f_{i^*}\}$
        $F \longleftarrow F \setminus \{f_{i^*}\}$
    Return ($S$)
```

According to the [35] MIFS-U criterion makes a better estimation of MI between input attributes and the class, than MIFS. The Famous Max-Relevance and Min-Redundancy algorithm (MRMR) [27, 106] suggests a penalization expression based on proportionality between redundancy rate and the subset size (*see eq.* 2.15).

$$\forall f_i \in F : maximize[I(f_i;C) - \frac{1}{|S|}\sum_{f_s \in S} I(f_s;f_i)] \tag{2.15}$$

NMIFS [35] proposed another variant of the MIFS criterion by normalizing the penalization term (*see eq.* 2.16).

$$\forall f_i \in F : maximize[I(f_i;C) - \frac{1}{|S|}\sum_{f_s \in S} \frac{I(f_s;f_i)}{min(H(f_i),H(f_s))}] \tag{2.16}$$

Readers could find in [35], a recent and detailed review of FS methods based on information theory criteria.

Huan Liu proposed a similar approach ***FCBF*** [130] which is based on the *S.U.* criterion which is used as well as relevance measure as criterion for removing redundancy. The algorithm includes two stages (*see Algo.* 3). The first selects the best correlated features to the class. Such features are considered as *predominant*.

Within the second stage, redundant features are discarded according to the following condition ($SU(f_p, f_q) \geq SU(f_q, C)$). Indeed, a feature is considered as redundant when it is more dependent to the current predominant feature than the class. Although, the relative low complexity and the multivariate nature of the approach, the algorithm could be seen as a hill climbing that discards redundancy through the search. Consequently, FCBF could be trapped in a local minima.

The same author devised another multivariate filter in [135]. The approach was maned ***Interact*** algorithm. It looks for interacting features among features sorted using SU criterion. It could be considered as a backward elimination strategy since it removes features that have low interaction potential with retained features.

Alternative multivariate approaches for selecting relevant and not redundant attributed could be found in [35, 82, 88].

Algorithm 3: FCBF: A Fast Correlation-Based Filter algorithm [130]

```
Input:
$F$: Initial Feature set
$C$: Target class Attribute
$\beta$: Threshold
Output:
$S$ : Selected Features
begin
    $S \leftarrow \emptyset$
    foreach $(f_i \in F)$ do
        Calculate $SU(f_i, C)$
        if $(SU(f_i, C) \geq \beta)$ then
            $S_{list}.append(f_i)$
    Sort($S_{list}, DESC_{order}$)
    int $p \leftarrow 0, q \leftarrow 0$
    while $(p < |S_{list}|)$ do
        $q \leftarrow p + 1$
        $f_p \leftarrow S_{list}.getElementAt(p)$
        while $(q < |S_{list}|)$ do
            $f_q \leftarrow S_{list}.getElementAt(q)$
            if $(SU(f_p, f_q) \geq SU(f_q, C))$ then
                $S_{list}.removeElementAt(q)$
            else
                $q + +$
        $p + +$
    $S \leftarrow S_{list}$
    Return $(S)$
```

2.3.2 Wrapper methods

When feature selection process is based on a wrapper, attributes are not evaluated individually, but together and the search is not limited to feature scores sorting, and the selection of *top-k* attributes. The resulting search space covers the $2^n - 1$ possible combinations of features subsets for a data set of n attributes. The exploration of the feature space is driven by both classification accuracy returned by the selected subset of features and the involved search technique. Typically, a classifier is used as a part of the evaluation process by awarding the retained subsets according to its predictive performance. The wrapper methods often provide better results than filter ones because they are tuned to the specific interaction between an induction algorithm (classifier) and its training data [54, 90]. The computational cost of wrapper is induced by the feature subset evaluation stage. It involves building and evaluation of a classification model for each attributes subset [56, 132]. Figure 2.4 illustrates evaluation process and shows how the classifier is involved in solution assessment. The process starts by the generation of a dataset represnting only selected features. Next, the classifier builds a model trying to predict, for each data set instance, the class from features values. The model is built on a training data. Once the learning process terminates, the evaluation procedures starts with a new data set (test data).

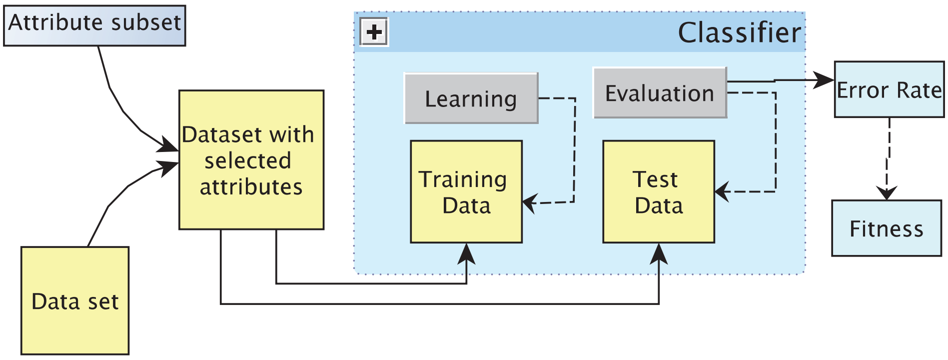


Figure 2.4: Classifiers and attribute subset evaluation

The error rate corresponds to proportion of instances correctly classified (predicted value equal to class value). The feature selection problem could be formulated as a maximization problem when we consider classification rate and as a minimization problem with error rate. The relation between classification and error rate is formulated by the following equation 2.17:

$$Classif_{rate} = 1 - Error_{rate} \tag{2.17}$$

Kohavi *et al.* [68] were the first to advocate the wrapper as a general framework for feature selection in machine learning. Numerous studies have adopted the above methodology by either changing the classifier or the search technique (*i.e.* greedy, randomized, stochastic methods *etc*). Nevertheless with wrappers the risk of overfitting [6] in classification [7] is higher than with filter approaches.

Feature selection methods based on wrappers are more computationally expensive than filters. It is due to the cost of iterative running of the classification algorithm for solutions evaluation and the number of solutions to examine within the search space [90]. For a data set of N attributes, the resulting search space is made of $2^N - 1$ subsets. Aiming to reduce the number of evaluations and to enhance search reliability some heuristic strategies (*i.e.* Tabu search, Genetic algorithms, Simulated annealing, *etc*) have been investigated to address the problem of finding the best subsets of features [53, 54, 90]. The following sections review and categorize the wrapper plethora techniques.

Sequential wrappers

In this section, we review basic sequential wrappers, combined alternatives, and recent advances for greedy methods. Sequential Forward Selection (SFS) and Sequential Backward Elimination (SBE) are two well known basic sequential approaches recognized as "bottom up" and "top down" methods. SFS starts, as shown in Algorithm 4, with an empty set S, and the most relevant feature (from the non-selected ones)

---

[6] Learning from noise

[7] Supervised learning algorithms

is added to S. The iterative process adds the attributes which maximizes the classification accuracy, until the desired number of attributes is reached or classification performances could not be improved. The SBE uses the same sequential process (*see Algo.* 5) but it starts with full feature set, and the less relevant feature (according to the currently selected subset) is removed at each iteration.

We should note that for the sake of simplicity we opt for the below equivalence notation (*see eq.* 2.18) to facilitate the readability of solution comparison. Here, fitness denotes the classification performance of the classifier used in solutions assessment.

$$S_A \succ S_B \Leftrightarrow \begin{cases} S_A.fitness \geq S_B.fitness, & \text{(lower error rate)} \\ & \vee \quad (or) \\ \|S_A\| < \|S_B\|, & if(S_A.fitness = S_B.fitness) \end{cases} \tag{2.18}$$

It is clear that the search is not exhaustive (*i.e* greedy methods), and the number of explored solutions is in the order of $\Theta(N^2)$. Both SFS and SBE apply the hill climbing procedure. In fact, the solutions explored at a given iteration represent the neighborhood of the current solution. The iterative neighborhood exploration and the stopping criterion made SFS and SBE, two local search procedures that could be easily trapped in local minima. With typical subset evaluator, SFS is faster than SBE because it evaluates smaller sets than SBE [54]. Indeed, once a feature is selected with SFS (*resp.* removed with SBE) it could not be removed (*resp.* added for SBE) during the search. The search is, hence, biased. The problem is well known in FS as the *"nesting effect"* [54, 120]. Multivariate filters, also, suffer from the nesting effect. Attempts to alleviate the nesting effect let to the so called *"Plus-l-Minus-r"* method [79]. Here, rather that adding or removing one feature at a time, in each iteration l features are added then r features discarded ($l \geqslant r$ for the forward procedure). By doing so, all attributes could be added or/and removed during the search. This method requires the definition of two additional parameters l and r. In comparison with basic sequential approaches, *"Plus-l-Minus-r"* method requires an additional computational effort. Some studies considered the search in the opposite direction to overcome of the problem of nesting effect as form of backtracking [54].

Algorithm 4: S.F.S. : Sequential Forward Search

```
Input:
$F$: Initial Feature set
$C$: Target class Attribute
$Cla$: a classifier for solution evaluation
Output:
$S_{best}$ : Final Feature set
begin
    $S \leftarrow \emptyset$ , $S_{best} \leftarrow S$
    Stop $\leftarrow false$
    repeat
        $Sol_{list} \leftarrow \{X, \forall f_i \in F, f_i \notin S, X = S \bigcup \{f_i\}\}$
        foreach ($X \in Sol_{list}$) do
            $Evaluate(X, Cla)$
        $S \leftarrow getBest(Sol_{list})$
        if ($S \succ S_{best}$) then
            $S_{best} \leftarrow S$
        else
            Stop $\leftarrow true$
    until ($Stop = true$);
    Return ($S_{best}$)
```

Pudil et al. [109], proposed one of the most effective sequential search methods for the FS problem [65]: the sequential forward floating Search SFFS and SBFS for the backword search.

In fact, the floating search combines forward and backward search without the need to specify any parameter. The forward floating search, starts with empty set and applies the selection procedure of the sequential forward search. Between two sequential forward iterations, a sequential backward iteration is conditionally applied. The backward procedure is taken into consideration only if it contributes to improve the current solution. The trade-off between the forward/backward iterations is set dynamically during the search. Algorithm 6 illustrates forward floating search (SFFS) approach as well as the conditional application of the backward search.

Algorithm 5: S.B.E. : Sequential Backward Elimination

```
Input:
$F$: Initial Feature set
$C$: Target class Attribute
$Cla$: a classifier for solution evaluation
Output:
$S_{best}$ : Final Feature set
begin
    $S \leftarrow F$ , $S_{best} \leftarrow S$
    Stop $\leftarrow false$
    repeat
        $Sol_{list} \leftarrow \{X, \forall f_i \in S, X = S \setminus \{f_i\}\}$
        foreach ($X \in Sol_{list}$) do
            $Evaluate(X, Cla)$
        $S \leftarrow getBest(Sol_{list})$
        if ($S \succ S_{best}$) then
            $S_{best} \leftarrow S$
        else
            Stop $\leftarrow true$
    until ($Stop = true$);
    Return ($S_{best}$)
```

Recently, a new improvement of the basic SFFS, was proposed in [101]. The improved forward floating selection search (IFFS) not only backtracks but also tries to replace weak feature in the current set. Once the forward step terminates, we firstly, attempt to improve solution with a backward stage. If the backward procedure fails, the second improvement attempt replaces the weak feature with one of the remaining features. In the case where the replacement enhances the solution, the backward procedure is applied again to the new solution, otherwise we restart the cycle with the forward procedure. In comparison to the basic floating search, this variant tries to overcome nesting effect and to escape the local minima by diversifying the search around the current solution with different neighborhood structures (*i.e.* forward backward procedures, features replacement, add/remove of more than one feature). The local minima risk is reduced, but not completely discarded.

Several enhancements of the basic sequential search were also proposed in [52].

Wrappers based on heuristics

In this section, we explore, in depth, another active research field of the FS modeling, namely the combinatorial optimization nature of the problem. Surveyed approaches fall into three classes: local search based methods, evolutionary and swarm approaches.

Heuristics based on Local search

The above presented sequential approaches approaches could, be also, considered as local search procedures. Nevertheless, the approaches presented here belong to the stochastic optimization wheareas the above presented ones were deterministic.

A common and important aspect that is shared with the previous apporaches is the notion of *neighborhood*. In fact, the search starts by defining a set of solutions around the current one: *neighborhood*. Then, the search evolves iteratively, by selecting the best solution in the neighborhood.

Simulated Annealing (SA): SA is a stochastic approach which is based on statistical thermodynamics for finding near optimal equilibrium. SA simulates the energy

Algorithm 6: S.F.F.S.: Sequential Forward Floating Search

```
Input:
$F$: Initial Feature set
$C$: Target class Attribute
$Cla$: a classifier for solution evaluation
Output:
$S_{best}$: Final Feature set
begin
    $S1 \leftarrow \emptyset$ , $S_{best} \leftarrow S1$
    Stop $\leftarrow false$
    repeat
        $Sol_{list} \leftarrow \{X, \forall f_i \in F, f_i \notin S1, X = S \bigcup \{f_i\}\}$
        $\forall X \in Sol_{list}, Evaluate(X, Cla)$
        $S1 \leftarrow getBest(Sol_{list})$
        $Sol_{list} \leftarrow \{X, \forall f_i \in S1, X = S \setminus \{f_i\}\}$
        $\forall X \in Sol_{list}, Evaluate(X, Cla)$
        $S2 \leftarrow getBest(Sol_{list})$
        if $(S2 \succ S1)$ then
            $S1 \leftarrow S2$
        if $(S1 \succ S_{best})$ then
            $S_{best} \leftarrow S1$
        else
            Stop $\leftarrow true$
    until $(Stop = true)$;
    Return $(S_{best})$
```

changes in a system subject to a cooling process until ti converges to an equilibrium state. The temperature is used to reflect the tolerance degree of the non improving solution during the search. The SA [75] search procedure starts with random solution (the features are randomly selected). The search process generates a neighborhood from the current solution. The solutions are evaluated and a fitness (subset merit that is usually assessed in terms of classification accuracy) is assigned to each solution. If the best solution of the neighborhood is better than the current one, the current solution is replaced by the best. When the result of neighborhood exploration could not improve the current solution, the value of the temperature which decreases during the search, decides whether the solution (best among neighborhood) is accepted to replace the current one. In fact, the temperature parameter is associated to a probability of acceptance of non improving solution. With high temperature such a solution is more likely to be accepted than in low temperature. Thus, the search will not get stuck in local minima at the beginning of the search. According to Liu *et al* [93], SA was applied to the FS problem in 1992 by Doak [28].

In [86], the SA was used for both feature selection and neural network structure (topology) optimization. Another application of SA was proposed by [97], where feature selection was applied to marketing data to build large-scale regression model. A hybrid and cooperative FS method using SA and Genetic algorithms was proposed in [43]. SA was also proposed for embedded [8] feature selection [85], where the feature selection is done and optimized within the classification process.

Tabu Search (TS): Proposed in 1989 by Glover [31, 44], as an effective local search that was able to escape local minima by going beyond. This optimization scheme applies an iterative local search based on classical neighborhood exploration. The specificity of TS is the ability to exchange the current solution with one that is less fittest when the search is trapped in a local minima. TS is endowed with a list that saves informations about the solutions recently visited. The list is called tabu because it prohibits backtracking to all of its solutions and hence prevents cycling. The study of the tabu list developed the concept of memory or adaptive memory which guides the search. It has received a wide spread attention since the introduction of TS.

---

[8] Embedded feature selection will be detail in section 2.3.3 (*p.* 44)

In [134], Zhang and Sun proposed a wrapper based on TS. It can be viewed as a direct application of the basic TS, where the tabu list saves best solutions resulting from the recent neighborhood explorations. The neighborhood, is generated by adding or removing one feature to/from the current solution. The tabu list is used as short term memory. The list is of a fixed size l. The TS was compared to several deterministic sequential schema. It outperformed genetic algorithms on small and medium sized problems. The fitness function, which evaluates the feature subset accuracy, is based on the aggregation of two objectives: classification accuracy and a penalization term in direct relation with subset size. In [123], The TS was extended to the optimization of both input features subset and the classifier. It used also a short term memory.

As previously stated the short term memory limited by the size of the tabu list did not definitely relieve the cycling problem, but imposes the use of an additional parameter that should be carefully defined (tabu list size). The same point of view was shared by Wang *et al.* [128] which, recently, proposed a Tabu search using long term memory for a small FS benchmarks. In fact, it avoids the problem of search cycling around the same local optima, and the empirical tuning of the size of the tabu list, by using a long term memory implemented as a linked list. Such a memory is not only limited to direct and recent neighbors but also to solutions with significant changes from the current one.

Greedy Randomized Adaptive Search Procedure (GRASP):

GRASP [37, 38] is recent optimization scheme applying an iterative local search process based on neighborhood exploration. GRASP relies on a multistart schema to diversify the search and escape local minima. The application to the FS problem was, recently done by Yusta in [133], where the GRASP was compared to genetic, memetic, and Tabu search. The results of this empirical study showed the superiority of GRASP. It is based on two main stages, namely solution construction and local search procedure using the neighborhood structure ($NH(.)$) defined by the equation 2.19. It explores all possible attribute permutations. The construction phase generates a solution that will be improved in the second step by an iterative local search procedure. The result of a given GRASP iteration, is a solution that reached a local minima. The process restarts with the construction of a new solution. This stage is based on a guided random solution generation. In fact, it requires the construction of a restricted candidates list (RCL) containing the most accurate attributes. Once the RCL generated, several solutions are randomly constructed and the best one goes through the second stage.

$$NH(S) = \{X | X = S \cup \{f_i\} \setminus \{f_j\}, \forall f_i \in X, \forall f_j \notin X\} \tag{2.19}$$

Evolutionary strategies

This section is devoted to the study of a well known stochastic optimization scheme that is mainly based on populations methods. Such heuristics are generally known as population based methods, because the optimization process does not rely on one current solution but on a set of candidate solutions called *population*. Indeed, the search is not limited to a particular region of the search space but extended to a more wide space covered by the population solutions.

Three techniques will be presented here, namely, *genetic algorithm*, *differential evolution*, and *estimation of distribution algorithms*.

Genetic Algorithms (GA)

GA [46] is considered as one of the leading stochastic optimization schema, reputed to be one of the most robust meta heuristics dealing with *np-hard* problems [31]. GA tries to make the analogy between the natural evolution and the optimization process using a set of concepts like selection and crossover, mutation. Siedlecki and Sklansky [119] was the first that have used GA as a feature subset selector.

Several papers and comparative studies consider that wrappers based on GA are among the most interesting approaches for tackling high dimensionalities and suggest its use in FS modeling [43, 62, 79, 99, 102, 119]. Algorithm 7 details the basic steps of a commonly used GA in FS modeling. In such approaches, a given feature subset represeinting a solution is coded in a binary string (chromosome) of length N (total number of features). Zero or one are possible values, respectively denoting the absence or the presence of the attributes at the i^{th} position. A fitness, reflecting the classification accuracy of the solution is assigned to each solution. The process starts by the random generation of the initial population of solutions. The population evolves through generations by replacing less fittest solutions by enhanced ones. The evolution process iteratively applies a set of operators to select solutions, to derive new ones from the combination of existing ones and to explore new regions of the search space. The evolution process continues until a convergence criterion is met or a given max iterations number is reached. In this section, we only provide an introduction of the genetic algorithms limited to main evolution concepts. A detailed description of both concepts and associated technical aspects will be adressed in the next chapter.

Several works have used GA in FS modeling either as specific optimization schema based on GA or as reference method for comparison and assessment of the proposed FS method [8, 54, 62, 65, 79, 90, 133, 134, 140].

In the comparative study conducted by *Kudo and Sklansky* in [79], a set of sequential approaches were compared to GA and authors conclude that the floating approaches and GA outperforms sequential approaches but sometimes GA found better solution. Besides the authors suggest the use of genetic algorithm for problem dimensions exceeding 100 features. Nevertheless, earlier works, [65] that had confirmed comparable efficiency between GA and SFFS, stated that GA becomes worse than SFFS as the dimensionality increases and reported a GA tendency to premature convergence. *Emmanouilidis et al.* [30] proposed an adapted Crossover operator and it claims that *Subset Size Oriented Common Feature Subset Crossover operator* (SSOCF) helps to preserve building blocks with promising performance.

The authors stated, that this procedure allows more flexible neighborhood exploration than sequential search methods. The operator was also applied by the following evolutionary FS approaches [42, 124].

Multi-Objective GA

Different formulations and evaluation criteria of the FS problem were devised. The introduction of new criteria in fitness validation have made the FS a multi-objective problem. Emmanouilidis *et al.* [30], consider that feature selection problem is well suited to multi objective optimization. The simplest form involves two objectives: minimization of the number of features and the maximization of the classification

Algorithm 7: A Genetic Algorithm for FS

```
Input:
$Size$: population size; $Cla$: Classifier;
$p_{mut}$: mutation probability; $p_{cross}$: crossover probability
$Maxgen$: Total number of iterations;
D: Dataset
Output: $S'$ : Population of the last generation
begin
    $S_0 \leftarrow$generateInitialSolutionSet($Size$)
    $P \leftarrow S_0$, P$_{tmp} \leftarrow \emptyset$, $i \leftarrow 0$
    while ($i < Maxgen$) do
        P$_{tmp}$ $\leftarrow$Select (P)
        Crossover(P$_{tmp}$, $p_{cross}$)
        Mutate(P$_{tmp}$, $p_{mut}$)
        Evaluate(P$_{tmp}$, $Cla$, D)
        Replace(P$_{tmp}$, P)
        $i \leftarrow i + 1$
    Return ($S'$$\leftarrow$P)
```

accuracy. Kudo and Sklansky [79], suggest the use of a penalty term with classification accuracy (*see eq.* 2.20).

$$fitness(X) = J(X) - \epsilon|X| \tag{2.20}$$

where X, $|X|$ and $J(X)$ respectively denote feature subset to be evaluated, the number of features, and the classification accuracy $(1 - ErrorRate)$. The parameter ϵ is defined by the equation 2.21 as follows:

$$\epsilon = \beta + \frac{(J_{max} - J_{min})}{N} \tag{2.21}$$

J_{max} and J_{min} correspond to the estimation of the lower and upper bound of the classification function. N is the total number of attributes.

In [35], a more simple formulation was proposed using a penalty term and a control parameter λ (*see eq.* 2.22):

$$fitness(X) = J(X) - \lambda \frac{|X|}{N} \tag{2.22}$$

where N is the number of features. λ is a parameter set to the range of [0..1], to control the trade-off between the subset size and accuracy. *Freitas* discussed, in [41], the diversification the GA objective function by adding a filter oriented criterion. In such formulation, the GA has both aspects of wrapper and filter approaches. He reported the fitness function proposed by Bala *et al.* [14] (*see eq.* 2.23):

$$Fitness(X) = \underbrace{Info(X)}_{filter} - \underbrace{|X|}_{cardinality} + \underbrace{J(X)}_{classif.\ \ rate} \tag{2.23}$$

where $Info(X)$ is a filter criterion estimating discriminatory power of X attributes.

We think that, in any case, the three involved criteria are not comparable because ranges are different. Besides, normalization, the filter criterion could be used more effectively when it reflects attribute properties that could not be assessed by a wrapper (*i.e.* attribute dependency, redundancy, *etc*).

Recently, a new multi-objective formulation was proposed in [42], where more specific classification criteria were involved. In fact, assessment procedure does not only rely on global error rate but it is extended to error measures based on *specificity* (*see eq.* 2.25) and *sensitivity* (*see eq.* 2.26). For a classification problem with two classes (positive and negative labels), classification prediction might generate a *true positive* (respectively true negative) with successful classification, or inverted class labels with a wrong prediction (false positive/false negative). The following Table illustrates possible prediction combinations.

		Data class label	
		Positive	Negative
Prediction	Positive	true positive (TP)	false positive (FP)
	Negative	false negative (FN)	true negative (TN)

Table 2.2: Classification outcomes (confusion matrix)

Sensitivity, returns the proportion of the successfully classified instances from the first class (positive class label), whereas *specificity* tests how well the classification model identifies instances of the second class (instances with negative class labels).

$$f_1(X) = \frac{|X|}{N} \tag{2.24}$$

$$f_2(X) = \frac{\#TP}{\#TP + \#FN} \tag{2.25}$$

$$f_3(X) = \frac{\#TN}{\#FP + \#TN} \tag{2.26}$$

The proposed multi-objective scheme uses a vector evaluation function $F(s_i) = \langle f_1(s_i), f_2(s_i), f_3(s_i) \rangle$ (*see eqs.* 2.24, 2.25 and 2.26). Solution are compared in terms of *pareto dominance*. A solution is dominant over another one only if it has better performance in at least one criterion and non-inferior performance with the remaining criteria. We should note that multi-objective formulation is not limited to fitness objective aggregation, but requires the adaptation of the evolution operators to make them able to handle different objectives.

A summary of the several fitness function used in multi-objective context, could be found in [41].

Weighting methods: Attribute selection could be seen as a particular case of attribute weighting. Two values are possible (0 and 1) with binary string encoding for solution representation. Attribute weighting assigns to each feature a weight in the range of [0..1]. Features are considered as selected only if its associated weighted is above a given threshold. Evolutionary approaches that have adopted weighted representation could be found in [41, 63, 107].

Differential Evolution (DE)

Differential Evolution [121] (DE) is a population based approach like GA, and applies similar operators (*i.e.* crossover and mutation). The main difference is that GA relies on crossover as an intensification mechanism, while DE uses a specific mutation scheme. This main operator is based on the difference between two random population solutions. DE is able to add the difference to a third member and hence, generate new solution (*see eq.* 2.27).

$$\widehat{x_i} = x_{r1} + f * (x_{r2} - x_{r3}) \tag{2.27}$$

where $\widehat{x_i}$ denotes the resulting solution and $r1 \neq r2 \neq r3$ three distinct random indexes of population members. f is a scaling factor that controls the rate at which the population evolves. The iterative process starts, after the random generation of the initial population, by the mutation operator followed by a uniform crossover [31] between the current solution and the mutation result. Once the new solution evaluated, the offspring is compared to the current one and the less fittest solution is replaced. Besides, all the solutions have a chance to be selected without any reference to fitness. In [73], Khushaba *et al.* proposed an adaptation of the original DE, (initially devised for continuous contexts), to the FS problem, by the use of non-binary solution representation. The solution is of a fixed length[9] and encodes features indexes [10]. Redundant indexes are replaced with a specific wheel selection mechanism. Empirical results stated improvement over both GA and PSO (Particle Swarm Optimizer). The empirical soundness of the approach was pointed out on different datasets [8]. A hybrid alternative based on *Ant Colony Optimization* and DE was also proposed in [74].

Estimation of Distribution Algorithms (EDA)

EDA [11, 81] is a recent evolutionary paradigm that is considered as attractive alternative to GA [95]. The method is based on the generation of an initial population. Next, a number of solution are selected to form a sample. Then a population model estimating the distribution of the selected individuals generates a new population by sampling the estimated distribution. The process iterates until convergence. In fact, EDA builds a probabilistic model to learn from explored solutions and guides the search process.

Several applications of the EDA in FS with encouraging results in gene selection could be found in the following studies [11, 17, 64, 116].

---

[9] Fixed number of selected attributes

[10] An array of selected attributes (indexes)

Swarm approaches

Swarm intelligence is an innovate distributed intelligent optimization paradigm that took its inspiration from social behaviors by swarming, flocking and herding phenomena [4]. In fact, the swarm techniques, imitate foraging behavior for real ants in *Ant Colony Optimization* (ACO), and swarming behaviors of schools of fish, bees colony, or even social human behaviors. Such an optimization paradigm is based on a fundamental concept: *the stigmergy* which is illustrated by collective behavior and implicit or even explicit communication of optimization components (particles or ants) through environment. Collective behavior, distributed nature, locality and stigmergy make swarm approaches a real attractive alternative to classical evolutionary optimization.

Particle Swarm optimization (PSO)

PSO [72] is a population based search technique. The population is made of random solutions called *particles*. Each particle flies over the search space with specific *velocities*. When they move, respective positions (eash position correspond to a solution) change and they try to find out better positions by following leading particles and its own experience. To that end, particles iteratively adjust its velocities according to both swarm (best among neigborhood) and personal behaviors. Once velocities are adjusted, solutions (positions) are updated according to new velocities. Details about technical aspects and algorithmic issues of the PSO process as well as velocity updating rules will be provided in chapter 5. The application of PSO to the FS problem requires the use of the Binary PSO variant (BPSO) [72] and velocities were mapped into boolean values using a transformation function (*i.e.* logistic regression function).

Firpi et al. introduced the swarm FS [40] and its performances were compared to GA. Another comparison of PSO with GA was done in [12], where the FS has been applied to, a relatively small sized problem (23 features). The classification accuracy of SVM (Support Vector Machines) was improved in comparison to genetic wrapper (GA) and SVM classifier without FS. An improved BPSO (IBPSO) [21] was applied to high dimensional gene expression data ($\#features > 2000$). The IBPSO tries to escape local minima by reseting the values of best solution when it was not improved after a given number of iterations. By doing so, the velocity updates would only rely on particle experience. Then the collective behavior would generate another optimum.

Al-Ani compared in [8], PSO with both evolutionary GA and DE schema. PSO was outperformed by both evolutionary schema with small ($\#features < 50$) and medium sized problems ($\#features > 50$).

Ant Colony Optimization (ACO)

ACO approaches focus on the ability of ants to find shortest paths from nest to food. ACO models the social behavior of ants in both information sharing and decision making. They are, also based on population scheme, representing a colony of ants. Ants construct its solutions in a n incremental way, and leave a chemical pheromone on the ground to remember the trail on the next iterations. The collective behavior makes the ants the ability to adjust their paths according to the available pheromone concentration. Since ants are guided by pheromone smell, indirect communication enable them to find short paths.

In opposition to the PSO paradigm, the ACO are well sweeten for FS, since, ACO was originally designed for combinatorial optimization problems and adopts solution construction scheme. In fact, features are represented by a network of nodes and the ants try to find suitable paths. The ant behavior would be comparable to a FS sequential algorithm where it starts with a given feature and constructs solution by adding attributes (*i.e.* each time the ant selects the next node). In [7], an ACO approach was proposed to tackle FS problem. The iterative swarm process starts by the selection of random starting point for each ant (initial feature added to the solution subset), and then uses pheromone to guide network exploration.

In [5], the ACO was proposed to tackle the FS problem in text categorization where subset size was taken into account with classification accuracy in the pheromone update stage. The selection of the next feature to add to the subset which was materialized with ant move, used the classification accuracy of the subset.

2.3.3 Embedded methods

Since wrapper methods employ a heuristic search guided by the accuracy of the classification method, embedded methods use of the classification process learning itself to both perform feature selection and construct an optimized classifier. In fact, the ability of some classifiers to discard irrelevant input features, during the learning process, has been exploited in several researches and especially in bioinformatics (*i.e.* decision trees, Support Vector Machines: SVM).

Considered as one of the prominent approaches to FS [54], *recursive features elimination* (RFE-SVM) [53] extends the basic SVM classification scheme to a classifier endowed with a FS ability. RFE iteratively removes least relevant feature and re-estimates the resulting classifier on the remaining features. RFE procedure tries to select $n < N$ that lead to largest margin in class separation.

This combinatorial problem was solved by a greedy procedure that iteratively removes the feature which minimizes the margin decrease of *hyperplane* classification boundaries. The procedure could be accelerated by removing more than one feature in each iteration. According to [54] and [90], RFE has shown good performance on micro-array data and gene selection problems. The combinatorial problem of embedded selection of subsets of n features was also addressed with meta-heuristics. Examples include the use of SA [85] and GA [61, 125] to optimize SVM parameters.

An alternative embedded method replacing SVM classifier by Random Forest (RF) classifier [19] and extending the binary classification (problem with 2 classes) ability of the initial RFE, to multi-class problems was proposed by Granitto *et al.* [48, 49, 50]. Reported results [49], pointed out superiority of RF-RFE over SVM-RFE.

The advantage of embedded RFE based methods is its effective classification accuracy and moderate computational cost compared to wrapper scheme based on heuristics. However the gain in classification for a given classifier is not guaranteed for a different classifier with the selected features. Such methods are known to be effective for the targeted classifier. Numerous classification schema based on parameters optimization and input pruning during the learning process were proposed. For further details readers could be redirected to the following references [13, 86, 126].

2.3.4 Hybrid and boosted approaches

Recently, numerous studies started to pay more attention to some of the complementary aspects of feature selection. The motivation to a such orientation is the exhibited multi-disciplinary of the FS problem property and the lack of clear adapted methodology for the search space exploration. The majority of the new FS alternatives that are being proposed are hybrid approaches [54, 90]. In fact, numerous approaches consider more than one aspect and aim to design reliable and accurate (*i.e.* unbiased) models [87, 127].

Direct combination: filter then wrapper

The simplest form of combination is to use both filters and wrappers. The common scheme of combination entails two steps. The first one applies a filter to reduce the number of attributes, and hence, the search space. The second step explores with a wrapper the subsets built from features returned by the first step. In [127], four combinations of filter/wrapper methods are proposed using KNN as classifier, and both forward and backward search procedures (SFS and SBS) with different filters.

Another featured hybrid method was proposed in [114]. BIRS algorithm was designed as an incremental method that is able to tackle high dimensionality (*i.e.* gene selection from micro-array data). BIRS involved two stages. The first one ranks attributes according to their usefulness to the class using a filter or wrapper criterion. The second stage, starts from an empty set and incrementally adds ranked features using wrapper subset evaluation. The feature selection process uses first stage ranking to replace the exhaustive neighborhood search in high dimensional space. In fact, attributes were added to the current solution by exploring non selected attributed following the order provided by the first stage. Besides, the selection of a given feature requires a significant improvement of fitness or classification accuracy otherwise the next ranked feature is considered. The significance of the improvement were statistically validated with $t-$test.

Ensemble Feature Selection

As the selection of relevant features is optimized for a particular learning paradigm (*i.e* classifiers used for subsets evaluation), the whole process (feature selection and classification) can be seen, in some ways, biased. For this reason, some of proposed approaches have focused on the ensemble learning and its relation with feature selection. Ensemble learning techniques entail the use of more than one classifier which could cooperate to enhance classification performance. Therefore, the output of a given FS scheme exploring the search space with an ensemble classifier, is not optimized for a unique classifier. Furthermore, some recent studies have shown that ensemble FS increases the stability and the robustness of the FS process [54, 90]. In [115], a prediction model was designed using feature selection based on ensemble learning. The exploration of the search space of subsets was done using a genetic algorithm, whereas the both evaluation and validation involve a set of SVM classifiers.

In addition to the idea of exploring feature subset spaces using an ensemble classifier, *Sayes et al.* investigated in [115] the aggregation of feature selection approaches in the same way as the classifier combined with ensemble learning. According to the authors, the ensemble FS might reduce the risk of unstable results and gives better approximation of the optimal subset since individual feature selectors could lead to different suboptimal solutions.

The ensemble FS was based on feature ranking aggregation (SU, Relief and RFE-SVM). The feature selectors were feeded with different subset instances, generated with bootstrap aggregation (bagging). The aggregation scheme of the second stage is based on weighted voting. Recently, a similar ensemble scheme was applied to the identification of bio-marker from micro-array data [3]. Reported results confirm the stability of the ensemble FS based on RFE-SVM for high dimensional problems.

Memetic Approaches (MA)

In [58], authors considered that the use of memetic approaches was among the more sophisticated recombination and hybridization issues. In a MA, the local search evolves as a component of the whole evolutionary process. These boosting methods are being shown as promising solutions in more than one combinatorial optimization research [58]. By this way, global search of genetic [11] processes will be endowed with the intensification mechanism of the local search.

Seok et al. [103], suggested to embed local search procedures to within GA using deterministic sequential search procedures as local search operators. Reported results showed improvement over floatting search (SFFS) and GA. *Zhu et al.* suggested, in [140], the use of filter as local search operator. The devised memetic algorithm is based on genetic wrapper where solutions are refined using filter criterion. Empirical study showed that the designed memetic schema was able to improve classification accuracy and reduce the number of selected features, and the best results were obtained with memetic schema based on Relief filter. Besides, the memetic GA based on Relief filter outperforms the memetic approach proposed in [103].

A similar memetic schema (MBEGA) was proposed in[139], and applied to high dimensional problems ($\#features > 1000$) using micro-array datasets. The neighborhood structure is almost identical to the proposed in [140], except the use of an adapted add/delete operation. In fact the add is based on SI, and the remove operation delete redundant feature using the Markov blanket approximation used by the FCBF algorithm. Empirically, the memetic scheme outperforms GA and BIRS, but comparable accuracy was obtained with FCBF. In [138], the two last memetic schema were empirically assessed with micro-array data and both approaches showed similar results. Another memetic scheme based on mutual information and genetic algorithm was devised by [62].

In [35], a boosted memetic schema was devised using a multivariate Filter. In fact, The evolutionary process and its operators were guided by NMIFS filter [12]. The filter was involved within initial population generation, chromosomes mutation, and local search.

Recently [133], a local search operator was deployed in three different optimization paradigms (GA, GRASP and TS). The local search procedure was not based on filters but inspired from sequential search. It relied on a neighborhood structure based on

---

[11] the concept of memetic computing was recently extended to several optimization paradigms and was not only limited to evolutionary processes or GA

[12] the filter was proposed within the same paper [35] and was based on information theory measures

attribute exchange. Results, provided for small dimensional problems ($\#features < 100$), pointed out superiority of the GRASP and TS over memetic and basic GA.

Hybrid heuristics

This section covers another kind of hybridization where the combination where not limited to wrappers and filters or the use of local search to enhance exploitation performance but extended to metaheristics combination. In addition to the combination of the local search with evolutionary process, other hybridization schema were proposed to tackle FS problems. The common point between the following approaches are the seek of behavioral complementarity in search space exploration and the trade-off between intensification and exploration capabilities.

In [74], a hybrid system based on ACO and DE were devised where DE evolves solutions provided by ants. The DE crossover and mutation were applied at the end of each iteration. The newly generated solutions have replaced those resulting from the ants search. The resulting subsets, are then used to update pheromone trails and the process restarts. Empirical study pointed out the superiority of the hybrid scheme over GA, ACO, BPSO, and DE.

Recently, another hybridization of evolutionary and swarm approaches was devised to tackle a functional genomic problem: the prediction of protein function [102]. The hybrid ACO-GA evolved in parallel both feature selection meta-heuristics. At the end of each iteration, solutions are evaluated and fittest subsets are selected to update pheromone trails for the ACO and GA population. The ants use the pheromone updates to look for new paths and adjust previous subsets found. The empirical results showed the superiority of hybrid scheme toward GA and ACO. Nevertheless, the evaluation procedure was in some way biased, since the compared algorithms did not have the same computational complexities, running time, and number of fitness function calls.

The third recent hybrid approach combined a SA with GA in [43]. The devised hybrid scheme involves three components: SA, GA, and local search based on hill climbing. The search starts with a SA, then followed by a GA, and terminates with local search solution refinement. The GA starts the evolution process with the best solution returned by SA. Intensive empirical study was conducted and the proposed approach was compared to sequential procedures (SFS, SBE and SFFS), ACO and PSO. The first stage of the optimization process showed comparable results to, almost, all the compared approaches, whereas second and final stage have outperformed all the approaches. The mixed behaviors (global-local search) of the approach made the optimization approach able to avoid premature convergence.

2.3.5 Distributed feature selection

Parallel and distributed implementations of meta-heuristics seem quite naturally as an effective alternative to speed up the search for combinatorial optimization problems. Moreover, the multi-instances scheme might explore different region of the search space and provides more flexible implementation by testing different combination of methods and parameter settings.

A sequential and parallel scatter search were devised in [94]. Scatter search (SS) is a meta-heuristics based on evolutionary scheme. Such an approach generates a population of solutions and subset called *reference set*. The reference set is not only limited to fittest solutions but includes the most diverse solutions in the population. Once the reference set is built, solutions are selected combined and improved iteratively. Improved solution could update the reference set via replacement procedure. Solutions are combined with a variant of uniform crossover that preserves common selected features. The resulting new solutions were refined by a local search schema. The parallel version of the scatter search applies different parameters to a set of sequential instances evolving simultaneously. Both sequential and parallel SS were assessed on on small and medium size benchmark problems of the UCI repository [16]. Sequential SS pointed out slight improved over GA, however parallel SS showed comparable results to sequential one. Suh result could be explained by the absence of interaction between parallel instances.

In [136], a multi-population GA approach was proposed. The GA evolves two populations which cooperate using solution migration. A generalized version of the multi-population schema was proposed in [33], using a genetic island model. Each population is assigned to an island and simultaneously evolves with the remaining ones. A collaboration protocol was devised to manage solution exchange. Empirical results showed the effectiveness of distribution as well as island collaboration to enhance final feature subsets accuracy and avoid premature convergence.

Subpopulation applies crossover mutation and local search operators. The populations communicate and share informations about explored search space via an agent. The authors defined both collaboration and competition policies between populations.

Finally, the proposed distributed model has attempted to enhance the ability of sequential and centralized algorithms by a simultaneous multi-start scheme. However, search diversification issues (*i.e.* hybrid filter-wrapper schema, hybrid heuristics, different local search schema evolving simultaneously within unique distributed model) were not yet explored.

Figure 2.5 summarizes the plethora of the approaches devised as wrappers and surveyed throughout the previous the previous sections. Each approach is illustrated by featured reference.

As a conclusion, we provide a table (Table 2.3) which lists books and papers that have focused on either comparative studies or state of the art reviews [13]. References are compared according to their state of the art coverage and the type and the depth of the empirical study.

2.4 Feature selection modeling challenges

Feature selection challenging problems could be summarized as follows:

- *effective modeling*: enhancing accuracy of the proposed approaches as well as the complexity of the search process and its underling mechanisms.
- *assessment methods*: the reliability of the devised approaches mainly depends on the evaluation methods. The particularity, of the FS problems is that the assessment procedure requires in addition to evaluation of classification accuracy

---

[13] Theses references was, in a part, used to build feature selection survey provided by the current chapter

of the selected attribute subsets, a validation stage which is in most of the cases independent of the search process. Besides, a plethora of evaluation criteria exist for both subset evaluation and individual attribute relevance.

- *high dimensional data*: existing feature selection approaches, are not able to cope with combinatorial nature of high dimensional problems (*i.e.* thousand of attributes). Even though, recent studies started to devise new alternatives, the adaptability of the optimization paradigms to the high dimensional FS modeling requires further investigations.

2.5 Conclusion

In this chapter we introduced basic material of feature selection research field and surveyed main modeling trends. As data evolve, new challenges arise and, hence the expectation of feature selection are elevated. The three following chapters investigates performance improvement issues through the study of new and hybrid optimization paradigms.

Title	Type of pub.	Year of pub.	Survey	Empirical Study			Ref.
				Comparative	Wrapper - filter comparison	Behavioral study	
Wrappers for feature subset selection	Journal paper	1997	++	+	-	-	[76]
Comparison of algorithms that select features for pattern classifiers	Journal paper	2000	++	++	-	+	[79]
Feature Selection: Evaluation, Application, and Small Sample Performance	Journal paper	1997	+	++	-	-	[65]
Benchmarking Attribute Selection Techniques for Discrete Class Data Mining	Journal paper	2003	++	++	+	-	[57]
Feature Extraction, Foundations and Applications	Book	2006	+++	+++	++	-	[54]
Toward Integrating Feature Selection Algorithms for Classification and Clustering	Journal paper	2005	+++	-	-	-	[93]
Feature Selection Using Mutual Information: An Experimental Study	Conference communication	2005	+	++	-	-	[88]
Computational methods of feature selection	Book	2008	+++	++	++	-	[90]
A review of feature selection techniques in bioinformatics	Journal paper	2007	+++	-	-	-	[117]
Data Mining and Knowledge Discovery with Evolutionary Algorithms	Book	2002	++	-	-	-	[41]

Table 2.3: State of the art and comparative studies references

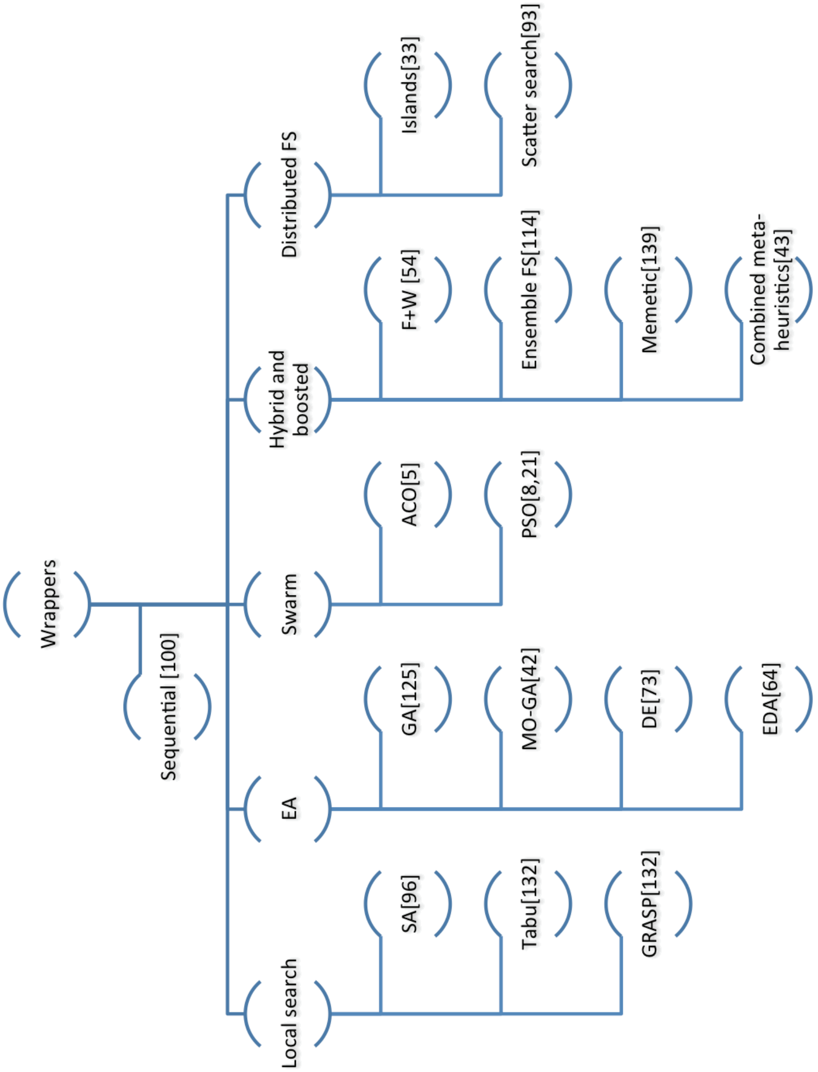


Figure 2.5: Wrapper models

Chapter 3

Memetic Feature Selection: Local search and hybridization issues

3.1 Introduction

Evolutionary algorithms have been successfully applied to the problem of feature selection [54, 90]. Different models, techniques and evolutionary approaches have been explored, with particular interest to the application and the adaption of genetic algorithms to the FS problem. On the other hand, several recent works have developed a number of hybrid and boosted techniques ranging from the heuristic adaptation to the combination of different optimization schema. The aim is to improve, both feature selection classical modeling tools which struggle to gain attended reliability, especially when they face high dimensional data [90]. As a result, some trends in feature selection have attempted to tackle this challenge by proposing hybrid approaches based on the combination of the local search with genetic algorithms. The resulting memetic scheme seems to be an interesting hybridization alternative since it offers the possibility to use specific problem knowledge as well as local search design and integration alternatives. Section 2 reviews fundamental concepts of the genetic and memetic design. Next, we discuss some of the featured memetic approaches that were recently applied to the FS problem. Section 4, details the proposed local search operators and evaluates and discusses its empirical effectiveness.

3.2 Genetic and Memetic Algorithms

In this section we set the scene for the rest of the chapter. We briefly overview and compare technical and behavioral concepts of evolutionary and memetic algorithms. This is motivated by the fact that suggested approaches as well as featured and reference ones -used in this chapter- are based on either evolutionary or memetic mechanisms.

3.2.1 Basic concepts

Evolutionary computing is considered as one of the more popular optimization schema [31, 58]. Reputed to be successful in the application of *Darwinian* principles in problem solving, different paradigms have emerged (genetic algorithms, genetic programming, evolution strategies, *etc*). A genetic algorithm tries to make the analogy between the natural evolution and optimization process. For a given population of individuals, mechanisms inspired form natural selection are used to evolve individuals according to their fitness.

Memetic Algorithms (MAs) refer to an attractive and growing research field, of a class of stochastic heuristics combining the global search nature of Evolutionary Algorithms (EA) with local search techniques, improvement procedures, mechanisms of search guidance and learning. The local search targets the improvements of the solutions quality. Tools and mechanisms used by the local search involves, in most of the cases, techniques of neighborhood exploration (*i.e.* similar solutions: solutions with common features). MA are based on the concept of *meme* [58]. Conceptually, a meme can be defined as an information unit or a pattern for cultural evolution and transmission. Indeed, memes are to culture as genes to biology. Cultural evolution can be understood through the same basic biological and natural selection mechanisms. The fact that a solution moves to a similar one to improve its fitness could be seen as transmission that enhances the solution. The MA paradigm is more inspired by the social concepts in relation with culture than genetic ones. Indeed, culture is generated when individuals becomes more similar due to the mutual social learning. the sweep of culture allows individuals to move toward more adaptive patterns of behaviors.

Common concepts

As all population based metaheuristics three common concepts in direct relation with design issues will be introduced in the following paragraphs.

Representation The solutions of the problem being solved are represented as chromosome where positions and values respectively correspond to genes and alleles. A solution representing a set of decision variables are metaphorically encoded on the same way as chromosomes encode genetic material within genes. Many synonyms exist like *candidate solution* or *individual* [58]. The proposed encoding schema should be able to represent any solution of the targeted search space. Often, the adopted encoding schema require further *interpretation* to get true decisions variables values. In this case we can talk about *phenotype* and *genotype* spaces. The representation could be defined by the relation $R = (P, G, M)$, where P, G respectively denote phenotype, genotype spaces and M the mapping function with domain in G and range in P which provides interpretation of the representation. In the case of feature selection problem, the target is to select optimal features subset. Hence, features are either selected or not. The binary string representation (one variable state per attribute) has been widely adopted in FS modeling [62, 87, 90].

Solution assessment Solutions are made comparable according to a value assigned to each chromosome: the *fitness*. It measures the solution quality. The evaluation function measures solution interestingness, and it is commonly called *fitness function*. Problems typically solved by evolutionary algorithms are optimization problems, which are formulated with an *objective function* maximizing or minimizing the fitness.

Population Both Genetic and memetic algorithms are meta heuristics based on populations. By opposition to some optimization strategies focusing on best solution found or the current solution, evolutionary strategies evolve a set of solutions denoted as a *population*. The concept of population is fundamentally determinant for evolutionary strategies. Firstly, the population is *dynamic*, it changes or *evolves* over generations. Fittest solutions have the advantage to be maintained on the next generation, whereas less fittest or weak ones are threatened to be replaced, or even eliminated. Secondly, the population evolves different genotypes or phenotypes stemming from different regions of the search space. Such representations of solutions offer the global optimization process, different evolution alternatives through population diversity.

Search intensification and diversification Intensification encourage the move to attractive regions of the neighborhood whereas, diversification makes the moves so far, in the aim to explore new regions of the search space or to avoid local minima. The exploration potential is in direct relation with the ability of the proposed heuristic to explore different regions of the search space and to maintain an acceptable diversity level. Operators implementing such mechnisms will be detailed in the next section.

3.2.2 Components and algorithms

This section is devoted to algorithmic aspects of both genetic and memetic processes. For each process we, first, introduce main components, then we illustrate their usage within evolutionary design.

Common GA and MA components

A set of components are used in both evolutionary processes. They try to imitate some of the featured genetic and natural evolution behaviors. The common point between these components (operators) is its tight relation to the population (selection, crossover, mutation). In fact, all of them are applied to a subset or a targeted solutions of the population. The components are designed as operators. All these operators, materialize stochastic behavior of the evolutionary processes.

Selection-Replacement operators Population evolves throughout the update and replacement of population candidates. The evolution is guided by a couple of selection mechanisms: *selection* and *replacement*. The ***selection operator***, also known as *parent selection* or *mating selection* mechanism [31], allows and prepares better individuals to become parents of the next generation. New solutions are derived from the selected ones. The selected solutions are not limited to the fittest ones. They are randomly selected according to a given elitism policy (*selection pressure*). The selection operator forms a subset to undergo variation in order to create new solutions. The resulting subset is also called *mating pool*. Once the new solutions generated, another selection mechanism defines candidates that will be present in the population of the next generation. The process is not deterministic and is controlled by an elitism policy. Since some of the existing solution might be replaced, this operator is usually called **replacement** or **survivor selection** mechanism.

Mutation operator In general, the role of the variation operator is to generate new solutions from the existing ones. The mutation operator imitates the *genetic mutation* process by arbitrarily changing the values of solution elements (decision variables). In the case of binary presentation, values of selected positions are switched from 1 to 0 and inversely. The operator is applied to the mating pool candidates with a probability p_{mut} (parameter of the mutation operator). Problem-specific operator could be designed, to fix the problem of mutation effectiveness and population diversity. The role of mutation operators regarding search space, is to move existing solutions to new regions that might be interesting for exploration or to escape the evolutionary process from solutions locally optimal.

Recombination operator The second variation operator is **crossover**. It is also inspired from the genetic crossover applied to chromosomes. It merges information from two solutions (parents) of the mating pool to generate new offspring solutions. For example, this could be done by exchanging a portion of a given chromosome according to a given cutting position (*one point crossover*) or by exchanging different portions of the chromosome. Like mutation, the crossover operator is applied, with stochastic mechanisms, to mating pool solutions with p_{cross} probability.

Such variation operators materialize inheritance mechanisms which depend on similarity degree between solutions. If the best solution elements are combined, the offsprings might improve the search.

Genetic algorithm process

The above presented stochastic operators are, generally, applied within a genetic algorithm. Basic and commonly used GA steps are described by Algorithm 8.

Algorithm 8: An example of Genetic Algorithm

```
Input:
$Size$: Population size;
$p_{mut}$: Mutation probability; $p_{cross}$: Crossover probability
$Maxgen$: Total number of iterations
Output:
$S_{best}$ : Best solution found
begin
    Population $P \leftarrow$ GenerateInitialSolutionSet($Size$)
    foreach ($s \in P$) do
        $Evaluate(s)$
    $S_{best} \leftarrow getBest(P)$
    P$_{tmp} \leftarrow \emptyset$ ; $i \leftarrow 0$
    while ($i < Maxgen$) do
        P$_{tmp}$ ←Select ($P$)
        Crossover(P$_{tmp}$, $p_{cross}$)
        Mutate(P$_{tmp}$, $p_{mut}$)
        foreach ($s \in P_{tmp}$) do
            $Evaluate(s)$
        Replace($P_{tmp}$, $P$)
        $S_{best} \leftarrow getBest(P)$
        $i \leftarrow i + 1$
    Return $S_{best}$
```

The process starts by the random generation of the initial population of chromosomes ($Size$ parameter defines the number of solutions to generate). The individuals evolve during the search resulting in different population generations. The iterative procedure consists five of steps, typically, scheduled in five steps: (i) selection of a subset of candidate solutions for combination (*line* 8) in the mating pool P_{tmp}; (ii) random pairs are selected from the mating pool, and new pairs are generated by exchanging some parts of the selected solutions (*line* 9) using the crossover operator which is applied with a probability p_{cross} (iii) according to the mutation probability p_{mut} a subset of the mating pool solutions are candidates for mutation, in this step each solution is randomly perturbed to generate a new solution. (iv) all the new solutions are evaluated and a fitness value is assigned to each solution. (v) another selection mechanism called replacement operator designates the candidates that will be replaced by the new solutions and decides which ones of the mating pool that have to quit the evolution process.

According to the fitness value, solutions are assessed, selected and replaced during the search. The more the fitness of a given solution is high, the more it have a chance to survive within population. The four involved operators (selection, crossover, mutation and replacement) are either stochastic or applied in a stochastic manner. Hence, they could be applied in different ways. The adaptation of GA for a given problem involves the adequacy of the operators with nature of the problem and the search space [31].

Memetic operators: Local search possibilities

The idea of applying memetic algorithms aims at the improvement of the search capabilities with *Local Search* (LS). In fact, the genetic process as well as its components will be endowed with intensification possibilities materialized by the add of a new ***operator*** [1].

According to Krasnogor *et al.* [58], three components characterize the behavior of a local search procedure [2]:

- *Neighborhood generating function:* defines the set of solutions that could be reached by the application of the local search to a given solution. Such a function

---

[1] the term *operator* is used on same way as evolutionary operators (mutation, selection and crossover operators)

[2] throughout the thesis, LS procedure and LS operator will be used interchangeably

is also known as a *move operator*. The resulting neighborhood structure depends mainly on the nature of the move. In fact, for a given solution a different set of neighborhoods could be considered depending on the nature of the move procedure. The effectiveness of the local search depends, in a part, on the structure of the neighborhood.

- *Depth:* the depth parameter (d) defines how the LS will be applied, it controls the scope of the LS. Indeed, it could be applied once $d = 1$ or iteratively, the process is restarted with the best solution found until no improvement is found (*Hill Climbing* scheme). It could, also, be used as a mean to control the cost of the neighborhood exploration.
- *Pivot rule:* defines the criterion of accepting an improving solution. Such a criterion is used to prune the solutions of the neighborhood to explore. Pivot rule could, also, aggregate a set of criteria, particularly, in multi-objective optimization problem, and composite neighborhood structure.

The pseudo-code of the local search operator is illustrated by the Algorithm 9

Algorithm 9: Pseudo-code of a L.S. Algorithm

```
Input:
$S$: Solution
Output: $S_{best}$: Improved Solution
begin
    $s \leftarrow S$ , $S_{best} \leftarrow s$
    repeat
        $NH_s \leftarrow GenerateNeighborhood(s)$
        foreach ($s_i \in NH_s$) do
            if ($pivot\_condition\_satisfied \quad \wedge s_i > S_{best}$) then
                $S_{best} \leftarrow s_i$
        $s \leftarrow S_{best}$
    until $depth\_condition\_satisfied$;
    Return $S_{best}$
```

Memetic algorithm: general scheme

MAs are derived from GA since they inherit main components and enhance intensification mechanism with LS. Besides, MA could be considered as a hybrid system, since it allows the combination of global and local search mechanisms. Design issues of a MA depends on:

- the LS components and associated configurations (*i.e.* depth, neighborhood structure)
- the integration of the LS within genetic process
- the rationale for the use of specific refinement procedure with EA (*i.e.* adapted operators)
- the use of problem specific knowledge to adapt the design of the LS operator.

Figure 3.1 illustrates the different levels of local search integration alternatives. The impacts on evolution process as well as on final results vary according to the integration level and the design of the LS operators. Besides, adaptive effort, hybridization strategies as well as use of specific problem knowledge could be implemented outside of the local search operators at different levels of the evolutionary process.

Some authors consider that global search mechanism of GA is the dominant behavior. On the other hand, Memetic Algorithms endow GA components, with more effective intensification and hybridization mechanisms. The memetic design is more effective when we look for a trade-off between evolutionary components and local search capabilities, than to simply boost GA with a local search. Such conclusion, will be confirmed at different levels of the empirical study.

We should also note that recent memetic design was extended to several optimization paradigms. Memetic algorithms cover all hybridization issues as well as the use of valuable problem-knowledge within the implementation of mechanisms enhancing the search.

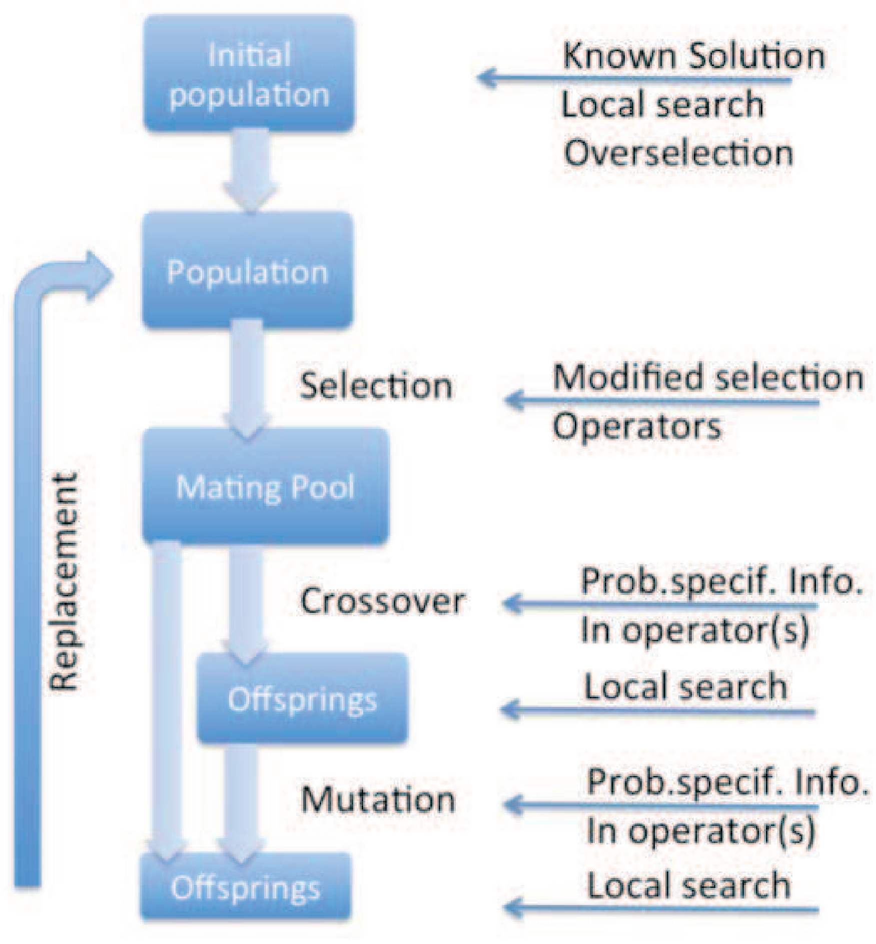


Figure 3.1: Memetic hybridization issues

3.3 Memetic algorithms for feature selection

Memetic approaches as well as genetic ones and other heuristic strategies were surveyed in the previous chapter. In this chapter we recall some of the concepts introduced within the state of the art, and we detail them by providing technical and algorithmic materials in relation with memetic modeling. More specifically, we review and discuss local search operators and their impacts on the memetic process. Some of the detailed operators will be used as reference approaches for the empirical study.

3.3.1 MA of "Yusta09"

In [133], a recent MA was proposed and compared to different methaheuristics. The proposed local search operator is based on attributes exchange (replacing a selected attribute by an unselected one). The neighborhood is based on all pair combinations between selected and discarded attributes (*see eq.* 3.1). Intuitively, the operator could be considered adapted to the FS problem, since it looks for best possible replacement which maximizes the fitness. The local search was iterative ($d \geq 1$), and stops when a local minima is found. The same operator was applied to both *Tabu* and *Grasp* heuristics. The adopted refinement procedure seems to be expensive since all exchange combinations should be explored. The complexity of the LS operator is the order of $\Theta(N^2) * d$ where N is the attributes number, and d is the search depth. Although the LS seems to boost GA results, the adopted memetic design represents a possible alternative among several ones using the same LS operator (section 3.4.1 [3] studies different integration alternatives as well as their impacts on final results).

On the other hand, with such complexity, the LS operator could not be applied to benchmarks with significant attributes number.

$$NH(S) = \{X | X = S \cup \{f_i\} \setminus \{f_j\}, \forall f_j \in S, \forall f_i \notin S\} \qquad (3.1)$$

---

[3] p.70

3.3.2 WFFSA and MBEGA

This paragraph is devoted to a new class of local search operators. WFFSA [140] and MBEGA [139] respectively denote *Wrapper-Filter Feature Selection Algorithm* and *Markov Blanket-Embedded Genetic Algorithm*. The particularity of the respective LS operators is that they allow the resulting memetic algorithms not only to add the local search to the evolutionary process, but allow the hybridization of filter-wrapper scheme. From the algorithmic point of view, both local search operators rely on the same refinement procedure. They are based on two basic operations: $Add(.)$ and $Del(.)$ (delete of one attribute). The neighborhood of the LS operator is based on the application of $Add(.)$ a times and $Del(.)$ procedure d times to a given solution S. The size of the neighborhood is controlled by a parameter l. The total number of combinations is limited to l^2 (*see eq.* 3.2).

$$0 \leq a, d \leq l \tag{3.2}$$

The local search operators of WFFSA and MBEGA are illustrated by Algorithm 10. Only the filter F to apply is different.

We recall that for the sake of simplicity we opted for the below equivalence notation (*see eq.* 3.3) to facilitate the readability of solution comparison.

$$S_A \succ S_B \Leftrightarrow \begin{cases} S_A.fitness \geq S_B.fitness, & \text{(lower error rate)} \\ & \vee \quad (or) \\ \|S_A\| < \|S_B\|, & if(S_A.fitness = S_B.fitness) \end{cases} \tag{3.3}$$

All $Add(.)$ and $Del(.)$ combinations are respectively bounded by a and d. The unique random generation of the couple $< a, d >$ (*Line* 4) guarantees the exploration of all combinations after l^2 iterations.

Both $Add(.)$ and $Del(.)$ uses filter ranking to respectively add and eliminate attributes from current solution. In addition, attributes are selected in a stochastic manner through the use of a random selection (*wheel selection*) based on a given level

Algorithm 10: L.S. of the WFFSA and MBEGA Algorithms

```
Input:
S: Solution
l: search depth
F: Filter
Output: $S_{best}$: Best solution found
begin
    $s \leftarrow S$ ;$S_{best} \leftarrow s$
    for $i = 1$ to $l^2$ do
        Generate a unique random pair $(a, d)$ where $0 \leq a, d \leq l$
        $j \leftarrow 0$;$k \leftarrow 0$
        while $(j < a)$ do
            $Add(s, F)$
            $j++$
        while $(k < d)$ do
            $Del(s, F)$
            $k++$
        $Evaluate(s)$
        if $(s \succ S_{best})$ then
            $S_{best} \leftarrow s$
        $s \leftarrow S$
    Return $S_{best}$
```

of elitism [139]. *Add*(.) procedure for both LS operators is detailed by Algorithm 11.

The only difference between MBEGA and WFFSA is the filter used in the delete *Del*(.) operation. The *Del*(.) procedure for MBEGA operator is based on the principle of *Blanket Markov* [77, 138, 139]. The *Blanket Markov* approximation [138, 139] for redundancy removal is illustrated by equations 3.4 which recall the definition of *symmetrical uncertainty* SU, and equation 3.5 which defines approximation constraints. MBEGA[4] *Del*(.) procedure deletes redundant attributes whereas WFFSA applies *Del*(.) procedure within local search to discards irrelevant features.

$$SU(f_i, f_j) = 2\left[\frac{IG(f_i|f_j)}{H(f_i)+H(f_j)}\right] \tag{3.4}$$

---

[4] Local search operator of MBEGA

Algorithm 11: Add procedure

```
Input:
$S$: Solution;
$R$: relevance measure (filter scores)
Output: $S_{best}$ : new solution
begin
    Rank unselected features in $S$ according to $R$
    Select a feature $f_i$ with high score using wheel selection
    /*features (selected) with higher relevance score are more likely to be
    selected */
    $S_{best} \leftarrow S \bigcup \{f_i\}$
    Return $S_{best}$
```

For a given two features f_i and f_j, $(i \neq j)$, f_j is said to be an approximate *Markov blanket* of f_i only with the following two conditions:

$$\begin{cases} SU(f_j, C) \geq SU(f_i, C) & \wedge \text{ (and)} \\ SU(f_i, f_j) \geq SU(f_i, C). \end{cases} \tag{3.5}$$

Algorithms 12 and 13 summarize the delete procedure of the WFFSA and MBEGA respectively. WFFSA was initially proposed as a hybrid system introducing filter scores within genetic algorithms. Next, MBEGA was derived from WFFSA to handle high dimensional problem, particularly micro-array data sets. In [138], a comparative study assesses empirically the effectiveness of MBEGA and WFFSA on large data set. Globally, the results showed that the two operators provided comparable accuracies with slight advantage to MBEGA. In 2010, Zexuan *et al.*, provide a unifying scheme for memetic feature selection modeling [137]. In this work, MBEGA and WFFSA were considered as two inherited instances of the proposed model. The model abstracts the tools leading to the identification of attribute redundancy as well as the tools used to select added features.

The neighborhood structures of these operators were made of solutions with different *hamming distances*[5]. We consider that possible improvement returned by such

---

[5] Hamming distances measures the the similarity between solution (binary strings)

Algorithm 12: Del procedure of the WFFSA L.S.

```
Input:
S: Solution
R: relevance measure (filter scores)
Output:
S_best: new solution
begin
    Rank selected features in $S$ according to $R$
    Select a feature $f_i$ (from the ranked list) using wheel selection
    /*features (selected) with lower relevance score are more likely to be selected */
    $S_{best} \leftarrow S \backslash \{f_i\}$
    Return $S_{best}$
```

search mechanisms (neighborhood structures) is appreciated but it would be more interesting to reach locally optimal neighborhood solution through the design of effective composite local search (composite neighborhood structures).

3.3.3 MA-C

Correlation based memetic algorithm (MA-C) is a recent memetic approach proposed by *Kannan and Ramaraj* [70]. It was compared to a genetic algorithm and to WFFSA. The comparison to the GA showed that it succeeded to reduce the size of the final subset of attributes, with comparable classification levels. Nevertheless, the performances are globally comparable to WFFSA. The results are not surprising since the devised memetic scheme (MA-C) is based on SU filter ranking and *Blanket Markov* approximation (*see eqs.* 3.4, and 3.5). From one hand, MA-C is not technically different from MBEGA since they use the same assessment tools. On the other hand, they are not conceptually, similar because MA-C local search operator is deterministic, while MBEGA local search limits the neighborhood size and relies on stochastic selection, hence its lower computational complexity.

Algorithm 13: Del procedure of the MBEGA L.S. [139]

```
Input:
S: Solution
R: relevance measure (SU filter scores)
Output:
S' : new solution
begin
    $X \leftarrow S$
    Rank selected features in $X$ in a descending order based on $SU$
    Select a feature $f_i$ (from X) using wheel selection
    /*features (selected) with higher relevance score are more likely to be selected*/
    Eliminate all features in $X \setminus \{f_i\}$ which are in the approximate Markov blanket of $f_i$
    If no feature eliminated remove $f_i$
    $S' \leftarrow X$
    return $S'$
```

3.3.4 Memetic FS design challenges

From one hand, the memetic scheme is motivated by the hybrid design modeling, the use of specific problem knowledge, and the empirical effectiveness. On the other hand, the local search component of memetic approaches seems to be not adapted due to its evaluation cost and its associated computational complexity.

Moreover, some other factors having direct impact on final results and behavioral evolution, like LS integration issues, valuable memetic design, and trade-offs between global and local search require further investigations. The local search challenges in feature selection modeling could be summarized as follows:

- most of the existing local search operators were designed to tackle small FS problems. Such operators require additional adaptation effort to become applicable to high dimensional data sets.
- memetic design: with existing LS operators the evolutionary processes are dominated by the intensification mechanisms. An appropriate trade-off should be found between global and local searches.

- one of the problems of the existing LS operators is the size of the neighborhood. More guided neighborhood exploration and effective refinement procedures are required.

3.4 Proposed Memetic Schema

In this section, we propose a set of local search operators which try to respond to some of the previously formulated requirement. We develop design and algorithmic aspects. Next, we discuss and assess empirical results. These operators, are organized in three classes. The first class covers basic local search operators, that could be applied to any problem based on binary representation[6]. The second class is devoted to LS operators based on composite neighborhoods. The third class is devoted to the adaptation of existing refinement procedures to high dimensional spaces. The local search operators try to preserve neighborhood structures of two previous LS classes and provide some alternatives of effective intelligent exploration. In addition, this class of operators materialize effective hybridization schema between filter and wrapper modeling.

3.4.1 Basic local Search operators

In this section, we study the behavioral aspects of two basic local search operators: Attribute Flip (AF) and Bit Flip (BF). These local search procedures, were initially proposed, in [128] and [133] respectively.

The aim of this introductory section, is to show the importance of the effectiveness of the memetic design and the impact of LS integration at different levels of the evolutionary process. In fact, memetic behaviors as well as results are not necessarily the same even for a memetic configuration based on the same local search operator. According to the previously presented local search operators and to the possible local search integration alternatives within genetic processes, this section is devoted to the study of some possible memetic schema. Local search procedures will be formalized as local search operators, integrated to genetic processes and assessed according to

---

[6] Not necessarily adapted to FS modeling

their contribution and the enhancement of the initial evolutionary processes.

The first operator is based on the BF local search. It explores the neighborhood of a given solution, by exchanging the state of one feature at a time (selected to unselected and inversely). The hamming distance is equal to 1. The resulting neighborhood is formalized by equations 3.6, 3.7 and 3.8.

$$NH_{BF}(S) = \{X | X = SF(S) \cup SB(S)\} \tag{3.6}$$

where $SF(S)$ and $SB(S)$ denote respectively neighborhoods issued from the attribute add and removal from the current solution. It could be also seen as the union of two search heuristics: sequential forward (SFS) and backward (SBE) procedures. (*see p.* 30-31)

$$SF(S) = \{X | X = S \cup \{f_i\}, \forall f_i \in F, f_i \notin X\} \quad \text{(F: feature set)} \tag{3.7}$$

$$SB(S) = \{X | X = S \setminus \{f_i\}, \forall f_i \in X\} \tag{3.8}$$

On the other hand, Attribute Flip operator (AF) constructs the neighborhood using permutation between selected and non-selected features (*see eq.* 3.9). All combinations are considered. Two properties characterize such neighborhood structure: (i) the hamming distance is equal to 2 which requires more exploration effort; (ii) and the operator preserves the feature subset size.

$$NH_{AF}(S) = \{X | X = S \cup \{f_i\} \setminus \{f_j\}, \forall f_j \in S, \forall f_i \notin S\} \tag{3.9}$$

The impact of local search depth will be assessed empirically. The pivot rule did not discard any solution of the considered neighborhood. The two operators explore different regions of the current solution neighborhood. There is no overlapping region ($NH_{BF}(S) \cap NH_{AF}(S) = \emptyset$) and the second neighborhood structure is much larger than the first one which would require more computational time for exploration.

Algorithm 14 illustrates an example of iterative LS procedure that could be applied to FS problem. $NH(.)$ refers in this example to a neighborhood function (*i.e.* NH_{AF},

Algorithm 14: Example of Iterative Local Search Operator applied to FS problem

```
Input:
$S$: Solution
$Cla$: Classifier for fitness evaluation
Output:
$S_{best}$: Best solution found
begin
    $S1 \leftarrow S$ , $S_{best} \leftarrow S1$
    $Stop \leftarrow false$
    repeat
        $Sol_{list} \leftarrow NH(S1)$    /*Neighborhood generation*/
        foreach ($S_i \in Sol_{list}$) do
            $Evaluate(S_i, Cla)$
        $S1 \leftarrow getBest(Sol_{list})$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            $Stop \leftarrow true$
    until ($Stop = true$);
    Return ($S_{best}$)
```

NH_{BF}).

Hybridization issues

The proposed memetic schema target the study and assessment of integration alternatives of the above presented operators within various configurations and processes. The local search could be applied to GA, at different stages of the evolution process. A simplest way of integration, is to apply LS to the best solution at the end of each iteration. Besides, the application of the refinement operator to the new solutions of the mating pool and even to all the population individuals is an exiting alternative since it allows the solutions to evolve simultaneously, over generations. However such alternative requires intensive exploration effort, in comparison to the application of LS to one solution.

These two memetic schema are respectively applied with both AF and BF local search operators. Besides, any local search operator may vary its depth parameter. It could be applied once ($depth = 1$), or iteratively, by the reapplication of the local search to the newly improved solution until no improvement will be found [7] ($depth \geq 1$). The eight combination alternatives will be evaluated to assess the behavior of the resulting memetic processes in FS modeling.

The assessment of the proposed schema covers theorical and empirical comparison through the evaluation of the computational cost and the study of experimental results.

Complexity

Table 3.1 compares the computational complexity of the proposed local search operators. In fact, the operators complexities depend on *three factors*: neighborhood structure, LS depth and the application mode (the solutions to which it will be applied). The first operator, sequential BF applied to the best solution of the generation involves the smallest complexity, whereas the iterative AF operator applied to the newly added solutions generates the highest computational burden.

---

[7] Locally optimal solution

LS-Operator		Order of Complexity	Parameters
Sequential ($depth = 1$)	BitFlip(Best)	$\Theta(N)$	N: number of features m : mating pool size
	AttribFlip(Best)	$\Theta(N^2)$	
	BitFlip(all)	$\Theta(N.m)$	
	AttribFlip(all)	$\Theta(N^2.m)$	
Iterative ($depth \geq 1$)	BitFlip(Best)	$\Theta(N.d)$	d: local search depth
	AttribFlip(Best)	$\Theta(N^2.d)$	
	BitFlip(all)	$\Theta(N.m.d)$	
	AttribFlip(all)	$\Theta(N^2.m.d)$	

Table 3.1: Complexity of local search operators

Empirical results[8]

In this section, we report mean values, standard deviation, and statistical t-test for the assessment of the statistical validity of the obtained results toward the baseline method (GA). For each experiment we present best solution fitness (lowest generalization error rate %), test accuracy on independent dataset, average CPU runtime, cardinality of best solution (*#Attributes*) and the gain in comparison to GA fitness. A ranking based fitness is provided for each dataset. We should note that the negative *t-test*[9] values correspond to improvement over the baseline method (because we tackle a minimization problem), and the confidence level of 99% requires absolute $t - value$ greater than 2.528 for 20 independent runs.

The experiments mainly involve benchmarks belonging to small, medium, and large problems (dimensionality size is ranging from 57 to 2000 attributes).

Globally, we can point out, from the four Tables (3.2, 3.3, 3.4, and 3.5), the superiority of memetic schema over GA at the expense of computation cost resulting from additional evaluations. Another interesting result confirmed by different datasets, is the highest and the lowest gain of memetic algorithms, the best improvements were obtained with the iterative AF operator applied to all the solutions and the smallest ones were obtained by the non-iterative version of the BF operator when it is only

---

[8] Empirical study evaluation criteria and assessment procedures are detailed by the protocol validation section (Annex 1 *see p.* 165)

[9] Stundent Test

applied to the best solution. Such a result, could be explained by the relative impact of neighborhood size and computational complexity of local search operators. Furthermore, when we compare results improvements over the four datasets, we can depict a remarkable enhancement of memetic schema with colon cancer data set. The results are interesting because it is the dataset with the largest search space ($2000 \quad attributes$), and the relatively good GA performance. The obtained results in Table 3.4 are attractive and encouraging for tackling high dimensional search spaces and, particularly, genomic data.

On the other hand, the memetic improvement are not proportional to operators computational complexity. In fact, some sequential LS operators are more interesting than some iterative ones. For example, sequential AF applied to all solutions of the generation is always ranked at the second position, and performing better than some iterative LS schema.

Moreover, the top-3 operators, involve the same neighborhood structure: AF operator. Such neighborhood could be, relatively, adapted to FS problem, particularly, with sequential LS.

Conclusion

The empirical study as well as the computational complexity assessement of LS operators showed the effectiveness of some memetic shema; and give idea about the trade-off that could be found between cost, accuracy and LS problem adaptation.

3.4.2 Composite Local Search Operators

Since, sequential forward (SFS) and backward (SBS) approaches could be considered and formalized as local search procedures (*see eqs.* 3.7 and 3.8), any combination of these two heuristics could be easily implemented as LS operator with composite neighborhood structure. The simple application of these two heuristics either separately or together might not contribute to a significant improvement. Its major drawback is the *nested effect* problem [54]. The attribute added with SFS or removed

				VALIDATION ERROR%						
LS applied to GA		Measures	Fitness	ANN	NB	CPU (s)	# Attrib.	# Eval	Gain%GA	RANK
No LS (GA)		Mean:	9,32%	10,74%	7,81%	14774,13	15,04	1089		
		Sd:	0,95%	1,45%	1,10%	14759,99	3,15	0		
		t-test	0	0	0	0	0	-		
SEQ	BitFlip(Best)	Mean:	**8,54%**	11,10%	8,24%	35672,26	**13,87**	3089	8,37%	8
		Sd:	1,02%	2,04%	1,27%	36614,23	2,94	0	-	
		t-test	-18,92	3,71	4,95	22,08	-12,67	-	-	
	AttribFlip(Best)	Mean:	**6,88%**	**9,61%**	**7,26%**	32705,52	15,96	3089	26,18%	3
		Sd:	1,15%	2,25%	1,49%	28955,4	4,88	0	-	
		t-test	-37,43	-8,41	-5,85	18,74	5,31	-	-	
	BitFlip(all)	Mean:	**8,19%**	**10,58%**	8,23%	91471,7	15,45	11289	12,12%	6
		Sd:	0,98%	1,98%	1,50%	40881,06	3,2	0	-	
		t-test	-29,95	-1,7	7,22	31,64	1,49	-	-	
	AttribFlip(all)	Mean:	**5,75%**	**8,95%**	**6,73%**	95083,75	17,6	11289	38,30%	2
		Sd:	0,69%	1,93%	1,50%	45427,35	4,47	0	-	
		t-test	-73,7	-10,69	-23,21	44,9	85,05	-	-	
Iterative	BitFlip(Best)	Mean:	**8,44%**	**9,98%**	**7,49%**	36313,91	15,61	3171,17	9,44%	7
		Sd:	1,17%	1,75%	1,56%	34137,85	4,38	32,04	-	
		t-test	-20,55	-6,81	-4,72	54,25	8,08	-	-	
	AttribFlip(Best)	Mean:	**7,01%**	**9,62%**	**7,27%**	35944,87	**14,87**	3209,43	24,79%	4
		Sd:	1,01%	2,10%	1,44%	38229,41	3,63	33,37	-	
		t-test	-45,78	-11,65	-9,22	23,64	-3,54	-	-	
	BitFlip(all)	Mean:	**7,93%**	**10,35%**	7,70%	128806,75	15,35	14180,5	14,91%	5
		Sd:	0,93%	1,50%	1,18%	112289,05	3,1	418,7	-	
		t-test	-54,13	-3,5	-2,06	55,51	2,61	-	-	
	AttribFlip(all)	Mean:	**5,52%**	**7,69%**	**6,30%**	106967,85	17,6	12911	40,77%	1
		Sd:	0,49%	1,22%	1,31%	49985,26	2,91	137,52	-	
		t-test	-151,28	-31,84	-22,96	51,35	85,05	-	-	

Table 3.2: Data set: SpamBase (57 Attrib.)

with SBS could not be removed or added by the same operator. In addition, the empirical study, of the previous section, clearly, showed that AF operator outperformed BF one in different memetic schema. The problem of effective refinement procedure could be handled from two perspectives. The first, considers that the simple application of SFS, SBS or even BF could be enhanced more effectively (*i.e.* looking for redundancy among selected attributes), whereas the second, tries to find appropriate of add/remove combinations. The problem was in a part, fixed and some of the requirements of the two points of view were satisfied by some heuristics. The nesting effect problem was addressed by the floating search strategies [54, 120]. The empirical studies showed that floating strategies outperform classical sequential search procedures but are not so effective than evolutionary strategies [54, 90]. The effectiveness toward an evolutionary strategy could be explained by the relative luck of diversification mechanisms of floating heuristics. The idea here is to bring the appropriate use of the floating search within a genetic process. The integration of such heuristics could be designed as local search operators since they explore neighborhood of the

				VALIDATION ERROR%						
LS applied to GA		Measures	Fitness	ANN	NB	CPU (s)	# Attrib.	# Eval	Gain%GA	RANK
No LS (GA)		Mean:	13,80%	6,88%	12,01%	7891,74	13,09	1092		
	-	Sd:	1,67%	1,94%	4,18%	7468,36	3,22	0	-	
	-	*t-test*	0	0	0	0	0	-	-	
SEQ	BitFlip(Best)	Mean:	14,05%	7,53%	13,07%	16106,09	12,96	3092	-1,81%	8
		Sd:	1,28%	2,73%	3,23%	15283,92	2,95	0	-	
		t-test	1,83	5,94	4,4	24,54	-0,48	-	-	
	AttribFlip(Best)	Mean:	**12,24%**	6,72%	12,71%	16355,78	15,04	3092	11,30%	4
		Sd:	1,58%	2,45%	3,92%	14378,28	4,43	0	-	
		t-test	-9,19	-1,36	3,23	23,98	6,81	-	-	
	BitFlip(all)	Mean:	13,55%	7,14%	12,70%	61595,86	14,09	11292	1,81%	7
		Sd:	1,59%	2,70%	3,85%	59565,85	3,39	0	-	
		t-test	-1,96	2,14	2,34	474,37	2,95	-	-	
	AttribFlip(all)	Mean:	**10,65%**	6,66%	12,04%	60927,74	15	11292	22,83%	2
		Sd:	1,36%	2,05%	3,84%	57579,19	3,33	0	-	
		t-test	-15,68	-1,84	0,13	173,43	5,22	0	-	
Iterative	BitFlip(Best)	Mean:	**13,49%**	7,11%	11,54%	17913,52	14,78	3147,65	2,25%	6
		Sd:	1,55%	1,74%	3,05%	17418,26	3,67	22,33	-	
		t-test	-2,37	2,92	-1,51	23,05	6,59	0	-	
	AttribFlip(Best)	Mean:	**11,74%**	**6,15%**	11,77%	18753	18,05	3193,36	14,93%	3
		Sd:	1,33%	1,35%	2,94%	16008,79	3,54	25,87	-	
		t-test	-16,1	-6,86	-1,2	23,52	19,29	0	-	
	BitFlip(all)	Mean:	**13,15%**	7,19%	11,65%	70264,39	14,96	13143,3	4,71%	5
		Sd:	1,19%	2,12%	3,02%	66756,71	3,2	275,87	-	
		t-test	-4,82	2,19	-1,93	37,65	6,87	-	-	
	AttribFlip(all)	Mean:	**9,73%**	**6,33%**	**10,70%**	73319,09	17,17	12521,57	29,49%	1
		Sd:	1,28%	1,45%	2,45%	74439,68	2,53	105,89	-	
		t-test	-31,76	-7,29	-6,94	34,36	14,01	-	-	

Table 3.3: Data set: Sonar (60 Attrib.)

solution to be enhanced. The resulting new local search operators based on floating heuristics (SFFS and IFFS [54, 101]) could be also considered as an improved version of the LS operators of the previous section. Besides the neighborhood structure is more diversified, due to the alternated forward add and backward remove. Three memetic schema were derived from floating search approaches.

Floating Local search: FLS1

The first operator is inspired from the SFFS and SBFS heuristics [120]. The exploration is based on two search directions: forward and backward searches. The new neighborhood structure is generated from the search directions and updating rules. The search continues if at least one of the two search directions leads to an improvement. We recall that SFFS and SBFS respectively start the search from a known predefined set of features (empty set and full set) and define a main and alternative search direction (*main*: can stop the entire search; *alternative*: explored after the

				VALIDATION ERROR%						
LS applied to GA		Measures	Fitness	ANN	NB	CPU (s)	# Attrib.	# Eval	Gain%GA	RANK
No LS (GA)		Mean:	6,52%	6,58%	12,19%	31079,73	23,18	3032		
		Sd:	2,92%	3,14%	4,41%	22980,17	7,45	0		
		t-test	0	0	0	0	0	-		
SEQ	BitFlip(Best)	Mean:	**4,46%**	**6,28%**	**10,32%**	38454,95	**20,55**	5032	31,60%	4
		Sd:	2,05%	2,97%	4,59%	30297,6	6,6	0	-	
		t-test	-20,87	-2,11	-7,8	5,73	-5,45	-	-	
	AttribFlip(Best)	Mean:	**4,64%**	7,06%	**10,06%**	41058,14	**21,41**	5032	28,83%	5
		Sd:	2,25%	2,96%	3,23%	37162,49	8,55	0	-	
		t-test	-14,38	5,19	-15,01	7,69	-4,35	-	-	
	BitFlip(all)	Mean:	**4,78%**	**5,96%**	**11,49%**	200004,95	**22,05**	13232	26,69%	6
		Sd:	2,96%	2,98%	2,96%	102160,5	6,34	0	-	
		t-test	-17,19	-5,06	-4,23	22,81	-3,15	-	-	
	AttribFlip(all)	Mean:	**3,59%**	**5,74%**	**11,49%**	198320,55	**22,05**	13232	44,94%	2
		Sd:	1,62%	2,59%	3,12%	102691,3	6,78	0	-	
		t-test	-24,16	-4,14	-8,47	40,91	-3,15	-	-	
Iterative	BitFlip(Best)	Mean:	**5,58%**	6,29%	11,47%	39703,43	**20,91**	5073,74	14,42%	7
		Sd:	2,25%	3,17%	4,39%	30596,31	7,36	30,84	-	
		t-test	-5,31	-1,42	-1,37	12,2	-3,23	-	-	
	AttribFlip(Best)	Mean:	**4,35%**	6,94%	**11,84%**	39066,78	**20,09**	5068,09	33,28%	3
		Sd:	2,17%	2,81%	3,44%	31127,5	6,31	10,33	-	
		t-test	-15,9	2,48	-3,93	6,94	-6,98	-	-	
	BitFlip(all)	Mean:	**5,65%**	7,02%	**11,38%**	249361,3	**18,45**	15199,5	13,34%	8
		Sd:	2,86%	3,59%	3,86%	142791,7	6,51	711,49	-	
		t-test	-4,35	1,71	-5,02	57,21	-12,59	-	-	
	AttribFlip(all)	Mean:	**2,83%**	6,28%	**9,89%**	213912,1	**20,2**	13790	56,60%	1
		Sd:	1,88%	3,27%	3,86%	122469,18	7,03	66,38	-	
		t-test	-21,43	-1,33	-23,24	29,99	-6,06	-	-	

Table 3.4: Data set: Colon cancer (2000 Attrib.)

main one and taken into considerations only if it improves the search). In comparison to original versions of the floating search heuristics, FLS1 transforms the heuristic in a local search:

- defines a composite neighborhood structure from search directions
- LS operator is able to refine solution without any size constraint.
- accepts an improvement provided by forward or backward search without considering main or alternative search direction.

Algorithm 15 details the main steps of the FLS1 operator where $SF(.)$ and $SB(.)$ respectively denote forward and backward search procedures.

Besides, there is no risk of cycling, because the Neighborhood procedures are only applied to improved solutions. Note that FLS1 is not comparable to either AF or BF operators, although that they are based on SFS and SBS neighborhoods. Nevertheless, it preserves some features of the two operators. Indeed, BF applies

Algorithm 15: Floating LS: FLS1

```
Input:
$Cla$: Classifier for solution evaluation
$S$: Input Solution
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $imp \leftarrow true$    /* Flag*/
    while ($imp$) do
        $Sol_{list} \leftarrow SF(S, Cla)$
        foreach ($X \in Sol_{list}$) do $Evaluate(X, Cla)$
        $S1 \leftarrow getBest(Sol_{list})$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            $imp \leftarrow false$
        $Sol_{list} \leftarrow SB(S_{best}, Cla)$
        foreach ($X \in Sol_{list}$) do $Evaluate(X, Cla)$
        $S2 \leftarrow getBest(Sol_{list})$
        if ($S2 \succ S_{best}$) then
            $S_{best} \leftarrow S2$
            $imp \leftarrow true$
    Return $S_{best}$;
```

				VALIDATION ERROR%						
LS applied to GA		Measures	Fitness	ANN	NB	CPU (s)	# Attrib.	# Eval	Gain%GA	RANK
No LS (GA)		Mean:	17,14%	13,38%	17,33%	158683,53	85,79	1311		
		Sd:	0,90%	1,13%	1,31%	72380,76	17,63	0		
		t-test	0	0	0	0	0	-		
SEQ	BitFlip(Best)	Mean:	**16,81%**	13,86%	17,09%	490897,85	93,7	3311	1,93%	8
		Sd:	0,91%	1,11%	1,30%	222623,21	14,37	0	-	
		t-test	-23,47	6,69	-2	26,47	11,23	-	-	
	AttribFlip(Best)	Mean:	**15,52%**	14,03%	17,29%	519275,2	92,25	3311	9,45%	3
		Sd:	1,08%	1,67%	1,59%	347110,37	46,42	0	-	
		t-test	-65,25	2,69	-0,42	17,8	2,1	-	-	
	BitFlip(all)	Mean:	**16,46%**	14,02%	17,36%	1785624,68	**82,74**	11511	3,97%	6
		Sd:	1,24%	1,91%	1,75%	898120,57	20,56	0	-	
		t-test	-36,87	7,77	0,22	37,76	-3,46	-	-	
	AttribFlip(all)	Mean:	**14,76%**	14,22%	**16,83%**	1892140,75	89,4	11511	13,89%	2
		Sd:	0,80%	2,25%	1,30%	1141857,26	42,76	0	-	
		t-test	-71,41	7,94	-5,5	33,02	11,46	-	-	
Iterative	BitFlip(Best)	Mean:	**16,80%**	14,40%	**16,59%**	525793,95	91,2	3516	1,98%	7
		Sd:	0,91%	1,50%	1,67%	290212,63	26,49	55,2	-	
		t-test	-16,38	11,1	-8,17	43,48	5,22	-	-	
	AttribFlip(Best)	Mean:	**16,03%**	14,16%	**16,76%**	468881,85	93,9	3512,5	6,48%	4
		Sd:	1,03%	1,78%	1,87%	315436,97	56,97	40,3	-	
		t-test	-26,71	5,14	-6,36	28,42	1,19	-	-	
	BitFlip(all)	Mean:	**16,29%**	13,79%	**16,67%**	2591560,62	**83,1**	15471,95	4,96%	5
		Sd:	1,15%	1,50%	1,52%	1343356,45	22,25	600,1		
		t-test	-24,61	5,56	-7,31	51,56	-8,75	-	-	
	AttribFlip(all)	Mean:	**14,19%**	14,91%	**16,61%**	1384673,95	**51,65**	13731,5	17,21%	1
		Sd:	0,92%	1,30%	1,38%	763942,71	32,57	272,29	-	
		t-test	-138,18	20,26	-7,89	172,14	-11,26	-	-	

Table 3.5: Data set: Arrhythmia (279 Attrib.)

$SF(.)$ and $SB(.)$ to the same initial solution while, with FLS1, $SB(.)$ is applied to the improved solution after the application of $SF(.)$.

The neighborhood structure of the FLS1 operator is made of diversified sub-neighborhoods which are adapted dynamically according to the ability of basic operators to improve the current solution. The design of different neighborhood structures with LS operators is highly recommended, since it alleviates the risk of the local minima's and diversifies the search [58, 60].

Floating Local Search with Iterative Backtrack (FLS2)

The second local search operator derived from the floating search heuristics applies the same principle for attributes add and removal. Once an attribute X_i is added to a given solution, some existing (selected) attributes could be correlated or redundant to X_i. The idea is to eventually remove all redundant attributes which do not decrease classification accuracy of current solution.

In FLS2, the same floating search scheme as in FLS1 is adopted, except the fact that backward search procedure $SB(.)$ is not applied once but the backtrack is applied iteratively until no improvement could be reached. Comparatively to FLS1, FLS2 requires more computational time than FLS1 but might lead to more compact subset size. Algorithm 16 illustrates main steps of the FLS2 operator.

Improved floating Local search: IFLS

The idea of IFLS is inspired from a recent approach which attempted to improve the floating search heuristic (IFFS) [101]. The Improved Floating search heuristic is made of the set of steps of the SFFS search but adds a new step to handle the case where the current solution is not improved by the backward stage. The new step, according to the authors of $IFFS$, consists of replacing the weakest feature. In other words, the new stage involves the application of the AF operator. The local search operator, inspired from the IFFS heuristic, could be designed in a manner to enhance composite neighborhood structure. In fact, when backtrack fails to improve solutions AF operator is called. Algorithm 18, summarizes the three main stages of the local search operator (forward search, backward search, and attributes exchange) as well as the updating rules governing transitions between associated neighborhoods. The third stage might be interesting when it allows neighborhood structure diversification, and also to eventually escape local minima. We should note that the particularity of $IFLS$ operator is that its complexity is greater than AF in the case where the third stage is iteratively performed. Figure 3.2 illustrates different local search operators proposed in this section as well as their relation with previously presented LS operators and reference approaches.

Empirical results[10]

In this section, we assess and compare the three composite local search operators and compare them to both memetic approaches based on AF.

---

[10] Empirical study evaluation criteria and assessment procedures are detailed by the protocol validation section (Annex I *see p.* 165)

Algorithm 16: Floating LS procedure with iterative backtrack: FLS2

```
Input:
$Cla$: Classifier for solution evaluation
$S$: Input Solution
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $imp \leftarrow true$    /* Flag*/
    while ($imp$) do
        $Sol_{list} \leftarrow SF(S)$
        foreach ($X \in Sol_{list}$) do
            $Evaluate(X, Cla)$
        $S1 \leftarrow getBest(Sol_{list})$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            $imp \leftarrow false$
        while ($true$) do
            $Sol_{list} \leftarrow SB(S_{best})$
            foreach ($X \in Sol_{list}$) do
                $Evaluate(X, Cla)$
            $S2 \leftarrow getBest(Sol_{list})$
            if ($S2 \succ S_{best}$) then
                $S_{best} \leftarrow S2$
                $imp \leftarrow true$
            else
                Break    /* Flag*/
    Return ($S_{best}$)
```

Algorithm 17: Simplified version of FLS2

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $imp \leftarrow true$    /* Flag*/
    while ($imp$) do
        $S1 \leftarrow SF(S_{best}, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            $imp \leftarrow false$
        while ($true$) do
            $S2 \leftarrow SB(S_{best}, Cla)$
            if ($S2 \succ S_{best}$) then
                $S_{best} \leftarrow S2$
                $imp \leftarrow true$
            else
                Break
    Return $S_{best}$
```

Algorithm 18: Improved floating local search: IFLS

```
Input: $S$: Input Solution
$Cla$: Classifier for solution evaluation
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $imp \leftarrow true$    /* Flag*/
    while ($imp$) do
        $S1 \leftarrow SF(S_{best}, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            $imp \leftarrow false$
        while ($true$) do
            $S2 \leftarrow SB(S_{best}, Cla)$
            if ($S2 \succ S_{best}$) then
                $S_{best} \leftarrow S2$
                $imp \leftarrow true$
            else
                $S2 \leftarrow AF(S_{best}, Cla)$
                if ($S2 \succ S_{best}$) then
                    $S_{best} \leftarrow S2$
                    $imp \leftarrow true$
                else
                    Break
    Return $S_{best}$
```

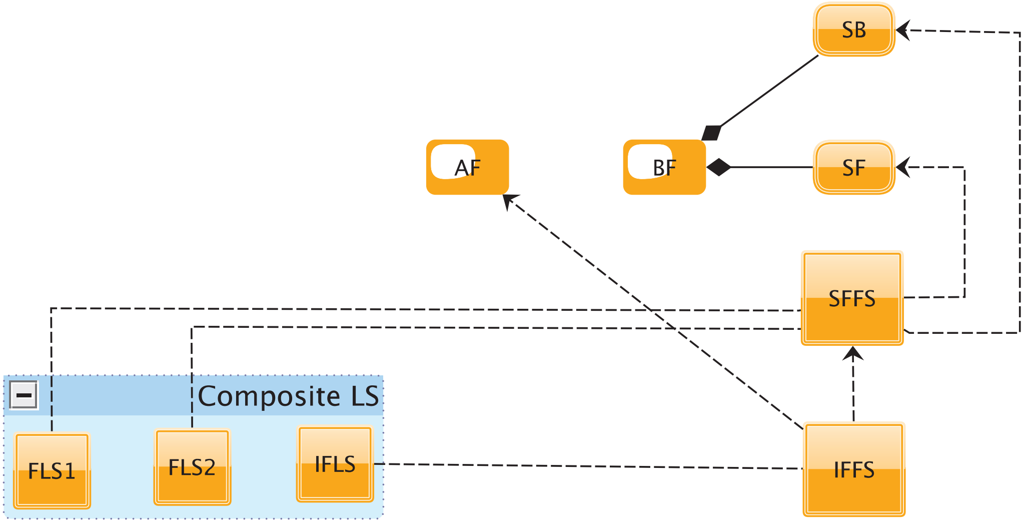


Figure 3.2: Composite LS operators

Tables C.4, C.5 (*see ANNEX III, p.* 176-177) detail evaluation and validation accuracies of the best solution found by the application of memetic algorithms respectively endowed with the $FLS1$, $FLS2$ and $IFLS$ local search operators to memetic approaches based on AF. According to the reported results, the floating operators globally outperform the AF operator on fitness and at least one of the validation criterion. For the operator FLS, fitness results were improved for all the six benchmarks. Whereas, both *Soybean* and *Arrhythmia* benchmarks, confirmed the superiority of all the composite operators in comparison to AF.

Besides, at least one of the proposed composite operators succeeded to improve all validation and evaluation qualitative criteria (fitness, and two validation classifiers).

When comparing the three floating LS operators, FLS seems to be the more effective and it is followed by the $IFLS$. Improved results in favor of floating LS operators, could be explained and argued by the structure of the neighborhood. In fact, the composite neighborhood structure provides to the refinement procedures implemented by the LS operators more intensification issues. In addition, when comparing computational complexities, the more effective operators among composite ones, requires less exploration effort.

Although that basic and composite operators, studied within current and previous sections succeeded to enhance evolutionary processes, the computational effort required by respective local searches and induced by the neighborhoods exploration, should be adapted to be able to cope with high dimensional datasets.

3.4.3 Adapting LS for high dimensional spaces

From one hand, local search operators contribute to refine solutions and to globally enhance evolutionary process.

Results of the two previous sections confirmed local search effectiveness in memetic context. On the other hand, hoping to tackle high dimensional problems with local search operators facing thousand of attributes seems to be unrealizable. More formally, LS complexities in the order of $\Theta(N^2)$ [11] or even $\Theta(N)$ require more computational effort than the evolutionary process itself. Two questions formulate the problems associated to such challenge: *(i)* how to make LS independent or less dependent to the total number of features? *(ii)* is the exhaustive neighborhood exploration mandatory? The main idea behind the LS adaptation to the high dimensional spaces relies on the alleviation of the neighborhood exploration cost. The design of the new operators is based on the use and the valuable integration of a specific problem knowledge to the local search operators. Such a knowledge can be provided by filters and the resulting attribute ranking. The ranking is not systematically used to select or discard attributes but brings out interesting neighborhood exploration issues. Several hybrid approaches have attempted to integrate filter through global search (classical filter-wrapper scheme) [54, 87, 90, 114].

Nevertheless they faced the problem of attributes initially discarded by a filter at the beginning of the search and the wrapper couldn't reintegrate them. The problem is similar to the *nested effect*. We think that the use of filter scores with local search is appropriate for design of hybrid approaches, since all the attributes are considered within the global search process, and only refinement procedures act with a reduced

---

[11] case of AF (N: number of attributes)

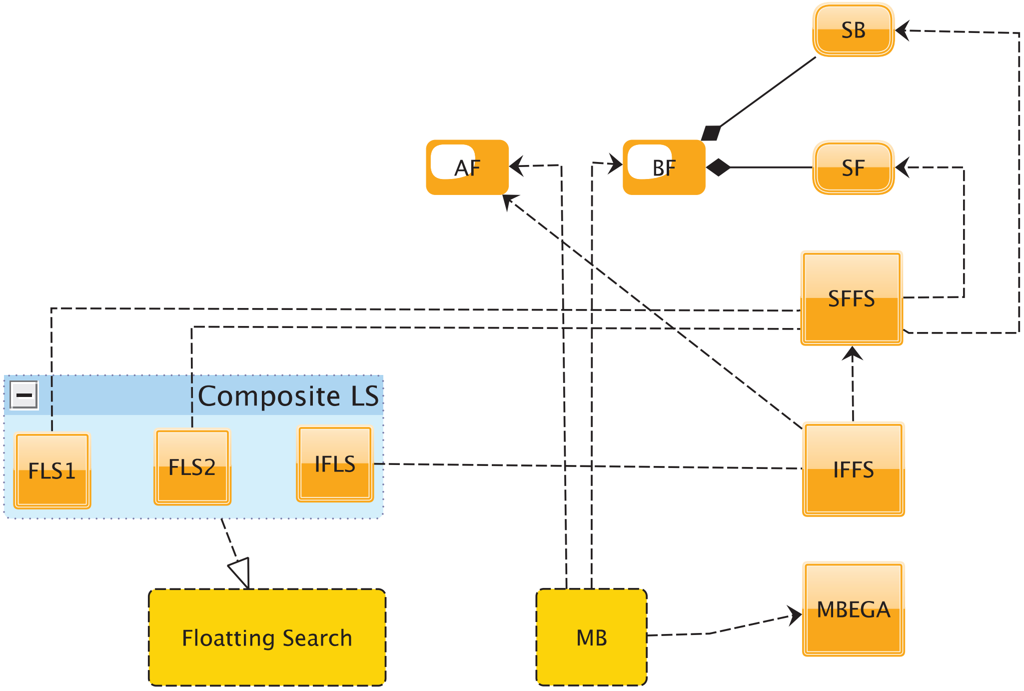


Figure 3.3: Schema

or targeted subset of features.

More than ten local search operators are devised and will be presented in this section. They are organized in families and schema. Families involve deterministic and stochastic operators whereas schema refer to some common steps adapted by the local search processes. Schema could be considered as way to preserve some neighborhood structures, particularly composite ones.

Memetic schema

Devised memetic schema refer to an abstract description of some proposed local search operators. They could also be seen as templates describing local search main steps. The operators derived from these schema illustrate some of the possible alternatives of local search improvement or adaptation to high dimensional spaces. Two schema are proposed. The first one is mainly based on the MBEGA local search operator [139] which is illustrated by the Algorithm 10. The attributes are added or/and removed according to the filter scores with an upper bound for the maximal number of attributes to add or remove. The neighborhood structure is diversified (a different hamming distance for each solution), when it is compared to classic local search operator. Nevertheless, we consider that the eventual improvement resulting from this operator is not optimal and could be enhanced. In fact, the LS neighborhood has the advantage to be of a reduced size and controlled by a parameter but the resulting solution could be refined since it does not represent a local minima for different neighborhood structures. For this reason, the scheme that we propose here, endows the MBEGA local search with an additional step involving AF and BF operators. In fact, if we assume that MBEGA operator provides the best local minimum that could be reached from a given solution (before the application of the LS of MBEGA), the solution is not necessarily optimal for a different neighborhood structure. The new stage adapts the refinement procedure to the result of the MBEGA local search. If the solution were not improved, it seems suitable to opt for a neighborhood structure with a higher probability to enhance the solution like AF. On the other hand, successful application of the MBEGA local search procedure, leads necessarily to an improvement. Consequently, the additional stage adopted with the scheme did not require an extensive search since the initial solution was already enhanced. Hence, we suggest for this stage, BF operator which is less costly than the AF operator. We should also note, that the AF is more effective and time consuming than the BF operator. Indeed, the global-local search scheme design seems to be relatively equilibrated by offering more chance to the non improved solutions and less expensive operator for the improved ones. This scheme will be called MB.

The second scheme is based on floating operators. Such a choice is motivated by the fact that the floating heuristics overcome the problem of nested effect. Besides, from the design point of view of LS operator, the composite neighborhood and dynamic switch between them bring more diversification possibilities to LS. This scheme is mainly based on operators derived from FLS and IFLS operators. We should note that some of the LS operators and the adaptation alternatives proposed in the few next paragraphs cover LS operators that adopt as scheme and other that are not depending on them. Besides the schema impact on results and on behaviors of memetic approaches will be discussed and studied empirically. Figure 3.3 illustrates schema as well as their relation with previously presented LS operators and reference approaches.

Pruning alternatives: deterministic components

In this section, we study pruning issues in relation of the adaptation of the local search operators to high dimensional spaces. The main ideas behind the adaptation are three folds: (i) maintain the previously presented local search operators which proved its effectiveness (ii) focusing the adaptation effort on the alleviation of the local search neighborhood exploration cost (iii) integration of valuable knowledge to guide local search processes (*i.e.* neighborhood pruning).

Pruned AF search space: AF^+

Since AF local search is not applicable for even few runs, because of the cost of the exhaustive neighborhood exploration (*i.e.* the application of AF operator to a solution of 20 selected attributes of the colon dataset requires 39600 evaluations for only one iteration), a new way of neighborhood structure definition and exploration is required. AF^+ is the improved version of AF operator. The AF neighborhood is represented as a matrix composed of regions. Our target is to look for the most interesting regions to explore and to avoid exhaustive exploration by keeping acceptable performances. Lets start by mapping matrix areas and the illustration of the neighborhood structure. As the AF operator replaces an existing attribute by an unselected one, the neighborhood covers all the combinations of selected and unselected attributes.

The neighborhood space is represented by a matrix where the matrix rows refer to the selected attribute indexes (one per row) and the columns refer to the indexes of unselected attributes. Each matrix cell illustrates a possible combination of attribute exchange, and hence to a solution of the neighborhood. Attributes on both columns and rows were sorted according to their filter scores. The more relevant attributes are set on the first rows/columns, and attributes with lowest scores are on the latest lines/columns. We set two thresholds a and b to divide matrix horizontally and vertically in four regions. The region boundaries are defined the axes materializing the thresholds. Figures 3.4 illustrates neighborhood space organization and how the regions are limited with thresholds.

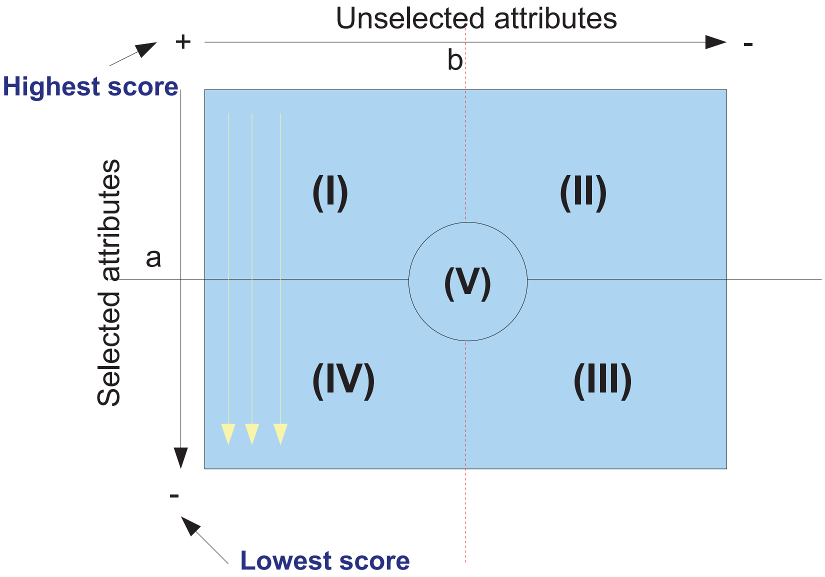


Figure 3.4: AF neighborhood segmentation

The aim of such organization is to localize improving solutions, and if possible the best one. According to a dedicated experimental study targeting several high dimensional datasets, the AF^+ was applied to 1000 solutions randomly generated. The empirical results confirm that both first improving solutions and best ones reached by the local search operators belongs to regions I and IV in more than 76% of the cases. Parameters a and b refer to cells having scores greater than the mean of respective scores (for the selected and the unselected attributes list).

Such interesting result paves the way to a pruning strategy limiting the search to regions I and IV and discarding the remaining ones. The local search operators (*see Algo.* 19) builds from the input solution two sorted lists one for the selected attributes and the second for the remaining ones (Lines 1-5). Next the neighborhood exploration starts by exploring the targeted regions. Line 8 returns the region identifier of given attribute exchange combination. All combinations belonging to a relatively non interesting regions are discarded (Lines 9-10).

Algorithm 19: Pruned AF: AF^+

```
   Input:
   $S$: Input Solution
   $F$: Filter
   $Cla$: Classifier for solution evaluation
   Output:
   $Sol_{best}$: Best solution within neighborhood
1  begin
2      $UnSel_{list} \leftarrow getUnSelectedAtt(S)$    /* unselected attributes*/
3      $Sol_{best} \leftarrow S$
4      $Sort(UnSel_{list}, F, Desc)$
5      foreach ($y \in UnSel_{list}$) do
6          foreach ($x \in S$) do
7              /* x: attribute */
8              $r \leftarrow getRegion(x, y, F)$    /* see fig. 3.4 */
9              if ($r = 2 \vee r = 3$) then
10                 continue
11             $S1 \leftarrow S \cup \{y\} \setminus \{x\}$
12             $Evaluate(S1, Cla)$
13             if ($S1 \succ Sol_{best}$) then
14                 $Sol_{best} \leftarrow S1$
15     Return $Sol_{best}$
```

Why MB scheme?

The MB scheme extends from the MBEGA local search procedure. The last stage adds to the initial local search a new stage to improve the resulting solution. This is motivated by the composite structure of the neighborhood. Solution which are not improved by the first neighborhood structure might be improved by another neighborhood structure. This idea is widely accepted, as a way to escape local minima as well as in other optimization paradigms like *Variable Neighborhood Search* (VNS) which diversifies the search by changing the neighborhood structure. From the other hand, the additional stage requires more of exploration efforts and exploration costs. One can consider that the two operators are not comparable, since the MB scheme involves two stages. For this reason, we empirically compare the impact of the application of the MB scheme regarding MBEGA local search. Besides, we compare the MB operator to the MBEGA applied twice. Table 3.6 reports results of mean improvements in comparison to the initial solution fitness. The improvements refer to the percentage of gain compared to initial solutions. The experiments were repeated ten times [12], and in each run 1000 solutions were randomly generated, then the local search operators were applied to each solution. The results clearly show how the MB scheme outperforms MBEGA applied once and twice. Besides, the experiments have been applied by varying the length parameter l of the local search of the MBEGA operator ($l = 3$;$l = 5$). The value of l denotes the threshold for allowed add/remove operations. In both configuration the MB scheme confirms its superiority. Such result confirms the advantages of exploration with composite neighborhoods. The operator derived from the MB scheme used for these experiment is the MB^+ which be developed in the next section and experimented within evolutionary context.

Operators	$l = 5$			$l = 3$		
Data(# attrib.)	**MBEGA**	**MBEGA*2**	**MB**	**MBEGA**	**MBEGA*2**	**MB**
Sonar(60)	$39,33\%$	$42,72\%$	$46,04\%$	$31,99\%$	$36,54\%$	$43,1\%$
Soybean(35)	$38,97\%$	$46,81\%$	$49,35\%$	$29,9\%$	$37,87\%$	$43,16\%$
Arrhythmia(279)	$18,38\%$	$25,28\%$	$27,12\%$	$10,91\%$	$15,32\%$	$22,06\%$
Semeion(256)	$16,28\%$	$19,53\%$	$25,06\%$	$9,67\%$	$14,17\%$	$19,76\%$
Colon(2000)	$62,11\%$	$66,22\%$	$78,03\%$	$53,47\%$	$58,84\%$	$74,48\%$
Lymphoma(4026)	$23,27\%$	$34,71\%$	$36,95\%$	$12,15\%$	$19,66\%$	$25,95\%$

Table 3.6: Mean improvement of LS operators applied to random solutions

A Blanket Markov LS enhanced by pruning capabilities: MB^+

The previously proposed operator AF^+ will be now used in a scheme based on Markov blanket local search [139]. In fact, the MBEGA local search operator will be enhanced

---

[12] mean values were reported

with operators based on pruning capabilities: AF^+ and BF^+. Algorithm 20 illustrates main steps of the LS operator MB^+ which is derived from the MB scheme. The BF operator could be seen as composite operator merging neighborhoods of the forward and the backward search procedures. BF^+ prunes LS neighborhood using the same threshold technique. Thresholds are set to the mean of the selected and the unselected attribute scores. The neighborhood of BF^+ (*see Algo.* 21) operator is pruned through the adjustment of the search on both directions.

Algorithm 20: MB operator based on pruning: MB^+

```
Input:
$S$: Input Solution
$F$: Filter
$Cla$: a classifier for solution evaluation
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow MB(S)$ /* MBEGA LS operator*/
    if ($S \succ S_{best}$) then
        $S_{best} \leftarrow AF^+(S_{best}, Cla, F)$ /* Algo. 19 */
    else
        $S_{best} \leftarrow BF^+(S_{best}, Cla, F)$ /* Algo. 21 */
    Return $S_{best}$
```

Empirical results

Tables C.6, C.7 and C.8 (*p*.178-180) detail evaluation and validation accuracies of the best solution found by the application of memetic algorithms respectively endowed with the AF^+ and MB^+ local search operators.

On the other hand, AF^+ and MB^+ were compared to the MBEGA algorithm and results were reported in Tables C.9, C.10, and C.11 (*p.* 181-183). All the proposed memetic approaches apply the local search operators to the best solution of the current generation. They were applied only once. Such limitation could be argued by the cost of the LS neighborhood exploration.

Algorithm 21: Improved version of BF: BF^+

```
Input:
$S$: Input Solution
$F$: Filter
$Cla$: a classifier for solution evaluation
Output:
$S_{best}$: Best solution within neighborhood
begin
    $UnSel_{list} \leftarrow getUnSelectedAtt(S)$
    $m \leftarrow getMeanScores(UnSel_{list}, F)$ , $n \leftarrow getMeanScores(S, F)$
    foreach ($x \in Sel_{all}$) do
        if ($Score(x, F) < Min(m, n)$) then
            $continue$
        if ($x \in S$) then
            $S1 \leftarrow S \backslash \{x\}$
        else
            $S1 \leftarrow S \cup \{x\}$
        $Evaluate(S1, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
    Return $S_{best}$
```

The two first operators AF^+ and MB^+ provide closes results to the basic genetic algorithms and to the memetic approach MBEGA. According to the fitness values the level of 0% was reached by the three compared approaches (GA and memetic based on AF^+ and MB^+) on different benchmarks. Nevertheless, error classification rates reported by the fitness column were not outperformed by either memetic algorithms. On the other hand, the some improvements were obtained with validation data sets and classifier although the gaps between evaluation and validation errors. Improvements were more frequent with best subsets size and running time. When we compare results of the AF^+ and MB^+ operators with two reference approaches, MB^+ seems to be more efficient than AF^+, while both operators provide slightly better results than GA but are not more effective than MBEGA. The major remarkable improvements confirmed by both operators were its abilities to reduce the number of attributes while keeping acceptable performances (in comparison to reference approaches).

FLS^+: Floating search enhanced by pruning capabilities

The floating search operator presented here is derived from the above presented floating schema. In fact, the FLS^+ preserves the main steps of the FLS operator (search direction, switching search direction rules and best solution updates). The changes is only operated on the process components. The forward and the backward searches are not explored in an exhaustive way. They couldn't be dependent on the number of attributes. The proposed forward and backward components are based on filter knowledge. Filter scores can guide the search to improve current solution. Algorithm 22 details main steps of the local search operator with the new components SF^+ and SB^+. In comparison to the classical forward and backward procedures, both SF^+ and SB^+ respectively described by Algorithms 23 and 24 add filter as a parameter. Next, they sort candidate attributes for add/delete. Thresholds are set to mean scores. Finally, the search starts with an order in relation with filter scores. The search stops when the threshold condition is not satisfied. Indeed, the filter knowledge is not only used to define the new neighborhood but also to define an order for solutions exploration. Such consideration is recommended specifically when the LS operator looks for a first improving solution.

$IFLS^+$: Improved floating local search enhanced by pruning capabilities

The second scheme based on pruning capabilities is inspired from the IFLS operator. As presented previously the improved floating search adds a new stage to the neighborhood exploration to diversify the search. The last step defines a neighborhood structure relying on attribute exchange mechanism. We propose for $IFLS^+$ three stages based on pruned neighborhood structures. The two first stages use SF^+ and SB^+, whereas the last stage adopts the pruning rules of the AF^+ operator. The Algorithm 25 illustrates the new operator derived from the floating scheme. Remark that the effective design of the basic operators make them reusable as pluggable components according the adopted scheme.

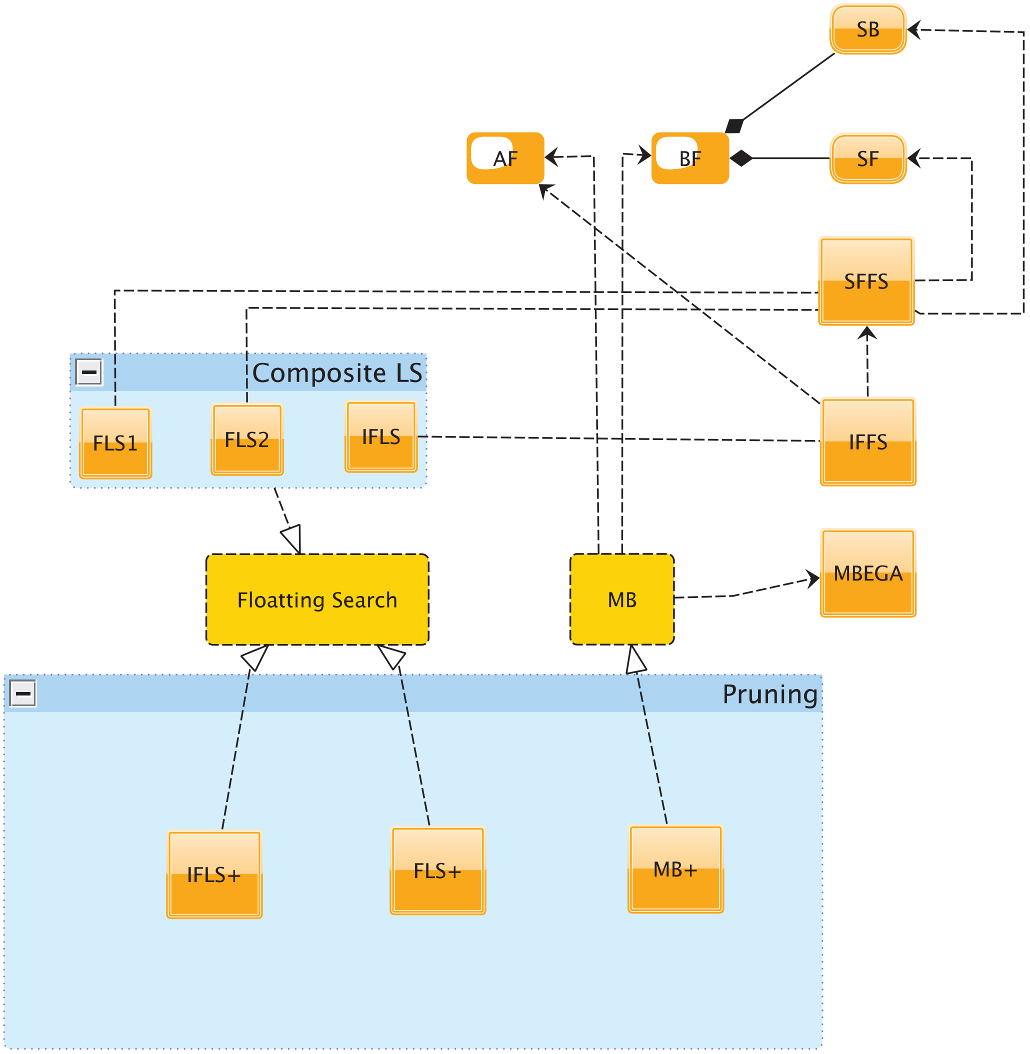


Figure 3.5: LS operators based on pruning

Algorithm 22: Pruned Floating LS operator: FLS^+

```
Input: $S$: Input Solution
$Cla$: Classifier for solution evaluation
$F$: Filter for attribute ranking
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$ ; $imp \leftarrow true$
    while ($imp$) do
        $S1 \leftarrow SF^+(S_{best}, Cla, F)$    /*Algo. 23*/
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            $imp \leftarrow false$
        while ($true$) do
            $S2 \leftarrow SB^+(S_{best}, Cla, F)$    /*Algo. 24*/
            if ($S2 \succ S_{best}$) then
                $S_{best} \leftarrow S2$
                $imp \leftarrow true$
            else
                Break
    Return $S_{best}$
```

Figure 3.5 illustrates local search operators proposed in this section as well as their relation with previously presented LS operators and reference approaches.

Empirical results

Tables C.12, C.13 and C.14 (*p.* 184-186) detail results of the FLS^+ and $IFLS^+$ comparison to GA, whereas, Tables C.15, C.16, C.17 and C.18 (*p.* 187-190) compare the same LS operators to MBEGA.

In comparison to GA the memetic schema based on the two local search operators FLS^+ and $IFLS^+$ outperform GA in more than one criterion. From the evaluation perspective the significant improvement is mainly related to the reduction of the

Algorithm 23: Pruned Forward Search LS operator: SF^+

```
**Input**: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
**Output**:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $UnSel_{list} \leftarrow getUnSelectedAtt(S)$
    $Threshold \leftarrow getMeanScores(UnSel_{list}, F)$
    foreach ($x \in UnSel_{list}$) do
        if ($Score(x, F) < Threshold$) then
            continue
        $S1 \leftarrow S \cup \{x\}$
        $Evaluate(S1, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
    Return $S_{best}$
```

number of features. Whereas the classification error rates reported by fitness values has been in most of cases outperformed by the reference approach with a slight gap. Besides, with some benchmarks optimal fitness values [13] were obtained by reference and proposed memetic schema (*i.e.* Ovarian, MLL and SRBCT dataset). Floating schema, based on FLS^+ and $IFLS^+$, applied to *Breast*, *Brain_Tumor* and *MLL* data sets succeeded to improve their accuracies on four criteria (both validation classifiers, CPU time and attribute number). Besides, at least one of the proposed floating schema succeeds to improve its results (with statistical validation) on eight benchmarks (almost half on the benchmarks). On the other hand, memetic floating schema, based on FLS^+ and $IFLS^+$, struggle to gain the same effectiveness when they are compared to the MBEGA approach. Mainly, both running time and intentionality reduction confirm a relative improvement. With *CNS* benchmark, both proposed operators improved the fitness, while with *Prostate* data set four criteria were improved

---

[13] Error rate of 0% for the fitness criterion

Algorithm 24: Pruned Backward Search LS operator: SB^+

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $Threshold \leftarrow getMeanScores(S, F)$
    foreach ($x \in S$) do
        if ($Score(x, F) \geq Threshold$) then
            continue
        $S1 \leftarrow S \backslash \{x\}$
        $Evaluate(S1, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
    Return ($S_{best}$)
```

by FLS^+. In comparison to the previously proposed operators, based on pruning techniques (AF^+ and MB^+), floating local search operators outperforms AF^+ and are at the same efficiency level of MB^+.

Globally, the four proposed operators outperform GA but provide relatively comparable results to MBEGA. The main advantages of the pruning approaches are its ability to reduce subset size and to provide some improvements on at least one validation criterion. The major pitfalls of the pruning approaches, are its convergence speed and the neighborhood exploration costs which are closely related. In fact, the more the exploration is expensive, the more the evolutionary search is perturbed, less effective and the trade-off between exploration and exploitation is broken. We should also note that the operators based on schema, are more effective than the ones which are not based on scheme AF^+.

Algorithm 25: Improved Floating LS operator with pruning: $IFLS^{+}$

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$ ; $imp \leftarrow true$
    while ($imp$) do
        $S1 \leftarrow SF^{+}(S_{best}, Cla, F)$    /*Algo. 23*/
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            $imp \leftarrow false$
        while ($true$) do
            $S2 \leftarrow SB^{+}(S_{best}, Cla, F)$    /*Algo. 24*/
            if ($S2 \succ S_{best}$) then
                $S_{best} \leftarrow S2$
                $imp \leftarrow true$
            else
                $S2 \leftarrow AF^{+}(S_{best}, Cla, F)$    /*Algo. 19*/
                if ($S2 \succ S_{best}$) then
                    $S_{best} \leftarrow S2$
                    $imp \leftarrow true$
                else
                    Break
    Return $S_{best}$
```

Stochastic LS operators

This section is devoted to the adaptation alternatives of the LS neighborhood exploration. In fact, the sequential exploration could be replaced by a stochastic selection selection mechanism. The stochastic selection might provide a tool to reduce the size of the neighborhood more effectively (in comparison to the pruning operators). In fact, the stochastic mechanism endows exploration process with diversification issues. In addition, the application of the LS operator could provide different enhanced solutions each time the operator is called.

Stochastic MB operator: MB^*

The operator is derived from the MB scheme. The local search involves stochastic components even for the last stage. The AF^* and BF^* operators replace pruning behavior by stochastic ones. The exploration order suggested by filter scores are replaced by stochastic selection mechanism which is not necessity obliged to explore attributes to add or delete in a predefined order. The neighborhood could be limited to a given number of solutions to explore. The LS operator is described by Algorithm 26. The last stage refers to AF^* and BF^*. AF^* and BF^* are based on $Add(.)/Del(.)$ procedures of the MBEGA local search (*see Algos.* 11 and 13)[14]. In order to delimit the neighborhood search space, threshold parameters are added to both AF^* and BF^*. Empirical study involves two variant of this operator. The first one explores all neighborhood solutions (limited by the threshold parameters), while the second stops after a first improvement.

Empirical results

Tables C.19, C.20, and C.21 (*p.* 191-193) detail evaluation and validation accuracies found by the application of memetic algorithms respectively endowed with two variants of the MB^* local search operator: the first explores the entire associated neighborhood and the second returns the solution found after the first improvement.

---

[14] p. 67 and p. 69

Algorithm 26: Fully Stochastic MB Operator: MB^*

```
Input:
$S$: Input Solution
$F$: Filter
$Cla$: a classifier for solution evaluation
$Nh$: Neighborhood size
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow MB(S)$    /* MBEGA LS operator*/
    if $(S \succ S_{best})$ then
        $S_{best} \leftarrow AF^*(S_{best}, Cla, F, Nh)$    /* Algo. 31 */
    else
        $S_{best} \leftarrow BF^*(S_{best}, Cla, F, Nh)$    /* Algo. 32*/
    Return $S_{best}$
```

Algorithm 27: Stochastic Forward LS operator: SF^*

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
$Nh$: Neighborhood size
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $i \leftarrow 0$
    while $(i < Nh)$ do
        $S1 \leftarrow S$
        $Add(S1, F)$    /*Algo. 11*/
        $Evaluate(S1, Cla)$
        if $(S1 \succ S_{best})$ then
            $S_{best} \leftarrow S1$
        $i++$
    Return $S_{best}$
```

Algorithm 28: Stochastic Backward LS operator: SB^*

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
$Nh$: Neighborhood size
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $i \leftarrow 0$
    while ($i < Nh$) do
        $S1 \leftarrow S$
        $Del(S1, F)$    /*Algo. 13*/
        $Evaluate(S1, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        $i++$
    Return $S_{best}$
```

According to evaluation criterion which is fitness value, both memetic schema provided close or equal results even for benchmarks with a fitness reaching 0%. In some cases (*i.e.* *Colon*, 9_*Tumors*, and *Challenge*_2004) the fitness values of GA were outperformed by memetic algorithms. The number of attributes was significantly reduced as well as running times. For these two criteria the performances were improved for all benchmarks with at least one of the two memetic variants. The results of 9 benchmarks were improved on four criteria. When we compare the two variants, the resulting performances are on the same level and sometimes the variant with first improvement outperforms those which explore the neighborhood entirely. Globally, the proposed stochastic MB schema outperforms GA.

The two stochastic variants of the MB scheme have been also compared to, the MBEGA algorithm and results were reported in Tables C.22, C.23, and C.24 (*p.* 194-196). When comparing fitness values, MBEGA slightly outperforms the two MB schema. Nevertheless, optimal solutions (according to the fitness criterion) were found by both reference and proposed memetic schema. Our best solutions were more compact and include smaller subset sizes. In some other cases, the fitness levels were close but the proposed memetic schema and LS operators succeeded to smaller subsets. Both subset size and running time criteria confirm the superiority of the stochastic MB schema. Six benchmarks results showed that the stochastic MB variant outperforms MBEGA on four criteria, while 10 benchmarks confirm its superiority on at least 3 criteria with one of the two variants.

Stochastic floating LS operator: FLS^* and $IFLS^*$:

The proposed operators inherit the main steps of the floating scheme (forward and backward search), with a slight change aiming to reduce the number of solutions to explore. In fact, the search stops when forward search did not succeed to improve the solution at hand. Indeed, backward stage is only applied when forward stage succeed to add an attribute improving feature subset fitness. The search components are fully stochastic and rely on basic $Add(.)/Del(.)$ procedures. The exploration procedures are guided by both stochastic selection and a threshold limiting the size of the neighborhood. New forward and backward procedures are respectively illustrated by algorithms 27 and 28. FLS^* and $IFLS^*$ are two variants of the scheme are described by Algorithms 29 and 30. $IFLS^*$ requires a stochastic attribute exchange operator AF^* for the last stage. The stochastic version of AF is illustrated by Algorithm 31.

Empirical results

Tables C.25, C.26 and C.27 (*p.* 197-199) detail evaluation and validation performances obtained by the application of memetic algorithms respectively endowed with the FLS^* and $IFLS^*$ local search operators. The reported results showed the superiority of the memetic alternatives based on FLS^* and $IFLS^*$ in comparison to GA on more than one criteria. According to the attributes number and running time criteria, results were improved in almost all the cases. For the validation criteria, classification accuracies were improved on several benchmarks. Results of four criteria were improved on seven benchmarks. Except two benchmarks, the proposed floating

Algorithm 29: Stochastic Floating LS operator: FLS^*

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
$Nh$: Neighborhood size
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    while ($true$) do
        $S1 \leftarrow SF^*(S_{best}, Cla, F, Nh)$    /*Algo. 27 */
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            Return $S_{best}$
        while ($true$) do
            $S2 \leftarrow SB^*(S_{best}, Cla, F, Nh)$    /*Algo. 28 */
            if ($S2 \succ S_{best}$) then
                $S_{best} \leftarrow S2$
            else
                Break
```

schema outperforms the GA on at least 3 criteria.

On the other hand, FLS^* and $IFLS^*$ were compared to the MBEGA algorithm and results were reported in Tables C.28, C.29 and C.30 (p. 200-202). Globally, proposed floating search alternatives is less effective than MBEGA. Nevertheless, the gap in results is not so important, and in numerous cases performances are comparable. Only the results of three benchmarks confirm the superiority of stochastic floating alternatives on four criteria.

In comparison, to both reference approaches, stochastic floating search operators outperform GA and are close to MBEGA. Globally, among the proposed stochastic operators, schema based on MB behave more accurately than those based on floating

Algorithm 30: Stochastic alternative of the IFLS operator: $IFLS^*$

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
$Nh$: Neighborhood size
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    while ($true$) do
        $S1 \leftarrow SF^*(S_{best}, Cla, F, Nh)$    /*Algo. 27 */
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        else
            Return $S_{best}$
        while ($true$) do
            $S2 \leftarrow SB^*(S_{best}, Cla, F, Nh)$    /*Algo. 28 */
            if ($S2 \succ S_{best}$) then
                $S_{best} \leftarrow S2$
            else
                $S2 \leftarrow AF^*(S_{best}, Cla, F, Nh)$    /*Algo. 31 */
                if ($S2 \succ S_{best}$) then
                    $S_{best} \leftarrow S2$
                else
                    Break
```

scheme, although that the floating search operators are comparable to reference approaches and outperform them on several benchmarks and criteria. The comparison of the stochastic operators to the operators based on neighborhood pruning showed that stochastic alternatives are more effective than pruning ones. This could be explained by the improvement on the neighborhood exploration.

Stochastic LS operator without scheme: AF^* and BF^*

Stochastic $Add(.)/Del(.)$ procedures (*see Algos.* 11 and 13) of the MBEGA LS could be used to design LS operators independent of the previously presented schema. Algorithms 31 and 32 implement Attribute flip and Bit flip operators in a stochastic way with a limited neighborhood size. Two versions of the AF^* and BF^* operators were deployed with memetic algorithms. For each LS operator a simple and iterative version were tested. The four combinations were also compared to the previously proposed stochastic LS operators. By this way, the scheme contribution would be easy to assess.

Empirical results

Tables C.31, C.32 and C.33 (*see p.* 203-205) detail evaluation and validation accuracies of the best solution found by the application of memetic algorithms respectively endowed with the AF^* and BF^* local search operators.

MAs based on AF^* and BF^* were, also compared to the MBEGA algorithm and results were reported in Tables C.34, C.35 and C.36 (*see p.* 206-208). Globally the MA based on both the AF^* and BF^* are more effective than the GA, however it behaves less better than the MBEGA. In comparison to local search operators based schema, the AF^* and BF^* are not well performing than operators based to stochastic MB schema. This could be explained by the composite neighborhood structure as well as stochastic components of the proposed operators. On the other hand, stochastic LS operators without schema are more effective than pruning local search alternatives. Iterative version of the local search operators the AF^* and MB^*

Algorithm 31: Stochastic AF LS operator: AF^*

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
$Nh$: Neighborhood size
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $i \leftarrow 0$
    while ($i < Nh$) do
        $S1 \leftarrow S$
        if ($RandBoolean()$) then
            $Add(S1, F)$
            $Del(S1, F)$
        else
            $Del(S1, F)$
            $Add(S1, F)$
        $Evaluate(S1, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        $i++$
    Return $S_{best}$
```

Algorithm 32: Stochastic BF LS operator: BF^*

```
Input: $S$: Input Solution
$Cla$: a classifier for solution evaluation
$F$: Filter for attribute ranking
$Nh$: Neighborhood size
Output:
$S_{best}$: Best solution within neighborhood
begin
    $S_{best} \leftarrow S$
    $i \leftarrow 0$
    while ($i < Nh$) do
        $S1 \leftarrow S$
        if ($RandBoolean()$) then
            $Del(S1, F)$
        else
            $Add(S1, F)$
        $Evaluate(S1, Cla)$
        if ($S1 \succ S_{best}$) then
            $S_{best} \leftarrow S1$
        $i++$
    Return $S_{best}$
```

were also assessed and compared to the GA in Tables C.37, C.38 and C.39 (*p.* 209-211), and to MBEGA in Tables C.40, C.41 and C.42 (*p.* 212-214). Reported results showed a slight improvements over on iterative version of the same operators. They are clearly more effective than GA, and the performances in comparison to MBEGA are acceptable but the number of attributes was not reduced significantly.

3.4.4 Summary of experimental results

This section is devoted to the review of the performance of a set of memetic approaches involving the local search operators proposed in section 3.4.3. The assessement procedure is based on the five criteria used throughout the empirical study (*i.e.* fitness, validation classifier1, validation classifier2) which involve metrics in relation with evaluation as well as validation processes. The comparison takes into account only improvements (compared to reference approach) statistically validated. In fact, for each local search and criterion we count the number of times the improvement has been validated. Besides, the memetic alternatives were also assessed on all the benchmarks [15] used in this thesis, including small, medium and large benchmarks. Table 3.7 summarises the local search operators (proposed in this section) as well as their common properties 3.4.3.

	LS-criteria			
LS operators	**MB**	**Floatting search**	**Pruning**	**Stochastic add/Del**
AF^+			•	
MB^+	•		•	
FLS^+		•	•	
$IFLS^+$		•	•	
MB^*	•			•
$MB^*(FI)$	•			•
FLS^*		•		•
$IFLS^*$		•		•
AF^*				•
BF^*				•
$Iter.AF^*$				•
$Iter.BF^*$				•

Table 3.7: Local search operators classification

---

[15] 23 benchmarks

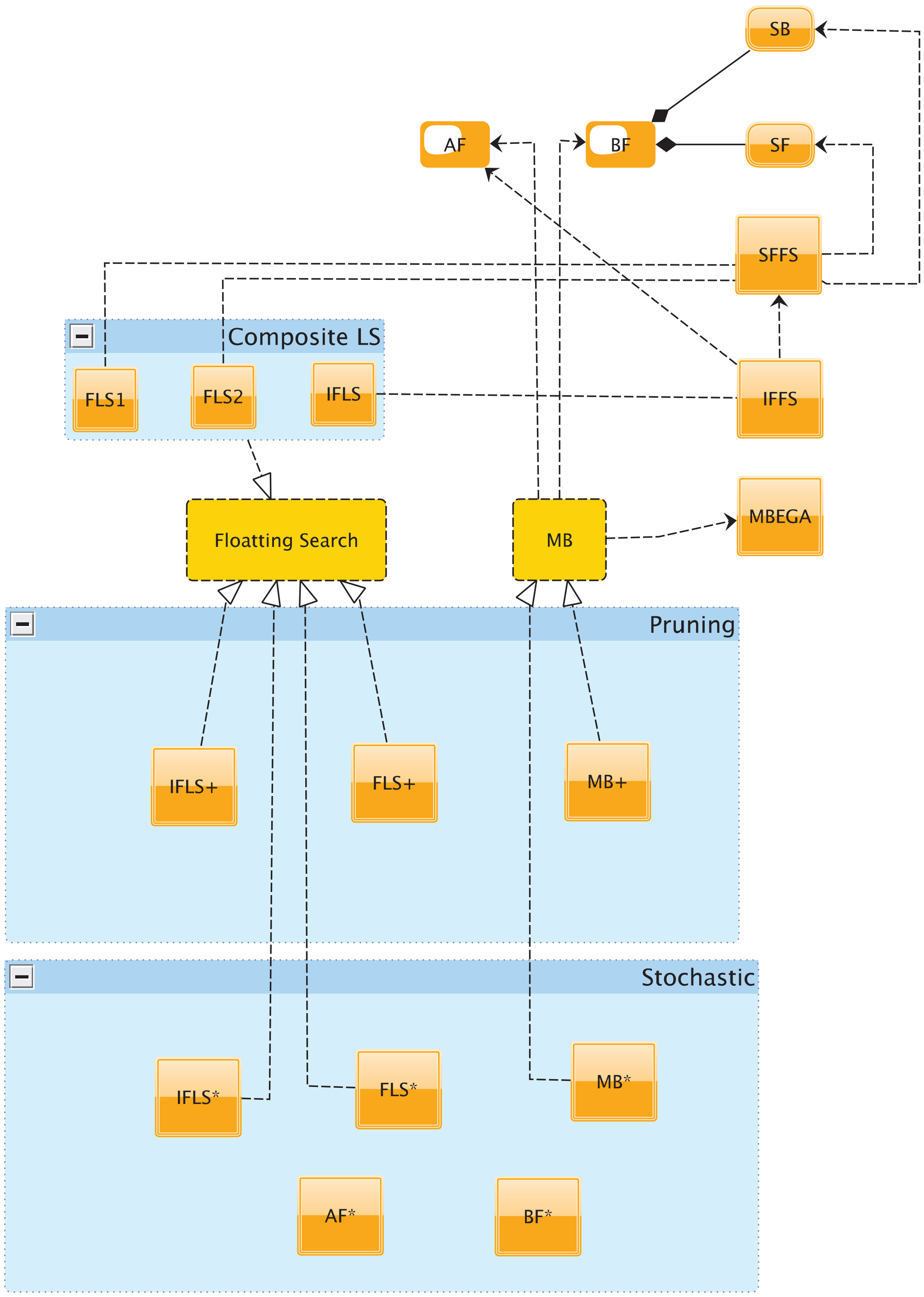


Figure 3.6: Proposed LS operators and its relation with reference approaches

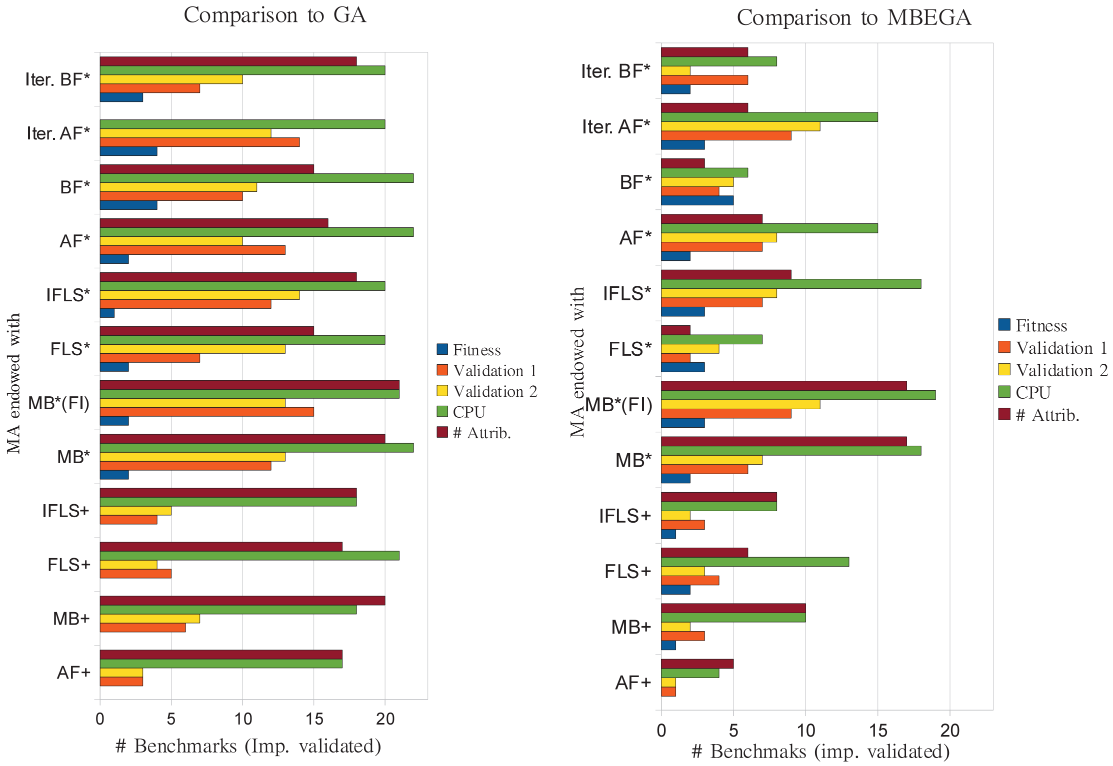


Figure 3.7: MA comparison to GA and MBEGA

Figure 3.7 graphically illustrates performance improvements compared to reference approaches GA and MBEGA.

The big picture, shows the superiority of almost all proposed memetic alternatives over GA. Some of these memetic approaches clearly outperform the MBEGA. However the gap is less important with MBEGA than GA. MA based on MB^+ seems to be the more effective memetic alternative and memetic algorithms based on a scheme are respectively more effective than those belonging to the same family (*i.e.* pruning, stochastic) without any schema (*i.e.* AF^+ *vs* FLS^+, MB^* *vs* BF^*, FLS^* *vs* $It.BF^*$). Such result confirms the fact that composite neighborhood structure for local search operators is very interesting. In fact, even local search requires diversification mechanisms. Besides, when we compare the pruning to stochastic LS operators, the first seems to be more effective on more than one criterion.

On the other hand, we summarize the experiments of the local search operators proposed throughout this chapter, by only considering two criteria: fitness and the number of selected attributes. In fact, each local search devised and endowed within a memetic scheme is compared to a reference approach (*GA or MBEGA*). Operators behaviors or tendencies are assessed according to their statistical validated results for all benchmarks. The results provided by the following Table (*see Fig.* 3.8) illustrates for each Ls operators the percentages of benchmarks for which the its better [16], worse[17] and have a comparable results[18] with its respective reference approach. For example, first row compares GA to an MA using FLS1 as local search. for the finess criterion, 83% of the benchmarks confirms the superiority of GA and the remaining 17% provides comparable results, whereas MA suceeds to reduce selected attribute numbers in 50% percent of the cases.

Globally with fitness criterion, better results are in favor of reference approaches. In many cases, proposed apporaches provides results comparable to reference approaches (*i.e.* stochastic operators). However, when we look only to the results of the second criterion (selected subset size) we can clearly see the superiority of almost all the proposed operators. Now when we consider both criteria and we compare results, we can depict that operators which succeed to provide comparable fitness, and at the same time outperform reference approaches on the second criterion are necessarily quite better and more interesting to investigate in depth. Such tendency covers almost the second part of the table (stochastic operators). Such result is particularly interesting for large benchmarks when the fitness is close to optimal value (very low classification error rate). We have previously shown that for many large benchmarks, when the reference approach provides fitness equal to 0%, some of the LS operators succeeds to provide same fitness with a reduced number of attributes.

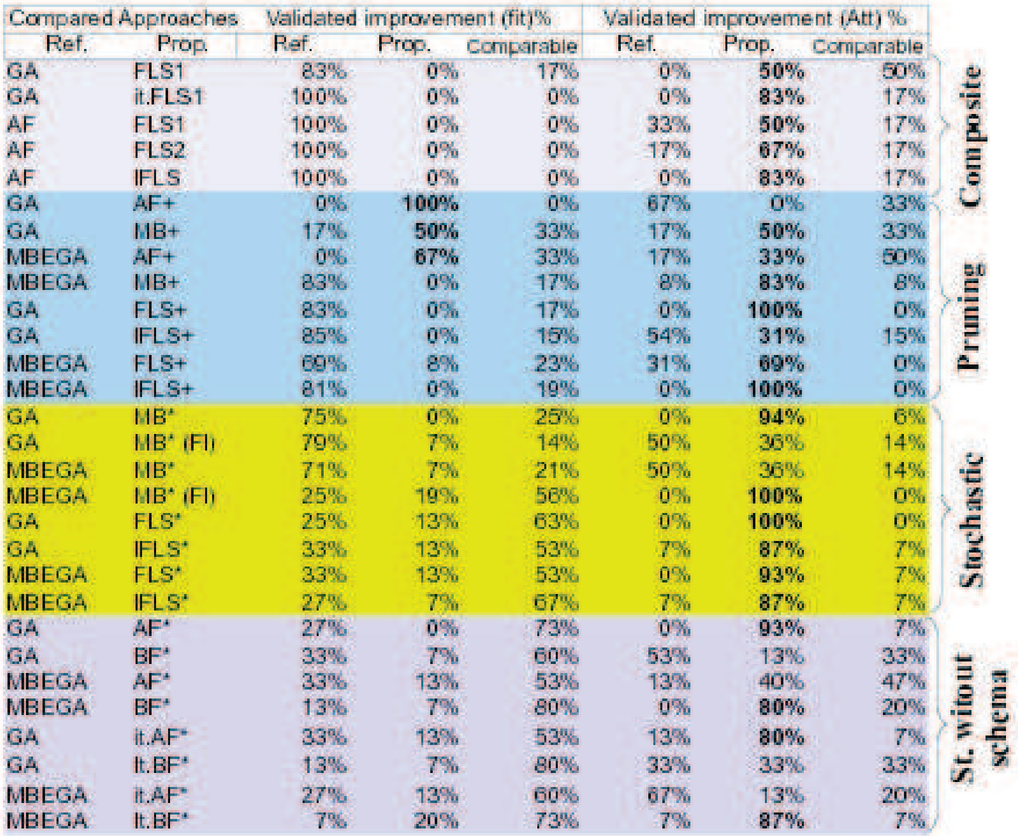
The second part of the figure (*see Fig.* 3.8) graphically illustrates local search operators results aggregated by family. It is evident that the four families confirm the superiority of LS operators for subset size reduction. In addition, it is interesting to see, with fitness criterion the evolution of portions of comparable results and those

---

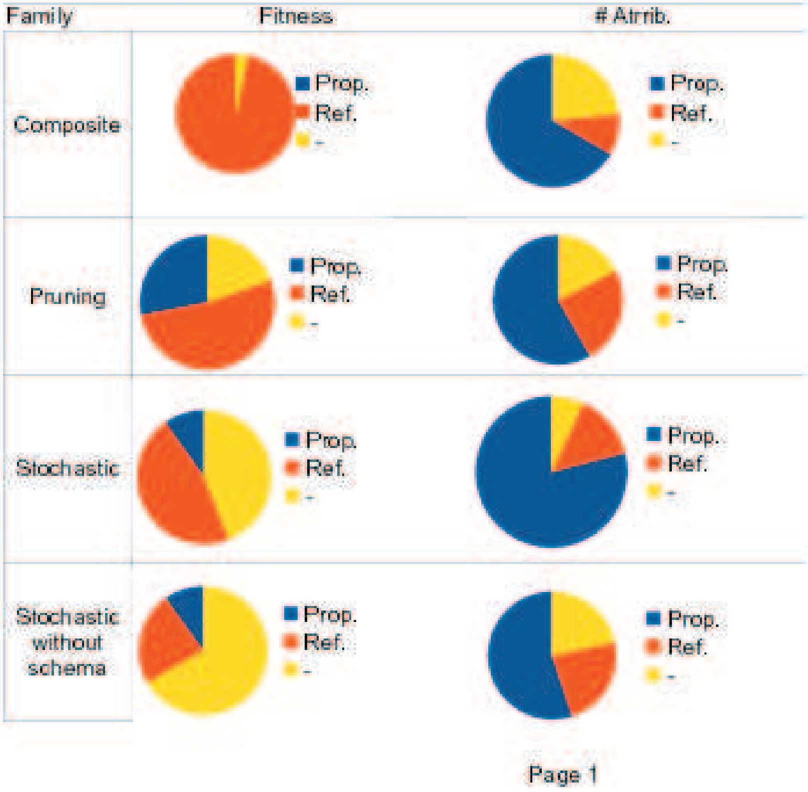
[16] improvement over 10 runs (at least) is statistically validated

[17] superiority of reference approach over 10 runs (at least) is confirmed statistically

[18] statistical test is not in favor of any approach



Compared Approaches		Validated improvement (fit)%			Validated improvement (Att) %			
Ref.	Prop.	Ref.	Prop.	Comparable	Ref.	Prop.	Comparable	
GA	FLS1	83%	0%	17%	0%	**50%**	50%	Composite
GA	it.FLS1	100%	0%	0%	0%	**83%**	17%	
AF	FLS1	100%	0%	0%	33%	**50%**	17%	
AF	FLS2	100%	0%	0%	17%	**67%**	17%	
AF	IFLS	100%	0%	0%	0%	**83%**	17%	
GA	AF+	0%	**100%**	0%	67%	0%	33%	Pruning
GA	MB+	17%	**50%**	33%	17%	**50%**	33%	
MBEGA	AF+	0%	**67%**	33%	17%	33%	50%	
MBEGA	MB+	83%	0%	17%	8%	**83%**	8%	
GA	FLS+	83%	0%	17%	0%	**100%**	0%	
GA	IFLS+	85%	0%	15%	54%	31%	15%	
MBEGA	FLS+	69%	8%	23%	31%	**69%**	0%	
MBEGA	IFLS+	81%	0%	19%	0%	**100%**	0%	
GA	MB*	75%	0%	25%	0%	**94%**	6%	Stochastic
GA	MB* (FI)	79%	7%	14%	50%	36%	14%	
MBEGA	MB*	71%	7%	21%	50%	36%	14%	
MBEGA	MB* (FI)	25%	19%	56%	0%	**100%**	0%	
GA	FLS*	25%	13%	63%	0%	**100%**	0%	
GA	IFLS*	33%	13%	53%	7%	**87%**	7%	
MBEGA	FLS*	33%	13%	53%	0%	**93%**	7%	
MBEGA	IFLS*	27%	7%	67%	7%	**87%**	7%	
GA	AF*	27%	0%	73%	0%	**93%**	7%	St. witout schema
GA	BF*	33%	7%	60%	53%	13%	33%	
MBEGA	AF*	33%	13%	53%	13%	40%	47%	
MBEGA	BF*	13%	7%	80%	0%	**80%**	20%	
GA	it.AF*	33%	13%	53%	13%	**80%**	7%	
GA	It.BF*	13%	7%	80%	33%	33%	33%	
MBEGA	it.AF*	27%	13%	60%	67%	13%	20%	
MBEGA	It.BF*	7%	20%	73%	7%	**87%**	7%	



for which selected devised approaches are better. The third family seems to provide the better compromise. Finally the comparison to the the third and the fourth family shows that the use of a scheme which enlists neighborhood diversification participates, in a part to enhance dimensionality reduction. Figure 3.6 illustrates local search operators proposed in this section as well as their relation with previously presented LS operators and reference approaches.

3.5 Conclusion

Throughout this chapter which was devoted to the memetic modeling we designed and assessed a set of memetic approaches. The associated local search operators were proposed according to the requirement of the problem of feature selection modeling. A set of operators were adapted to high dimensional spaces. Some of them have proven their efficiency in comparison to memetic reference approaches. The aim was to develop evolutionary algorithms that are able to find good trade-offs between exploration and intensification through hybridization schema. Although that some of the LS operators provided comparable fitness results we succeed, in many cases, to confirm the ability to reduce the number of selected attributes. We also succeed to adapt and hybridize some heuristics for which the applications were limited to small benchmarks. The next chapter focuses on another hybridization issue targeting the effective combination of the wrapper-filters approaches using another optimization paradigm.

Chapter 4

Effective Wrapper-Filter Integration Through a GRASP Modeling

4.1 Introduction

In this chapter, we propose, a new hybrid search technique through the adaptation of GRASP approach to the FS problem. The devised approach investigates the effective wrapper-filter combination by exploiting the intrinsic properties of the GRASP heuristic. The main motivations for this proposal are three folds: (i) filter-wrapper collaboration might enhance the relevance of the selected feature subsets. Effective combination may lead to a schema or a framework allowing the use of both approach advantages (ii) local search approaches have shown their effectiveness in FS as well with sequential deterministic procedures (*i.e.* SFFS [120], IFFS[101], *etc*) as with stochastic approaches (*i.e.* memetic [34], Simulated Annealing [97] and Tabu search [128]). (iii) endowing, respectively, filters and wrappers with stochastic diversification and guidance mechanisms to alleviate FS, challenging problems like local minima and nesting effect [54, 90].

The main contributions of this chapter are the investigation of the GRASP metaheuristics as a scheme for the FS modeling as well as the study of the behavioral aspects in relation with adaptation to high dimensional FS problems.

Section 2 formalizes and reviews the optimization paradigm behind the GRASP and gives an overview of representative components, associated approaches, and applications. Section 3 details the proposed GRASP-FS apporaches and compares it to a reference approach. Section 4 is devoted to the adaptation of the GRASP-FS approaches to the high dimensional problems, since the local search is considered as a main component of the investigated metaheuristic. We study some of the behavioral aspects in relation to the adequation of the construction phase to the local search operators. Finally, Section 5 concludes the chapter and provides some directions of future research in relation to GRASP modeling.

4.2 GRASP Metaheuristic Comprehensive Overview

This section introduces GRASP heuristic principles, details components and sheds some lights on featured enhancements and variants. Next, we survey recent application of GRASP to the FS problems as well as the study and the discussion of the FS modeling challenges with GRASP.

4.2.1 The metaheuristic big picture

The Greedy Randomized Adaptive Search Procedure (GRASP) is meta-heuristic for combinatorial optimization problems [37, 38]. It was initially proposed as a probabilistic heuristic for the set covering problem [36]. Usually, known as multi-start procedure, GRASP is based on an iterative process which constructs a solution then fine-tune it, through the exploration of its neighboring solutions.

Algorithm 33 illustrates the basic steps of the GRASP. The iterative process is made up of (i) a *construction phase*, where a feasible solution is greedily built (*i.e.* incrementally), (ii) next, a *local search phase* starts at the constructed solution and iteratively updates the solution until a locally optimal solution is found. The best overall solution is kept as result.

Algorithm 33: GRASP: basic steps

```
Input:
Specific problem parameters
Output:
$S_{best}$: Best solution
begin
    $Solution \quad S \leftarrow \emptyset$
    $S_{best} \leftarrow S$
    ReadInputs();
    while ($Stopping\_condition\_not\_met$) do
        $S \leftarrow GreedyRandomizedSolutionConstruction()$
        $S \leftarrow LocalSearch(S)$
        UpdateSolution(S, $S_{best}$)
    Return ($S_{best}$)
```

The multi-start property enlarges the search coverage by exploring different regions of the search space without being influenced by the previous solutions found. The GRASP heuristic is endowed with both global search mechanisms (multi-start scheme and the stochastic construction stage) allowing diversity during the search and intensification mechanisms implemented by the local search operators.

GRASP was successfully applied to numerous problems ranging from fundamental and classical optimization problems (*i.e.* routing [20], timetabling [55] and scheduling [6]), to industrial applications (transportation [10], telecommunications [47], manufacturing [18]).

Readers can refer to [38], for a recent GRASP survey and to [39] for annotated literature bibliography.

4.2.2 GRASP components

This section examines, in depth, components of the GRASP optimization scheme [37] as well as its main features and behavioral aspects. GRASP involves an iterative process which consists of two stages: *the construction* of a feasible solution and the *local search*. The two next paragraphs detail construction mechanisms and investigate refinement principles.

Construction Stage

The aim of this stage is to build a solution S from a set of candidate C elements. Solutions are iteratively constructed, starting from an empty set. Elements are incrementally added to the solution (*i.e.* one element at each iteration). A greedy function $g(.)$ could be used to measure the benefit of the selection of an element (or the cost induced by the add of an element to a solution, in the case of a minimization problem). Such a greedy mechanism used as a GRASP construction stage might provide the same solution for the local search procedure. In other words, a deterministic greedy function would generate always the same initial solution for the second stage. One can think of a construction stage based on a totally random generation process. Too much randomness would transform the GRASP in an iterative LS procedure. Besides, random solutions would require higher intensification effort (second stage) in comparison to solutions greedily generated. Consequently the convergence is relatively slower. A trade-off between diversity level and solution generation greediness should be guaranteed to allow effective construction stage design.

A non deterministic procedure is used by GRASP to select elements among best candidates. Hence, solution elements are selected from a Restricted Candidate List (RCL), and the greedy *random construction mechanism* is implemented according to one of these two alternatives [111]:

- either, by using greediness to generate RCL and randomness to select elements from it,
- or, by using randomness to build RCL and greediness for selection.

An example of a construction mechanism, for a minimization problem based on greediness and randomness is illustrated by algorithm 34 [38]. A threshold α is used to define the restricted list (Line 8). Once an element is added to the solution S, both candidate list C and RCL are updated.

Algorithm 34: Construction of a Greedy Randomized Solution (minimization problem)

```
Input:
$\alpha$: Threshold
Output:
$S$: Feasible solution
begin
    $S \leftarrow \emptyset$
    Initialize the candidate set $C$ by all elements
    foreach ($i \in C$) do
        Evaluate the incremental cost $g(i)$
    while (|C| > 0) do
        $g_{min} \leftarrow min_{i \in C} g(i)$; $g_{max} \leftarrow max_{i \in C} g(i)$;
        $RCL \leftarrow \{i \in C \mid g(i) \leq g_{min} + \alpha(g_{max} - g_{min})\}$
        $v \leftarrow SelectElement(RCL)$    /*v is removed from RCL */
        $S \leftarrow S \cup \{v\}$
        $Update(C)$
        foreach ($i \in C$) do
            Evaluate the incremental cost $g(i)$
    Return ($S$)
```

Refinement Stage

The first stage resulting solution will be used as initial solution for LS refinement, its neighborhood is investigated until a local minimum is found. The best overall solution is kept as result.

Throughout the second stage solutions are iteratively refined by local search procedures. It could be made of a unique or multiple local search operators. The intensification mechanism could also rely on a composite neighborhood exploration to escape solutions locally optimal. We recall the main steps of the iterative local search (LS) procedure is with Algorithm 35.

The multi-start property of GRASP allows the search process to be not trapped in a local optimum and to explore different regions of the search space, without being constrained or influenced by the best solution found.

Algorithm 35: Pseudo-code of a L.S. Algorithm

```
Input:
$S$: Solution
Output: $S_{best}$: Improved Solution
begin
    $s \leftarrow S$ ; $S_{best} \leftarrow s$
    repeat
        $NH_s \leftarrow GenerateNeighborhood(s)$
        foreach $s_i \in NH_s$ do
            if $(pivot\_condition\_satisfied \wedge s_i > S_{best})$ then
                $S_{best} \leftarrow s_i$
        $s \leftarrow S_{best}$
    until $depth\_condition\_satisfied$;
    Return $S_{best}$
```

4.2.3 GRASP Variants

In this section, we survey some featured approaches derived from the basic GRASP as well as successful components enhancements. In fact, GRASP lacks of learning mechanisms since it not use the history of solutions found in the previous iterations. Such memory can be used to avoid redundant neighborhood exploration. For example, a hash table was suggested to save the solutions of the construction phase. Consequently, only new solutions that were not present on the hash table were added to it, and were considered on the GRASP second stage.

In *Reactive GRASP* [108], The RCL parameter α (*see Algo.* 34) was adjusted according to the evolution of the process. In fact, a learning mechanism was implemented in the constructed phase. The single value of the α parameter was replaced by a set of a discrete possible values. A probability was associated to each α_i value. Throughout iterations, α_i is being selected according to its associated probability. Probabilities are updated, in a manner, to favor values that have led to better solution in the previous iterations. Another, intensification procedure based on the best explored solutions was proposed by *path-relinking* mechanism. *Path-relinking* was originally proposed by *Glover* [45] as a strategy connecting elite solutions obtained by Tabu and Scatter search. The idea is based on the exploration of solutions leading from one locally optimal solution to the another one(s) obtained on the previous iterations. The new solutions connecting two elite solutions forms a *path*. According to [38] and [111], such mechanism leads to significant enhancement in solution quality. For two elite solutions e_1 and e_2, different paths generation alternatives (connecting e_1 to e_2) were proposed (*i.e.* forward relinking, backward relinking, mixed relinking, *etc*) [111]. It was applied to locally optimal solutions provided by the GRASP local search stage. The first use of path relinking mechanism as GRASP components was in 1999 [80]. Since then, several enhancements and hybridization alternatives were devised. Two main strategies were adopted: (i) path-relinking was applied at the end of each GRASP iteration between current locally optimal solution and best solution found; (ii) or, it was applied to all elite solutions pairs either periodically or at the end of the GRASP process [111].

In addition to *Reactive*-GRASP alternatives and *path-relinking* intensification procedures, the GRASP scheme was also combined with other optimization paradigms (*i.e.* GA, Tabu, VNS, *etc*) [111]. For example, it was combined with Tabu search in [26] where the second stage was replaced by the Tabu heuristic.

4.2.4 GRASP for FS

The first application of GRASP to the FS problem was, recently proposed by *Yusta* in [133]. The proposed GRASP was compared to effective FS search techniques like GA, Tabu search and floating search (SFFS).

The GRASP proposed in [133], is illustrated by Algorithm 36. The algorithm is based on two main stages, namely solution construction (Lines 6-21) and local search procedure. The first stage constructs n_{max} solutions, and the best one will be selected as a candidate for the second stage. Solutions are constructed according to the attributes selected within the RCL list. The RCL is based on the *In-Group Variability* criterion (*see eq.* 4.1).

$$IGV(f_j, C) = \sum_i (f_j^i - \mu_{C(i)})^2 \tag{4.1}$$

Where f_i^j and $\mu_{C(i)}$ denote respectively the *i-th* value of the attribute f_j and the mean $\mu_{C(i)}$ of f_j values for the instances (data sets rows) belonging to the same class as the instance i. Besides, the attribute selection, is controlled by the parameter α (Lines 11-17). In fact, it controls the degree of randomness of the procedure.

The second stage iteratively applies AF operator to the solution provided by the first stage[1].Each iteration generates the neighborhood of the current solution and exchanges current solution with best neighbor if it can improve classification accuracy (solution fitness). The neighborhood structure proposed, by Yusta in [133], is based on attribute replacement. The local search procedure iteratively applies AF operator. (*see eq.* 3.1 *p.* 64)

Reported experiment results have shown the superiority of GRASP in comparison to GA and meta-heuristics based on the same LS operator (GA, Tabu).

4.2.5 GRASP modeling challenges

Since the GRASP metaheuristic is based on two main stages, several aspects should be taken into consideration for each stage to the adapt them to the requirements and the specificities of the FS modeling.

From one hand the construction stage could be designed more effectively using incremental wrappers or filter scores. The local search operators could be replaced by more enhanced local search operator like those used in the previous chapter. Furthermore, the GRASP second stage proposed in [133] could be only applied to small dimensional spaces.

The neighborhood structure $NH(S)$, of [133], considers all combinations of attribute exchange. Consequently, LS is sensitive to the number of selected features. The neighborhood exploration becomes prohibitive even for moderate value of n. The computational complexity is in the order of $\Theta(p * m)$ [2] (for a non iterative LS).

On the other hand, the GRASP should guarantee or preserve a kind of trade-off between search diversification and intensification capabilities. In other words, a GRASP scheme based on a LS operator requiring intensive neighborhood exploration

---

[1] LS is repeated until a locally optimal solution is found

[2] p and m respectively denote the number of selected and non-selected features ($p + m = n$).

Algorithm 36: The G.R.A.S.P. proposed in [133]

```
Input:
$F$: Initial Feature set
$C$: Target class Attribute
$\alpha$: Threshold
$d$: number of attributes to select
$n_{max}$: number of solutions to generate
Output:
$S_{best}$: Selected Features
begin
    $S \leftarrow \emptyset$
    $S_{best} \leftarrow S$
    while (Stopping Criterion not Satisfied) do
        //Construction stage
        foreach ($f_i \in F$) do
            $g_i \leftarrow IGV(f_i, C)$
        $Sol_{list} \leftarrow \emptyset$
        repeat
            $S \leftarrow \emptyset$
            repeat
                $min \leftarrow argmin_i(g_i)$, $max \leftarrow argmax_i(g_i)$
                $RCL_{list} = \{v_j| \quad g_j \leq \alpha.g_{max} + (1-\alpha)g_{min}\}$
                Randomly select $v_j \in \{v_j \in RCL_{list}, v_j \notin S\}$
                $S \leftarrow S \cup \{v_j\}$
                $RCL \leftarrow RCL \setminus \{v_j\}$
            until ($|S| = d$);
            $Evaluate(S, Cla)$
            $Sol_{list} \leftarrow \{S\} \cup Sol_{list}$
        until ($|Sol_{list}| = n_{max}$);
        $S \leftarrow$ getBest($Sol_{list}$)
        // iterative local search
        $S \leftarrow LocalSearch(S)$    /*$AF$ Operator */
        if ($S \succ S_{best}$) then
            $S_{best} \leftarrow S$
    Return ($S_{best}$)
```

may result in a behavior comparable to a hill climbing or one GRASP iteration, particularly when the process is bounded a number of evaluations. GRASP modeling challenges as well as metaheuristic design requirements can be summarized as follows:

- effective neighborhood investigation[3],
- Construction stage providing acceptable diversification level,
- Guidance mechanism for the construction stage.

4.3 An effective GRASP scheme for FS: GRASP-FS

In this section, we investigate, the proposed new GRASP schema for FS. We focus on a set of a devised local search operators [4] and filters in the aim to adapt and deploy them as components within GRASP schema.

Since the GRASP scheme is based on a restricted list of candidates, this list could be represented by features that seem to be, individually, relevant or those that might provide incremental usefulness to the selected feature subset. For the GRASP construction stage we opt for selection scheme capable of generating attribute ranking. Hence, features scores will serve as selection criterion for the RCL generation. The construction stage generates solutions from the RCL using a random selection mechanism. The second stage of GRASP enhances solutions by an iterative neighborhood exploration. The quality of solution fine-tuning, mainly, depends on the nature of the involved neighborhood structure of the LS operators.

We devise a number of LS procedures based on different neighborhood structures inspired from well known sequential search procedures. The following two sections, detail different design alternatives for both RCL and local search GRASP components.

---

[3] Not necessarily intensive exploration

[4] effective operators of the previous chapter

4.3.1 Construction stage: RCL generation

Comparatively to the GRASP approach proposed by Yusta in [133], construction phase steps (*see Algo.* 37) are based on the selection of the best solution among n_{max} ones generated. Each solution randomly selects candidates from an RCL which is made up of attributes of acceptable filter scores.

Any filter criterion could be, instead, used to build RCL. In this chapter, we opt for four well known and different selection metrics: χ^2, *ReliefF* [112], *Symmetrical Uncertainty* (SU) [54], and *FCBF* [130].

Typically, filters return solutions based on the selection of features with the highest scores. A threshold is used (th) to define the score level allowing attributes to be selected within RCL (*Algo* 37 *Line* 4-6). Once the initial RCL is generated[5], the variables are randomly selected to build GRASP first stage solutions. Such a selection scheme has, at least, three benefits: (i) reducing the risk of selection of, only, highly correlated relevant features; (ii) the combination of features with moderate usefulness, which are not highly relevant to the target, might promote interaction among selected attributes; (iii) in comparison to the construction stage of the GRASP of [133], the parameter α which was used to define selection randomness is replaced by a stochastic mechanism in direct relation with specific problem knowledge implemented by random selection of attributes with scores above a given threshold.

Besides, the construction mechanism is bounded by a parameter n_{max} to limit the evaluation cost. Such limitation is less sensitive to the number of the data set attributes, in comparison to greedy construction mechanisms and particularly incremental wrappers and sequential search procedures. The RCL list is rebuilt in [133], in each iteration, according to the condition of *line 13* of Algorithm 36, whereas in our approach it's constructed once then attributes are randomly selected. The update or rebuild computational cost is reduced in comparison to the reference approach.

---

[5] using filter criterion

Algorithm 37: Construction stage

```
Input:
$th$: Threshold
$F$: Filter
$A$: Attribute list
$n_{max}$: Iterations number
$Cla$: Classifier
Output:
$S_{best}$ : Feasible solution
begin
    /* building RCL*/
    $RCL \leftarrow \emptyset$
    foreach ($d \in A$) do
        if ($F.score(d) \geq th$) then
            $RCL \leftarrow RCL \cup \{d\}$
    /* Solution generation*/
    $i \leftarrow 0$
    while ($i < n_{max}$) do
        /*Random(n) generates a random int value in $[0..n-1]$*/
        $n \leftarrow min(Random(RCL.size()) + 1, A.size()/2)$
        /*Selects randomly n distinct elements from RCL*/
        $S \leftarrow RandomSelection(RCL, n)$
        $Evaluate(S, Cla)$
        if ($S \succ S_{best}$) then
            $S_{best} \leftarrow S$
        $i++$
    Return ($S_{best}$)
```

4.3.2 Local search procedures

The local search (LS) is applied at the second stage of the GRASP. It aims at the improvement of the solution provided by the GRASP first stage process. An interesting aspect that could motivate the wrapper choice as component of the GRASP second stage, is the successful application of local search methods in FS modeling (*i.e.* Tabu search, Simulated annealing, Memetic approaches) [54].

In this chapter, we devise a GRASP model based on various LS operators. Such choice could be argued by the fact that the reference approaches were applied to small benchmarks and the provided operators had complexities comparable to *AF* local search.

The following Table (Table 4.1) details the neighborhood structures that will be deployed within the local search procedures (second GRASP stage). For the first part of this chapter we opt for four local search operators which include basic and composite local search operators (used with memetic schema)[6].

Local search	Operator	Reference	Description
Attribute Flip	AF	*eq. 3.9*	Replaces selected attribute by a non selected one (best of all combination pairs)
Bit Flip	BF	*eq. 3.6*	Attribute state inverted (one at a time)
Floating Search	FLS1	*Algo. 15*	Composite operator: forward and backward search.
Floating Search With iterative back track	FLS2	*Algo. 16*	FLS1 + possible iterative backward search

Table 4.1: Local search operators applied to GRASP

---

[6] The remaining local search operators will be assessed with high dimensional data sets

4.3.3 Empirical results

In this section, we empirically assess the behavior of proposed GRASP-FS schema as well as a selection of the devised components. They will be, also, compared to the baseline GRASP $\langle IGV, AF \rangle$ proposed by Yusta in [133], where reported results have confirmed the superiority of GRASP in comparison to Tabu search, Genetic and Memetic algorithms, and SFFS heuristic.

Five benchmark datasets were used to assess GRASP-FS instances: *Sonar*, *Ionosphere*, *SpamBase*, *Audiology* and *Arrhythmia* with respectively 60, 34, 57, 69 and 279 attributes.[7]

Since reference approach was based on the selection of a fixed number of attributes defined by the parameter d (*see Algo.* 36), we opt for an extended version of the algorithm which randomly generated the number of attributes to select. With such modification it becomes comparable to our GRASP-FS which is not limited by the attribute number constraint. Reported results, correspond to the average values of at least 50 trial runs. Means, Standard deviation and statistical test validation (t-Test with confidence level of 97.5%) are also provided.

Three analyses are provided. The first one focuses on the assessement of the components of the construction stage, whereas the second compares effectiveness of local search apporaches. Finally, the third one compares the reference approaches to instances of well performing GRASP-FS components on both stages.

Construction Phase

The first part of the empirical study is devoted to the assessment, of the behaviors of the baseline GRASP with the devised GRASP-FS scheme which is based on Filters to both built RCL and construct solutions. We keep the same LS operators for all the experiments. The aim, here, is to be able to compare the construction mechanisms. Table 4.2 provides results for each data set. Globally, according to the gain compared to the reference approach. (last table column) obtained with a GRASP-FS instances which generate the RCL with filters, the baseline method is outperformed in most of

---

[7] Datasets from by the UCI repository [16]

the cases.

In addition, the improvement obtained with fitness values is confirmed with validation criteria (independent data, and different classifiers for validation). The overall improvement, points out the reliability of the construction stage, particularly the filters enlisted in the selection of suitable features. Surprisingly, Relief scores used in the RCL build, seems to be the less relevant filter used in the first stage. On the other hand, GRASP alternatives based on FCBF confirm superiority over the remaining filters. We should also note that the mean running time is kept comparable for both reference and proposed apporaches. In some cases, the GRASP-FS succeeded to reduce running time.

Data	Model		Fitness (%)	Validation Error (%)		CPU Time(s)	Gain % [133]
	RCL	LS		ANN	NB		
Sonar	IGV	AF	15,89(1,71)	32,83(3,40)	40,30(3,01)	14183,88(7196)	
	Relief	AF	**14,29(1,88)**$^{+}$	**31,31(3,66)**$^{+}$	**39,85(2,71)**$^{+}$	14915,13(7493)	10,07%
	SU	AF	**12,79(1,13)**$^{+}$	**30,54(3,06)**$^{+}$	**39,53(2,51)**$^{+}$	15867(8524)	19,51%
	FCBF	AF	**13,46(0,00)**$^{+}$	**31,32(1,30)**$^{+}$	**37,27(1,70)**$^{+}$	14920(7450)	15,29%
Audiology	IGV	AF	49,12(1,96)	52,4(3,14)	54,05(0,15)	343915(280683)	
	Relief	AF	**46,74(3,92)**$^{+}$	**51,54(4,47)**$^{+}$	54,09(0,22)	337756(248912)$^{-}$	4,85%
	SU	AF	**33,36(3,2)**$^{+}$	**40,52(4,55)**$^{+}$	54,16(0,25)	350761(267582)	32,08%
	FCBF	AF	**36,08(4,72)**$^{+}$	**40,53(6,91)**$^{+}$	54,06(0,14)	338687(258913)	26,55%
Arrhythmia	IGV	AF	39,72(1,57)	41,98(2,04)	43,61(1,65)	183959(132900)	
	Relief	AF	40,17(1,76)	42,7(2,05)	44,22(1,74)	**170925(114978)**$^{+}$	-1,13%
	SU	AF	**36,15(1,89)**$^{+}$	**39,73(2,42)**$^{+}$	44,56(1,79)	**173505(115749)**$^{+}$	8,99%
	FCBF	AF	**33,82(1,26)**$^{+}$	**39,33(2,11)**$^{+}$	**43,54(1,76)**$^{+}$	176065(117012)	14,85%
Ionosphere	IGV	AF	5,63(0,91)	16,34(1,98)	17,9(2,13)	22316(12626)	
	Relief	AF	5,95(1,00)	15,51(2,26)	**16,97(2,04)**$^{+}$	21855(12081)	-5,68%
	SU	AF	5,76(0,98)	**15,21(2,48)**$^{+}$	**17,31(2)**$^{+}$	24031(14531)	-2,31%
	FCBF	AF	**3,51(0,32)**$^{+}$	16,33(0,92)$^{-}$	15,73(0,89)	**21973(11873)**$^{+}$	37,66%
SpamBase	IGV	AF	16,47(1,04)	19,91(1,50)	20,23(1,57)	347062(190196)	
	Relief	AF	16,43(1,05)$^{-}$	**19,59(2,19)**$^{+}$	**19,58(1,72)**$^{+}$	**338671(185750)**$^{+}$	0,24%
	SU	AF	**14,18(1,12)**$^{+}$	**15,89(1,66)**$^{+}$	**17,13(2,27)**$^{+}$	311037(156931)	20,77%
	FCBF	AF	**13,05(0,84)**$^{+}$	**15,96(2,18)**$^{+}$	**15,31(1,88)**$^{+}$	331498(181414)	20,97%

[8]result format: $[m(sd)^{+/-}]$; m: Mean; sd: Standard deviation;$(+/-)$: T-test validity

Table 4.2: GRASP with RCL based on filters

Local search enhancement

The local search of the baseline method uses Attribute Flip operators whereas the proposed GRASP-FS uses local search procedures inspired from heuristics successfully applied to the FS problem. In fact, four components could be deployed on the

GRASP-FS second stage: AF, BF, FLS and FLS with iterative backtrack. The devised local search procedures are deployed within new GRASP instances using the IGV criterion on the First stage. Table 4.3 compares and evaluates the four GRASP instances. Even though, the solutions provided by the first GRASP stage are based on IGV criterion, numerous devised local search procedures have succeed to outperform the baseline algorithm. In comparison to the results obtained by the previous empirical analysis, the enhancement on the second stage is more significant than impact of changes of the first stage.

Furthermore, local search alternatives adopting floating search, have empirically confirmed their superiority over Yusta GRASP. On the other hand, the neighborhood structure based on the selection or removal of one attribute (NH_{BF}) is the less effective fine tuning scheme. The same result was confirmed with memetic algorithms. The overall improvement could be explained by the enhancement of the neighborhood structure design, since the initial solutions are provided by the same construction mechanism.

The adapted new GRASP-FS scheme instances have empirically shown that enhancements could be afforded by filters in first stage as well as wrappers in second stage.

Both stage improvements

Both of the previous sections studied the enhancement of the proposed procedures, separately, in each stage. This section assesses the GRASP-FS global behavior when we rely on the successfully used components: the construction stage based on attribute filter ranking as selection criterion, and the composite neighborhood local search operators, for the second stage. Table 4.4 compares two instances of the GRASP-FS to the reference approach proposed in [133].

It is not surprising to remark the significant improvement of the GRASP-FS over the reference approach. Both fitness and validation criteria results were enhanced in most of cases. In opposition to the results provided by the two previous tables, there is no negative gain. The gain (fitness criterion) is ranging from $15,03\%$ to $63,89\%$. The improvement of classification accuracies, was not followed by a significant increase

Data	Model		Fitness (%)	Validation Error (%)		CPU Time(s)	Gain %
	RCL	LS		ANN	NB		[133]
Sonar	IGV	AF	15,89(1,71)	32,83(3,40)	40,30 (3,01)	14183(7196)	
	IGV	BF	28,68(1,58)	33,59(4,90)	41,22(4,19)	15215(7922)	-80,49%
	IGV	FLS1	**5,92(2,05)**$^+$	**31,14(3,47)**$^+$	**40,05(3,08)**$^+$	14481(6652)	62,74%
	IGV	FLS2	**6,6(1,9)**$^+$	**31,26(3,38)**$^+$	**38,75(3,62)**$^+$	**12208(5244)**$^+$	58,46%
Audiology	IGV	AF	49,12(1,96)	52,4(3,14)	54,05(0,15)	343915(280683)	
	IGV	BF	68,78(1,34)	69,59(2,87)	72,64(2,4)	**322789(248574)**$^+$	-40,02%
	IGV	FLS1	**29,41(1,47)**$^+$	**41,78(3,96)**$^+$	54,09(0,22)	**234579(98248)**$^+$	40,13%
	IGV	FLS2	**30,99(1,21)**$^+$	**41,12(2,81)**$^+$	54,08(0,17)	**209670(111471)**$^+$	36,91%
Arrhythmia	IGV	AF	39,72(1,57)	41,98(2,04)	43,61(1,65)	183959(132900)	
	IGV	BF	49,47(1,02)	44,8(1,98)	46,34(1,59)	**160659(96674)**$^+$	-24,55%
	IGV	FLS1	**25,38(2,57)**$^+$	**38,27(2,9)**$^+$	**43,41(1,81)**$^+$	**156343(73463)**$^+$	36,10%
	IGV	FLS2	**24,42(2,67)**$^+$	**36,64(2,99)**$^+$	**42,64(1,71)**$^+$	**152719(89546)**$^+$	38,52%
Ionosphere	IGV	AF	5,63(0,91)	16,34(1,98)	17,9(2,13)	22316(12626)	
	IGV	BF	12,63(0,67)	**15,78(3,01)**$^+$	**17,38(2,03)**$^+$	21374(11603)$^-$	-124,33%
	IGV	FLS1	**2,27(0,52)**$^+$	**14,77(1,64)**$^+$	**17,35(1,2)**$^+$	**18561(8464)**$^+$	59,68%
	IGV	FLS2	**2,48(0,56)**$^+$	**15,42(1,44)**$^+$	**17,66(1,03)**$^+$	**16035(6597)**$^+$	55,95%
SpamBase	IGV	AF	16,47(1,04)	19,91(1,50)	20,23(1,57)	347062(190196)	
	IGV	BF	23,72(1,1)	22,38(3,46)	21,59(2,43)	328341(17611)	-44,02%
	IGV	FLS1	**6,85(0,73)**$^+$	**12,28(1,27)**$^+$	**14,90(2,66)**$^+$	532610(281813)	58,41%
	IGV	FLS2	**6,87(0,84)**$^+$	**12,05(1,26)**$^+$	**15,11(2,75)**$^+$	496728(206866)	58,29%

[9] result format: $[m(sd)^{+/-}]$; m: Mean; sd: Standard deviation;$(+/-)$: T-test validity

Table 4.3: GRASP with different local search procedures

on the mean number of selected attributes. In most of the cases the GRASP-FS has preserved the same level of dimensionality reduction.

4.4 GRASP-FS and high dimensional spaces

The second study is devoted to the investigation of the behavior of the GRASP-FS when it tackles high dimensional problems. We recall that, to the best of our knowledge, the few first GRASP attempts [32, 133], in feature selection modeling, have limited their the experimental study to benchmarks of small and medium size. This section studies the behaviors of some of adapted LS operators [10], within the GRASP scheme. At the same time, we assess the effectiveness of the component of the construction stage, and how they behave with the new operators and the tremendous search space nature. Detailed results of this section are provided by

---

[10] *see section* 3.4.3 *p.* 86

Data	Model		Fitness(%)	Validation Error(%)		CPU (ms)	Gain % [133]
	RCL	LS		ANN	NB		
Sonar	IGV	AF	15,84(1,81)	32,72(3,30)	40,28(2,91)	15378(7798)	
	FCBF	FLS1	**13,46(0)**$^+$	**31,36(1,18)**$^+$	**37,4(1,68)**$^+$	**11820(4720)**$^+$	15,03%
	SU	FLS2	**5,72(1,26)**$^+$	**29,07(3,43)**$^+$	40,25(2,48)	**14295(7014)**$^+$	63,89%
Audiology	IGV	AF	49,10(1,94)	52,41(3,22)	54,04(0,14)	334229(256748)	
	FCBF	FLS1	**28,54(1,22)**$^+$	**44,27(3,15)**$^+$	54,09(0,25)	**231281(115434)**$^+$	41,87%
	SU	FLS2	**28,17(0,6)**$^+$	**46,42(2,79)**$^+$	54,1(0,21)	**245239(140910)**$^+$	42,63%
Arrhythmia	IGV	AF	39,62(1,56)	41,81(1,96)	43,53(1,6)	180487(117167)	
	FCBF	FLS1	**23,61(1,8)**$^+$	**37,23(2,84)**$^+$	**43,03(1,62)**$^+$	**164488(74931)**$^+$	
	SU	FLS2	**24,41(2,43)**$^+$	**37,35(2,69)**$^+$	**43,15(2,04)**$^+$	**163011(81754)**$^+$	38,39%
Ionosphere	IGV	AF	5,67(0,95)	16,18(2,06)	17,90(2,06)	22968(12636)	
	FCBF	FLS1	**2,29(0,20)**$^+$	**15,6(0,64)**$^+$	**17,57(0,81)**$^+$	**17743(8624)**$^+$	59,61%
	SU	FLS2	**2,5(0,55)**$^+$	**15,46(1,55)**$^+$	**17,66(1,18)**$^+$	**16927(7417)**$^+$	55,91%
SpamBase	IGV	AF	16,46(1,06)	19,81(1,52)	20,18(1,54)	358741(186075)	
	FCBF	FLS1	**6,6(0,41)**$^+$	**11,93(1,16)**$^+$	**13,33(2,23)**$^+$	548332(283068)	59,90%
	SU	FLS2	**7,10(0,79)**$^+$	**12,51(1,53)**$^+$	**15,32(2,81)**$^+$	548493(286901)	56,87%

Table 4.4: Basic Grasp *vs* enhanced variants

associated chapter *Annex IV* (*p.* 237-239)

4.4.1 Behaviors of the construction mechanisms

In one of the previous experimental studies, we have seen that the construction mechanism based on the IGV was less effective than those based on filters. In this section we try to shed some light on the relation between the construction mechanisms and the different local search operators. In this preliminary analysis, we compare for each local search operator the behavior and results of IGV [11] in comparison to both construction mechanisms based on filters. Figures 4.1 and 4.2 respectively illustrate, for both validation and evaluation criteria, the number of times where the GRASP-FS instances based on the Relief (respectively SU) succeeded to outperform IGV on a set of 16 high dimensional benchmarks [12]. Reported results were only limited to results, for which a statistical test was validated. Figure 4.1, shows that both LS operators MB^* and FLS^* are more effective with construction stage based on RELIEF. On the other hand, GRASP-FS instances based on MB^+ are less sensitive to the construction stage mechanisms. The same conclusions could be drawn for the comparison between IGV and SU based construction mechanisms. In fact, operators MB^* and FLS^* behave

---

[11] GRASP-FS instances based on the IGV construction mechanism

[12] attributes number is ranging from 2000 to 26000

more accurately with filter based construction processes than those based on IGV. In addition, MB^{+} is, relatively, less sensitive to the usage of a particular construction mechanism. For the remaining LS operators used within GRASP instances we can remark any relative sensitivity to construction mechanisms.

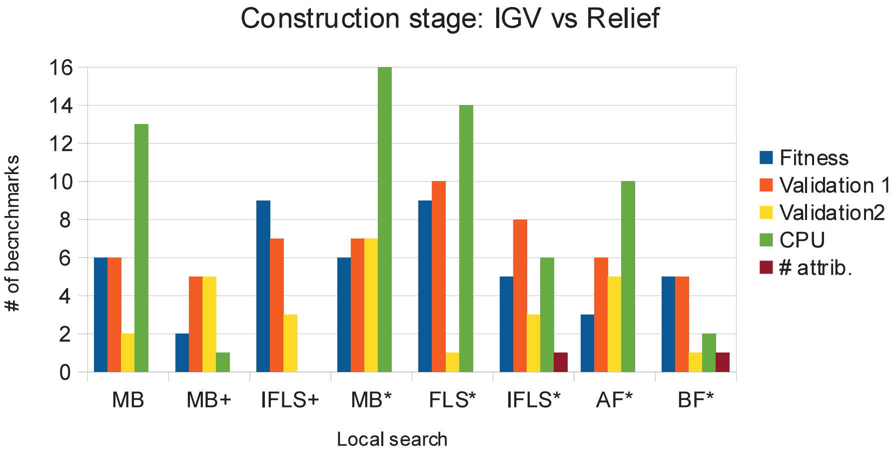


Figure 4.1: Construction stage assessment: IGV *vs* Relief

4.4.2 Local search analysis

In this section, we mainly focus on the behaviors as well as the impact of the LS procedures on the best solutions found. For each data set, we compare accuracy levels of the $GRASP - FS$ instances by varying the LS operator. The best solution fitness was used as comparison criterion. Figures 4.3 and 4.4 graphically compare different instances of the GRASP-FS. Mainly, three tendencies could be identified. The first one concerns data sets, where almost the same results were obtained. Benchmarks like $Leukemia3C$, $Challenge2004$ and 14_Tumors provided the same results and the LS operators behaved on the same way with each construction mechanism. For such data sets, the GRASP instances are insensitive to LS components.

The second tendency regards data sets and LS operators providing the lowest

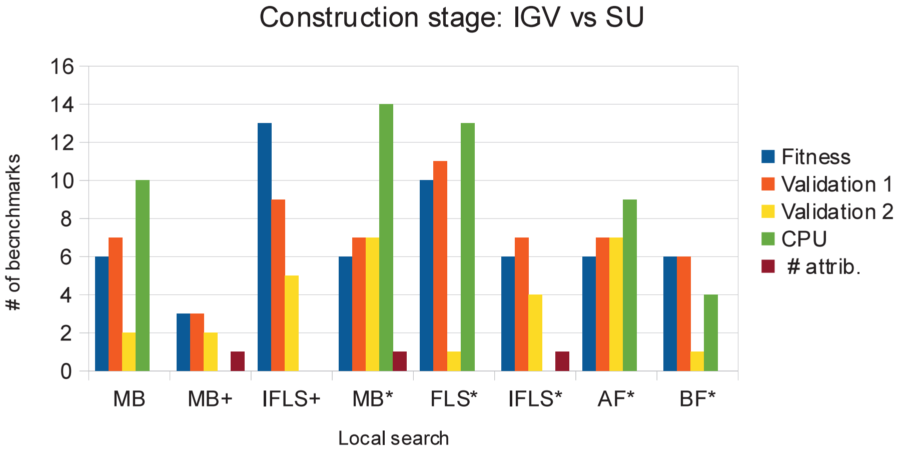


Figure 4.2: Construction stage assessment: IGV *vs* SU

classification error rate [13]. GRASP-FS instances applied to the CNS, MLL and $ProstateTumor$ data sets provide feature subsets allowing perfect classification with different local search operators. We note that comparable results were obtained with memetic algorithms. The third tendency is in relation with benchmarks covered by the two previous tendencies. In fact, $IFLS^+$ the local search was always less effective than the remaining ones. Such operator is not recommended as a component of the GRASP second stage. The same consequence could be drawn for some data sets with the FLS^* operator.

When we compare, for each local search operator, the construction mechanism of instances that succeeded to reach the best performances, we can clearly remark the superiority of GRASP-FS instances endowed with construction mechanism based on SU.

As concluding remark, the luck of sensitivity to LS with some GRASP instances could be explained, partly by, the nature of solution provided by the construction stage.

---

[13] evaluation stage

4.4.3 Summary of empirical analysis

We summarize the GRASP experiments, by considering the following two criteria: fitness and the number of selected attributes. The particularity of these comparisons is that we compare our proposed approaches, because to the best of our knowledge there was not a GRASP applied or adapted to high dimensional FS benchmarks. In fact, GRASP instances will be compared by varying either construction or Local search processes. Resulting behaviors or tendencies are assessed according to their statistical validated results for all benchmarks. The following Table (*see Table* 4.5) illustrates for each GRASP instance the percentages of benchmarks for which the its better [14], worse and have a comparable results[15].

Globally results of fitness criterion are in most of the cases comparable , and not in favor of one of favor of one of the GRASP instances. In some cases, approaches based on floating local search operators (*i.e.* $IFLS^+$ and FSL^*) provides better fitness results with inferior results on the number of selected attributes criterion.

On the other hand, when we look to the results of the second criterion (selected subset size) we can clearly see the superiority of almost all the GRASP schema based on IGV for the construction stage.

4.5 Conclusion

In this chapter, we investigate the GRASP metaheuristic. The approach has a number of attractive features that have allowed us to design a hybrid model combining Wrapper and filters as GRASP components. The second part of the chapter was devoted to exploration of high dimensional spaces and the study of the deployment a set of adapted LS operators. Reported results confirmed the successful application of the GRASP-FS. As perspective, The GRASP search capabilities could be endowed with path relinking intensification mechanism and the impact on global search process behaviors could be studied. Besides, the first GRASP stage could be diversified with

---

[14] improvement over 10 runs (at least) is statistically validated

[15] statistical test is not in favor of any GRASP instance

Compared Approaches		Validated improvement % (fitness)			Validated improvement % (att)		
A	B	A	B	Comparable	A	B	Comparable
$< IGV; MB >$	$< Relief; MB >$	0,00%	37,50%	62,50%	**75,00%**	6,25%	18,75%
$< IGV; MB >$	$< SU; MB >$	0,00%	37,50%	62,50%	**81,25%**	0,00%	18,75%
$< IGV; MB^{+} >$	$< Relief; MB^{+} >$	25,00%	25,00%	50,00%	**68,75%**	0,00%	31,25%
$< IGV; MB^{+} >$	$< SU; MB^{+} >$	31,25%	18,75%	50,00%	**81,25%**	6,25%	12,50%
$< IGV; IFLS^{+} >$	$< Relief; IFLS^{+} >$	0,00%	**62,50%**	37,50%	**68,75%**	31,25%	0,00%
$< IGV; IFLS^{+} >$	$< SU; IFLS^{+} >$	0,00%	**87,50%**	12,50%	**68,75%**	31,25%	0,00%
$< IGV; MB^{*} >$	$< Relief; MB^{*} >$	0,00%	37,50%	62,50%	**56,25%**	12,50%	31,25%
$< IGV; MB^{*} >$	$< SU; MB^{*} >$	0,00%	37,50%	62,50%	**68,75%**	18,75%	12,50%
$< IGV; FLS^{*} >$	$< Relief; FLS* >$	12,50%	**56,25%**	31,25%	**100,00%**	0,00%	0,00%
$< IGV; FLS^{*} >$	$< SU; FLS^{*} >$	12,50%	**62,50%**	25,00%	**100,00%**	0,00%	0,00%
$< IGV; IFLS^{*} >$	$< Relief; IFLS^{*} >$	0,00%	37,50%	62,50%	**62,50%**	12,50%	25,00%
$< IGV; FLS^{*} >$	$< SU; IFLS^{*} >$	0,00%	37,50%	62,50%	**81,25%**	6,25%	12,50%
$< IGV; AF^{*} >$	$< Relief; AF^{*} >$	0,00%	31,25%	68,75%	**87,50%**	0,00%	12,50%
$< IGV; AF^{*} >$	$< SU; AF^{*} >$	0,00%	31,25%	68,75%	**93,75%**	0,00%	6,25%
$< IGV; BF^{*} >$	$< Relief; BF^{*} >$	6,25%	31,25%	62,50%	**81,25%**	6,25%	12,50%
$< IGV; BF^{*} >$	$< SU; BF^{*} >$	6,25%	37,50%	56,25%	**100,00%**	0,00%	0,00%

Table 4.5: Empirical Study Synthesis of GRASP

component based on more robust construction mechanisms like those used within embedded methods [54] or incremental wrappers. It could be also endowed with more effective design of the RCL list (*i.e.* Reactive GRASP [38]).

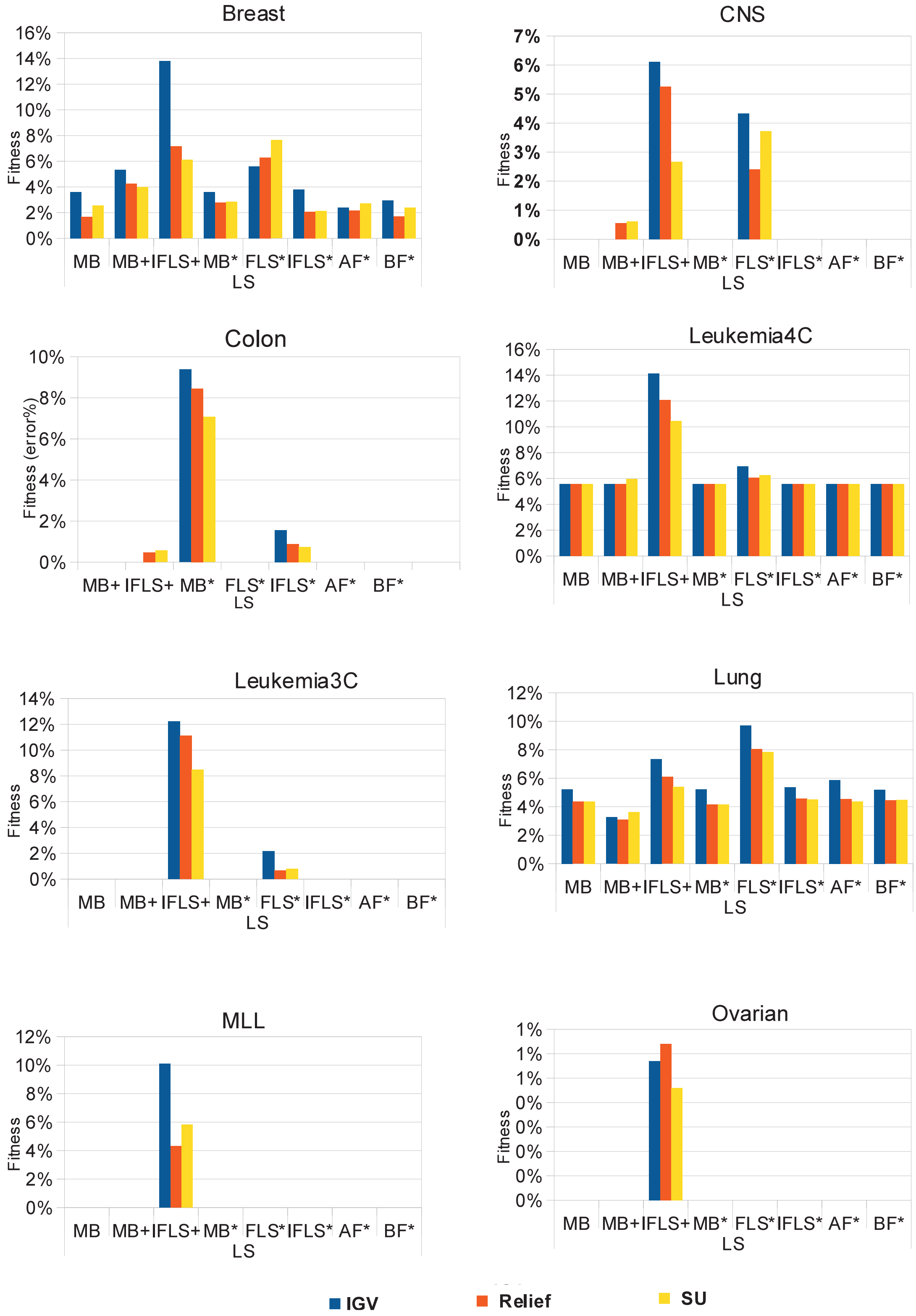


Figure 4.3: Best results found (Fitness criterion) (1)

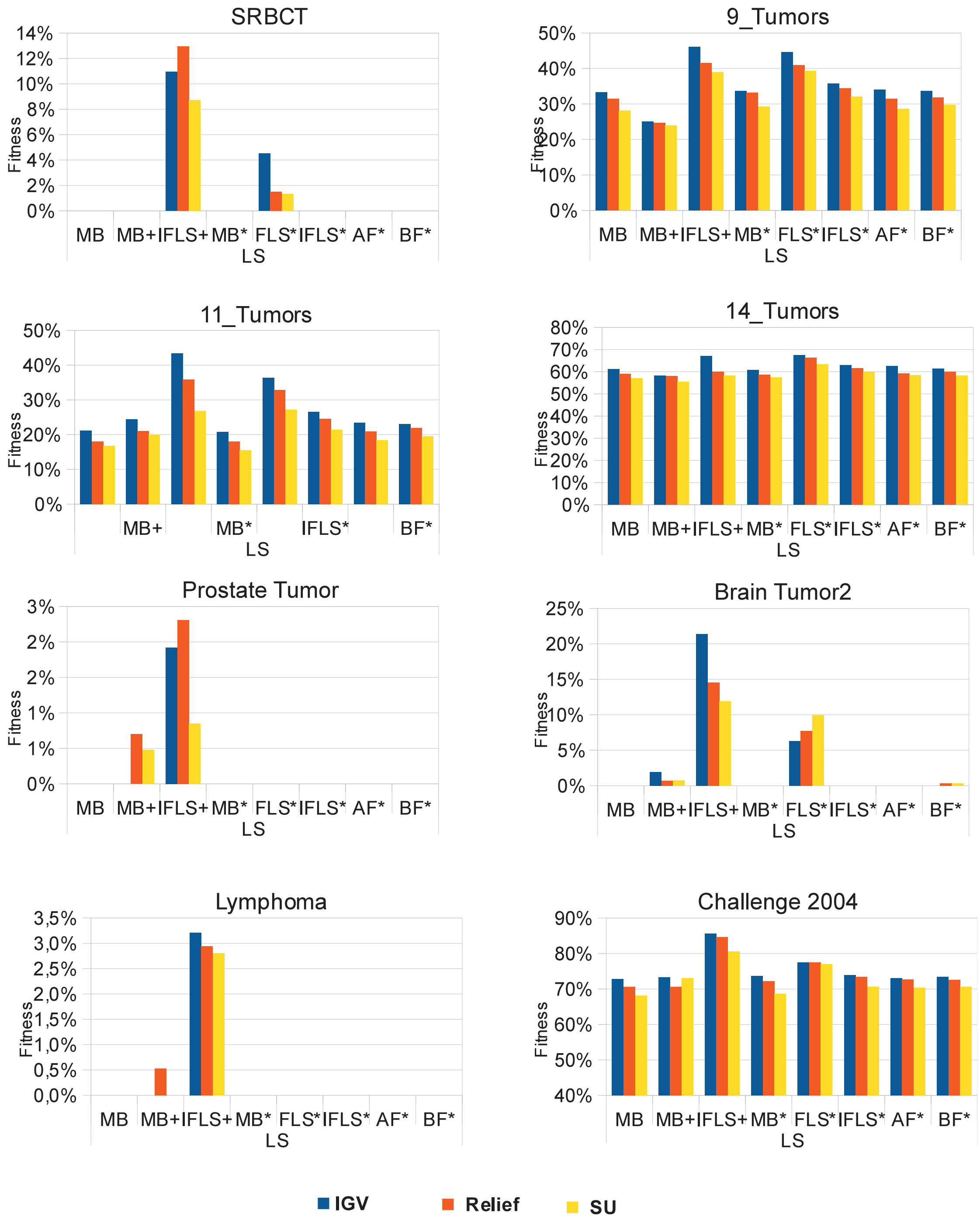


Figure 4.4: Best results found (Fitness criterion) (2)

Chapter 5

Swarm Feature Selection: A continuous PSO for the FS Problem

5.1 Introduction

Swarm intelligence refers to the problem solving behaviors that emerge from the interaction of agents or entities that communicate to each other, by acting on their environments [31]. Particle swarm optimization (PSO), part of the swarm intelligence family, is known to effectively solve large-scale nonlinear optimization problems [25]. This chapter is devoted to the investigation of FS modeling with swarm behavior. We devise a model based on the continuous PSO that is able to aggregate filter scores and hybridize filters with wrapper within the same framework. The reminder of this chapter is organized as follows: an overview of PSO modeling concepts as well as its featured properties is given in Section 2. Section 3 is devoted to the devised PSO-FS approach. Enhancement and more elaborate improvements are investigated in Section 4. Main contributions and future research topics are summarized by Section 5.

5.2 Particle Swarm optimization

This section sheds some lights on the main concepts of the swarm optimization paradigm, details basic PSO. Next, we cover some of the featured enhancements of the PSO. Finally, we discuss some of the challenging design and implementation aspects of the PSO in general and FS modeling specifically.

5.2.1 Swarm universe: Overview

Formally, Swarm intelligence is the property of a system whereby a collective behaviors of unsophisticated agents interacting locally with the environments [31]. Computational swarm intelligence studies the algorithmic and modeling aspects of such behaviors. Swarm intelligence was also been referred as collective intelligence. In fact, swarms entities are endowed with basic and simple behaviors, but the emerging social behaviors of the swarm are more complex and motivate researchers to develop computational models. A number of models were inspired from the study of social behaviors of both animals and insects. Among these biological swarm systems, simple in structure, and generating collective social behaviors we can cite: *ants*, *bees*, *termites*, *fish scholes*, *bacteria*, *etc*.

The objective of swarm intelligence is to model the simple behaviors of individuals and to provide them means of commutation and interaction with environment to solve complex problems. In our case, we study the Particle Swarm Optimization, which models two simple behaviors. The first one is to move toward best (or locally best) solution (or neighbor). The second models the attractiveness to the best memorized solution found by the particle itself. The resulting collective behaviors of the swarm (population of particles) is that particles are looking for and updating best solutions. In several cases commutation is done indirectly and through the environment. Such a concept plays a key role in the evolution of swarm and it is always referred to as stigmergy. The idea of PSO is based a set of particles flying through a hyper-dimensional search space. The population of particles forms a swarm that simulates the social behaviors of birds within a flock [72]. Particle position changes emulate social tendencies. They are influenced by its own experience or/and the knowledge of its neighbors. In real number space, the position of each particle is determined by a vector $\overrightarrow{x}_i$ and its movement by its velocity $\overrightarrow{v}_i$. The position update rule (at time t) is illustrated by equation 5.1:

$$\overrightarrow{x}_i(t) \leftarrow \overrightarrow{x}_i(t-1) + \overrightarrow{v}_i(t) \tag{5.1}$$

On the other hand, velocity is adjusted according to the informations available to each individual. Three components characterize velocity update (*see eq.* 5.2):

$$\overrightarrow{v}_i(t) \leftarrow \underbrace{\overrightarrow{v}_i(t-1)}_{Habit} + \underbrace{\varphi_1 * rand_1 * (\overrightarrow{p}_i - \overrightarrow{x}_i(t-1))}_{self \quad knowledge} + \underbrace{\varphi_2 * rand_2 * (\overrightarrow{p}_g - \overrightarrow{x}_i(t-1))}_{social \quad knowledge} \tag{5.2}$$

- first component models the tendency of the particle to continue on the same direction. It is also known as *"habit"*, *"momentum"* or *"inertia"* component.
- second component illustrates attraction towards the best particle position p_i found. It is "referred" to as *"self-knowledge"*, *"memory"* or *"nostalgia"*.
- third component models attraction towards best position found among neighborhood or swarm. It is referred to as *"cooperation"*, *"social knowledge"* or *"information sharing"*.

Because of the relative importance of these factors, which can vary from one decision to another, random weights are applied with $rand_1$ and $rand_2$, and are scaled with φ_1 and φ_2.

5.2.2 Basic PSO

Pseudo-code of the basic PSO [72] is illustrated by Algorithm 38. After swarm initialization with nb_{part} particles, particles start the exploration of the search space by adjusting velocities (Line 9). Velocities updates take into account, for each problem dimension k, best particle position ($x^k_{best_i}$), and best swarm position found ($x^k_{best_g}$) as

Algorithm 38: Particle Swarm optimizer

```
Input:
$nb_{part}$: particles number;
$\omega$: inertia weight;
$\varphi_1, \varphi_2$: weight factors
Output: $x_{best}$ : Best solution
begin
    Swarm  $P \leftarrow CreateSwarm(nb_{part})$
    $x_{best} \leftarrow GetBest(P)$
    while (Stopping criterion satisfied) do
        foreach ($p_i \in P$) do
            $k \leftarrow 0, r_1 \leftarrow 0, r_2 \leftarrow 0$
            foreach (dimension $k$ in $p_i$) do
                $r_1 \leftarrow rand(); r_2 \leftarrow rand();$
                $v_i^k \leftarrow v_i^k + \varphi_1.r_1.(x_{best_i}^k - x_i^k) + \varphi_2.r_2.(x_{best_g}^k - x_i^k)$
                $x_i^k \leftarrow x_i^k + v_i^k$
            //—Update best personal position—
            if ($x_i > x_{best_i}$) then
                $x_{best_i} \leftarrow x_i$
            //—Update best global position—
            if ($x_{best_i} > x_{best_g}$) then
                $x_{best_g} \leftarrow x_{best_i}$
    $x_{best} \leftarrow x_{best_g}$
    Return ($x_{best}$)
```

well as respective random factors (φ_1 and φ_2) and weights (r_1 and r_2). Once velocities were adjusted, each swarm particle moves to its new position $\overrightarrow{x}_i$. The Resulting new solutions are compared to the best ones found locally and among the swarm. The velocity update for the social component could be made according to two strategies. The first one considers the global best solution of the swarm, while the second limits the social component update to best among current particle neighborhood. Kennedy [72], said that with global best the PSO converged fast, but may be trapped in a local minimum, while with the local best (among neighborhood) the swarm has more chance to find optimal solution, with slower convergence.

Numerous neighborhood topologies have been studied [72]. They include *ring*, *wheel*, and random topologies. *Kennedy and Mendes* in [71], have empirically shown that the suggested neighborhood size for a swarm of 20 particles was 5.

5.2.3 PSO variants

This section is a somewhat more technical look at what researchers have proposed to extend the basic particle swarm algorithm. We survey some of the more featured variants that have shown their empirical effectiveness. We should also note that the majority of the PSO issues are dressed by the velocity update equation (*see eq.* 5.2).

Parameter selection

Several considerations should be taken into account, when implementing the PSO. One of the major concerns of the PSO evolution is to look for mechanisms facilitating rapid convergence and preventing swarm ”explosion” [25]. Such considerations include velocity clamping, adequate selection of acceleration weights, or even the introduction of new parameters as the *constriction* factor or the *inertia* constant. According to [31], it was found that the velocity quickly explodes to large values, especially, for particles far from global and local best positions. To control global exploration of particles, an upper (v_{max}) and lower (v_{min}) limit of v_i^k were proposed in [72] as follows (*eq.* 5.3):

$$v_i^k \begin{cases} v_{max}, & if(v_i^k > v_{max}) \\ v_{min}, & if(v_i^k < v_{min}) \\ v_i^k, & otherwise. \end{cases} \tag{5.3}$$

If limits are set too high the particle movement may be beyond good solutions, whereas with too small interval the particle movement is limited as well as its exploration capability. In most of the cases, bounds are selected empirically. *Shi and Eberhart* proposed in [118] the *inertia* weight ω as a new parameter to control global and local search swarm abilities (*see eq.* 5.4).

$$\overrightarrow{v}_i(t) \leftarrow \omega.\underbrace{\overrightarrow{v}_i(t-1)}_{Habit} + \underbrace{\varphi_1.r_1.(\overrightarrow{p}_i - \overrightarrow{x}_i(t-1))}_{self \quad knowledge} + \underbrace{\varphi_2.r_2.(\overrightarrow{p}_g - \overrightarrow{x}_i(t-1))}_{social \quad knowledge} \tag{5.4}$$

The value of ω should be carefully defined, because a large value increases diversity and facilitate exploration, and a small value might eliminates swarm exploration ability. A similar coefficient was proposed in [22], where velocities were constricted by a parameter. The constriction factor could be considered as an extension or a generalization of the inertia parameter, in the sense that it is not only applied to the previous velocity value but to all velocity updating rule components (habit, self knowledge and social knowledge).

According to [25], constriction factor improved the convergence once the particle is focused on the best point in an optimal region. However, it may not converge when the particle best performance p_i is far from p_g. In addition to the above presented velocities regulation issues, basic PSO parameters (*i.e.* swarm size, neighborhood structure and size, acceleration coefficients and number of iterations) should be adjusted according to the characteristics of the problems.

Binary PSO

PSO was originally developed for continuous valued search space. Nevertheless a binary PSO was developed in [72]. In such a model particle velocities and movements were rather defined in terms of probabilities. For example, $v_i^k = 0.3$ implies 30% to be 1. The sigmoid function was used to normalize velocities in the range of $[0, 1]$ (*see eq.* 5.5):

$$sig(v_i^k) \leftarrow \frac{1}{1 + \exp\left(-v_i^k\right)} \tag{5.5}$$

Particle positions are updated in the following way:

$$x_i^k \begin{cases} 1, & \text{if } (rand < sig(v_i^k)) \\ 0, & \text{otherwise} \end{cases} \tag{5.6}$$

where $rand$ is random value with a uniform distribution in the range of $[0, 1]$ and x_i^k is the position of particle i on the k^{th} dimension.

Enhanced variants

One can think that instead of considering only the attraction to individual and global best solution, the particle velocities could be influenced by the success of all its neighbors. The idea was applied with FIPS in [98], where a fully informed PSO was devised. The acceleration weights φ_i were equally distributed across the entire neighborhood. CLPSO tried to prevent premature convergence by updating the velocities on the following way:

$$v_i^k \leftarrow \omega . v_i^k + \varphi . (pbest_{f_i(k)}^k - x_i^k) \tag{5.7}$$

where $f_i(k)$ defines the particles' best solution that particle i should learn from, for the k^{th} dimension. The decision depends on probability p_c, referred to a learning probability. If it is lower than a random generated value ($p_c < rand$), then the particle will follow its own p_{best}. Otherwise it will learn from the particle defined by $f_i(.)$. $f_i(.)$ applies a tournament selection to define the requested particle. Instead of limiting the velocity update to only particle experience and best swarm solution, as in PSO, any potentially interesting particle could guide the current one. Besides, instead of learning from the same exemplar for all dimensions, each dimension can learn from a different *pbest*. The adopted rules increase the swarm diversity and might lead to enhanced PSO accuracy.

The unified PSO [105], proposed a hybrid scheme between the PSO relying on attractiveness of best swarm solution for social component, and the PSO using the best solution within neighborhood to update its velocity. Consequently, the resulting scheme combines both global and local search capabilities within the same PSO scheme. In fact, it defines two velocity update rules: G_i^k and L_i^k. The first one, (*see eq.* 5.8) refers to the global variant of the PSO, while the second (*see eq.* 5.9) is for the variant relying on neighborhood as social component. Equation 5.10 aggregates the two updating rules via a unification factor u. The resulting velocity adjusts particle movement as by equation 5.11.

$$G_i^k \leftarrow \chi.[v_i^k + \varphi_1.(p_i^k - x_i^k) + \varphi_2.(p_g^k - x_i^k)] \tag{5.8}$$

$$L_i^k \leftarrow \chi.[v_i^k + \varphi_1'.(p_i^k - x_i^k) + \varphi_2'.(p_{g_i}^k - x_i^k)] \tag{5.9}$$

$$U_i^k \leftarrow u.G_i^k + (1-u).L_i^k \qquad u \in [0,1] \tag{5.10}$$

$$x_i^k \leftarrow x_i^k + U_i^k \tag{5.11}$$

Hybrid approaches Some of the PSO variants incorporated evolutionary capabilities. The main goal was to increase the diversity of the population and to gain better performances. Many studies proposed operators as crossover, mutation, and selection [25, 72]. Such hybridization includes combination of GA with PSO, evolutionary PSO (EPSO) and differential evolution PSO (DEPSO). Detailed surveys could be found in [25, 31].

5.2.4 Swarm Feature selection modeling

Swarm feature selection modeling research mainly covers PSO and ACO approaches. ACO modeling for a combinatorial problem like feature selection, seems to be more intuitive because of the discrete nature of decision variables and the possibility of incremental solution construction. Featured and recent approaches that adopted ACO in FS include [69], where a decision support system was devised for face recognition, and a feature selection approach combining ACO and differential evolution [74] [1]

On the other hand, feature selection approaches based PSO favored the use of discrete, or even binary, version of the PSO as an alternative to adapt the original continuous PSO to the FS modeling requirement. We focus, on a recent approach based on an improved version of binary PSO (IBPSO [21]), which was applied to high dimensional data sets. The model was applied to gene expression data (*bioiformatics*). The proposed swarm explores the search space of possible attribute combinations. Each particle position refers to a solution[2]. The velocity update rule used was comparable to basic PSO (*see eq.* 5.4) with upper and lower bounds respectively sets to $V_{min} = -6$, and $V_{min} = 6$. Next, the velocities were normalized by *sigmoid function* (*see eq.* 5.5). Finally, particles positions were set to binary values[3], according to the following equation (eq. 5.12):

$$Attribute(k)_i \begin{cases} 1, & \text{if } (rand < sig(v_i^k)) \\ 0, & \text{otherwise} \end{cases} \tag{5.12}$$

where i denotes the particle index and k is the k^{th} dimension of the particle position. The binary PSO variant proposed in [21], tried to tackle the problem of best solution (p_g) trapped in local minima. It suggested to reset p_g when the fitness was not improved for a given number of iterations. [4]

We think that it is acceptable to use binary alternative as in IBPSO to adapt the metaheuristic to the basic requirement of FS modeling. However the question that should be asked is: *was it so effective with FS problems particularly, for search space exploration?*

The evolution ability of the swarm as well as the resulting stigmergy could be devised more effectively with combinatorial problems like feature selection.

In addition, several PSO variants providing alternatives for velocities update, definition of neighborhood structures, and suggesting mechanisms for the control of the exploration/exploitation trade-off motivate us to investigate its adaptations issues

---

[1] The study of the ACO variants applied to the FS problem may be the subject of further research

[2] Each feature is represented by a dimension

[3] for each dimension

[4] $n_{it} = 3$ was used for empirical evaluation [21]

with FS problems.

Besides, either basic or advanced alternatives, should study the behaviors of the devised approaches toward high dimensional datasets, since such datasets are characterized by redundant attributes.

5.3 PSO-FS: Swarm Feature selection

Since the basic PSO was originally designed to cope with continuous problems, the adaptation to the context of feature selection did not necessarily require the systematic move to binary or discrete alternatives of basic PSO. We devise a PSO based the original PSO scheme [72], and adapt it to the requirement of the feature selection problem. The Particles explore the search space, adjust their velocities and update their positions according to valuable knowledge provided by filter scores.

The motivation for the adapted continuous PSO design could be summarized as follows: (i) relying of the original PSO paradigm (ii) improving particle position update rules. Rather than using the sigmoid function to define probabilities, a weighting scheme is proposed and attributes are selected only if they are above a threshold (iii) particles initial positions are not set randomly but defined according to filter scores. (iv) the PSO is not only used as optimization schema but we extend it to an integration framework combining wrappers and filters strengths.

The devised PSO is based on a weighting scheme. Solutions are represented by a vector of weights and subsets of selected features are derived from associated weight vectors using a threshold. A particle position corresponds to a vector of weights. When particles of the swarm fly over the search space, positions are updated according to respective velocities and current position. Positions are initially set to the values of feature scores. Once the particles have moved to the new positions, the resulting subset of selected features is deduced according to weights and threshold values. One can think whether the score changes make a sense or if there is semantic that could be associated to the filter score adjustment? When we consider filter as an approach to tackle feature selection problems, the resulting ranking scores reflect the individual relevance of features to the class. It does not reflect the relevance among selected features, neither its contribution to improve classification. One of the advantages of the proposed PSO is its ability to overcome such filters limitation. In fact, the PSO initializes its particles with individual feature scores and adjust them, over generation, to reflect relative relevance or contribution to the subset to which it belongs. In addition, particles adjust their velocities according to relative relevance of the best solution found and the best in memory. Such dynamic adjustment, could alleviate the problem encountered with filters which assess only individual feature relevance.

Besides, most of the filter approaches provide scores and not the best subset. The problem, is how to select the best subset from a given ranking particularly when they did not take into account feature redundancy. Usually, *top-k* ranked attributes are selected [54]. The later filter problem was addressed by our PSO approach by the generation of initial solutions using filter ranking and stochastic selection of features among best ranked attributes.

Intuitively, the combination of wrapper and filter might enhance performances and guide the search to the exploration of interesting regions. In this section, the swarm feature selection scheme will be studied according to two variants.

The following paragraphs detail algorithmic, technical aspects as well as empirical results of the devised FS-PSO alternatives.

5.3.1 Swarm based on one filter

The section is devoted to the devised, FS-PSO approach. The detailed alternative is based on the scores of one filter.

The swarm particles start with weights reflecting feature-class dependency levels and, next, the PSO process adjusts its weights, by looking for relevant features combination to generate improved subset. Since each solution is represented by an individual vector of weights, the particle positions updates are done according to equations 5.1 and 5.2 of the basic PSO scheme [72]. In comparison to, the position update formula of the binary PSO, we can note that particles move are more dependent on velocities. In fact, during swarm evolution, velocities tend to reflect contribution of the attributes rather than selection probabilities.

Algorithm 39: FS-PSO

```
Input:
$nb_{part}$: particles number; $Cla$: Classifier;
$\omega$: inertia weight;
$\upsilon_{min}, \upsilon_{max}$: velocity bounds
$\varphi_1, \varphi_2$: weight factors
F: filter
$th$: weight Threshold
Output: $S_{best}$: Best solution found
begin
    $S \leftarrow CreateParticlesSet(nb_{part})$
    $S_{best} \leftarrow \emptyset$
    foreach $(particle \quad i \in S)$ do
        $x_i \leftarrow GenerateSolution(F)$
        $Evaluate(x_i, Cla)$
        $x_{best_i} \leftarrow x_i$
        if $(x_{best_i} \succ S_{best})$ then
            $S_{best} \leftarrow x_{best_i}$
    while (Stopping criterion satisfied) do
        foreach $(particle \quad i \in S)$ do
            /* particles move */
            foreach $(Attribute \quad k)$ do
                $r_1 \leftarrow rand(); r_2 \leftarrow rand();$
                $\upsilon_i^k \leftarrow \omega * \upsilon_i^k + \varphi_1 * r_1 * (x_{best_i}^k - x_i^k) + \varphi_2 * r_2 * (S_{best}^k - x_i^k)$
                if $(\upsilon_i^k \notin [\upsilon_{min}, \upsilon_{max}])$ then
                    $\upsilon_i^k \leftarrow max(min(\upsilon_{max}, \upsilon_i^k), \upsilon_{min})$
                $x_i^k \leftarrow x_i^k + \upsilon_i^k$
            $Evaluate(x_i, Cla)$
            /* Local and global best solutions Update */
            if $(x_i \succ x_{best_i})$ then
                $x_{best_i} \leftarrow x_i$
            if $(x_{best_i} \succ S_{best})$ then
                $S_{best} \leftarrow x_{best_i}$
    Return $(S_{best})$
```

Algorithm 39 details main steps of the proposed approach. The first stage, creates the swarm and initialize its particles (Lines 2-8). The second stage, involves iterative steps, which define, for each particle p_i and attribute weight k the new velocity (Lines 13-14), and update current particle position vector $p_i.Sol_{cur}$ (Line 17). The last stage defines the list of attributes retained by the current solution (Line 18), evaluates its fitness and updates particle and swarm bests (Lines 21-24). The selected attributes list is derived from the current particle position Sol_i according to the threshold th (*see eq.* 5.13):

$$Attribute(k)\begin{cases} Selected & \text{if } (Sol_i^k > th) \\ unselected, & \text{otherwise} \end{cases} \tag{5.13}$$

Initial Solution generation Initial solutions as well as particle positions are generated according to the filter scores and with stochastic selection mechanism [5].

Algorithm 40: Solution generation based on filter

```
Input:
F: filter
Output:
S: a solution
begin
    $S \leftarrow \emptyset$
    $list \leftarrow F.GetSortedList()$
    $k \leftarrow rand(list.length)$ /* top-k attribute $k > n$ */
    $n \leftarrow rand(k)$ /* n: number of attributes to select */
    $i \leftarrow 0$
    while ($i \leq k$) do
        $S.add(List(i))$
        $i++$
    while ($i > n$) do
        $S.remove(List(rand(List.length)))$
        $i--$
    $S.updateWeights()$
    Return ($S$)
```

---

[5] with a given elitism level

Algorithm 40 illustrates the generation process of intial particle solutions. Particles are initialized with the top-k attributes. For each particle, the value of k is randomly generated (Line 4). Besides, a subset of n attributes, is randomly selected among k best attributes. Such mechanism allows the swarm particles to start from different search space positions and prevents the selection of similar solutions based on the top-k attributes, with variable k. Consequently, among k best attributes, provided by filter ranking, any combination of $n \in]0, k]$ attributes would be accepted. The associated weights for the non-retained attributes are set below the threshold th. The final vector of weights will define the initial position of the particle.

Any filter criterion could be used to generate scores for PSO. In this paper, We opt for five well known filters (scoring methods):

- Relief [112] attempts to assess features according to their discriminative power. A weight $W[i]$ is assigned to each feature. Weights are updated in a manner to reflect feature ability to distinguish between class values. This reference approach, remains one of the more representative filters used in FS.
- Information Gain and Gain Ratio filter [54] are based on information theory measures derived from information entropy and mutual information criteria.
- Symmetrical Uncertainty (SU) is another measure of the information theory. SU criterion is widely used and considered as a robust measure for attribute ranking [54, 57].
- χ^2 filter [54]: is based on χ^2 statistics which compute the difference between attribute values distribution.

The scores provided by filters are normalized to the range of $[-1, 1]$. If the threshold th is set to zero, only features with positive values are considered in feature subset solutions.

Empirical study

In this section, we empirically assess, different instances of the proposed swarm approach. Note that experiments are based on the Validation protocol (*see Annex I* *p.* 165). The behavior of the PSO-FS is studied through its comparison with the reference approach IBPSO proposed in [21]. In addition, when particles of the devised approaches are initialized with filter scores, once can think that the comparison to a continuous PSO with randomly generated swarm would be of a great interest. This empirical section involves two sets of experiments. The first one compares the binary PSO and the continuous PSO randomly initialized to PSO-FS[6] while the second compares the continuous PSO, randomly initialized, to instances of PSO-FS with different filters.

Tables E.1, E.2, E.3 and E.4 (*see ANNEX V*, *p.* 241-244) detail evaluation and validation accuracies of the respective best solutions obtained with BPSO, continuous PSO [7] and PSO-FS. The PSO-FS instance relies on Relief filter as initialization scheme. Reported results, showed significant improvement of continuous PSO over Binary PSO. The enhancements mainly include *Fitness* and *CPU time* and, in some cases attributes numbers. It covers small, medium and large benchmarks. The more significant results were obtained with high dimensional problems.

The more impressive results that capture the attention were indisputably provided by PSO-FS which outperformed reference approach on all benchmarks according to the fitness criterion [8]. For example, the classification error rate (Fitness) obtained with Breast benchmark was reduced from 28.40% with binary PSO to 6.46% with PSO-FS.

We also note, the significant reduction of the selected attributes which can be explained by the swarm initialization with filter scores.

Globally, continuous PSO based on weighted solution representation, showed that they were better performing than binary PSO. Among the compared approaches, PSO-FS was the more effective in terms of results, computational cost and dimensionality reduction.

Tables E.5, E.6, E.7, E.8, E.9 and E.10, (*see ANNEX V*, *p.* 245-250) summarize the second set of experiments where the impact of filter scores on swarm initialization were assessed. Evaluation and validation accuracies of the respective best solutions

---

[6] PSO-FS initialized with relief scores

[7] Randomly initialized

[8] With statistical test validation

obtained with three PSO-FS instances using χ_2, *Relief* and *SU* filter scores are compared to continuous PSO randomly initialized.

The interesting results observed within the last experiment set with PSO-FS relying on *Relief* filter, could be extended to other instances of PSO-FS. The use of filters in swarm initialization confirms the effectiveness of PSO-FS with Relief, χ^2 and SU. The continuous PSO which showed better results over binary PSO, is now outperformed by all PSO-FS instances with almost all data sets. The significant improvements were obtained with high dimensional problems which confirm the same finding observed with previous experiment sets. On numerous benchmarks, PSO-FS succeeded to provide statistically validated improvement for all criteria (evaluation and validation). The three PSO-FS instances, returned comparable results for more than one criterion with a slight advantage to swarms initialized with χ^2.

5.3.2 Multiple filters impact on swarm

When selection of features are based on weight values, one can think that the more these values are representative of the feature-class dependence and the classification context, more the search process would be effective.

As shown in the previous empirical study, the PSO initialization scheme based on filter scores is suitable. Neverthetheses there are many filters and each one of them is based on a different criterion. Besides, from one filter to another, both ranking and scores are different. In general, merging different information sources is a challenging task. Such consideration would motivate the use of numerous filters within PSO not only to merge filters scores but also to evolve them within an optimization scheme. From the optimization perspective, this could enhance the search diversification and cooperation between particles initialized with different filters. All the steps of the FS-PSO approach are preserved, and only the generation process of initial solutions will change. In fact, each swarm particle p_i randomly selects the filter that will be used to build initial solution and starting particle solution. Although that the filters generate scores in different ranges, all ranges are normalized into $[-1, 1]$ range [9].

---

[9] $min - max$ normalization

Such PSO alternative provided not only a way to combine filter and wrapper, but a framework to merge filters scores.

Empirical study

Tables E.11, E.12, E.13, E.14, (*see ANNEX V*, *p.* 251-254) detail evaluation and validation performances of the respective best solutions obtained with PSO-FS initialized with 3 filters. It was compared to a randomly initialized PSO and PSO-FS based on one filter (*Relief*).

The comparison of the PSO based on swarm initialized with three filters to continuous PSO[10] confirms the superiority of PSO endowed with filters. The results enhancement cover all benchmarks as well as problem dimensions. Now, when we compare PSO-FS instance based on one filter to the instance using three filters[11], we can remark comparable results in terms of improvements. Moreover, multi-filter instance enhances performances of both CPU time and number of attributes. In addition, the number of statistically validated results with enhancements obtained with Multi-filter PSO scheme is superior to the number provided by one filter PSO scheme. For example, when PSO-FS based on Relief succeeds, for a given problem, to improve results with two or three criteria, the multi filter instance provides more validated criteria. Such results confirm, the stability and the robustness of the multi-filters PSO scheme, although comparable improvements provided by both PSO-FS approaches based on filters.

We should also note that the multi filter PSO scheme is more accurate and effective with high dimensional problem than with small and medium benchmarks. We think that for small/medium sized problems, filters enhance the search but PSO based on one filter are sufficient.

---

[10] with no filter

[11] labeled as PSO-FS2 within table results

5.4 Enhanced PSO-FS: EPSO-FS

CLPSO is considered as a recent alternative which attempted to alleviate the problem of premature convergence. The application of CLPSO to the FS problem is first, motivated by the need to encourage search diversification and to reduce the relative dominance of the best swarm particle when social component velocity is updated. In fact the CLPSO particles are attracted by good solutions which are not necessarily the best one. Next, we are interested by the study of the behavioral aspects of both basic PSO and CLPSO. Finally, we study the impact of the learning parameter on the PSO variant. In addition to the velocity update rule (*see eq.* 5.7), three main points characterize CLSPO [84]:

- instead of limiting the velocity adjustment to the best particle, all particles are potentially candidates to be selected as exemplars to guide the current particle.
- instead of learning from the same solution (swarm best particle) for all dimensions, each dimension of a particle could learn from a different solution.
- velocity update is based on only one particle instead of two with the basic PSO.

5.4.1 Swarm based on one filter

Here, the devised PSO is based on CLPSO variant and its is endowed with filter scores that serve as initializations schema. The next section, assesses empirically the impact of filter used on CLPSO as well as the impact of different filters on final results.

Empirical results

As reference approach based on binary PSO was outperformed by the proposed continuous PSO alternatives, it would be interesting to compare CLPSO to performances of PSO-FS[12] based on a filter and CL-PSO with a random initialization scheme. Tables E.15, E.16, E.17 and E.18, *p.* (E.15- E.18) detail empirical results of these experiments. From one hand, the CLPSO based on filter clearly outperforms CLPSO

---

[12] PSO-FS based on Relief filter is more accurate than random PSO and binary PSO

randomly initialized. The improvements were not limited to fitness criterion and attribute number. At least 3 criteria (among) were enhanced for each benchmark. For one third of the datasets the EPSO-FS succeeded to enhance CLPSO results with all criteria. Similar results were obtained by the comparison of continuous PSO and PSO-FS. Such results could confirm the superiority of swarm based filter approaches and the impact of filter scores initialization on final PSO performances. On the other hand, when we compare enhancements of both PSO based filter alternatives, we can remark comparable results with slight advantage to CLPSO. Besides, CLPSO based filter succeed to improve results of small and medium size benchmarks more effectively than PSO-FS.

The second set of experiments compares performances of four instances of CLPSO: three instances using different filters and a PSO instance randomly initialized (Tables E.19, E.20, E.21 and E.22, *p.* 259-262). As with PSO-FS, all CLPSO based filter instances confirmed their superiority over CLPSO randomly initialized. Here, we have another empirical results approving the effectiveness of hybrid wrapper-filter modeling for schema based on PSO. We should, note that use of filter scores within wrappers does not only contribute to enhance classification results and attribute number but also, reduces EPSO-FS running time.

Figure 5.1 illustrates the superiority of CLPSO instances based on filters, through the distribution of benchmarks and the number of validated criteria for each filter. For example, Swarm initialized with filter χ^2 succeed to enhance results for 5 criteria on 38% of benchmarks, and more than a half of the benchmarks with exactly 4 criteria [13]. This figure also shows the slight superiority of EPSO based χ^2.

5.4.2 Multiple filters impact on swarm

In this section, we empirically assess the CLPSO version of the swarm with more than one filter initializations scheme. Three filters were used: χ^2, Relief and SU.

Tables E.24, E.25, E.26 and E.27 (*see ANNEX V*, *p.* 264-267) detail evaluation and validation performances of the respective best solutions obtained with EPSO-FS

---

[13] Non cumulated

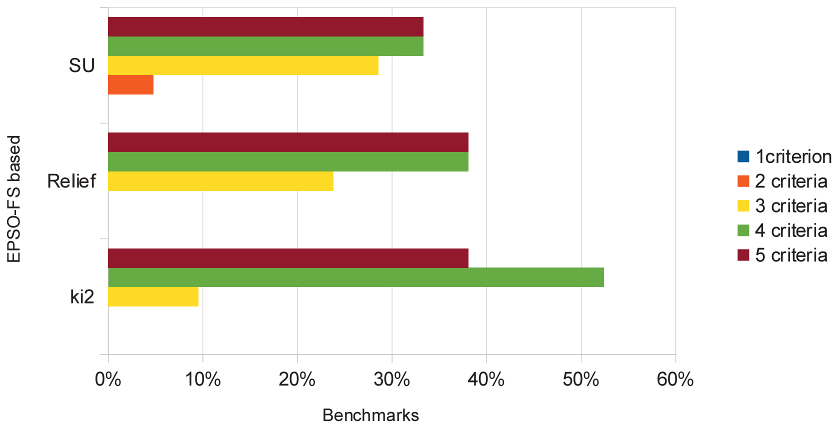


Figure 5.1: Enhancements distribution

initialized with 3 filters [14]. It was compared to a CLPSO based on (*Relief*) filter [15] and PSO-FS instance based on multi-filters scheme. The comparison of the two EPSO-FS instances showed the superiority of multi-filters scheme for almost all data sets. Similar results were found with continuous PSO-FS based on three filters. Hence, both PSO variants are more effective with swarm initialized by more than one filter scores. Finally, the comparison of PSO variants based on multi-filters initialization provides similar results with slight advantage of CLPSO variant on some data sets.

5.5 Summary of empirical analysis

We summarize the PSO experiments, by considering the following two criteria: fitness and the number of selected attributes. These comparisons use as references an existing recent binary PSO apporach and an PSO instances randomly initialized[16]. Resulting behaviors or tendencies are assessed according to their statistical validated results for all benchmarks. The following Table (*see Table* 5.1) illustrates for each PSO

---

[14] Table lines with EPSO-FS 2

[15] EPSO-FS(R)

[16] no filter used

instance the percentages of benchmarks for which the its better [17], worse and have a comparable results[18].

Globally results of fitness criterion are better than the reference approaches for both studied PSO alternative (continious PSO and CLPSO). In fact, the Binary PSO is outperformed by all our continuous PSO even in the case of the swarm not initialized by filter scores. Next, the use of PSO, randomly initialized, as reference approaches confirms the effect of filter scores on swarm evoltion. Comparable results were confirmed by CLSPO. Finally, PSO-FS based on multi-filters initialization scheme performs better than CLPSO based on one filter.

Compared Approaches		Validated improvement % (fitness)			Validated improvement % (Att)		
A	B	A	B	Comparable	A	B	Comparable
PBSO	PSO	10,53%	**84,21%**	5,26%	36,84%	**47,37%**	15,79%
BPSO	PSO-FS	10,53%	**89,47%**	0,00%	5,26%	**84,21%**	10,53%
PSO	PSO-FS(X2)	0,00%	**90,48%**	9,52%	0,00%	**100,00%**	0,00%
PSO	PSO-FS(R)	9,52%	**90,48%**	0,00%	0,00%	**100,00%**	0,00%
PSO	PSO-FS(SU)	0,00%	**100,00%**	0,00%	4,76%	**95,24%**	0,00%
PSO	PSO-FS2	0,00%	**76,19%**	23,81%	0,00%	**95,24%**	4,76%
CLPSO	PSO-FS(R)	14,29%	**71,43%**	14,29%	4,76%	**95,24%**	0,00%
CLPSO	CLPSO(R)	9,52%	**80,95%**	9,52%	0,00%	**100,00%**	0,00%
CLPSO	CLPSO(X2)	0,00%	**90,48%**	9,52%	0,00%	**100,00%**	0,00%
CLPSO	CLPSO(SU)	0,00%	**85,71%**	14,29%	0,00%	**95,24%**	4,76%
CLPSO(R)	PSO-FS2	4,76%	**71,43%**	23,81%	0,00%	**95,24%**	4,76%
CLPSO(R)	CLPSO2	4,76%	**85,71%**	9,52%	0,00%	**100,00%**	0,00%

Table 5.1: Empirical Study Synthesis of PSO

5.6 Conclusion

This chapter was devoted to feature selection modeling with swarm approaches. We devised a number of continuous PSO variants using weighted solutions representation. PSO based approaches succeed through the different proposed alternatives to provide a framework allowing the hybridization of filters and wrappers. In fact, empirical study has confirmed the superiority of models based on swarm initialized with filters.

---

[17] improvement over 10 runs (at least) is statistically validated

[18] statistical test is not in favor of any PSO instance

Such encouraging results encourage us to investigate more PSO variants as well as the integration of information fusion modeling tools to cope with the diversity of problem knowledge sources provided by filter scores.

Chapter 6

Conclusion And Perspectives

6.1 Conclusion

After a detailed review of both featured and recent trends of feature selection modeling, we investigated a set of approaches that have covered a set of concepts in relation with combinatorial optimization as well as feature Selection problem specificities. These approaches were ranging from local search to evolutionary and swarm paradigms. The second chapter covered a set of local search operators devised to endow genetic algorithm with intensification and guidance mechanisms. The resulting memetic alternatives were studied from different perspectives and aspects. The effectiveness and accuracies were empirically studied and compared to reference approaches using a set of benchmarks and validation criteria. The experiments showed the superiority of the devised memetic alternatives for small and medium sized benchmarks. In addition, the adaptation efforts of LS operators to high dimensional data sets have shown its effectiveness in comparison to GA and a recent reference memetic approaches. The third chapter was devoted to investigation of the second contribution of this thesis. The proposed GRASP scheme was adapted to the requirement of the FS problem. In fact, effective and composite LS operators proposed within the memetic schema were deployed as refinement component of the GRASP. Besides, the devised GRASP-FS alternatives allowed us to combine both wrapper and filter in a natural way where filters participate to the construction of initial solutions and the second stage refines them through local search. The fourth chapter explored a swarm optimization paradigm. The investigated approach was designed as a continuous PSO based on weigthing scheme for attribute selection. The design of continuous PSO based on normalized real valued encoding allows us to hybridize wrappers with filters through the usage of filter scores as particle initialization schema.

Throughout the different contributions both quantitative and qualitative criteria were studied, and an important aspect that was taken into consideration: the exploration-exploitation trade-off. Such aspects mainly determined the effectiveness and the accuracy of the devised algorithms for the optimization paradigm investigated. From the design perspective, a good optimization algorithm balances these contradictory objectives effectively.

6.2 Perspectives

Although that the proposed and studied feature selection approaches could be considered as alternatives to the existing featured and reference approaches, numerous aspects, ranging from design issues to the empirical behavioral assessment, should be investigated. In fact, memetic approaches could be enhanced to schema deploying different local search operators. The diversification of the neighborhood structure might be beneficial for both local search and the whole evolutionary process. Adaptive LS behaviors (*i.e.* meme selection mechanism), complexity, and exploration-exploitation trade-off issues should be investigated. Besides, the deployment of different local search operators on distributed and cooperative instances of evolutionary strategies (*i.e.* island models) might be an interesting research perspective for the feature selection problem.

The GRASP-FS could be endowed with *path-relinking*. Such mechanism could be implemented in different ways and the resulting behaviors would be interesting to investigate. Another, idea that deserves to be studied is to use filter knowledge, within construction stage to generate new solutions from feature subsets of best solutions found.

Also, PSO could be extended to more enhanced PSO alternatives like UPSO

[25, 105] or FIPS [98]. It is also interesting to study the behavior of PSO when it is endowed with LS operators and to analyze the impact of LS on different PSO variants.

In addition, variable and self adjustable threshold for attribute selection in PSO-FS could be helpful for researches looking for adaptive PSO for FS.

On the other hand the work done in this thesis, might pave the way to the adoption of new formulation issues in relation with both machine learning and combinatorial optimization. In fact, both validation and evaluation criteria could be investigated in order to design more reliable fitness function. In addition, the proposed LS operators, could be also considered in the design of Variable Neighborhood Search (VNS) tackling the problem of feature selection. Other swarm optimization methods could be investigated to hybridize filters scores as well as wrappers and filters. Such methods could be based on Artificial bees or ant colonies.

Finally, all approaches that could be devised should take into consideration, in addition to the adaptation effort, the trade-off between exploration and exploitation as well as the investigation of the behaviors of the resulting processes.

Appendix A

Annex I: Validation Protocol

This annex covers main steps of the assessment procedures, validation, experiments settings and empirical comparison of the proposed feature selection approaches.

Literature of FS endeavors, provides empirical assessment of their approaches either as application to specific problem (*i.e.* fault detection, financial modeling, gene bio-marker identification, *etc*) or using benchmark data sets with a given range of problem dimensionalities. Therefore there is no clear evidence about the effectiveness of references or new approaches for its ability to handle different problem dimensionality scales with the same robustness. The need to such study becomes of first requirement, particularly for the behavioral study of the proposed approaches.

The proposed assessment methodology that was used throughout the thesis, covers evaluation requirements of both search space exploration and validation stage of the final solution [1].

The first stage needs an evaluation procedure which is able to assign a fitness value to a solution representing a set of selected attributes. In the case of wrappers , usually, a classifier is used to assess the solution accuracy. Once the classifier terminates learning from data, the classification accuracy is evaluated on a different data set (test data). Therefore, we opt for generalization (test) error as a criterion for fitness evaluation, and hold out (HO). The same classifier is used during the search process.

---

[1] Best solution found

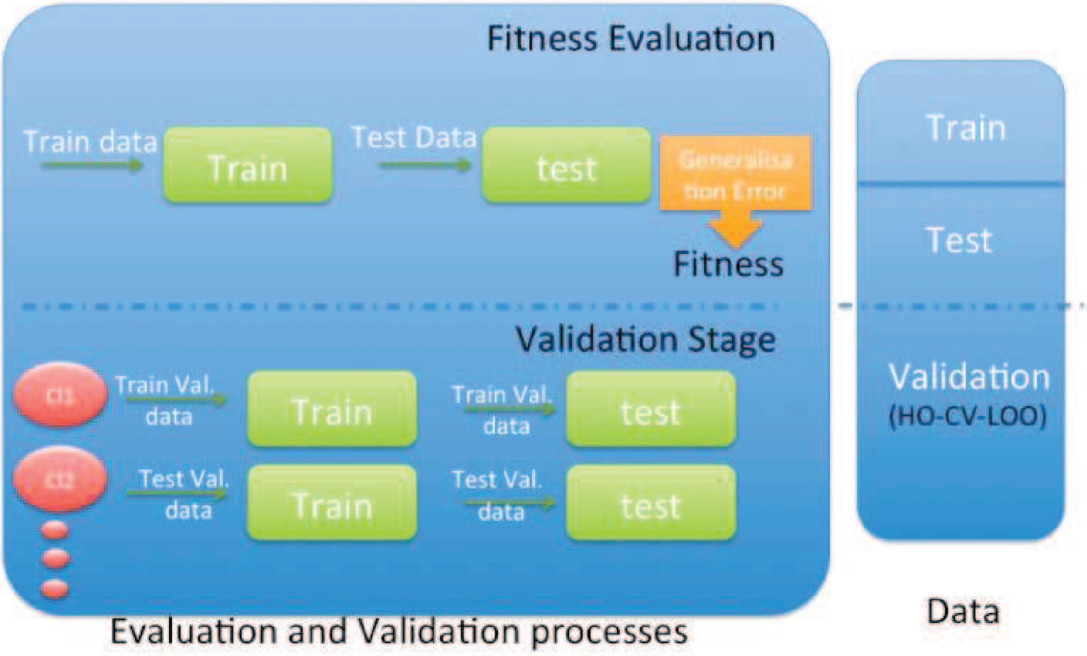


Figure A.1: Assessment procedures

The second stage requires more robust and reliable evaluation process to evaluate final solution returned by the search process. The selected attributes could be optimized for the classifier used in first stage, and the data used. It could be in some way biased by the search process and the classifier, when we try to minimize the error rate on test data. In order to, provide reliable assessment procedures of selected attributes, we decided to evaluate them on different classifiers and data than those used during the search process. Consequently, initial data set will be split in two subsets, one for the search and fitness evaluation and the second for validation. We should note, that the two data set generated after the split maintains class distribution of the original data set dataset. It is the same for the data used with the first stage for fitness evaluation. Besides, the classification performances of the selected features will be assessed on different classifiers. In comparison to the classical wrapper approaches, often, validation stage is omitted and when it is done it is limited to the test on different data set. Figure A.1 illustrates the assessment procedures for both fitness evaluation and validation stage.

Experiments involving search and validation stages are repeated at least ten times. Indeed, the comparison to the respective reference approaches was based on mean and standard deviation values of experiments of the same types. Statistical tests were provided with T-*test*. Experiences of a given approach or model instance are considered statistically validated with a confidence level of $97,5\%$ only if the value of T-*test* is approximatively less than $-2,086$ [2]. The set of adopted procedures and metrics in relation with experiments search and validation define the *validation protocol*.

The validation protocol relies on a set of featured benchmarks used by different feature selection reference approaches as well as for classification matters. We should note, that to the best of our knowledge, there was not a study that has assessed the behaviors of its approaches on different problem dimensionality. To this end, we provides three sets of benchmarks (*see Table* A.1). The small data sets provides less than 100 attributes, while those of medium size ranges from 100 to 1000 attributes. Finally, large data sets correspond to high dimensional data sets with more than 1000 attributes.

---

[2] in comparison to reference approach, and for experiments repeated at least 10 times

Data Set	Attributes	Instances	Classes	Type	Source	Description
Sonar	60	208	2	Small	[16]	Classification of sonar signals
Spam Base	57	4601	2	Small	[16]	Email spam detection
Soybean	35	307	19	Small	[16]	Soybean disease
Arrhythmia	279	452	2	Medium	[16]	Cardiac arrhythmia
Secom	590	1567	2	Medium	[16]	Semi-conductor manufacturing process
Semeion	256	1593	10	Medium	[16]	Handwritten digits recognition
Breast	24481	97	2	Large	[1]	Microarrays of breast cancer dataset
CNS	7128	60	2	Large	[1]	Patients outcome prediction for central nervous system embryonal tumor
Colon	2000	60	2	Large	[1]	Colon tumor
Leukemia_3C	7129	72	3	Large	[1]	blood cell cancer
Leukemia_4C	7129	72	4	Large	[1]	blood cell cancer
Lung_Cancer	15154	253	2	Large	[1]	Lung Cancer
MLL	12582	72	3	Large	[1]	MLL: Mixed Lineage Leukemia
Ovarian	15154	253	2	Large	[1]	Ovarian cancer
SRBCT	2308	83	4	Large	[1]	Microarray gene expression profiles of small, round blue cell tumors
9_tumors	5726	60	9	Large	[2]	Transcript profiles of 9 common human tumors
11_tumors	12533	174	11	Large	[2]	Transcript profiles of 11 common human tumors
14_tumors	15009	308	26	Large	[2]	Transcript profiles of 14 common human tumors
Brain_Tumor	10367	50	4	Large	[2]	Brain Tumor
Prostate_Tumor	10509	102	2	Large	[2]	Prostate Tumor
Lymphoma	4026	62	3	Large	[1]	Lymphatic cells cancer (immune system)
Challenge_2004	27679	90	34	Large	ECML PKDD'04 challenge [3]	http://www.upo.es/eps/bigs/datasets.html

Table A.1: Benchmark data sets

Appendix B

Annex II: FS-Framework Overview

In this section, we shed some light on the implemented framework that was used for the empirical validation of the proposed approaches. We detail the software architecture as the common and reusable services. Figures B.1 and B.2, illustrate the big picture of the of the software that we have developed and extended throughout the thesis.

Figure B.1, shows the main search strategies implemented. Each metaheuristic is considered as a component of the FS-framework which provides them low level routines and basic services. Common routines bring basic services which include: Data preparation for search and validation stages, output stream redirection of experiments to respective log files, and assessments procedures required by the validation protocol. Such services were designed to be reusable by search strategies. Consequently, any new heuristic would reuse common routines as services. Besides the framework interfaces the Weka package [129] was used for classification tasks and to retrieve filter scores.

Once an experiment *batch* [1] terminates, the analyzer module generates compact and readable results through the transformation of experiments log-file. The transformation process is illustrated by Figure B.3 The first stage provides inputs to the transformation process through stream redirection. The pruning stage selects experiments that will be analyzed while the aggregation stage prepares the data for

---

[1] Set of experiments repeated n times

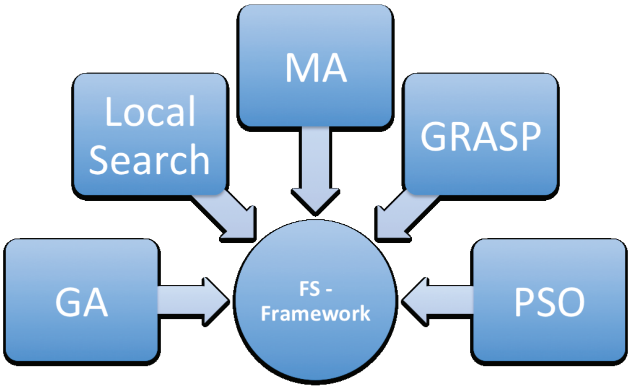


Figure B.1: Framework Architecture: Big picture

statistical analysis by extracting requested results from log-files. Statistical Analysis stage computes according to the number of experiments runs, means, standard deviation and statistical tests (T-*test*).

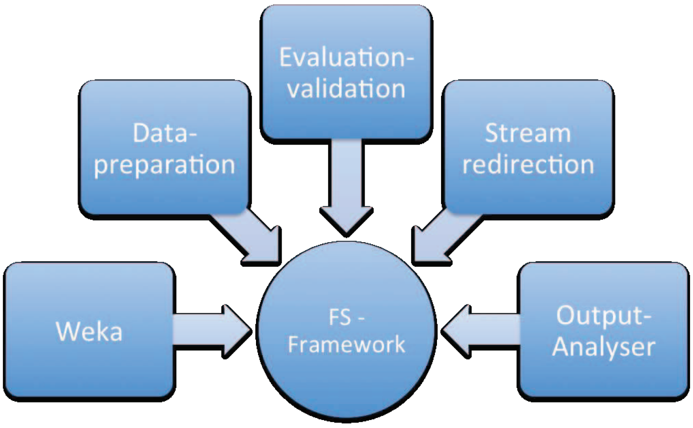


Figure B.2: Framework orthogonal services

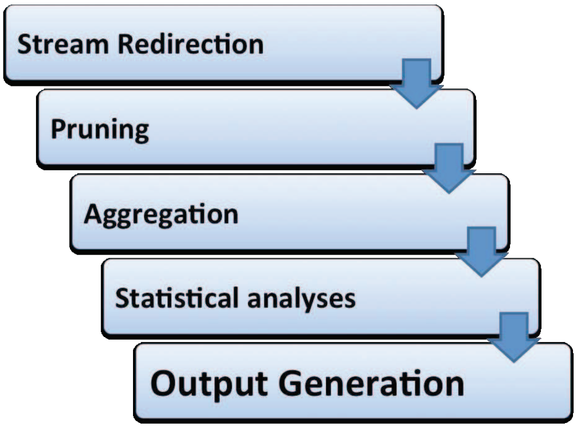


Figure B.3: Analyzer process: main Steps

Appendix C

Annex III: Detailed Empirical Results of Chapter III

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Sonar	-	M :	4,20%	25,17%	41,38%	4945022,09	22,27
		Sd:	1,68%	1,73%	2,22%	1811950,54	4,78
		T-:	-	-	-	-	-
	FLS1	M :	5,81%	25,97%	42,47%	**2057890,07**	**18,51**
		Sd:	1,60%	2,68%	3,41%	941186,38	3,61
		T-:	29,04	7,83	56,88	-19,5	-16,23
	It.FLS1	M :	5,89%	26,38%	42,16%	**1898746,08**	**17,46**
		Sd:	1,51%	3,27%	2,81%	783274,58	4,53
		T-:	34,42	10,65	9,06	-20,52	-16,03
Spambase	-	M :	5,52%	9,92%	15,34%	34180282,63	23,88
		Sd:	0,39%	0,80%	1,93%	11210986,2	3,6
		T-:	-	-	-	-	-
	FLS1	M :	5,88%	9,91%	15,59%	**17992626,02**	24,41
		Sd:	0,36%	0,66%	1,58%	6840868,62	3,58
		T-:	3,92	-0,11	3,69	-8,08	0,62
	It.FLS1	M :	5,98%	10,03%	**15,13%**	**16175281,71**	23,33
		Sd:	0,58%	0,72%	1,47%	6890785,28	3,7
		T-:	5,03	0,95	-2,73	-9	-0,63
Soybean	-	M :	3,80%	6,61%	60,93%	132256274,25	16,63
		Sd:	0,44%	0,75%	2,20%	22491792,41	0,74
		T-:	-	-	-	-	-
	FLS1	M :	4,16%	6,95%	61,42%	**65302386,89**	**16,34**
		Sd:	0,42%	0,77%	2,27%	18691227,55	1,35
		T-:	9,34	4,01	1,94	-12,48	-5,08
	It.FLS1	M :	4,18%	6,85%	61,31%	**63108997,28**	**16,05**
		Sd:	0,55%	0,92%	2,17%	17533184,75	1,4
		T-:	9,45	2,87	1,48	-12,89	-10,7
Arrythmia	-	M :	17,85%	40,09%	42,12%	72989654,33	42,67
		Sd:	1,62%	1,74%	1,34%	20906189,17	5,16
		T-:	-	-	-	-	-
	FLS1	M :	22,63%	40,38%	42,42%	**36445116,03**	41,03
		Sd:	2,28%	2,57%	1,26%	13510740,06	9,14
		T-:	112,24	1,21	1,68	-7,29	-1,55
	It.FLS1	M :	23,19%	40,73%	42,73%	**32049749,13**	**34,67**
		Sd:	2,41%	3,49%	1,52%	12710273,72	12,47
		T-:	145,33	2,23	3,41	-8,11	-6,29
Secom	-	M :	5,15%	7,05%	8,28%	10388936,67	12
		Sd:	0,47%	0,41%	0,28%	4498211,79	4
		T-:	-	-	-	-	-
	FLS1	M :	6,33%	**6,73%**	8,80%	**2567049,82**	**4,13**
		Sd:	0,61%	0,16%	1,56%	1443081,99	3,63
		T-:	36,63	-4,42	9,61	-12,5	-15,64
	It.FLS1	M :	5,98%	**6,71%**	9,37%	**2668796,41**	**4,51**
		Sd:	0,70%	0,16%	1,56%	1298633,22	3,1
		T-:	25,96	-4,62	34,82	-12,35	-14,92
Semeion	-	M :	10,90%	16,52%	81,15%	172086309,86	49
		Sd:	1,13%	1,34%	0,33%	29063662,28	1,41
		T-:	-	-	-	-	-
	FLS1	M :	14,77%	16,97%	81,25%	**86968845,26**	49,6
		Sd:	1,59%	1,21%	0,32%	22934450,76	1,88
		T-:	35,99	3,82	6,12	-15,32	13,13
	It.FLS1	M :	15,09%	17,51%	81,20%	**89357655,23**	**48,5**
		Sd:	1,66%	1,53%	0,28%	29140532,37	3,16
		T-:	37,87	8,17	3,47	-14,88	-40

Table C.1: GA *vs* MAs respectively endowed with FLS and $It.FLS$

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib
Sonar	MBEGA	M :	5,16%	24,65%	42,56%	3298287,84	22,89
		Sd:	1,58%	3,29%	2,83%	1461028,35	5,57
		T-:	-	-	-	-	-
	FLS1	M :	5,81%	25,97%	42,47%	**2057890,07**	**18,51**
		Sd:	1,60%	2,68%	3,41%	941186,38	3,61
		T-:	7,85	15,92	-0,62	-22,11	-28,72
	it.FLS1	M :	5,89%	26,38%	**42,16%**	**1898746,08**	**17,46**
		Sd:	1,51%	3,27%	2,81%	783274,58	4,53
		T-:	9,23	17,85	-2,32	-24,49	-22,24
Spambase	MBEGA	M :	5,59%	9,36%	15,53%	18637193,26	18,79
		Sd:	0,37%	1,08%	1,90%	5962776,42	4,77
		T-:	-	-	-	-	-
	FLS1	M :	5,88%	9,91%	15,59%	**17992626,02**	24,41
		Sd:	0,36%	0,66%	1,58%	6840868,62	3,58
		T-:	8,21	10,07	0,49	-3,26	28,66
	it.FLS1	M :	5,98%	10,03%	**15,13%**	**16175281,71**	23,33
		Sd:	0,58%	0,72%	1,47%	6890785,28	3,7
		T-:	10,74	12,71	-3,22	-14,5	26,87
Soybean	MBEGA	M :	3,91%	6,51%	61,16%	81449381,88	16,31
		Sd:	0,35%	0,67%	2,35%	16299303,56	1,4
		T-:	-	-	-	-	-
	FLS1	M :	4,16%	6,95%	61,42%	**65302386,89**	16,34
		Sd:	0,42%	0,77%	2,27%	18691227,55	1,35
		T-:	15,19	10,87	1,61	-9,03	0,32
	it.FLS1	M :	4,18%	6,85%	61,31%	**63108997,28**	**16,05**
		Sd:	0,55%	0,92%	2,17%	17533184,75	1,4
		T-:	13,31	8,48	0,89	-10,24	-3,05
Arrythmia	MBEGA	M :	18,90%	38,04%	42,22%	39118432	33,86
		Sd:	1,72%	2,81%	0,73%	15702992,78	14,92
		T-:	-	-	-	-	-
	FLS1	M :	22,63%	40,38%	42,42%	**36445116,03**	41,03
		Sd:	2,28%	2,57%	1,26%	13510740,06	9,14
		T-:	40,31	39,61	1,96	-2,17	5,62
	it.FLS1	M :	23,19%	40,73%	42,73%	**32049749,13**	34,67
		Sd:	2,41%	3,49%	1,52%	12710273,72	12,47
		T-:	47,65	16,09	4,95	-5,24	0,55
Secom	MBEGA	M :	5,31%	7,01%	8,77%	6773548,46	8,85
		Sd:	0,50%	0,36%	0,89%	3119318,3	4,04
		T-:	-	-	-	-	-
	FLS1	M :	6,33%	**6,73%**	8,80%	**2567049,82**	**4,13**
		Sd:	0,61%	0,16%	1,56%	1443081,99	3,63
		T-:	16,58	-26,97	0,65	-38,37	-30,88
	it.FLS1	M :	5,98%	**6,71%**	9,37%	**2668796,41**	**4,51**
		Sd:	0,70%	0,16%	1,56%	1298633,22	3,1
		T-:	10,95	-28,59	21,71	-38,23	-29,32
Semeion	MBEGA	M :	10,74%	15,20%	81,04%	123425926,22	51,44
		Sd:	0,96%	1,33%	0,16%	18016344,16	2,79
		T-:	-	-	-	-	-
	FLS1	M :	14,77%	16,97%	81,25%	**86968845,26**	**49,6**
		Sd:	1,59%	1,21%	0,32%	22934450,76	1,88
		T-:	66,77	15,08	27,63	-38,52	-6,7
	it.FLS1	M :	15,09%	17,51%	81,20%	**89357655,23**	**48,5**
		Sd:	1,66%	1,53%	0,28%	29140532,37	3,16
		T-:	66,03	19,16	26,7	-35,83	-10,83

Table C.2: MBEGA *vs* MAs respectively endowed with $It.FLS1$ and $It.FLS1$

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Sonar	-	M :	4,20%	25,17%	41,38%	4945022,09	22,27
		Sd:	1,68%	1,73%	2,22%	1811950,54	4,78
		T-:	-	-	-	-	-
	it.FLS1	M :	5,89%	26,38%	42,16%	**1898746,08**	**17,46**
		Sd:	1,51%	3,27%	2,81%	783274,58	4,53
		T-:	34,42	10,65	9,06	-20,52	-16,03
	IFLS	M :	5,24%	27,45%	**40,04%**	**1625966,05**	**16,5**
		Sd:	1,98%	3,57%	3,72%	712344,82	5,65
		T-:	14,76	31,76	-47,52	-22,56	-22,14
Spambase	-	M :	5,52%	9,92%	15,34%	34180282,63	23,88
		Sd:	0,39%	0,80%	1,93%	11210986,2	3,6
		T-:	-	-	-	-	-
	it.FLS1	M :	5,98%	10,03%	**15,13%**	**16175281,71**	23,33
		Sd:	0,58%	0,72%	1,47%	6890785,28	3,7
		T-:	5,03	0,95	-2,73	-9	-0,63
	IFLS	M :	5,83%	9,94%	15,76%	**15257565,92**	24
		Sd:	0,52%	0,83%	1,65%	4705756,86	3,74
		T-:	3,22	0,22	6,15	-9,46	0,14
Soybean	-	M :	3,80%	6,61%	60,93%	132256274,25	16,63
		Sd:	0,44%	0,75%	2,20%	22491792,41	0,74
		T-:	-	-	-	-	-
	it.FLS1	M :	4,18%	6,85%	61,31%	**63108997,28**	**16,05**
		Sd:	0,55%	0,92%	2,17%	17533184,75	1,4
		T-:	9,45	2,87	1,48	-12,89	-10,7
	IFLS	M :	4,22%	6,82%	61,64%	**56129293,65**	**16,3**
		Sd:	0,39%	0,81%	2,18%	6264024,58	1,18
		T-:	10,08	2,5	2,73	-14,19	-5,75
Arrythmia	-	M :	17,85%	40,09%	42,12%	72989654,33	42,67
		Sd:	1,62%	1,74%	1,34%	20906189,17	5,16
		T-:	-	-	-	-	-
	it.FLS1	M :	23,19%	40,73%	42,73%	**32049749,13**	**34,67**
		Sd:	2,41%	3,49%	1,52%	12710273,72	12,47
		T-:	145,33	2,23	3,41	-8,11	-6,29
	IFLS	M :	21,84%	41,54%	43,14%	83127748,53	**28,63**
		Sd:	2,66%	2,36%	1,93%	55025810,76	11,2
		T-:	16,07	5,09	4,38	2,02	-12,84
Secom	-	M :	5,15%	7,05%	8,28%	10388936,67	12
		Sd:	0,47%	0,41%	0,28%	4498211,79	4
		T-:	-	-	-	-	-
	it.FLS1	M :	5,98%	**6,71%**	9,37%	**2668796,41**	**4,51**
		Sd:	0,70%	0,16%	1,56%	1298633,22	3,1
		T-:	25,96	-4,62	34,82	-12,35	-14,92
	IFLS	M :	6,13%	**6,80%**	8,67%	**2136618,26**	**3,58**
		Sd:	0,77%	0,38%	1,57%	1335901,87	3,08
		T-:	11,15	-3,05	12,45	-12,69	-13,95
Semeion	-	M :	10,90%	16,52%	81,15%	172086309,86	49
		Sd:	1,13%	1,34%	0,33%	29063662,28	1,41
		T-:	-	-	-	-	-
	it.FLS1	M :	15,09%	17,51%	81,20%	**89357655,23**	**48,5**
		Sd:	1,66%	1,53%	0,28%	29140532,37	3,16
		T-:	37,87	8,17	3,47	-14,88	-40
	IFLS	M :	16,17%	19,50%	81,47%	258732416,17	**44,58**
		Sd:	1,94%	1,67%	0,44%	85018732,66	3,26
		T-:	27,71	12,26	6,82	14,24	-14,79

Table C.3: GA *vs* MAs respectively endowed with $It.FLS1$ and $IFLS$

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib
Sonar	AF	M :	7,88%	28,50%	41,42%	628848,1	18,3
		Sd:	1,42%	4,15%	3,97%	263522,47	6,18
		T-:	-	-	-	-	-
	FLS1	M :	**6,83%**	**26,84%**	42,15%	705239,22	20,16
		Sd:	1,78%	3,90%	2,73%	306780,3	5,22
		T-:	-41,28	-13,77	2,26	4,1	4,2
	FLS2	M :	8,48%	**26,59%**	41,93%	618405,35	**16,41**
		Sd:	2,12%	2,95%	3,43%	297707,6	3,92
		T-:	7,3	-79,08	1,66	-0,59	-4,26
Spambase	AF	M :	6,91%	10,85%	16,24%	5477152,6	24,2
		Sd:	1,01%	1,29%	1,98%	905929,68	2,82
		T-:	-	-	-	-	-
	FLS1	M :	**6,36%**	**10,11%**	**15,36%**	6058079,71	24,76
		Sd:	0,66%	0,78%	1,47%	2460281,39	3,27
		T-:	-6,1	-30,27	-10,72	9,19	2,01
	FLS2	M :	**6,71%**	**10,12%**	**15,74%**	6130995,57	**23,2**
		Sd:	0,56%	0,78%	1,95%	2835806,16	3,49
		T-:	-2,3	-27,83	-4,16	8,56	-3,54
Soybean	AF	M :	5,56%	8,98%	62,76%	17173025,4	15,2
		Sd:	0,79%	2,24%	3,95%	3835722,11	1,75
		T-:	-	-	-	-	-
	FLS1	M :	**4,68%**	**6,88%**	**60,38%**	18231616	16,33
		Sd:	0,60%	1,20%	2,24%	4443395,67	0,8
		T-:	-10	-15,63	-6,06	9,95	6,29
	FLS2	M :	**4,83%**	**7,36%**	**60,33%**	20324036,1	15,6
		Sd:	0,43%	1,39%	2,10%	7534630,92	1,61
		T-:	-7,91	-12,15	-6,12	23,56	2
Arrythmia	AF	M :	27,14%	42,19%	43,73%	21437945,27	28,27
		Sd:	2,05%	3,10%	2,35%	15475436,96	14,19
		T-:	-	-	-	-	-
	FLS1	M :	**23,61%**	**40,85%**	**42,84%**	**11465370,72**	38,21
		Sd:	3,53%	3,28%	1,65%	4422342,17	12,53
		T-:	-20,08	-7,03	-5,04	-8,19	9,01
	FLS2	M :	**26,48%**	42,48%	43,82%	**10367383,7**	36,7
		Sd:	2,54%	2,96%	2,04%	4480849,19	10,59
		T-:	-3,48	1,24	0,45	-9,03	6,25
Secom	AF	M :	6,51%	6,77%	8,59%	2470382,24	5,18
		Sd:	0,48%	0,28%	1,90%	4875783,25	6,71
		T-:	-	-	-	-	-
	FLS1	M :	**6,21%**	6,78%	8,86%	**1119385,16**	6,84
		Sd:	0,69%	0,33%	1,62%	755082,59	8,16
		T-:	-11,08	1,68	3,05	-10,58	5,75
	FLS2	M :	6,47%	**6,68%**	8,62%	**857933,71**	**3,04**
		Sd:	0,43%	0,13%	1,51%	725750,43	2,91
		T-:	-0,74	-4,77	0,35	-12,74	-6,11
Semeion	AF	M :	17,17%	19,22%	81,46%	64127293,33	46,33
		Sd:	1,74%	2,19%	0,47%	14045264,33	4,62
		T-:	-	-	-	-	-
	FLS1	M :	**15,34%**	18,16%	**81,32%**	**26848309,69**	48,31
		Sd:	3,14%	1,79%	0,34%	6879201,84	2,25
		T-:	-2,73	-1,31	-6,18	-19,14	1,11
	FLS2	M :	18,01%	19,25%	81,46%	**25366502,22**	45,87
		Sd:	1,76%	2,21%	0,36%	9271511,53	4,53
		T-:	1,25	0,05	0,16	-19,86	-0,26

Table C.4: Memetic algorithms comparison: AF *vs* $FLS1$ and $FLS2$

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Sonar	AF	M :	7,88%	28,50%	41,42%	628848,1	18,3
		Sd:	1,42%	4,15%	3,97%	263522,47	6,18
		T-:	-	-	-	-	-
	FLS2	M :	8,48%	**26,59%**	41,93%	618405,35	**16,41**
		Sd:	2,12%	2,95%	3,43%	297707,6	3,92
		T-:	7,3	-79,08	1,66	-0,59	-4,26
	IFLS	M :	**6,57%**	**26,81%**	42,02%	**463557,25**	**15,67**
		Sd:	2,78%	3,14%	3,12%	133370,31	4,29
		T-:	-13,77	-22,69	1,41	-9,46	-5,93
Spambase	AF	M :	6,91%	10,85%	16,24%	5477152,6	24,2
		Sd:	1,01%	1,29%	1,98%	905929,68	2,82
		T-:	-	-	-	-	-
	FLS2	M :	**6,71%**	**10,12%**	**15,74%**	6130995,57	**23,2**
		Sd:	0,56%	0,78%	1,95%	2835806,16	3,49
		T-:	-2,3	-27,83	-4,16	8,56	-3,54
	IFLS	M :	6,96%	**10,48%**	**15,78%**	5159665,36	**20,36**
		Sd:	1,41%	0,81%	1,59%	2089628,95	5,85
		T-:	0,2	-5,08	-5,63	-1,75	-7,58
Soybean	AF	M :	5,56%	8,98%	62,76%	17173025,4	15,2
		Sd:	0,79%	2,24%	3,95%	3835722,11	1,75
		T-:	-	-	-	-	-
	FLS2	M :	**4,83%**	**7,36%**	**60,33%**	20324036,1	15,6
		Sd:	0,43%	1,39%	2,10%	7534630,92	1,61
		T-:	-7,91	-12,15	-6,12	23,56	2
	IFLS	M :	**4,78%**	**7,42%**	**61,10%**	16871408	15,5
		Sd:	0,82%	1,97%	3,35%	3974319,42	1,83
		T-:	-7,46	-11,52	-4,02	-1,79	1,62
Arrythmia	AF	M :	27,14%	42,19%	43,73%	21437945,27	28,27
		Sd:	2,05%	3,10%	2,35%	15475436,96	14,19
		T-:	-	-	-	-	-
	FLS2	M :	**26,48%**	42,48%	43,82%	**10367383,7**	36,7
		Sd:	2,54%	2,96%	2,04%	4480849,19	10,59
		T-:	-3,48	1,24	0,45	-9,03	6,25
	IFLS	M :	**23,01%**	42,42%	43,93%	82732595,38	33,13
		Sd:	4,20%	3,53%	1,77%	59253033,88	12,06
		T-:	-6,09	0,48	0,83	10,89	2,79
Secom	AF	M :	6,51%	6,77%	8,59%	2470382,24	5,18
		Sd:	0,48%	0,28%	1,90%	4875783,25	6,71
		T-:	-	-	-	-	-
	FLS2	M :	6,47%	**6,68%**	8,62%	**857933,71**	**3,04**
		Sd:	0,43%	0,13%	1,51%	725750,43	2,91
		T-:	-0,74	-4,77	0,35	-12,74	-6,11
	IFLS	M :	**6,25%**	6,80%	**7,97%**	3840185,88	4,5
		Sd:	0,70%	0,40%	1,22%	8213648,22	6,39
		T-:	-3,25	1,64	-8,13	3,08	-1,35
Semeion	AF	M :	17,17%	19,22%	81,46%	64127293,33	46,33
		Sd:	1,74%	2,19%	0,47%	14045264,33	4,62
		T-:	-	-	-	-	-
	FLS2	M :	18,01%	19,25%	81,46%	**25366502,22**	45,87
		Sd:	1,76%	2,21%	0,36%	9271511,53	4,53
		T-:	1,25	0,05	0,16	-19,86	-0,26
	IFLS	M :	18,05%	22,84%	**81,33%**	115284482,5	41,5
		Sd:	2,13%	1,49%	0,09%	29707595,88	4,95
		T-:	0,87	3,76	-3,3	4,79	-1,94

Table C.5: Memetic algorithms comparison: AF *vs* $FLS2$ and $IFLS$

Data	Local search	Measure	Fitness	**Validation**	**Validation**	CPU(ms)	# Attrib.
Breast	- (GA only)	M :	2,13%	45,19%	50,14%	5086055,47	38,33
		Sd:	2,56%	7,28%	4,87%	1719792,54	6,65
		T-:	-	-	-	-	-
	AF^+	M :	11,11%	49,49%	**46,48%**	**1054957**	**11,56**
		Sd:	5,21%	5,08%	5,71%	1002712,06	6,86
		T-:	8,45	3,6	-5,4	-36,74	-25,3
	MB^+	M :	3,64%	48,22%	**48,26%**	**830939,18**	**3,82**
		Sd:	2,16%	6,63%	4,24%	264272,02	0,87
		T-:	3,83	4,26	-3,33	-51,93	-53,2
CNS	- (GA only)	M :	0,00%	39,23%	43,95%	2951100,54	34,69
		Sd:	0,00%	10,70%	5,11%	1517342,83	11,47
		T-:	-	-	-	-	-
	AF^+	M :	4,24%	48,06%	**41,58%**	**884081,91**	**15,18**
		Sd:	6,16%	5,12%	6,13%	623905,41	7,11
		T-:	19,25	38,61	-2,33	-48,7	-63,71
	MB^+	M :	0,83%	44,25%	**41,67%**	**172647,13**	**3,75**
		Sd:	2,36%	7,00%	6,52%	43401,1	1,91
		T-:	8	14,61	-6,88	-111,89	-225,03
Colon	- (GA only)	M :	1,92%	27,79%	32,66%	2154451,46	28,15
		Sd:	3,00%	4,83%	5,79%	1209434,43	10,54
		T-:	-	-	-	-	-
	AF^+	M :	6,73%	**24,32%**	34,09%	**1443520,77**	**26,08**
		Sd:	4,75%	5,06%	7,36%	927548,22	12,31
		T-:	8,67	-7,27	1,7	-4,65	-2,53
	MB^+	M :	3,75%	**22,71%**	**30,77%**	**145156,1**	**3,2**
		Sd:	6,04%	9,33%	10,51%	56091,5	1,14
		T-:	3,64	-5,53	-4,74	-14,68	-53,61

Table C.6: GA *vs* MAs endowed with AF^+ and MB^+ (1)

Data	Local search	Measure	Fitness	**Validation**	**Validation**	CPU (ms)	# Attrib.
Ovarian	- (GA only)	M :	0,00%	3,08%	19,50%	13669924,58	26,67
		Sd:	0,00%	2,40%	10,12%	6257542,88	7,99
		T-:	-	-	-	-	-
	AF^+	M :	0,00%	**2,59%**	**16,13%**	29587089,5	31,5
		Sd:	0,00%	2,10%	11,00%	18545416,25	8,87
		T-:	-	-3,82	-2,56	27,6	4,64
	MB^+	M :	0,00%	**2,74%**	**5,74%**	**1154254,29**	**6,14**
		Sd:	0,00%	1,99%	2,82%	560388,95	3,85
		T-:	-	-2,6	-19,04	-66,5	-20,49
SRBCT	- (GA only)	M :	0,00%	13,30%	43,86%	7344262,55	41,45
		Sd:	0,00%	7,72%	5,68%	3084313,63	6,71
		T-:	-	-	-	-	-
	AF^+	M :	2,16%	19,11%	48,78%	**3167339,45**	**33,36**
		Sd:	3,27%	6,31%	7,12%	1235739,39	7,74
		T-:	9,17	7,79	8,6	-15,51	-9,39
	MB^+	M :	1,19%	23,17%	**40,30%**	**706259,75**	**11,63**
		Sd:	3,37%	7,20%	10,50%	409180,35	7,63
		T-:	8	14,07	-5,32	-29,38	-24,99
9_Tumors	- (GA only)	M :	20,00%	77,17%	94,39%	5829173,33	41,17
		Sd:	4,02%	9,07%	2,10%	2471650,62	7,03
		T-:	-	-	-	-	-
	AF^+	M :	34,07%	79,56%	**93,70%**	**4218204,33**	**27,67**
		Sd:	4,01%	9,19%	1,86%	2932298,71	10,93
		T-:	16,34	6,49	-2,55	-3,56	-6,31
	MB^+	M :	29,33%	78,80%	94,20%	5856291,9	**34,5**
		Sd:	5,62%	7,52%	1,89%	3415222,97	8,71
		T-:	15,15	1,96	-0,64	0,06	-7,55
11_Tumors	- (GA only)	M :	5,99%	28,36%	74,04%	21449781,82	47,18
		Sd:	1,53%	4,60%	4,19%	9370248,49	2,04
		T-:	-	-	-	-	-
	AF^+	M :	14,77%	33,66%	78,05%	**16282683,5**	**40**
		Sd:	5,28%	4,33%	3,09%	6927139,97	7,35
		T-:	15,06	14,49	9,86	-8,42	-22,54
	MB^+	M :	13,38%	32,26%	75,91%	22693365,67	**35**
		Sd:	1,78%	4,07%	3,05%	27640084,2	15,34
		T-:	23,86	5,56	7,98	0,59	-3,91

Table C.7: GA *vs* MAs endowed with AF^+ and MB^+ (2)

Data	Local search	Measure	Fitness	**Validation**	**Validation**	CPU (s)	# Attrib.
14_Tumors	- (GA only)	M :	38,59%	61,50%	87,55%	79843700,29	47,14
		Sd:	4,42%	3,47%	1,30%	12116726,59	2,48
		T-:	-	-	-	-	-
	AF^+	M :	50,81%	68,28%	88,64%	**45985386,25**	**34**
		Sd:	3,77%	5,55%	2,06%	13892955,99	10,01
		T-:	80,61	10,56	6,39	-7,2	-11,99
	MB^+	M :	48,24%	66,35%	86,94%	**54729945**	**36,86**
		Sd:	2,85%	1,86%	1,64%	28889404,02	12,32
		T-:	27,67	7,54	-1,76	-8,33	-4,74
Brain Tumor2	- (GA only)	M :	0,00%	48,80%	59,60%	4025836	42,67
		Sd:	0,00%	11,36%	10,86%	860787,07	5,16
		T-:	-	-	-	-	-
	AF^+	M :	4,27%	**40,80%**	57,69%	**2439530,78**	**21,56**
		Sd:	4,05%	9,70%	5,98%	1313326,49	7,67
		T-:	9	-9,98	-1,97	-5,84	-19,51
	MB^+	M :	5,77%	**42,20%**	59,40%	**1486722,38**	**16,13**
		Sd:	6,82%	6,19%	6,16%	1741569,24	15,42
		T-:	8	-7,22	-0,85	-11,71	-22,62
Prostate Tumor	- (GA only)	M :	0,00%	29,02%	34,84%	6442101,5	33,83
		Sd:	0,00%	6,65%	9,66%	2353803,75	7,19
		T-:	-	-	-	-	-
	AF^+	M :	0,85%	33,20%	39,78%	5738869,67	**23,89**
		Sd:	1,70%	5,38%	5,62%	4748139,33	10,15
		T-:	9	3,46	2,78	-1,72	-10,7
	MB^+	M :	0,43%	**20,70%**	**22,83%**	**853289,78**	**9,56**
		Sd:	1,28%	2,62%	7,66%	897978,41	7,13
		T-:	9	-12,13	-7,61	-71,05	-33,2
Lymphoma	- (GA only)	M :	0,00%	7,49%	17,96%	2955623,82	16,18
		Sd:	0,00%	6,24%	4,23%	1490486,33	7,65
		T-:	-	-	-	-	-
	AF^+	M :	0,00%	10,68%	19,77%	**948803,88**	16,25
		Sd:	0,00%	5,25%	4,50%	728389,16	8,38
		T-:	-	1,92	6,73	-11,82	0,16
	MB^+	M :	0,00%	16,52%	20,15%	**245424,25**	**2,88**
		Sd:	0,00%	2,76%	2,55%	139098,44	1,46
		T-:	-	6,12	6,3	-15,48	-51,61
Challenge 2004	- (GA only)	M :	60,87%	96,11%	98,56%	33369485,25	41
		Sd:	3,55%	2,28%	0,99%	6075771,06	7,44
		T-:	-	-	-	-	-
	AF^+	M :	66,67%	98,67%	99,11%	**19615130,67**	**23,67**
		Sd:	2,51%	1,18%	0,77%	13882130,36	10,6
		T-:	12	2,98	3,27	-7,04	-4,44
	MB^+	M :	68,12%	99,85%	99,26%	**12717393,33**	**3,67**
		Sd:	6,64%	0,26%	0,68%	4141638,3	0,58
		T-:	15	4,64	7,26	-22,6	-36,44

Table C.8: GA *vs* MA endowed with AF^+ and MB^+ (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	MBEGA	M :	2,22%	45,05%	47,45%	1306510,89	15,33
		Sd:	2,11%	3,43%	6,95%	657546,55	12,52
		T-:	-	-	-	-	-
	AF^+	M :	11,11%	49,49%	46,48%	**1054957**	**11,56**
		Sd:	5,21%	5,08%	5,71%	1002712,06	6,86
		T-:	8,83	4,37	-0,58	-2,6	-2,22
	MB^+	M :	3,64%	48,22%	48,26%	**830939,18**	**3,82**
		Sd:	2,16%	6,63%	4,24%	264272,02	0,87
		T-:	7,06	9,45	0,5	-7,47	-7,76
CNS	MBEGA	M :	2,22%	41,19%	40,44%	776660,44	16,44
		Sd:	3,33%	6,85%	5,35%	579215,05	16,86
		T-:	-	-	-	-	-
	AF^+	M :	4,24%	48,06%	41,58%	884081,91	15,18
		Sd:	6,16%	5,12%	6,13%	623905,41	7,11
		T-:	6,1	13,34	1,11	1,5	-0,83
	MB^+	M :	**0,83%**	44,25%	41,67%	**172647,13**	**3,75**
		Sd:	2,36%	7,00%	6,52%	43401,1	1,91
		T-:	-5,18	5,33	3,86	-9,63	-8,48
Colon	MBEGA	M :	0,78%	23,47%	28,79%	769904,38	10,13
		Sd:	2,21%	3,10%	8,53%	631281	13,17
		T-:	-	-	-	-	-
	AF^+	M :	6,73%	24,32%	34,09%	1443520,77	26,08
		Sd:	4,75%	5,06%	7,36%	927548,22	12,31
		T-:	13,09	1,98	4,86	6,68	13,01
	MB^+	M :	3,75%	22,71%	30,77%	**145156,1**	**3,2**
		Sd:	6,04%	9,33%	10,51%	56091,5	1,14
		T-:	7,66	-0,85	2,5	-8,4	-6,77
Leukemia3C	MBEGA	M :	0,00%	15,56%	15,76%	2073395,38	4,13
		Sd:	0,00%	8,59%	2,74%	445394,47	0,99
		T-:	-	-	-	-	-
	AF^+	M :	2,02%	37,98%	42,83%	3607465	24
		Sd:	2,80%	8,91%	10,74%	2938350,12	11,45
		T-:	11	16,9	40,44	15,79	31,22
	MB^+	M :	1,85%	15,74%	24,72%	**272075,33**	**4**
		Sd:	2,87%	14,75%	20,34%	89346,79	0,63
		T-:	6	0,06	2,42	-76,56	-8

Table C.9: MBEGA *vs* MAs respectively endowed with $AF^+$and MB^+ (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	MBEGA	M :	0,00%	14,26%	23,43%	6746779,67	28,33
		Sd:	0,00%	4,54%	2,32%	1982244,45	15,4
		T-:	-	-	-	-	-
	AF^+	M :	1,23%	14,75%	27,00%	20555165,5	30,88
		Sd:	1,46%	4,99%	2,70%	14096615,09	8,43
		T-:	13,33	1,5	40,37	7,45	1,19
	MB^+	M :	1,96%	19,73%	26,31%	6326244,5	**16,75**
		Sd:	1,05%	3,40%	3,72%	2900852,41	2,71
		T-:		22,27	5,25	-0,78	-10,6
MLL	MBEGA	M :	0,00%	20,71%	19,52%	1924337,86	3,71
		Sd:	0,00%	7,93%	6,80%	670178,67	0,49
		T-:	-	-	-	-	-
	AF^+	M :	0,00%	30,00%	38,40%	4015421,88	22,38
		Sd:	0,00%	7,81%	10,66%	2494680,08	7,13
		T-:	-	4,04	8,74	7,12	20,22
	MB^+	M :	0,00%	23,96%	32,99%	**924702,88**	9,13
		Sd:	0,00%	11,31%	14,12%	1046838,27	6,24
		T-:	-	1,78	4,39	-3,24	3,36
Orarian	MBEGA	M :	0,00%	2,22%	4,69%	4376831,29	2
		Sd:	0,00%	0,56%	0,43%	1098440,94	0
		T-:	-	-	-	-	-
	AF^+	M :	0,00%	2,59%	16,13%	29587089,5	31,5
		Sd:	0,00%	2,10%	11,00%	18545416,25	8,87
		T-:	-	2,61	10,34	45,83	39,33
	MB^+	M :	0,00%	2,74%	5,74%	**1154254,29**	6,14
		Sd:	0,00%	1,99%	2,82%	560388,95	3,85
		T-:	-	3,72	10,5	-43,32	5,97
SRBCT	MBEGA	M :	0,00%	11,29%	32,54%	2253283,43	5,71
		Sd:	0,00%	4,43%	5,64%	502583,62	2,36
		T-:	-	-	-	-	-
	AF^+	M :	2,16%	19,11%	48,78%	3167339,45	33,36
		Sd:	3,27%	6,31%	7,12%	1235739,39	7,74
		T-:	9,17	9,46	25,09	5,42	109,88
	MB^+	M :	1,19%	23,17%	40,30%	**706259,75**	11,63
		Sd:	3,37%	7,20%	10,50%	409180,35	7,63
		T-:	8	15,09	10,6	-18,48	6,84
9_Tumors	MBEGA	M :	21,11%	75,22%	95,22%	2831195,5	38,67
		Sd:	2,72%	7,99%	1,66%	514528,93	17,11
		T-:	-	-	-	-	-
	AF^+	M :	34,07%	79,56%	**93,70%**	4218204,33	**27,67**
		Sd:	4,01%	9,19%	1,86%	2932298,71	10,93
		T-:	18,95	5,05	-5,57	3,78	-4,39
	MB^+	M :	29,33%	78,80%	**94,20%**	5856291,9	**34,5**
		Sd:	5,62%	7,52%	1,89%	3415222,97	8,71
		T-:	25,33	3,14	-3,41	7,44	-2,64

Table C.10: MBEGA *vs* MAs respectively endowed with $AF^+$and MB^+(2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
11_Tumors	MBEGA	M :	3,64%	23,22%	74,25%	10439972,6	49
		Sd:	2,03%	3,59%	0,92%	2127251,91	2,83
		T-:	-	-	-	-	-
	AF^+	M :	14,77%	33,66%	78,05%	16282683,5	**40**
		Sd:	5,28%	4,33%	3,09%	6927139,97	7,35
		T-:	18,67	11,57	9,93	9,63	-24,96
	MB^+	M :	13,38%	32,26%	75,91%	22693365,67	**35**
		Sd:	1,78%	4,07%	3,05%	27640084,2	15,34
		T-:	29,15	8,35	8,73	5,86	-4,49
14_Tumors	MBEGA	M :	38,40%	61,93%	85,84%	32426149,29	47,43
		Sd:	3,74%	3,48%	1,48%	3598869,19	4,69
		T-:	-	-	-	-	-
	AF^+	M :	51,08%	68,48%	88,16%	40369678,17	**31,67**
		Sd:	4,24%	6,54%	2,03%	6542676	10,67
		T-:	17,31	226,19	5,13	9,49	-16,52
	MB^+	M :	48,24%	66,35%	86,94%	54729945	**36,86**
		Sd:	2,85%	1,86%	1,64%	28889404,02	12,32
		T-:	17,34	127,78	3,58	15,1	-4,68
Brain Tumor2	MBEGA	M :	0,00%	51,60%	59,73%	966079,67	10,67
		Sd:	0,00%	15,88%	5,25%	424005,49	7
		T-:	-	-	-	-	-
	AF^+	M :	4,27%	**40,80%**	57,69%	2439530,78	21,56
		Sd:	4,05%	9,70%	5,98%	1313326,49	7,67
		T-:	9	-6,77	-1,36	7,63	11,46
	MB^+	M :	5,77%	**42,20%**	59,40%	1486722,38	16,13
		Sd:	6,82%	6,19%	6,16%	1741569,24	15,42
		T-:	8	-5,68	-0,28	5,09	5,19
Prostate Tumor	MBEGA	M :	0,00%	26,39%	40,11%	2119086,86	27
		Sd:	0,00%	6,07%	7,89%	420355,88	14,73
		T-:	-	-	-	-	-
	AF^+	M :	1,28%	32,22%	41,24%	6103394,17	24,17
		Sd:	1,99%	4,20%	2,01%	5524555,19	12,02
		T-:	6	14,78	1,26	4,25	-0,79
	MB^+	M :	0,55%	**21,06%**	**23,81%**	**921320,29**	**9,14**
		Sd:	1,45%	2,89%	8,55%	1017983,82	7,4
		T-:	7	-11,93	-11,35	-13,95	-8,25

Table C.11: MBEGA *vs* MAs respectively endowed with $AF^+$and MB^+ (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	- (only GA)	M :	2,13%	45,19%	50,14%	5086055,47	38,33
		Sd:	2,56%	7,28%	4,87%	1719792,54	6,65
		T-:	-	-	-	-	-
	FLS^+	M :	4,22%	**42,18%**	**47,55%**	**1259323**	**14,44**
		Sd:	3,49%	7,99%	5,94%	577700,54	8,52
		T-:	4,58	-4,01	-19,22	-46,71	-34,5
	$IFLS^+$	M :	4,00%	**43,44%**	**49,22%**	**2046134,93**	**12,27**
		Sd:	3,70%	9,37%	6,18%	1164253,83	6,49
		T-:	3,94	-2,47	-2,02	-31,74	-27,07
CNS	- (only GA)	M :	0,00%	39,23%	43,95%	2951100,54	34,69
		Sd:	0,00%	10,70%	5,11%	1517342,83	11,47
		T-:	-	-	-	-	-
	FLS^+	M :	1,25%	43,79%	44,25%	**415207,13**	**10,69**
		Sd:	2,69%	10,41%	6,77%	199190,95	4,48
		T-:	16	8,68	1,57	-105,48	-234,26
	$IFLS^+$	M :	1,25%	45,67%	**42,08%**	**675038,06**	**12,06**
		Sd:	2,69%	10,65%	5,80%	564074,31	6,38
		T-:	16	28,85	-9,42	-58,63	-68,16
Colon	- (only GA)	M :	1,92%	27,79%	32,66%	2154451,46	28,15
		Sd:	3,00%	4,83%	5,79%	1209434,43	10,54
		T-:	-	-	-	-	-
	FLS^+	M :	7,42%	27,98%	33,79%	**381501,94**	**11,31**
		Sd:	5,21%	6,65%	7,16%	196739,94	6,41
		T-:	16,13	0,44	2,27	-12,96	-33,33
	$IFLS^+$	M :	4,91%	**23,50%**	30,37%	**385083,36**	**11,5**
		Sd:	4,37%	4,41%	8,36%	183139,97	4,83
		T-:	6,18	-12,45	-1,91	-12,9	-36,92
Leukemia3C	- (only GA)	M :	0,00%	19,96%	27,78%	4745137,5	38,64
		Sd:	0,00%	7,04%	6,89%	1778242,67	7,62
		T-:	-	-	-	-	-
	FLS^+	M :	0,98%	25,65%	32,81%	**1352039,24**	**23,12**
		Sd:	2,18%	13,91%	12,68%	789451,7	10,73
		T-:	3,64	6,7	7,04	-20,07	-22,48
	$IFLS^+$	M :	0,40%	26,39%	29,37%	**1954565,43**	**18,93**
		Sd:	1,48%	12,71%	16,59%	989121,03	7,02
		T-:	14	3,41	0,71	-14	-19,94
Leukemia4C	- (only GA)	M :	0,00%	33,85%	39,44%	4781085,5	40,5
		Sd:	0,00%	9,82%	8,04%	1428019,32	4,85
		T-:	-	-	-	-	-
	FLS^+	M :	0,62%	36,11%	**36,17%**	**1352848,39**	**22,94**
		Sd:	1,80%	9,90%	9,11%	464380,8	6,17
		T-:	18	1,87	-2,91	-97,65	-70,37
	$IFLS^+$	M :	0,93%	38,29%	38,80%	**2959564,42**	**25**
		Sd:	3,21%	7,22%	8,99%	1403044,17	4,75
		T-:	12	3,79	-0,45	-22,93	-114,19

Table C.12: GA *vs* MAs respectively endowed with $FLS^+$and $IFLS^+$ (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	- (only GA)	M :	0,00%	15,03%	26,98%	18327174,77	46,62
		Sd:	0,00%	3,25%	2,13%	7173803,05	2,75
		T-:	-	-	-	-	-
	FLS^+	M :	2,17%	15,58%	**25,74%**	**9139586,16**	**38,16**
		Sd:	1,45%	4,89%	2,69%	3245476,14	7,83
		T-:	199,5	3,28	-6,1	-76,76	-25,03
	$IFLS^+$	M :	2,38%	16,31%	27,58%	**12407942,43**	**34,86**
		Sd:	2,06%	3,62%	1,89%	5616958,13	10,17
		T-:	79,33	11,31	2,38	-12,64	-29,32
MLL	- (only GA)	M :	0,00%	25,69%	34,40%	4737424,25	34
		Sd:	0,00%	8,64%	6,22%	2256399,3	10,73
		T-:	-	-	-	-	-
	FLS^+	M :	0,28%	**24,36%**	**28,72%**	**1890214,2**	**26,95**
		Sd:	1,24%	8,42%	8,58%	923936,3	8,57
		T-:	20	-2,81	-7,27	-11,01	-8,79
	$IFLS^+$	M :	0,00%	25,14%	**28,16%**	**2411450,38**	**21,31**
		Sd:	0,00%	7,52%	9,55%	1973694,51	8,17
		T-:	-	-1,45	-17,37	-8,93	-21,58
Orarian	- (only GA)	M :	0,00%	3,08%	19,50%	13669924,58	26,67
		Sd:	0,00%	2,40%	10,12%	6257542,88	7,99
		T-:	-	-	-	-	-
	FLS^+	M :	0,00%	4,07%	25,14%	**4266492,61**	**21,72**
		Sd:	0,00%	2,72%	8,42%	2391928,59	9
		T-:		4,81	5,43	-50,75	-5,57
	$IFLS^+$	M :	0,00%	3,10%	**16,90%**	**4922699**	24,56
		Sd:	0,00%	1,92%	11,13%	2185573,11	8,68
		T-:		0,11	-2,83	-29,81	-1,98
SRBCT	- (only GA)	M :	0,00%	13,30%	43,86%	7344262,55	41,45
		Sd:	0,00%	7,72%	5,68%	3084313,63	6,71
		T-:	-	-	-	-	-
	FLS^+	M :	0,00%	21,52%	45,88%	**2009761**	**24,59**
		Sd:	0,00%	12,10%	9,91%	962201,27	7,17
		T-:		7,18	2,37	-23,79	-18,72
	$IFLS^+$	M :	0,00%	17,01%	44,30%	**2656986**	**26,5**
		Sd:	0,00%	7,27%	8,16%	1091701,62	6,98
		T-:		13,69	0,54	-21,18	-14,8

Table C.13: GA *vs* MAs respectively endowed with $FLS^+$and $IFLS^+$ (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
9_Tumors	- (only GA)	M :	20,00%	77,17%	94,39%	5829173,33	41,17
		Sd:	4,02%	9,07%	2,10%	2471650,62	7,03
		T-:	-	-	-	-	-
	FLS^+	M :	28,42%	77,51%	**93,89%**	**2266225,84**	**26,21**
		Sd:	7,96%	7,62%	1,86%	614236,75	8,2
		T-:	11,85	0,82	-2,48	-13,03	-12,83
	$IFLS^+$	M :	27,08%	79,42%	**93,33%**	6142094,69	**30**
		Sd:	4,53%	6,48%	1,52%	3554311,54	9,72
		T-:	9,97	12,06	-4,07	0,89	-11,71
11_Tumors	- (only GA)	M :	5,99%	28,36%	74,04%	21449781,82	47,18
		Sd:	1,53%	4,60%	4,19%	9370248,49	2,04
		T-:	-	-	-	-	-
	FLS^+	M :	14,20%	30,06%	74,36%	**13445854,35**	**45,4**
		Sd:	4,29%	4,87%	3,98%	2722913,13	5,66
		T-:	24,22	7,68	1,41	-26,2	-2,83
	$IFLS^+$	M :	11,06%	28,90%	74,04%	30548785,6	**42,47**
		Sd:	4,85%	5,58%	3,14%	20144408,67	7,52
		T-:	16,25	4,14	-0,03	5,58	-6,11
14_Tumors	- (Only GA)	M :	38,59%	61,50%	87,55%	79843700,29	47,14
		Sd:	4,42%	3,47%	1,30%	12116726,59	2,48
		T-:	-	-	-	-	-
	FLS^+	M :	48,33%	63,01%	87,27%	**40237743**	**43,74**
		Sd:	5,29%	2,63%	1,47%	8446464,22	7,69
		T-:	128,7	2,32	-1,66	-14,47	-7,56
	$IFLS^+$	M :	47,40%	64,85%	87,80%	**68516459,81**	**43**
		Sd:	3,88%	3,54%	0,86%	24362428,87	7
		T-:	42,02	5,04	1,43	-3,63	-8,06
Brain Tumor2	- (only GA)	M :	0,00%	48,80%	59,60%	4025836	42,67
		Sd:	0,00%	11,36%	10,86%	860787,07	5,16
		T-:	-	-	-	-	-
	FLS^+	M :	4,86%	50,11%	59,71%	**1464280,11**	**25,84**
		Sd:	4,59%	7,52%	7,60%	748811,33	9,47
		T-:	32,57	1,92	0,63	-13	-16,64
	$IFLS^+$	M :	4,73%	**45,05%**	**58,95%**	**1648181,08**	**17,46**
		Sd:	3,90%	11,57%	8,31%	1220942,38	6,49
		T-:	20,8	-2,52	-3,62	-12,1	-28,11
Prostate Tumor	- (only GA)	M :	0,00%	29,02%	34,84%	6442101,5	33,83
		Sd:	0,00%	6,65%	9,66%	2353803,75	7,19
		T-:	-	-	-	-	-
	21 FLS^+	M :	0,43%	**24,12%**	31,66%	**1364486**	**18,78**
		Sd:	1,24%	7,30%	10,34%	712767,66	6,94
		T-:	18	-6,75	-1,99	-83,06	-25
	23	M :	1,10%	**25,46%**	33,19%	**2753627,43**	**19,07**
		Sd:	1,80%	8,07%	11,26%	3229070,64	7,62
		T-:	5,6	-4,41	-0,92	-37,56	-27,75

Table C.14: GA *vs* MAs respectively endowed with $FLS^+$and $IFLS^+$ (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU	# Attrib
Lymphoma	GA	M :	0,00%	7,49%	17,96%	2955623,82	16,18
		Sd:	0,00%	6,24%	4,23%	1490486,33	7,65
		T-:	-	-	-	-	-
	21	M :	0,00%	12,08%	22,50%	**682270,63**	**14,44**
		Sd:	0,00%	5,36%	4,77%	319649,76	4,72
		T-:		2,92	6,69	-13,39	-3,6
	23	M :	0,00%	12,41%	21,89%	**774349,94**	**13,47**
		Sd:	0,00%	7,44%	6,26%	291551,59	4,56
		T-:		3,32	11,16	-12,84	-24,24
Challenge 2004	GA	M :	60,87%	96,11%	98,56%	33369485,25	41
		Sd:	3,55%	2,28%	0,99%	6075771,06	7,44
		T-:	-	-	-	-	-
	21	M :	66,59%	97,01%	98,81%	**12560270,26**	**23,16**
		Sd:	6,16%	2,44%	0,57%	3585413,42	8,7
		T-:	14,84	1,09	2,73	-53,48	-15,73
	23	M :	62,50%	94,39%	98,89%	65205452,63	**26,38**
		Sd:	5,66%	2,70%	0,67%	39601039,45	9,78
		T-:	4,8	-2,05	3,79	4,99	-11,26

Table C.15: MBEGA *vs* MAs respectively endowed with $FLS^+$and $IFLS^+$

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	MBEGA	M :	2,22%	45,05%	47,45%	1306510,89	15,33
		Sd:	2,11%	3,43%	6,95%	657546,55	12,52
		T-:	-	-	-	-	-
	FLS^+	M :	4,22%	**42,18%**	47,55%	1259323	14,44
		Sd:	3,49%	7,99%	5,94%	577700,54	8,52
		T-:	6,52	-6,84	0,06	-0,74	-0,59
	$IFLS^+$	M :	4,00%	**43,44%**	49,22%	2046134,93	12,27
		Sd:	3,70%	9,37%	6,18%	1164253,83	6,49
		T-:	5,36	-4,84	1,11	9,17	-1,86
CNS	MBEGA	M :	2,22%	41,19%	40,44%	776660,44	16,44
		Sd:	3,33%	6,85%	5,35%	579215,05	16,86
		T-:	-	-	-	-	-
	FLS^+	M :	**1,25%**	43,79%	44,25%	**415207,13**	**10,69**
		Sd:	2,69%	10,41%	6,77%	199190,95	4,48
		T-:	-3,75	3,73	23,07	-5,79	-3,85
	$IFLS^+$	M :	**1,25%**	45,67%	42,08%	675038,06	**12,06**
		Sd:	2,69%	10,65%	5,80%	564074,31	6,38
		T-:	-3,75	8,74	9,53	-1,46	-2,87
Colon	MBEGA	M :	0,78%	23,47%	28,79%	769904,38	10,13
		Sd:	2,21%	3,10%	8,53%	631281	13,17
		T-:	-	-	-	-	-
	FLS^+	M :	7,42%	27,98%	33,79%	**381501,94**	11,31
		Sd:	5,21%	6,65%	7,16%	196739,94	6,41
		T-:	54,4	11,84	5,88	-5,24	1,14
	$IFLS^+$	M :	4,91%	23,50%	30,37%	**385083,36**	11,5
		Sd:	4,37%	4,41%	8,36%	183139,97	4,83
		T-:	11,34	0,13	1,14	-5,15	1,35
Leukemia3C	MBEGA	M :	0,00%	15,56%	15,76%	2073395,38	4,13
		Sd:	0,00%	8,59%	2,74%	445394,47	0,99
		T-:		-	-	-	-
	FLS^+	M :	0,98%	25,65%	32,81%	**1352039,24**	23,12
		Sd:	2,18%	13,91%	12,68%	789451,7	10,73
		T-:	3,64	8,58	28,41	-31,49	350,39
	$IFLS^+$	M :	0,40%	26,39%	29,37%	1954565,43	18,93
		Sd:	1,48%	12,71%	16,59%	989121,03	7,02
		T-:	14	5,28	6,2	-1,1	20,87
Leukemia4C	MBEGA	M :	0,00%	28,33%	33,89%	2055455	15,17
		Sd:	0,00%	12,12%	14,65%	748936,92	12,16
		T-:		-	-	-	0
	FLS^+	M :	0,62%	36,11%	36,17%	**1352848,39**	22,94
		Sd:	1,80%	9,90%	9,11%	464380,8	6,17
		T-:	18	3,6	0,98	-17,38	4,15
	$IFLS^+$	M :	0,93%	38,29%	38,80%	2959564,42	25
		Sd:	3,21%	7,22%	8,99%	1403044,17	4,75
		T-:	12	4,66	1,97	11,04	5,28

Table C.16: MBEGA *vs* $FLS^+$and $IFLS^+$ (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	MBEGA	M :	0,00%	14,26%	23,43%	6746779,67	28,33
		Sd:	0,00%	4,54%	2,32%	1982244,45	15,4
		T-:	-	-	-	-	-
	FLS^+	M :	2,17%	15,58%	25,74%	9139586,16	38,16
		Sd:	1,45%	4,89%	2,69%	3245476,14	7,83
		T-:	199,5	4,68	17,96	8,94	9,01
	$IFLS^+$	M :	2,38%	16,31%	27,58%	12407942,43	34,86
		Sd:	2,06%	3,62%	1,89%	5616958,13	10,17
		T-:	79,33	8,04	20,72	10,76	5,87
MLL	MBEGA	M :	0,00%	20,71%	19,52%	1924337,86	3,71
		Sd:	0,00%	7,93%	6,80%	670178,67	0,49
		T-:	-	-	-	-	-
	FLS^+	M :	0,28%	24,36%	28,72%	1890214,2	26,95
		Sd:	1,24%	8,42%	8,58%	923936,3	8,57
		T-:	20	5,67	8,46	-0,45	29,1
	$IFLS^+$	M :	0,00%	25,14%	28,16%	2411450,38	21,31
		Sd:	0,00%	7,52%	9,55%	1973694,51	8,17
		T-:		7,64	10,31	6,04	30,16
ovarian	MBEGA	M :	0,00%	2,22%	4,69%	4376831,29	2
		Sd:	0,00%	0,56%	0,43%	1098440,94	0
		T-:		-	-	-	
	FLS^+	M :	0,00%	4,07%	25,14%	4266492,61	21,72
		Sd:	0,00%	2,72%	8,42%	2391928,59	9
		T-:		8,73	27,13	-1,65	38,26
	$IFLS^+$	M :	0,00%	3,10%	16,90%	4922699	24,56
		Sd:	0,00%	1,92%	11,13%	2185573,11	8,68
		T-:		6,67	21,37	2,3	29,03
SRBCT	MBEGA	M :	0,00%	11,29%	32,54%	2253283,43	5,71
		Sd:	0,00%	4,43%	5,64%	502583,62	2,36
		T-:	-	-	-	-	-
	FLS^+	M :	0,00%	21,52%	45,88%	**2009761**	24,59
		Sd:	0,00%	12,10%	9,91%	962201,27	7,17
		T-:		8,53	14,74	-3,08	51,79
	$IFLS^+$	M :	0,00%	17,01%	44,30%	2656986	26,5
		Sd:	0,00%	7,27%	8,16%	1091701,62	6,98
		T-:		12,74	13,59	5,75	35,53
9_Tumors	MBEGA	M :	21,11%	75,22%	95,22%	2831195,5	38,67
		Sd:	2,72%	7,99%	1,66%	514528,93	17,11
		T-:	-	-	-	-	-
	FLS^+	M :	28,42%	77,51%	**93,89%**	**2266225,84**	**26,21**
		Sd:	7,96%	7,62%	1,86%	614236,75	8,2
		T-:	15,22	2,6	-6,51	-8,88	-7,12
	$IFLS^+$	M :	27,08%	79,42%	**93,33%**	6142094,69	**30**
		Sd:	4,53%	6,48%	1,52%	3554311,54	9,72
		T-:	12,45	5,26	-7,18	14,3	-5,36

Table C.17: MBEGA *vs* $FLS^+$and $IFLS^+$ (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib
11_Tumors	MBEGA	M :	3,64%	23,22%	74,25%	10439972,6	49
		Sd:	2,03%	3,59%	0,92%	2127251,91	2,83
		T-:	-	-	-	-	-
	FLS^+	M :	14,20%	30,06%	74,36%	13445854,35	45,4
		Sd:	4,29%	4,87%	3,98%	2722913,13	5,66
		T-:	29,23	8,01	0,6	10,3	-5,53
	$IFLS^+$	M :	11,06%	28,90%	74,04%	30548785,6	**42,47**
		Sd:	4,85%	5,58%	3,14%	20144408,67	7,52
		T-:	22,08	6,81	-1,49	12,35	-8,27
14_Tumors	MBEGA	M :	38,40%	61,93%	85,84%	32426149,29	47,43
		Sd:	3,74%	3,48%	1,48%	3598869,19	4,69
		T-:	-	-	-	-	-
	FLS^+	M :	48,33%	63,01%	87,27%	40237743	**43,74**
		Sd:	5,29%	2,63%	1,47%	8446464,22	7,69
		T-:	21,86	12,28	42,64	10,19	-4,75
	$IFLS^+$	M :	47,40%	64,85%	87,80%	68516459,81	**43**
		Sd:	3,88%	3,54%	0,86%	24362428,87	7
		T-:	18,21	17,37	44,29	21,33	-5,44
Brain Tumor2	MBEGA	M :	0,00%	51,60%	59,73%	966079,67	10,67
		Sd:	0,00%	15,88%	5,25%	424005,49	7
		T-:		-	-	-	-
	FLS^+	M :	4,86%	50,11%	59,71%	1464280,11	25,84
		Sd:	4,59%	7,52%	7,60%	748811,33	9,47
		T-:	32,57	-0,97	-0,02	10,56	17,48
	$IFLS^+$	M :	4,73%	**45,05%**	58,95%	1648181,08	17,46
		Sd:	3,90%	11,57%	8,31%	1220942,38	6,49
		T-:	20,8	-3,22	-0,67	15,17	9,28
Prostate Tumor	MBEGA	M :	0,00%	26,39%	40,11%	2119086,86	27
		Sd:	0,00%	6,07%	7,89%	420355,88	14,73
		T-:		-	-	-	-
	FLS^+	M :	0,43%	**24,12%**	**31,66%**	**1364486**	**18,78**
		Sd:	1,24%	7,30%	10,34%	712767,66	6,94
		T-:	18	-4,62	-8,93	-19,42	-3,8
	$IFLS^+$	M :	1,10%	25,46%	**33,19%**	2753627,43	**19,07**
		Sd:	1,80%	8,07%	11,26%	3229070,64	7,62
		T-:	5,6	-1,53	-5,6	7,37	-3,7

Table C.18: MBEGA *vs* endowed with $FLS^+$and $IFLS^+$ (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	- (only GA)	M :	2,13%	45,19%	50,14%	5086055,46	38,33
		Sd:	2,56%	7,28%	4,87%	1719792,53	6,65
		T-:	-	-	-	-	-
	MB^*	M :	2,44%	**41,37%**	**44,44%**	**1195308,66**	**5,33**
		Sd:	2,01%	6,18%	6,54%	500152,55	3,51
		T-:	0,77	-4,71	-67,49	-43,87	-50,20
	$MB^*(FI)$	M :	2,74%	**39,06%**	**42,54%**	**962519,42**	**4,31**
		Sd:	1,91%	4,97%	6,95%	374736,11	2,96
		T-:	1,44	-6,18	-21,15	-49,41	-52,76
CNS	- (only GA)	M :	0,00%	39,23%	43,95%	2951100,53	34,69
		Sd:	0,00%	10,70%	5,11%	1517342,83	11,47
		T-:	-	-	-	-	-
	MB^*	M :	0,70%	45,05%	**40,04%**	**777555,26**	**5,89**
		Sd:	2,10%	9,08%	6,19%	309502,8968	4,8292
		T-:	19	12,89	-14,79	-80,89	-192,42
	$MB^*(FI)$	M :	1,48%	40,07%	**39,59%**	**723923,83**	**4,44**
		Sd:	2,85%	7,08%	6,66%	285196,60	3,51
		T-:	18	1,63	-24,07	-67,07	-228,08
Colon	- (only GA)	M :	1,92%	27,79%	32,66%	2154451,46	28,15
		Sd:	3,00%	4,83%	5,79%	1209434,42	10,54
		T-:	-	-	-	-	-
	MB^*	M :	**1,04%**	**23,58%**	**29,50%**	**478149,88**	**4,11**
		Sd:	2,40%	6,26%	3,90%	236159,93	2,54
		T-:	-2,6	-7,43	-9,18	-12,23	-51,73
	$MB^*(FI)$	M :	2,08%	**24,62%**	**29,71%**	**454105,77**	**3,66**
		Sd:	3,03%	9,26%	9,16%	304342,90	2,24
		T-:	0,45	-6,0057	-2,79	-12,34	-53,33
Leukemia3C	- (only GA)	M :	0,00%	19,96%	27,78%	4745137,5	38,64
		Sd:	0,00%	7,04%	6,89%	1778242,67	7,62
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**13,09%**	**14,54%**	**1330103,5**	**5**
		Sd:	0,00%	4,24%	2,54%	358034,03	2,44
		T-:		-8,85	-16,97	-20,13	-46,48
	$MB^*(FI)$	M :	0,00%	**12,93%**	**15,08%**	**1449598,90**	**3,95**
		Sd:	0,00%	4,99%	2,58%	328167,76	1,13
		T-:		-9,25	-17,78	-19,50	-50,24
Leukemia4C	- (only GA)	M :	0,00%	33,85%	39,44%	4781085,5	40,5
		Sd:	0,00%	9,82%	8,04%	1428019,31	4,84
		T-:	-	-	-	-	-
	MB^*	M :	0,33%	**19,61%**	**22,97%**	**1325878,41**	**6,23**
		Sd:	1,35%	6,89%	6,49%	412608,10	3,47
		T-:	17	-25,56	-19,87	-78,95	-202,01
	$MB^*(FI)$	M :	0,00%	**20,06%**	**23,39%**	**1389635,15**	**4,9**
		Sd:	0,00%	10,05%	8,35%	399281,07	1,74
		T-:		-23,5716	-22,049	-209,8386	-295,5951

Table C.19: GA *vs* MAs endowed with MB^* (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	- (only GA)	M :	0,00%	15,03%	26,98%	18327174,76	46,61
		Sd:	0,00%	3,25%	2,13%	7173803,05	2,75
		T-:	-	-	-	-	-
	MB^*	M :	0,21%	15,58%	**20,71%**	**5445415,73**	**21,52**
		Sd:	0,62%	3,26%	2,45%	1411729,70	10,76
		T-:	19	1,6482	-21,1869	-116,2167	-123,7608
	$MB^*(FI)$	M :	0,28%	15,85%	**21,95%**	**5949881,23**	**23,23**
		Sd:	0,70%	3,19%	2,07%	1507922,75	10,73
		T-:	21	16,47	-27,65	-92,20	-65,15
MLL	GA	M :	0,00%	25,69%	34,40%	4737424,25	34
		Sd:	0,00%	8,64%	6,22%	2256399,29	10,728
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**22,31%**	**19,97%**	**1230299,47**	**2,84**
		Sd:	0,00%	6,10%	4,64%	231286,02	0,68
		T-:		-7,51	-50,24	-14,11	-372,04
	$MB^*(FI)$	M :	0,00%	**18,51%**	**16,72%**	**1250757,04**	**3,2727**
		Sd:	0,00%	4,38%	4,53%	296787,73	0,76
		T-:		-13,05	-36,36	-14,02	-302,89
Orarian	- (only GA)	M :	0,00%	3,08%	19,50%	13669924,58	26,66
		Sd:	0,00%	2,40%	10,12%	6257542,87	7,9924
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**2,04%**	**4,81%**	**3476207,7**	**2**
		Sd:	0,00%	0,61%	1,03%	879987,11	0
		T-:		-13,36	-20,49	-56,97	-34,15
	$MB^*(FI)$	M :	0,00%	**2,41%**	**4,86%**	**3546339,36**	**2**
		Sd:	0,00%	1,16%	0,62%	1016864,05	0
		T-:		-8,69	-20,43	-56,6045	-34,15
SRBCT	- (only GA)	M :	0,00%	13,30%	43,86%	7344262,54	41,45
		Sd:	0,00%	7,72%	5,68%	3084313,63	6,71
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**10,42%**	**35,38%**	**1879785,21**	**4,84**
		Sd:	0,00%	3,29%	4,22%	391358,75	1,34
		T-:		-19,85	-14,80	-24,68	-42,32
	$MB^*(FI)$	M :	0,00%	**12,50%**	**35,19%**	**1819164,85**	**4,71**
		Sd:	0,00%	6,67%	5,89%	406342,94	1,05
		T-:		-6,42	-13,99	-24,78	-42,74
9_Tumors	- (only GA)	M :	20,00%	77,17%	94,39%	5829173,33	41,16
		Sd:	4,02%	9,07%	2,10%	2471650,62	7,0302
		T-:	-	-	-	-	-
	MB^*	M :	19,67%	80,03%	**94,03%**	**1928653,05**	**24,5**
		Sd:	5,50%	5,85%	1,95%	594225,66	14,54
		T-:	-0,52	14,84	-1,69	-14,51	-14,94
	$MB^*(FI)$	M :	**17,78%**	77,50%	94,08%	**2041232,29**	**25,08**
		Sd:	5,08%	6,96%	1,82%	647967,02	13,52
		T-:	-3,79	0,95	-1,53	-14,21	-15,66

Table C.20: GA *vs* MAs endowed with MB^* (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
11_Tumors	- (only GA)	M :	5,99%	28,36%	74,04%	21449781,81	47,18
		Sd:	1,53%	4,60%	4,19%	9370248,4903	2,0405
		T-:	-	-	-	-	-
	MB^*	M :	6,58%	**24,16%**	**70,95%**	**8805682,78**	**41,36**
		Sd:	1,99%	3,60%	2,37%	1782312,8593	5,3042
		T-:	3,4637	-59,0141	-8,8355	-84,7514	-15,9187
	$MB^*(FI)$	M :	6,39%	**25,96%**	**71,93%**	**8369241,76**	**39,90**
		Sd:	2,45%	5,25%	3,50%	1551235,5956	7,674
		T-:	2,4909	-27,1815	-12,8801	-106,6688	-32,6841
14_Tumors	- (only GA)	M :	38,59%	61,50%	87,55%	79843700,2857	47,1429
		Sd:	4,42%	3,47%	1,30%	12116726,5858	2,4785
		T-:	-	-	-	-	-
	MB^*	M :	40,12%	64,28%	**85,56%**	**26873119,55**	**41,77**
		Sd:	3,15%	4,78%	1,53%	5022935,7642	9,8013
		T-:	28,4543	3,7365	-11,6149	-19,3963	-11,5219
	$MB^*(FI)$	M :	39,53%	**60,58%**	**85,71%**	**29345883,12**	**45,12**
		Sd:	2,68%	3,33%	1,23%	2324851,7162	3,6856
		T-:	13,4253	-1,427	-9,6123	-18,4764	-4,1726
Brain Tumor2	- (only GA)	M :	0,00%	48,80%	59,60%	4025836	42,6667
		Sd:	0,00%	11,36%	10,86%	860787,0661	5,164
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	49,89%	**56,34%**	**800287,71**	**7,92**
		Sd:	0,00%	11,64%	3,79%	350967,2153	9,778
		T-:		0,6014	-15,1711	-16,7898	-37,2674
	$MB^*(FI)$	M :	0,00%	**44,49%**	**55,63%**	**970313,07**	**5,15**
		Sd:	0,00%	10,32%	4,66%	340633,0128	3,2621
		T-:		-5,7862	-9,8447	-15,8014	-41,4874
Prostate Tumor	- (only GA)	M :	0,00%	29,02%	34,84%	6442101,5	33,8333
		Sd:	0,00%	6,65%	9,66%	2353803,7475	7,1949
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**15,51%**	**16,96%**	**1341873,06**	**3,56**
		Sd:	0,00%	4,44%	6,82%	381545,5621	1,1529
		T-:		-12,4239	-8,486	-75,0699	-57,2284
	$MB^*(FI)$	M :	0,00%	**14,66%**	**18,31%**	**1465425,93**	**3,37**
		Sd:	0,00%	4,03%	4,27%	434687,551	1,2583
		T-:		-19,0374	-10,9153	-59,2398	-56,9604
Lymphoma	- (only GA)	M :	0,00%	7,49%	17,96%	2955623,8182	16,1818
		Sd:	0,00%	6,24%	4,23%	1490486,3299	7,6527
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	16,94%	19,97%	**878011,05**	**1,05**
		Sd:	0,00%	3,41%	4,34%	171691,0113	0,2294
		T-:		6,4339	7,2121	-12,2312	-140,771
	$MB^*(FI)$	M :	0,00%	17,69%	20,84%	**1041620,52**	**1**
		Sd:	0,00%	0,56%	3,75%	246454,2847	0
		T-:		6,9489	12,7268	-11,2638	-141,3077
Challenge 2004	- (only GA)	M :	60,87%	96,11%	98,56%	33369485,25	41
		Sd:	3,55%	2,28%	0,99%	6075771,0639	7,4386
		T-:	-	-	-	-	-
	MB^*	M :	**58,26%**	**94,58%**	98,84%	**9353813**	**21,1**
		Sd:	3,67%	3,63%	0,37%	1607506,3801	10,461
		T-:	-10	-1,6352	3,3904	-62,8186	-13,453
	$MB^*(FI)$	M :	**59,78%**	**93,93%**	98,74%	**10103488,08**	**25,25**
		Sd:	3,28%	4,21%	0,70%	2371460,8057	12,0614
		T-:	-4	-2,6346	2,2162	-60,4516	-15,7466

Table C.21: GA *vs* MAs endowed with MB^* (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	MBEGA	M :	2,22%	45,05%	47,45%	1306510,8889	15,3333
		Sd:	2,11%	3,43%	6,95%	657546,5512	12,52
		T-:	-	-	-	-	-
	MB^*	M :	2,44%	**41,37%**	44,44%	1195308,66	**5,33**
		Sd:	2,01%	6,18%	6,54%	500152,553	3,5147
		T-:	1,0307	-7,1051	-1,9751	-1,5425	-6,7243
	$MB^*(FI)$	M :	2,74%	**39,06%**	**42,54%**	**962519,42**	**4,31**
		Sd:	1,91%	4,97%	6,95%	374736,1127	2,9637
		T-:	2,105	-7,7791	-3,1412	-5,2464	-7,4364
CNS	MBEGA	M :	2,22%	41,19%	40,44%	776660,4444	16,4444
		Sd:	3,33%	6,85%	5,35%	579215,0463	16,8605
		T-:	-	-	-	-	-
	MB^*	M :	**0,70%**	45,05%	40,04%	777555,26	**5,89**
		Sd:	2,10%	9,08%	6,19%	309502,8968	4,8292
		T-:	-6,0901	5,9881	-1,6664	0,0141	-7,0429
	$MB^*(FI)$	M :	**1,48%**	**40,07%**	**39,59%**	723923,83	**4,44**
		Sd:	2,85%	7,08%	6,66%	285196,6096	3,5184
		T-:	-2,846	-1,6039	-5,6084	-0,7934	-8,0196
Colon	MBEGA	M :	0,78%	23,47%	28,79%	769904,375	10,125
		Sd:	2,21%	3,10%	8,53%	631280,9958	13,174
		T-:	-	-	-	-	-
	MB^*	M :	1,04%	23,58%	29,50%	**478149,88**	**4,11**
		Sd:	2,40%	6,26%	3,90%	236159,9377	2,5412
		T-:	2,2941	0,2215	0,9221	-3,9128	-5,8823
	$MB^*(FI)$	M :	2,08%	24,62%	29,71%	**454105,77**	**3,66**
		Sd:	3,03%	9,26%	9,16%	304342,9031	2,2492
		T-:	8,5983	2,3882	0,7336	-4,1621	-6,3327
Leukemia3C	MBEGA	M :	0,00%	15,56%	15,76%	2073395,375	4,125
		Sd:	0,00%	8,59%	2,74%	445394,4673	0,991
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**13,09%**	14,54%	**1330103,5**	5
		Sd:	0,00%	4,24%	2,54%	358034,0325	2,4495
		T-:		-2,1962	-1,8144	-27,9338	3,9278
	$MB^*(FI)$	M :	0,00%	**12,93%**	15,08%	**1449598,90**	**3,95**
		Sd:	0,00%	4,99%	2,58%	328167,7618	1,1329
		T-:		-2,3591	-1,1489	-28,1698	-3,4075
Leukemia4C	MBEGA	M :	0,00%	28,33%	33,89%	2055455	15,1667
		Sd:	0,00%	12,12%	14,65%	748936,9207	12,1559
		T-:	-	-	-	-	-
	MB^*	M :	0,33%	**19,61%**	**22,97%**	**1325878,41**	**6,23**
		Sd:	1,35%	6,89%	6,49%	412608,1015	3,4736
		T-:	17	-4,6571	-4,9471	-15,1593	-4,787
	$MB^*(FI)$	M :	0,00%	**20,06%**	**23,39%**	**1389635,15**	**4,9**
		Sd:	0,00%	10,05%	8,35%	399281,0765	1,7442
		T-:		-4,3981	-4,8375	-25,8736	-5,514

Table C.22: MBEGA *vs* MAs endowed with MB^* (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	MBEGA	M :	0,00%	14,26%	23,43%	6746779,6667	28,3333
		Sd:	0,00%	4,54%	2,32%	1982244,4482	15,397
		T-:	-	-	-	-	-
	MB^*	M :	0,21%	15,58%	**20,71%**	**5445415,73**	**21,52**
		Sd:	0,62%	3,26%	2,45%	1411729,7028	10,7619
		T-:	19	3,2843	-10,8408	-4,9301	-6,447
	$MB^*(FI)$	M :	0,28%	15,85%	**21,95%**	**5949881,23**	**23,23**
		Sd:	0,70%	3,19%	2,07%	1507922,7508	10,7327
		T-:	21	6,8057	-15,9419	-2,9019	-4,6466
MLL	MBEGA	M :	0,00%	20,71%	19,52%	1924337,8571	3,7143
		Sd:	0,00%	7,93%	6,80%	670178,6692	0,488
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	22,31%	19,97%	**1230299,47**	**2,84**
		Sd:	0,00%	6,10%	4,64%	231286,0228	0,6882
		T-:		2,551	0,5521	-34,1451	-20,9388
	$MB^*(FI)$	M :	0,00%	**18,51%**	**16,72%**	**1250757,04**	**3,27**
		Sd:	0,00%	4,38%	4,53%	296787,7343	0,7673
		T-:		-3,1451	-3,1196	-29,0851	-6,2366
Orarian	MBEGA	M :	0,00%	2,22%	4,69%	4376831,2857	2
		Sd:	0,00%	0,56%	0,43%	1098440,9428	0
		T-:		0	0	0	-
	MB^*	M :	0,00%	2,04%	4,81%	**3476207,7**	2
		Sd:	0,00%	0,61%	1,03%	879987,1143	0
		T-:		-1,9402	3,5186	-19,5565	-
	$MB^*(FI)$	M :	0,00%	2,41%	4,86%	**3546339,36**	2
		Sd:	0,00%	1,16%	0,62%	1016864,0575	0
		T-:		2,005	4,7662	-18,1199	
SRBCT	MBEGA	M :	0,00%	11,29%	32,54%	2253283,4286	5,7143
		Sd:	0,00%	4,43%	5,64%	502583,6223	2,3604
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**10,42%**	35,38%	**1879785,21**	**4,84**
		Sd:	0,00%	3,29%	4,22%	391358,7537	1,3443
		T-:		-2,2398	4,3806	-5,2944	-3,3114
	$MB^*(FI)$	M :	0,00%	12,50%	35,19%	**1819164,85**	**4,71**
		Sd:	0,00%	6,67%	5,89%	406342,9433	1,0556
		T-:		3,1824	3,84	-5,77	-4,077
9_Tumors	MBEGA	M :	21,11%	75,22%	95,22%	2831195,5	38,66
		Sd:	2,72%	7,99%	1,66%	514528,92	17,1075
		T-:	-	-	-	-	-
	MB^*	M :	**19,67%**	80,03%	**94,03%**	**1928653,05**	**24,5**
		Sd:	5,50%	5,85%	1,95%	594225,66	14,54
		T-:	-3,93	6,01	-5,56	-23,45	-8,25
	$MB^*(FI)$	M :	**17,78%**	77,50%	**94,08%**	**2041232,29**	**25,08**
		Sd:	5,08%	6,96%	1,82%	647967,02	13,52
		T-:	-12,72	2,67	-5,59	-43,66	-8,1821

Table C.23: MBEGA *vs* MAs endowed with MB^* (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
11_Tumors	MBEGA	M :	3,64%	23,22%	74,25%	10439972,6	49
		Sd:	2,03%	3,59%	0,92%	2127251,9073	2,8284
		T-:	-	-	-	-	-
	MB^*	M :	6,58%	24,16%	**70,95%**	**8805682,78**	**41,36**
		Sd:	1,99%	3,60%	2,37%	1782312,8593	5,3042
		T-:	13,9472	1,1403	-10,2474	-13,7566	-18,9712
	$MB^*(FI)$	M :	6,39%	25,96%	**71,93%**	**8369241,76**	**39,90**
		Sd:	2,45%	5,25%	3,50%	1551235,5956	7,674
		T-:	13,6203	3,2999	-25,7121	-24,9407	-32,56
14_Tumors	MBEGA	M :	38,40%	61,93%	85,84%	32426149,2857	47,4286
		Sd:	3,74%	3,48%	1,48%	3598869,1874	4,6853
		T-:	-	-	-	-	-
	MB^*	M :	40,12%	64,28%	**85,56%**	**26873119,55**	**41,77**
		Sd:	3,15%	4,78%	1,53%	5022935,7642	9,8013
		T-:	3,7968	6,2675	-8,753	-7,4335	-7,1959
	$MB^*(FI)$	M :	39,53%	**60,58%**	85,71%	**29345883,12**	**45,12**
		Sd:	2,68%	3,33%	1,23%	2324851,7162	3,6856
		T-:	2,4836	-25,4624	-1,5049	-4,0812	-2,8937
Brain Tumor2	MBEGA	M :	0,00%	51,60%	59,73%	966079,6667	10,6667
		Sd:	0,00%	15,88%	5,25%	424005,4887	7,0048
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	49,89%	**56,34%**	**800287,71**	**7,92**
		Sd:	0,00%	11,64%	3,79%	350967,2153	9,778
		T-:		-0,7543	-2,8894	-9,1777	-3,5338
	$MB^*(FI)$	M :	0,00%	**44,49%**	**55,63%**	970313,07	**5,15**
		Sd:	0,00%	10,32%	4,66%	340633,0128	3,2621
		T-:		-4,5309	-3,3571	0,1487	-7,4399
Prostate Tumor	MBEGA	M :	0,00%	26,39%	40,11%	2119086,8571	27
		Sd:	0,00%	6,07%	7,89%	420355,8836	14,7309
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	**15,51%**	**16,96%**	**1341873,06**	**3,56**
		Sd:	0,00%	4,44%	6,82%	381545,5621	1,1529
		T-:		-11,4862	-13,8723	-15,9023	-10,936
	$MB^*(FI)$	M :	0,00%	**14,66%**	**18,31%**	**1465425,93**	**3,37**
		Sd:	0,00%	4,03%	4,27%	434687,551	1,2583
		T-:		-22,0299	-27,2419	-9,405	-11,0161
Lymphoma	MBEGA	M :	0,00%	12,29%	14,81%	1328352	2,8571
		Sd:	0,00%	2,56%	2,99%	380575,5339	0,378
		T-:	-	-	-	-	-
	MB^*	M :	0,00%	16,94%	19,97%	**878011,05**	**1,05**
		Sd:	0,00%	3,41%	4,34%	171691,0113	0,2294
		T-:		111,3392	21,8592	-8,6881	-87,6176
	$MB^*(FI)$	M :	0,00%	17,69%	20,84%	**1041620,52**	**1**
		Sd:	0,00%	0,56%	3,75%	246454,2847	0
		T-:		212,9715	35,1661	-5,5095	-91

Table C.24: MBEGA *vs* MAs endowed with MB^* (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	- (only GA)	M :	2,35%	46,15%	49,78%	4647524,7647	36,5294
		Sd:	2,47%	7,42%	4,89%	2030048,161	8,0477
		T-:	-	-	-	-	-
	FLS^*	M :	2,00%	**44,67%**	**45,17%**	**1333514**	**14,5**
		Sd:	2,11%	7,52%	5,36%	610227,5602	14,4318
		T-:	-1,5882	-5,6538	-5,6168	-16,6203	-9,6926
	$IFLS^*$	M :	2,55%	**37,65%**	**41,29%**	**982879,18**	**8,54**
		Sd:	2,02%	6,10%	6,28%	432756,9458	12,6283
		T-:	1,1744	-25,707	-18,2188	-19,3461	-31,3389
CNS	- (only GA)	M :	0,00%	42,67%	44,27%	2819096,5294	33,3529
		Sd:	0,00%	11,91%	5,84%	1436436,1201	11,7789
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	49,60%	**40,80%**	**855501,7**	**19**
		Sd:	0,00%	4,82%	4,52%	488674,9818	15,42
		T-:		10,4742	-6,5834	-18,7466	-9,5828
	$IFLS^*$	M :	0,00%	43,94%	**39,33%**	**681142,63**	**15,18**
		Sd:	0,00%	8,45%	5,54%	280565,0892	12,616
		T-:		1,8025	-9,3925	-22,4027	-9,3574
Colon	- (only GA)	M :	2,21%	26,91%	33,32%	1910513,9412	26,5882
		Sd:	3,08%	5,96%	6,26%	1205013,4761	10,7008
		T-:	-	-	-	-	-
	FLS^*	M :	**1,39%**	**23,94%**	**28,75%**	**540645**	**8,44**
		Sd:	2,76%	5,71%	8,87%	386287,492	11,4139
		T-:	-4,0521	-5,0802	-3,4134	-16,516	-24,134
	$IFLS^*$	M :	1,88%	27,42%	31,23%	**546720,5**	**4,8**
		Sd:	3,02%	8,60%	11,78%	268943,5013	3,4254
		T-:	-0,7251	0,2995	-0,8616	-13,6359	-45,2702
Leukemia3C	- (only GA)	M :	0,00%	21,23%	28,54%	4875401,7368	38,5263
		Sd:	0,00%	7,26%	7,82%	1918293,1208	7,9398
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	**19,35%**	**20,23%**	**2049922,91**	**22,75**
		Sd:	0,00%	11,95%	12,29%	378502,5758	15,0944
		T-:		-4,5813	-21,5828	-23,141	-10,0605
	$IFLS^*$	M :	0,00%	**12,53%**	**14,95%**	**1388104,36**	**8,72**
		Sd:	0,00%	6,55%	2,78%	365546,272	11,7055
		T-:		-32,2099	-24,3926	-29,3285	-72,582
Leukemia4C	- (only GA)	M :	0,00%	33,54%	40,58%	4838907,0526	41,9474
		Sd:	0,00%	9,16%	7,66%	1583335,4614	5,6517
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	33,56%	**32,41%**	**2333865,25**	**30,41**
		Sd:	0,00%	8,81%	10,35%	572518,444	16,8871
		T-:		0,0244	-13,7828	-19,055	-5,7743
	$IFLS^*$	M :	0,00%	**15,33%**	**21,89%**	**1231547,6**	**7,5**
		Sd:	0,00%	5,12%	3,65%	300043,4306	2,7588
		T-:		-109,5619	-24,01	-29,4478	-76,6206

Table C.25: GA *vs* MAs respectively endowed with FLS^* and $IFLS^*$ (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	- (only GA)	M :	0,00%	15,32%	26,88%	17129218,9444	46,1111
		Sd:	0,00%	3,74%	1,96%	6676624,4762	3,0076
		T-:	-	-	-	-	-
	FLS^*	M :	0,53%	15,88%	27,20%	**7728295,90**	**30,27**
		Sd:	0,92%	3,22%	2,70%	2524027,1574	17,1178
		T-:	11	0,8393	2,3605	-28,6974	-7,8199
	$IFLS^*$	M :	0,53%	16,40%	**22,59%**	**5892768,09**	**28**
		Sd:	0,92%	3,51%	1,72%	1912886,4127	13,0996
		T-:	11	1,8942	-111,346	-112,4203	-366,75
MLL	- (only GA)	M :	0,00%	22,66%	32,72%	4762235,3158	34,6316
		Sd:	0,00%	8,68%	8,22%	2181348,1554	9,7821
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	24,85%	**27,78%**	**1659378,90**	**9,72**
		Sd:	0,00%	10,21%	8,64%	436280,5364	11,0643
		T-:		3,0063	-4,8921	-61,4583	-157,4055
	$IFLS^*$	M :	0,00%	**19,49%**	**18,52%**	**1323288,41**	**3,83**
		Sd:	0,00%	5,28%	5,54%	379447,7209	0,8348
		T-:		-4,2911	-48,4295	-89,221	-310,6633
Orarian	- (only GA)	M :	0,00%	3,40%	19,12%	12696723,3333	24,5
		Sd:	0,00%	2,36%	10,46%	5157035,8265	6,7584
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	4,68%	**12,71%**	**4742630,3**	**9,9**
		Sd:	0,00%	2,10%	8,69%	802927,2552	9,3268
		T-:		14,7361	-12,0306	-15,4832	-44,7172
	$IFLS^*$	M :	0,00%	**2,50%**	**4,86%**	**3703556,54**	**3**
		Sd:	0,00%	1,19%	1,06%	895457,5787	0,7746
		T-:		-4,6914	-41,0366	-17,7286	-86
SRBCT	- (only GA)	M :	0,00%	12,85%	44,99%	7476719,9444	41,7222
		Sd:	0,00%	6,34%	7,32%	3298359,6883	5,9092
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	16,78%	44,93%	**2696851,4**	**15,8**
		Sd:	0,00%	7,42%	8,91%	656798,3324	12,9168
		T-:		11,1021	-0,0611	-17,4868	-33,9197
	$IFLS^*$	M :	0,00%	13,66%	**37,92%**	**1977765,63**	**6,90**
		Sd:	0,00%	5,51%	7,92%	314170,8714	2,9818
		T-:		1,3131	-10,6959	-20,2372	-89,3649
9_Tumors	- (only GA)	M :	20,00%	76,74%	94,32%	6014650,1053	41,1053
		Sd:	3,85%	8,32%	1,85%	2794142,5604	7,5196
		T-:	-	-	-	-	-
	FLS^*	M :	22,42%	78,18%	**93,33%**	**3177301,72**	**42,27**
		Sd:	4,49%	7,21%	2,84%	562579,1609	7,5245
		T-:	5,8507	2,4561	-19	-12,3704	1,8657
	$IFLS^*$	M :	20,56%	79,39%	**93,56%**	**2376800,5**	**36,83**
		Sd:	5,29%	6,13%	2,04%	487727,1835	18,4136
		T-:	1,5697	4,1597	-3,0875	-16,1567	-4,8242
11_Tumors	- (only GA)	M :	5,74%	27,19%	73,43%	22335377,2105	47,5789
		Sd:	1,76%	4,35%	3,54%	8341713,8785	1,8048
		T-:	-	-	-	-	-
	FLS^*	M :	7,39%	27,68%	**72,45%**	**10935088,25**	**46,58**
		Sd:	2,93%	4,88%	1,61%	1454011,6074	3,4499
		T-:	4,8914	1,8622	-4,8423	-18,1639	-17,3537
	$IFLS^*$	M :	6,82%	26,48%	**73,01%**	**8521776,16**	**46,91**
		Sd:	2,17%	4,33%	2,65%	1351529,1854	3,6546
		T-:	5,4457	-1,1718	-7,7309	-21,236	-2,5596

Table C.26: GA *vs* MAs respectively endowed with FLS^* and $IFLS^*$ (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	- (only GA)	M :	39,26%	62,83%	87,26%	74594390,4615	46,3846
		Sd:	3,83%	3,68%	1,11%	17820808,0553	3,9059
		T-:	-	-	-	-	-
	FLS^*	M :	41,56%	62,13%	**85,68%**	**31305602,9**	47,9
		Sd:	2,67%	4,18%	1,30%	2811654,1339	0,9944
		T-:	7,9956	-2,0152	-7,681	-33,7107	2,0831
	$IFLS^*$	M :	40,91%	63,95%	**86,57%**	**25333907,58**	45,66
		Sd:	2,90%	2,75%	1,55%	2716144,3819	4,5394
		T-:	8,0942	3,1448	-5,3741	-37,9305	-0,8895
Brain Tumor2	- (GA only)	M :	0,00%	44,65%	57,31%	3975539,1818	39,6364
		Sd:	0,00%	10,83%	10,01%	1120421,9314	10,8099
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	54,67%	63,20%	**1182840,88**	**15,66**
		Sd:	0,00%	6,54%	5,63%	360827,1869	10,7005
		T-:		39,59	4,4369	-35,3786	-18,9393
	$IFLS^*$	M :	1,71%	**40,44%**	55,91%	**1101117,77**	**10**
		Sd:	3,39%	10,73%	5,51%	310353,7219	9,4604
		T-:	9	-17,0285	-1,1148	-32,3433	-28,88
Prostate Tumor	- (only GA)	M :	0,00%	27,48%	35,72%	5427461,75	33
		Sd:	0,00%	5,74%	8,72%	2153667,5087	6,396
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	**23,02%**	**31,84%**	**2361803,2**	**17,1**
		Sd:	0,00%	7,25%	11,24%	744562,2425	14,0194
		T-:		-5,0461	-7,8095	-73,5648	-8,7973
	$IFLS^*$	M :	0,00%	**17,22%**	**19,43%**	**1892145,72**	**4,27**
		Sd:	0,00%	5,50%	6,49%	632392,798	1,4206
		T-:		-30,1973	-31,5264	-61,8562	-104,2823
Lymphoma	- (only GA)	M :	0,00%	7,84%	18,42%	2970675,6667	18,2
		Sd:	0,00%	5,55%	3,68%	1320242,23	7,6737
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	12,48%	18,85%	**1309051,1**	**5,1**
		Sd:	0,00%	7,92%	4,77%	353752,8195	1,5951
		T-:		9,0214	2,7218	-36,8667	-24,3424
	$IFLS^*$	M :	0,00%	16,36%	18,79%	**1024485,91**	**2,5**
		Sd:	0,00%	6,66%	5,16%	279465,7381	0,7977
		T-:		16,6817	0,7049	-178,6872	-34,487

Table C.27: GA *vs* MAs respectively endowed with FLS^* and $IFLS^*$ (3)

Data	Local Search	Local Search	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	MBEGA	M :	2,22%	45,05%	47,45%	1306510,8889	15,3333
		Sd:	2,11%	3,43%	6,95%	657546,5512	12,52
		T-:	-	-	-	-	-
	FLS^*	M :	2,00%	44,67%	45,17%	1333514	14,5
		Sd:	2,11%	7,52%	5,36%	610227,5602	14,4318
		T-:	-0,7905	-1,0688	-1,3237	0,3022	-0,3192
	$IFLS^*$	M :	2,55%	**37,65%**	**41,29%**	**982879,18**	**8,54**
		Sd:	2,02%	6,10%	6,28%	432756,9458	12,6283
		T-:	1,3598	-18,1231	-3,8778	-5,0483	-4,3376
CNS	MBEGA	M :	2,22%	41,19%	40,44%	776660,4444	16,4444
		Sd:	3,33%	6,85%	5,35%	579215,0463	16,8605
		T-:	-	-	-	-	-
	FLS^*	M :	**0,00%**	49,60%	40,80%	855501,7	19
		Sd:	0,00%	4,82%	4,52%	488674,9818	15,42
		T-:	-9	11,1884	1,5887	1,0038	1,4653
	$IFLS^*$	M :	**0,00%**	43,94%	**39,33%**	681142,6364	15,1818
		Sd:	0,00%	8,45%	5,54%	280565,0892	12,616
		T-:	-9	3,4808	-5,0558	-1,4558	-0,5907
Colon	MBEGA	M :	0,78%	23,47%	28,79%	769904,375	10,125
		Sd:	2,21%	3,10%	8,53%	631280,9958	13,174
		T-:	-	-	-	-	-
	FLS^*	M :	1,39%	23,94%	28,75%	**540645**	8,44
		Sd:	2,76%	5,71%	8,87%	386287,492	11,4139
		T-:	3,3273	0,7397	-0,0293	-3,0511	-1,4216
	$IFLS^*$	M :	1,88%	27,42%	31,23%	546720,5	4,8
		Sd:	3,02%	8,60%	11,78%	268943,5013	3,4254
		T-:	2,44	2,2826	0,9585	-2,3833	-5,1626
Leukemia3C	MBEGA	M :	0,00%	15,56%	15,76%	2073395,375	4,125
		Sd:	0,00%	8,59%	2,74%	445394,4673	0,991
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	19,35%	20,23%	2049922,9167	22,75
		Sd:	0,00%	11,95%	12,29%	378502,5758	15,0944
		T-:		3,2738	6,6822	-0,6475	11,9194
	$IFLS^*$	M :	0,00%	12,53%	14,95%	**1388104,36**	8,72
		Sd:	0,00%	6,55%	2,78%	365546,272	11,7055
		T-:		-2,7105	-1,0434	-29,3666	11,8388
Leukemia4C	MBEGA	M :	0,00%	28,33%	33,89%	2055455	15,1667
		Sd:	0,00%	12,12%	14,65%	748936,9207	12,1559
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	33,56%	32,41%	2333865,25	30,4167
		Sd:	0,00%	8,81%	10,35%	572518,444	16,8871
		T-:		2,438	-0,6836	4,8595	5,6553
	$IFLS^*$	M :	0,00%	**15,33%**	**21,89%**	**1231547,6**	**7,5**
		Sd:	0,00%	5,12%	3,65%	300043,4306	2,7588
		T-:		-7,0195	-5,3933	-25,9768	-4,1061

Table C.28: MBEGA *vs* MAs respectively endowed with FLS^* and $IFLS^*$ (1)

Data	Local Search ()	Local Search	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	MBEGA	M :	0,00%	14,26%	23,43%	6746779,6667	28,3333
		Sd:	0,00%	4,54%	2,32%	1982244,4482	15,397
		T-:	-	-	-	-	-
	FLS^*	M :	0,53%	15,88%	27,20%	7728295,9091	30,2727
		Sd:	0,92%	3,22%	2,70%	2524027,1574	17,1178
		T-:	11	2,6232	23,0867	2,3659	0,8493
	$IFLS^*$	M :	0,53%	16,40%	**22,59%**	**5892768,09**	28
		Sd:	0,92%	3,51%	1,72%	1912886,4127	13,0996
		T-:	11	4,1274	-8,7765	-3,1229	-0,3158
MLL	MBEGA	M :	0,00%	20,71%	19,52%	1924337,8571	3,7143
		Sd:	0,00%	7,93%	6,80%	670178,6692	0,488
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	24,85%	27,78%	**1659378,90**	9,72
		Sd:	0,00%	10,21%	8,64%	436280,5364	11,0643
		T-:		7,106	6,3903	-6,9518	37,0616
	$IFLS^*$	M :	0,00%	19,49%	18,52%	**1323288,41**	3,83
		Sd:	0,00%	5,28%	5,54%	379447,7209	0,8348
		T-:		-2,0543	-1,1734	-30,4602	1,129
Orarian	MBEGA	M :	0,00%	2,22%	4,69%	4376831,2857	2
		Sd:	0,00%	0,56%	0,43%	1098440,9428	0
		T-:		0	0	0	
	FLS^*	M :	0,00%	4,68%	12,71%	4742630,3	9,9
		Sd:	0,00%	2,10%	8,69%	802927,2552	9,3268
		T-:		19,5432	18,5926	2,5216	37,619
	$IFLS^*$	M :	0,00%	2,50%	4,86%	**3703556,54**	3
		Sd:	0,00%	1,19%	1,06%	895457,5787	0,7746
		T-:		1,2938	1,1047	-5,6004	
SRBCT	MBEGA	M :	0,00%	11,29%	32,54%	2253283,4286	5,7143
		Sd:	0,00%	4,43%	5,64%	502583,6223	2,3604
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	16,78%	44,93%	2696851,4	15,8
		Sd:	0,00%	7,42%	8,91%	656798,3324	12,9168
		T-:		12,1859	10,6145	5,5035	13,9545
	$IFLS^*$	M :	0,00%	13,66%	37,92%	**1977765,63**	6,90
		Sd:	0,00%	5,51%	7,92%	314170,8714	2,9818
		T-:		3,4904	5,8433	-3,6768	3,9806
9_Tumors	MBEGA	M :	21,11%	75,22%	95,22%	2831195,5	38,66
		Sd:	2,72%	7,99%	1,66%	514528,9278	17,1075
		T-:	-	-	-	-	-
	FLS^*	M :	22,42%	78,18%	**93,33%**	3177301,72	42,27
		Sd:	4,49%	7,21%	2,84%	562579,16	7,5245
		T-:	4,5617	3,2518	-9,2727	6,9652	2,1765
	$IFLS^*$	M :	**20,56%**	79,39%	**93,56%**	**2376800,5**	36,83
		Sd:	5,29%	6,13%	2,04%	487727,1835	18,4136
		T-:	-2,9104	4,4198	-5,285	-19,2304	-1,035
11_Tumors	MBEGA	M :	3,64%	23,22%	74,25%	10439972,6	49
		Sd:	2,03%	3,59%	0,92%	2127251,9073	2,8284
		T-:	-	-	-	-	-
	FLS^*	M :	7,39%	27,68%	**72,45%**	10935088,25	**46,58**
		Sd:	2,93%	4,88%	1,61%	1454011,6074	3,4499
		T-:	9,9197	5,1783	-8,9986	5,6064	-11,7415
	$IFLS^*$	M :	6,82%	26,48%	**73,01%**	**8521776,16**	**46,91**
		Sd:	2,17%	4,33%	2,65%	1351529,1854	3,6546
		T-:	12,1193	3,1913	-26,3171	-9,9723	-6,3983

Table C.29: MBEGA *vs* MAs respectively endowed with FLS^* and $IFLS^*$ (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	MBEGA	M :	38,40%	61,93%	85,84%	32426149,2857	47,4286
		Sd:	3,74%	3,48%	1,48%	3598869,1874	4,6853
		T-:	-	-	-	-	-
	FLS^*	M :	41,56%	62,13%	85,68%	31305602,9	47,9
		Sd:	2,67%	4,18%	1,30%	2811654,1339	0,9944
		T-:	6,0645	0,8866	-0,8802	-1,4992	0,6038
	$IFLS^*$	M :	40,91%	63,95%	86,57%	**25333907,58**	45,66
		Sd:	2,90%	2,75%	1,55%	2716144,3819	4,5394
		T-:	5,2297	8,3288	7	-9,1846	-2,0596
Brain Tumor2	MBEGA	M :	0,00%	51,60%	59,73%	966079,6667	10,6667
		Sd:	0,00%	15,88%	5,25%	424005,4887	7,0048
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	54,67%	63,20%	1182840,8889	15,6667
		Sd:	0,00%	6,54%	5,63%	360827,1869	10,7005
		T-:		1,9907	2,0414	10,5187	3,8631
	$IFLS^*$	M :	1,71%	**40,44%**	**55,91%**	1101117,7778	10
		Sd:	3,39%	10,73%	5,51%	310353,7219	9,4604
		T-:	9	-7,246	-2,3289	2,9525	-0,6281
Prostate Tumor	MBEGA	M :	0,00%	26,39%	40,11%	2119086,8571	27
		Sd:	0,00%	6,07%	7,89%	420355,8836	14,7309
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	**23,02%**	**31,84%**	2361803,2	**17,1**
		Sd:	0,00%	7,25%	11,24%	744562,2425	14,0194
		T-:		-3,6754	-9,1321	4,2816	-3,5457
	$IFLS^*$	M :	0,00%	**17,22%**	**19,43%**	**1892145,72**	**4,27**
		Sd:	0,00%	5,50%	6,49%	632392,798	1,4206
		T-:		-22,0967	-22,5591	-3,2951	-10,5906
Lymphoma	MBEGA	M :	0,00%	12,29%	14,81%	1328352	2,8571
		Sd:	0,00%	2,56%	2,99%	380575,5339	0,378
		T-:	-	-	-	-	-
	FLS^*	M :	0,00%	12,48%	18,85%	1309051,1	5,1
		Sd:	0,00%	7,92%	4,77%	353752,8195	1,5951
		T-:		0,3737	22,6841	-0,2819	7,7149
	$IFLS^*$	M :	0,00%	16,36%	18,79%	**1024485,91**	**2,5**
		Sd:	0,00%	6,66%	5,16%	279465,7381	0,7977
		T-:		8,0475	7,6143	-5,7671	-7,6977

Table C.30: MBEGA *vs* MAs respectively endowed with FLS^* and $IFLS^*$ (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	- (only GA)	M :	2,35%	46,15%	49,78%	4647524,7647	36,5294
		Sd:	2,47%	7,42%	4,89%	2030048,161	8,0477
		T-:	-	-	-	-	-
	AF^*	M :	2,50%	**40,83%**	**39,74%**	**767518,5**	**3,37**
		Sd:	2,07%	2,07%	7,59%	198523,0225	1,1877
		T-:	0,6968	-21,3451	-18,3461	-20,4848	-44,7335
	BF^*	M :	2,18%	**42,20%**	**46,40%**	**1771606,54**	**23,18**
		Sd:	2,75%	4,29%	7,87%	798072,3754	16,7082
		T-:	-0,7752	-10,6888	-4,0988	-15,0897	-10,2447
CNS	- (only GA)	M :	0,00%	42,67%	44,27%	2819096,5294	33,3529
		Sd:	0,00%	11,91%	5,84%	1436436,1201	11,7789
		T-:	-	-	-	-	-
	AF^*	M :	2,00%	**38,47%**	**40,07%**	**717483,1**	**7,2**
		Sd:	3,22%	6,13%	5,44%	372142,6251	13,3317
		T-:	10	-9,7539	-5,7513	-22,3552	-20,6129
	BF^*	M :	1,21%	**41,09%**	**40,55%**	**762122,36**	**17,63**
		Sd:	2,70%	4,00%	6,12%	287622,6685	11,9271
		T-:	11	-3,787	-6,5016	-21,989	-13,0576
Colon	- (only GA)	M :	2,21%	26,91%	33,32%	1910513,9412	26,5882
		Sd:	3,08%	5,96%	6,26%	1205013,4761	10,7008
		T-:	-	-	-	-	-
	AF^*	M :	2,08%	**23,73%**	**25,88%**	**668997,55**	**3,33**
		Sd:	3,12%	9,05%	6,97%	451509,0209	1,6583
		T-:	-0,4618	-9,1892	-26,8163	-12,4935	-51,9198
	BF^*	M :	3,47%	**22,58%**	**27,46%**	**612665**	**9,33**
		Sd:	3,29%	4,71%	9,18%	377102,3515	11,8427
		T-:	3,1111	-49,7886	-5,9986	-15,0021	-32,2066
Leukemia3C	- (only GA)	M :	0,00%	21,23%	28,54%	4875401,7368	38,5263
		Sd:	0,00%	7,26%	7,82%	1918293,1208	7,9398
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**18,64%**	**18,23%**	**1984104,36**	**16,81**
		Sd:	0,00%	9,88%	3,84%	443160,1701	16,576
		T-:		-3,3534	-33,0794	-22,2743	-39,8035
	BF^*	M :	0,00%	**18,22%**	**22,39%**	**2066252,9**	**18,4**
		Sd:	0,00%	14,03%	12,25%	619923,2709	13,93
		T-:		-11,1578	-6,5553	-18,1362	-8,8901
Leukemia4C	- (only GA)	M :	0,00%	33,54%	40,58%	4838907,0526	41,9474
		Sd:	0,00%	9,16%	7,66%	1583335,4614	5,6517
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**17,28%**	**23,83%**	**1727720,6**	**13,5**
		Sd:	0,00%	6,23%	6,44%	390183,4013	10,5646
		T-:		-13,1261	-25,3841	-24,0981	-40,9696
	BF^*	M :	0,00%	**27,37%**	**26,62%**	**2367789,27**	**33,54**
		Sd:	0,00%	6,85%	9,67%	925866,8675	15,6548
		T-:		-5,9442	-16,0442	-20,1165	-3,0899

Table C.31: GA *vs* MAs respectively endowed with AF^* and BF^* (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	- (only GA)	M :	0,00%	15,32%	26,88%	17129218,9444	46,1111
		Sd:	0,00%	3,74%	1,96%	6676624,4762	3,0076
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**14,10%**	**21,78%**	**6257937,8**	**28,2**
		Sd:	0,00%	1,94%	3,13%	1277321,3431	10,6019
		T-:		-3,6656	-36,6858	-204,2192	-95,9604
	BF^*	M :	0,78%	**14,69%**	**25,64%**	**7284814,1**	**31,6**
		Sd:	1,01%	4,00%	2,42%	966277,491	6,5693
		T-:	6,6667	-1,2765	-2,765	-104,836	-39,9347
MLL	- (only GA)	M :	0,00%	22,66%	32,72%	4762235,3158	34,6316
		Sd:	0,00%	8,68%	8,22%	2181348,1554	9,7821
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**20,11%**	**17,06%**	**1630589,4**	**5,4**
		Sd:	0,00%	4,50%	2,55%	564417,6302	1,1738
		T-:		-3,3484	-26,3499	-79,4004	-206,8226
	BF^*	M :	0,00%	23,82%	**20,14%**	**1915797,62**	**10,87**
		Sd:	0,00%	4,47%	7,33%	313542,6983	10,494
		T-:		1,3202	-17,6738	-71,6312	-32,3381
Orarian	-	M :	0,00%	3,40%	19,12%	12696723,3333	24,5
		Sd:	0,00%	2,36%	10,46%	5157035,8265	6,7584
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**2,56%**	**4,50%**	**4135321,22**	**4**
		Sd:	0,00%	1,24%	0,47%	427836,3558	1
		T-:		-3,9948	-45,8326	-17,2914	-82
	BF^*	M :	0,00%	4,40%	**14,27%**	**4497325,2**	**7,4**
		Sd:	0,00%	4,46%	9,35%	1511936,25	4,4771
		T-:		23,3	-5,0712	-16,2286	-33,7902
SRBCT	- (only GA)	M :	0,00%	12,85%	44,99%	7476719,9444	41,7222
		Sd:	0,00%	6,34%	7,32%	3298359,6883	5,9092
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**8,13%**	**32,41%**	**2263354,33**	**8,66**
		Sd:	0,00%	4,44%	6,41%	682208,1547	2,7839
		T-:		-14,4721	-19,814	-19,1572	-83,7102
	BF^*	M :	0,00%	12,90%	**41,84%**	**2559662,55**	**9,33**
		Sd:	0,00%	4,86%	8,16%	492488,8605	3,8079
		T-:		0,0526	-11,2893	-18,0685	-47,1038
9_Tumors	- (only GA)	M :	20,00%	76,74%	94,32%	6014650,1053	41,1053
		Sd:	3,85%	8,32%	1,85%	2794142,5604	7,5196
		T-:	-	-	-	-	-
	AF^*	M :	**18,67%**	**74,40%**	94,40%	**2181853,7**	35,6
		Sd:	4,22%	10,08%	1,92%	481033,3212	16,6747
		T-:	-3,5522	-4,7127	0,5008	-16,8575	-1,986
	BF^*	M :	**18,18%**	**73,94%**	94,67%	**2810649**	42,6364
		Sd:	4,31%	3,34%	2,35%	373754,2593	4,9452
		T-:	-4,6877	-5,2858	4,4043	-14,1736	5,3555
11_Tumors	- (only GA)	M :	5,74%	27,19%	73,43%	22335377,2105	47,5789
		Sd:	1,76%	4,35%	3,54%	8341713,8785	1,8048
		T-:	-	-	-	-	-
	AF^*	M :	6,59%	29,20%	74,76%	**9120201,7**	47,3
		Sd:	2,92%	4,63%	2,54%	1049930,1925	1,8886
		T-:	4,0014	5,5393	12,0357	-20,503	-1,6151
	BF^*	M :	6,20%	**25,89%**	**72,23%**	**11407949,18**	48
		Sd:	1,79%	3,41%	2,41%	1830587,46	3,0984
		T-:	2,844	-6,0884	-6,2765	-17,3567	13,8182

Table C.32: GA *vs* MAs respectively endowed with AF^* and BF^* (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	- (only GA)	M :	39,26%	62,83%	87,26%	74594390,4615	46,3846
		Sd:	3,83%	3,68%	1,11%	17820808,0553	3,9059
		T-:	-	-	-	-	-
	AF^*	M :	38,96%	63,45%	88,41%	**27580478,77**	47,11
		Sd:	2,34%	2,86%	2,76%	2424059,0204	3,0185
		T-:	-1,5809	1,4105	3,2088	-36,228	0,9971
	BF^*	M :	**38,70%**	**61,01%**	**86,52%**	**30804245,7**	45,4
		Sd:	4,01%	2,73%	1,75%	2975937,6218	3,4059
		T-:	-4,4517	-6,8575	-5,186	-32,2095	-1,0206
Brain Tumor2	- (only GA)	M :	0,00%	44,65%	57,31%	3975539,1818	39,6364
		Sd:	0,00%	10,83%	10,01%	1120421,9314	10,8099
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	42,70%	59,40%	**814440,62**	**10,5**
		Sd:	0,00%	12,37%	5,40%	259856,6338	5,757
		T-:		-0,833	2,6886	-36,3411	-27,6764
	BF^*	M :	0,77%	52,40%	57,36%	**1491165,9**	**26,6**
		Sd:	2,43%	10,04%	5,39%	446324,1521	16,0638
		T-:	10	37,0906	0,0883	-23,3714	-9,1369
Prostate Tumor	- (only GA)	M :	0,00%	27,48%	35,72%	5427461,75	33
		Sd:	0,00%	5,74%	8,72%	2153667,5087	6,396
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**17,98%**	**17,09%**	**1790517,28**	**6**
		Sd:	0,00%	6,64%	9,53%	1003849,2937	0,8165
		T-:		-26,1885	-21,7578	-169,4535	-108
	BF^*	M :	0,00%	32,04%	35,97%	**2265346,14**	**13,71**
		Sd:	0,00%	12,99%	11,84%	471051,8344	11,4705
		T-:		3,4123	0,3265	-314,2185	-22,5894
Lymphoma	- (only GA)	M :	0,00%	7,84%	18,42%	2970675,6667	18,2
		Sd:	0,00%	5,55%	3,68%	1320242,23	7,6737
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	18,38%	20,15%	**1359312,33**	**3,08**
		Sd:	0,00%	3,93%	3,53%	223090,9887	1,4434
		T-:		56,8212	9,9187	-53,4512	-33,3417
	BF^*	M :	0,00%	12,66%	19,39%	**1447421**	**6,55**
		Sd:	0,00%	5,16%	3,48%	302116,4127	5,0525
		T-:		61,4429	2,7493	-50,8622	-13,7399

Table C.33: GA *vs* MAs respectively endowed with AF^* and BF^* (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	MBEGA	M :	2,22%	45,05%	47,45%	1306510,8889	15,3333
		Sd:	2,11%	3,43%	6,95%	657546,5512	12,52
		T-:	-	-	-	-	-
	AF^*	M :	2,50%	**40,83%**	**39,74%**	**767518,5**	**3,375**
		Sd:	2,07%	2,07%	7,59%	198523,0225	1,1877
		T-:	1,0199	-12,2007	-4,7745	-8,4141	-8,0607
	BF^*	M :	2,18%	**42,20%**	46,40%	1771606,5455	23,1818
		Sd:	2,75%	4,29%	7,87%	798072,3754	16,7082
		T-:	-0,1443	-6,4681	-0,6085	6,8939	4,2887
CNS	MBEGA	M :	2,22%	41,19%	40,44%	776660,4444	16,4444
		Sd:	3,33%	6,85%	5,35%	579215,0463	16,8605
		T-:	-	-	-	-	-
	BF^*	M :	2,00%	**38,47%**	40,07%	717483,1	**7,2**
		Sd:	3,22%	6,13%	5,44%	372142,6251	13,3317
		T-:	-0,6994	-4,8599	-0,682	-0,9315	-5,9574
	BF^*	M :	**1,21%**	41,09%	40,55%	762122,3636	17,6364
		Sd:	2,70%	4,00%	6,12%	287622,6685	11,9271
		T-:	-3,7358	-0,1719	0,3186	-0,2314	0,7952
Colon	MBEGA	M :	0,78%	23,47%	28,79%	769904,375	10,125
		Sd:	2,21%	3,10%	8,53%	631280,9958	13,174
		T-:	-	-	-	-	-
	AF^*	M :	2,08%	23,73%	**25,88%**	668997,5556	**3,33**
		Sd:	3,12%	9,05%	6,97%	451509,0209	1,6583
		T-:	5,1827	0,594	-3,6945	-1,0855	-6,6827
	BF^*	M :	3,47%	**22,58%**	27,46%	612665	9,3333
		Sd:	3,29%	4,71%	9,18%	377102,3515	11,8427
		T-:	6,7617	-3,1518	-1,09	-1,9888	-0,7483
Leukemia3C	MBEGA	M :	0,00%	15,56%	15,76%	2073395,375	4,125
		Sd:	0,00%	8,59%	2,74%	445394,4673	0,991
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	18,64%	18,23%	1984104,3636	16,8182
		Sd:	0,00%	9,88%	3,84%	443160,1701	16,576
		T-:		2,3129	3,9229	-1,565	23,9876
	BF^*	M :	0,00%	18,22%	22,39%	2066252,9	18,4
		Sd:	0,00%	14,03%	12,25%	619923,2709	13,93
		T-:		2,3856	6,102	-0,07	6,3162
Leukemia4C	MBEGA	M :	0,00%	28,33%	33,89%	2055455	15,1667
		Sd:	0,00%	12,12%	14,65%	748936,9207	12,1559
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**17,28%**	**23,83%**	**1727720,6**	13,5
		Sd:	0,00%	6,23%	6,44%	390183,4013	10,5646
		T-:		-4,9757	-4,5994	-6,3452	-0,8588
	BF^*	M :	0,00%	27,37%	**26,62%**	2367789,2727	33,5455
		Sd:	0,00%	6,85%	9,67%	925866,8675	15,6548
		T-:		-0,4535	-3,2197	9,4627	5,6243

Table C.34: MBEGA *vs* MAs respectively endowed with AF^* and BF^* (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	MBEGA	M :	0,00%	14,26%	23,43%	6746779,6667	28,3333
		Sd:	0,00%	4,54%	2,32%	1982244,4482	15,397
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	14,10%	**21,78%**	6257937,8	28,2
		Sd:	0,00%	1,94%	3,13%	1277321,3431	10,6019
		T-:		-0,652	-10,0502	-1,8798	-0,1245
	BF^*	M :	0,78%	14,69%	25,64%	7284814,1	31,6
		Sd:	1,01%	4,00%	2,42%	966277,491	6,5693
		T-:	6,6667	0,9943	4,8689	1,983	2,9291
MLL	MBEGA	M :	0,00%	20,71%	19,52%	1924337,8571	3,7143
		Sd:	0,00%	7,93%	6,80%	670178,6692	0,488
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	20,11%	**17,06%**	**1630589,4**	5,4
		Sd:	0,00%	4,50%	2,55%	564417,6302	1,1738
		T-:		-0,9673	-2,4666	-13,7062	11,5596
	BF^*	M :	0,00%	23,82%	20,14%	1915797,625	10,875
		Sd:	0,00%	4,47%	7,33%	313542,6983	10,494
		T-:		4,0802	0,5724	-0,3887	9,7357
Orarian	MBEGA	M :	0,00%	2,22%	4,69%	4376831,2857	2
		Sd:	0,00%	0,56%	0,43%	1098440,9428	0
		T-:		0	0	0	
	AF^*	M :	0,00%	2,56%	**4,50%**	**4135321,22**	4
		Sd:	0,00%	1,24%	0,47%	427836,3558	1
		T-:		1,4668	-2,9278	-5,0585	
	BF^*	M :	0,00%	4,40%	14,27%	4497325,2	7,4
		Sd:	0,00%	4,46%	9,35%	1511936,25	4,4771
		T-:		21,677	10,5957	1,082	12,2727
SRBCT	MBEGA	M :	0,00%	11,29%	32,54%	2253283,4286	5,7143
		Sd:	0,00%	4,43%	5,64%	502583,6223	2,3604
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**8,13%**	32,41%	2263354,3333	8,6667
		Sd:	0,00%	4,44%	6,41%	682208,1547	2,7839
		T-:		-7,3727	-0,1462	0,1318	9,6159
	BF^*	M :	0,00%	12,90%	41,84%	2559662,5556	9,3333
		Sd:	0,00%	4,86%	8,16%	492488,8605	3,8079
		T-:		1,5085	13,3422	4,0102	5,6442
9_Tumors	MBEGA	M :	21,11%	75,22%	95,22%	2831195,5	38,6667
		Sd:	2,72%	7,99%	1,66%	514528,9278	17,1075
		T-:	-	-	-	-	-
	AF^*	M :	**18,67%**	74,40%	**94,40%**	**2181853,7**	35,6
		Sd:	4,22%	10,08%	1,92%	481033,3212	16,6747
		T-:	-10,7123	-0,9636	-3,1743	-16,4648	-0,968
	BF^*	M :	**18,18%**	73,94%	**94,67%**	2810649	42,6364
		Sd:	4,31%	3,34%	2,35%	373754,2593	4,9452
		T-:	-11,8011	-1,4692	-2,614	-0,6627	2,5439

Table C.35: MBEGA *vs* MAs respectively endowed with AF^* and BF^* (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
11_Tumors	MBEGA	M :	3,64%	23,22%	74,25%	10439972,6	49
		Sd:	2,03%	3,59%	0,92%	2127251,9073	2,8284
		T-:	-	-	-	-	-
	AF^*	M :	6,59%	29,20%	74,76%	**9120201,7**	**47,3**
		Sd:	2,92%	4,63%	2,54%	1049930,1925	1,8886
		T-:	10,7959	6,6667	4,7293	-7,7077	-6,4765
	BF^*	M :	6,20%	25,89%	**72,23%**	11407949,1818	**48**
		Sd:	1,79%	3,41%	2,41%	1830587,46	3,0984
		T-:	10,8616	3,1581	-10,6729	9,5634	-5
14_Tumors	MBEGA	M :	38,40%	61,93%	85,84%	32426149,2857	47,4286
		Sd:	3,74%	3,48%	1,48%	3598869,1874	4,6853
		T-:	-	-	-	-	-
	AF^*	M :	38,96%	63,45%	88,41%	**27580478,77**	47,1111
		Sd:	2,34%	2,86%	2,76%	2424059,0204	3,0185
		T-:	1,1764	4,2875	7,3383	-6,2887	-0,4061
	BF^*	M :	38,70%	**61,01%**	86,52%	30804245,7	45,4
		Sd:	4,01%	2,73%	1,75%	2975937,6218	3,4059
		T-:	0,6577	-25,1374	5,5852	-1,8629	-2,0175
Brain Tumor2	MBEGA	M :	0,00%	51,60%	59,73%	966079,6667	10,6667
		Sd:	0,00%	15,88%	5,25%	424005,4887	7,0048
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**42,70%**	59,40%	**814440,62**	10,5
		Sd:	0,00%	12,37%	5,40%	259856,6338	5,757
		T-:		-3,1836	-0,2537	-3,6148	-0,1533
	BF^*	M :	0,77%	52,40%	57,36%	1491165,9	26,6
		Sd:	2,43%	10,04%	5,39%	446324,1521	16,0638
		T-:	10	0,5216	-1,9685	7,0847	10,971
Prostate Tumor	MBEGA	M :	0,00%	26,39%	40,11%	2119086,8571	27
		Sd:	0,00%	6,07%	7,89%	420355,8836	14,7309
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**17,98%**	**17,09%**	**1790517,28**	**6**
		Sd:	0,00%	6,64%	9,53%	1003849,2937	0,8165
		T-:		-19,3715	-20,1411	-7,4645	-9,8
	BF^*	M :	0,00%	32,04%	**35,97%**	2265346,14	**13,71**
		Sd:	0,00%	12,99%	11,84%	471051,8344	11,4705
		T-:		4,1676	-3,8694	3,6817	-5,7938
Lymphoma	MBEGA	M :	0,00%	12,29%	14,81%	1328352	2,8571
		Sd:	0,00%	2,56%	2,99%	380575,5339	0,378
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	18,38%	20,15%	1359312,3333	3,0833
		Sd:	0,00%	3,93%	3,53%	223090,9887	1,4434
		T-:		35,7872	27,4982	0,5184	10,4925
	BF^*	M :	0,00%	12,66%	19,39%	1447421	6,5556
		Sd:	0,00%	5,16%	3,48%	302116,4127	5,0525
		T-:		14,1449	12,6359	1,9972	5,1629

Table C.36: MBEGA *vs* MAs respectively endowed with AF^* and BF^* (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	- (only GA)	M :	2,35%	46,15%	49,78%	4647524,7647	36,5294
		Sd:	2,47%	7,42%	4,89%	2030048,161	8,0477
		T-:	-	-	-	-	-
	AF^*	M :	**0,89%**	**42,27%**	**38,75%**	**995287,88**	**4**
		Sd:	1,76%	4,77%	8,64%	460122,2788	1,118
		T-:	-4,0781	-7,8874	-9,1393	-19,2099	-44,1362
	BF^*	M :	**1,78%**	**42,31%**	**47,31%**	**2111715,55**56	**30**
		Sd:	2,11%	7,74%	5,48%	905140,3792	19,685
		T-:	-2,6142	-3,5912	-2,7099	-13,3122	-4,8971
CNS	- (only GA)	M :	0,00%	42,67%	44,27%	2819096,5294	33,3529
		Sd:	0,00%	11,91%	5,84%	1436436,1201	11,7789
		T-:	-	-	-	-	-
	AF^*	M :	0,83%	**38,00%**	**36,67%**	**638983,62**	**12,75**
		Sd:	2,36%	6,74%	9,33%	315675,5265	12,7588
		T-:	1,1429	-4,7509	-15,1962	-20,4229	-11,4478
	BF^*	M :	1,90%	**39,33%**	**42,38%**	**918696,57**	**14**
		Sd:	3,25%	8,36%	3,69%	521110,0648	12,7541
		T-:	7	-5,6237	-3,6979	-12,7087	-16,0509
Colon	- (only GA)	M :	2,21%	26,91%	33,32%	1910513,9412	26,5882
		Sd:	3,08%	5,96%	6,26%	1205013,4761	10,7008
		T-:	-	-	-	-	-
	AF^*	M :	**0,69%**	**22,65%**	**29,46%**	**336776,88**	**3,88**
		Sd:	2,08%	5,99%	7,91%	164966,6779	0,928
		T-:	-10,0118	-19,9908	-3,399	-19,0316	-50,8341
	BF^*	M :	2,34%	25,32%	28,47%	**871693,62**	**3,25**
		Sd:	3,23%	13,70%	12,45%	432682,7494	0,8864
		T-:	0,4303	-0,5495	-1,5865	-12,3716	-52,1572
Leukemia3C	- (only GA)	M :	0,00%	21,23%	28,54%	4875401,7368	38,5263
		Sd:	0,00%	7,26%	7,82%	1918293,1208	7,9398
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**11,98%**	**14,75%**	**1936374,77**	**10,88**
		Sd:	0,00%	8,40%	3,51%	489143,8056	13,2518
		T-:	-	-32,21	-47,3492	-23,907	-61,1321
	BF^*	M :	0,00%	**18,26%**	**21,67%**	**2048525,62**	**9,87**
		Sd:	0,00%	5,26%	7,19%	688241,6941	16,2519
		T-:	-	-11,7409	-6,4217	-20,1786	-32,9477
Leukemia4C	- (only GA)	M :	0,00%	33,54%	40,58%	4838907,0526	41,9474
		Sd:	0,00%	9,16%	7,66%	1583335,4614	5,6517
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**20,49%**	**19,44%**	**1752124,75**	**8**
		Sd:	0,00%	6,07%	3,32%	404180,1855	2,1381
		T-:	-	-25,0289	-46,0981	-25,2517	-68,9899
	BF^*	M :	0,00%	32,36%	**30,62%**	**2208339,87**	**21**
		Sd:	0,00%	9,66%	10,44%	747196,5481	15,8024
		T-:	-	-0,852	-7,5682	-18,6135	-19,2867

Table C.37: GA *vs* MAs respectively endowed with iterative AF^* and BF^* (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	- (only GA)	M :	0,00%	15,32%	26,88%	17129218,9444	46,1111
		Sd:	0,00%	3,74%	1,96%	6676624,4762	3,0076
		T-:	-	-	-	-	-
	AF^*	M :	0,78%	**12,67%**	**21,78%**	**6881492,7**	**26,6**
		Sd:	1,01%	2,49%	1,90%	1960940,9098	12,6069
		T-:	10	-8,0793	-28,5519	-130,2916	-11,2088
	BF^*	M :	0,22%	17,49%	**25,43%**	**6594822,55**	**26,33**
		Sd:	0,65%	5,10%	2,55%	1395639,3074	14,0712
		T-:	9	5,3979	-41,0164	-122,6037	-11,6038
MLL	- (only GA)	M :	0,00%	22,66%	32,72%	4762235,3158	34,6316
		Sd:	0,00%	8,68%	8,22%	2181348,1554	9,7821
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**16,06%**	**14,72%**	**1579315,3**	**4,8**
		Sd:	0,00%	3,39%	3,39%	496630,9341	1,1353
		T-:	-	-9,0176	-34,0991	-31,1027	-362,4013
	BF^*	M :	0,00%	20,16%	**19,21%**	**1392950,71**	**8,14**
		Sd:	0,00%	10,93%	6,41%	284371,5625	4,3753
		T-:	-	-1,8204	-30,5593	-63,2021	-44,7328
Orarian	- (only GA)	M :	0,00%	3,40%	19,12%	12696723,3333	24,5
		Sd:	0,00%	2,36%	10,46%	5157035,8265	6,7584
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**2,13%**	**4,27%**	**3987748,2**	**4,6**
		Sd:	0,00%	0,83%	0,34%	865069,0292	1,075
		T-:	-	-9,2865	-47,3563	-17,4354	-78,6003
	BF^*	M :	0,00%	3,16%	**17,60%**	**6239140,22**	24,5556
		Sd:	0,00%	2,00%	9,47%	1741046,5277	13,5565
		T-:	c	-0,8496	-2,1473	-13,0276	0,187
SRBCT	- (only GA)	M :	0,00%	12,85%	44,99%	7476719,9444	41,7222
		Sd:	0,00%	6,34%	7,32%	3298359,6883	5,9092
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**11,32%**	**36,49%**	**2120388,5**	**8,5**
		Sd:	0,00%	4,96%	6,20%	456188,7659	2,3688
		T-:	-	-4,2222	-7,479	-19,4915	-94,2926
	BF^*	M :	0,00%	18,78%	**42,60%**	**2763555,66**	**7,66**
		Sd:	0,00%	11,52%	7,58%	284305,5357	2,8752
		T-:	-	8,1634	-2,8179	-17,3679	-96,4303
9_Tumors	- (only GA)	M :	20,00%	76,74%	94,32%	6014650,1053	41,1053
		Sd:	3,85%	8,32%	1,85%	2794142,5604	7,5196
		T-:	-	-	-	-	-
	AF^*	M :	**17,78%**	79,04%	**93,48%**	**2273256,77**	**34,55**
		Sd:	7,45%	6,95%	2,26%	486698,7644	17,3789
		T-:	-5,1794	1,7986	-8,0021	-16,5463	-2,0575
	BF^*	M :	**16,67%**	76,25%	94,67%	**2542024,87**	39,75
		Sd:	5,04%	6,27%	2,11%	600032,1476	14,0789
		T-:	-6,1193	-0,3805	1,0402	-15,442	-1,1441
11_Tumors	- (only GA)	M :	5,74%	27,19%	73,43%	22335377,2105	47,5789
		Sd:	1,76%	4,35%	3,54%	8341713,8785	1,8048
		T-:	-	-	-	-	-
	AF^*	M :	7,07%	28,45%	**73,03%**	**7719192,88**	41,8889
		Sd:	4,17%	5,75%	2,72%	2159676,3153	11,5157
		T-:	1,3511	1,2575	-2,2596	-21,5326	-1,9044
	BF^*	M :	7,58%	26,13%	73,49%	**10918564,11**	**46,88**
		Sd:	2,27%	6,39%	3,91%	1500935,5444	3,9511
		T-:	18,0758	-0,773	0,0662	-18,2162	-6,6763

Table C.38: GA *vs* MAs respectively endowed with iterative AF^* and BF^* (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	- (only GA)	M :	39,26%	62,83%	87,26%	74594390,4615	46,3846
		Sd:	3,83%	3,68%	1,11%	17820808,0553	3,9059
		T-:	-	-	-	-	-
	AF^*	M :	38,83%	64,05%	87,90%	**26427243,3**	46,7
		Sd:	3,54%	3,44%	1,48%	2425678,5106	4,4485
		T-:	-1,5582	3,9113	2,5921	-37,6098	0,4163
	BF^*	M :	40,55%	**60,92%**	**86,78%**	**30500610,11**	**43,55**
		Sd:	3,60%	2,70%	1,59%	3420530,2441	5,8119
		T-:	5,989	-7,0853	-2,6665	-34,2139	-3,6679
Brain Tumor2	- (only GA)	M :	0,00%	44,65%	57,31%	3975539,1818	39,6364
		Sd:	0,00%	10,83%	10,01%	1120421,9314	10,8099
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**40,50%**	**52,90%**	**928552,37**	**11,75**
		Sd:	0,00%	11,40%	4,16%	221960,0796	9,6028
		T-:	-	-5,4276	-9,5233	-38,8386	-28,3916
	BF^*	M :	0,96%	45,50%	56,30%	**979041,62**	**10,5**
		Sd:	2,72%	10,85%	4,32%	283275,6847	5,8554
		T-:	8	0,6485	-1,7617	-37,3824	-27,6764
Prostate Tumor	- (only GA)	M :	0,00%	27,48%	35,72%	5427461,75	33
		Sd:	0,00%	5,74%	8,72%	2153667,5087	6,396
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**16,55%**	**15,61%**	**1670356**	**6,9**
		Sd:	0,00%	2,78%	2,34%	500049,5387	1,2867
		T-:	-	-37,3096	-77,1694	-172,8398	-104,3166
	BF^*	M :	0,00%	**26,03%**	**30,54%**	**2110164**	**15,12**
		Sd:	0,00%	11,61%	13,69%	429317,4073	13,6532
		T-:	-	-4,0194	-3,8238	-184,5181	-8,4152
Lymphoma	- (only GA)	M :	0,00%	7,84%	18,42%	2970675,6667	18,2
		Sd:	0,00%	5,55%	3,68%	1320242,23	7,6737
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	16,06%	18,91%	**1282634,1**	**2,9**
		Sd:	0,00%	1,92%	2,52%	203379,0288	0,3162
		T-:	-	98,1332	4,585	-101,1695	-33,7418
	BF^*	M :	0,00%	10,04%	**14,72%**	**1444072,57**	**8,71**
		Sd:	0,00%	5,67%	3,53%	325350,2034	6,6762
		T-:	-	25,4963	-8,7923	-35,7386	-19,3931

Table C.39: GA *vs* MAs respectively endowed with iterative AF^* and BF^* (3)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	MBEGA	M :	2,22%	45,05%	47,45%	1306510,8889	15,3333
		Sd:	2,11%	3,43%	6,95%	657546,5512	12,52
		T-:	-	-	-	-	-
	AF^*	M :	**0,89%**	**42,27%**	**38,75%**	**995287,88**	**4**
		Sd:	1,76%	4,77%	8,64%	460122,2788	1,118
		T-:	-3,3489	-5,0752	-4,4842	-4,7057	-7,65
	BF^*	M :	1,78%	**42,31%**	47,31%	2111715,5556	30
		Sd:	2,11%	7,74%	5,48%	905140,3792	19,685
		T-:	-1,591	-2,4947	-0,0784	11,9858	7,92
CNS	MBEGA	M :	2,22%	41,19%	40,44%	776660,4444	16,4444
		Sd:	3,33%	6,85%	5,35%	579215,0463	16,8605
		T-:	-	-	-	-	-
	AF^*	M :	0,83%	**38,00%**	**36,67%**	638983,625	12,75
		Sd:	2,36%	6,74%	9,33%	315675,5265	12,7588
		T-:	-1,8041	-3,0476	-25,3628	-1,6956	-1,8387
	BF^*	M :	1,90%	**39,33%**	42,38%	918696,5714	14
		Sd:	3,25%	8,36%	3,69%	521110,0648	12,7541
		T-:	-0,864	-2,6762	10,5406	1,0719	-1,6289
Colon	MBEGA	M :	0,78%	23,47%	28,79%	769904,375	10,125
		Sd:	2,21%	3,10%	8,53%	631280,9958	13,174
		T-:	-	-	-	-	-
	AF^*	M :	0,69%	**22,65%**	29,46%	**336776,88**	**3,88**
		Sd:	2,08%	5,99%	7,91%	164966,6779	0,928
		T-:	-0,6975	-2,3842	0,4964	-5,7856	-6,1397
	BF^*	M :	2,34%	25,32%	28,47%	871693,625	**3,25**
		Sd:	3,23%	13,70%	12,45%	432682,7494	0,8864
		T-:	5,0596	0,6405	-0,1025	1,3346	-6,766
Leukemia3C	MBEGA	M :	0,00%	15,56%	15,76%	2073395,375	4,125
		Sd:	0,00%	8,59%	2,74%	445394,4673	0,991
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**11,98%**	14,75%	**1936374,77**	10,88
		Sd:	0,00%	8,40%	3,51%	489143,8056	13,2518
		T-:	-	-3,1902	-1,6319	-3,5151	15,6433
	BF^*	M :	0,00%	18,26%	21,67%	**2048525,62**	9,87
		Sd:	0,00%	5,26%	7,19%	688241,6941	16,2519
		T-:	-	2,4315	4,9125	-0,3202	6,6898
Leukemia4C	MBEGA	M :	0,00%	28,33%	33,89%	2055455	15,1667
		Sd:	0,00%	12,12%	14,65%	748936,9207	12,1559
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**20,49%**	**19,44%**	**1752124,75**	**8**
		Sd:	0,00%	6,07%	3,32%	404180,1855	2,1381
		T-:	-	-4,0939	-6,7682	-9,8806	-3,8165
	BF^*	M :	0,00%	32,36%	30,62%	2208339,875	21
		Sd:	0,00%	9,66%	10,44%	747196,5481	15,8024
		T-:	-	1,7478	-1,3241	1,9782	2,761

Table C.40: MBEGA *vs* MAs respectively endowed with iterative AF^* and BF^* (1)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	MBEGA	M :	0,00%	14,26%	23,43%	6746779,6667	28,3333
		Sd:	0,00%	4,54%	2,32%	1982244,4482	15,397
		T-:	-	-	-	-	-
	AF^*	M :	0,78%	**12,67%**	**21,78%**	6881492,7	26,6
		Sd:	1,01%	2,49%	1,90%	1960940,9098	12,6069
		T-:	10	-6,7586	-8,3021	0,5056	-0,8517
	BF^*	M :	0,22%	17,49%	25,43%	6594822,5556	26,3333
		Sd:	0,65%	5,10%	2,55%	1395639,3074	14,0712
		T-:	9	9,8212	21,2031	-0,5656	-0,9979
MLL	MBEGA	M :	0,00%	20,71%	19,52%	1924337,8571	3,7143
		Sd:	0,00%	7,93%	6,80%	670178,6692	0,488
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**16,06%**	**14,72%**	**1579315,3**	4,8
		Sd:	0,00%	3,39%	3,39%	496630,93	1,13
		T-:	-	-7,9259	-4,9883	-3,5631	12,0889
	BF^*	M :	0,00%	20,16%	19,21%	**1392950,71**	8,14
		Sd:	0,00%	10,93%	6,41%	284371,5625	4,3753
		T-:	-	-0,4263	-0,3457	-12,7185	7,465
Orarian	MBEGA	M :	0,00%	2,22%	4,69%	4376831,2857	2
		Sd:	0,00%	0,56%	0,43%	1098440,9428	0
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	2,13%	**4,27%**	**3987748,2**	4,6
		Sd:	0,00%	0,83%	0,34%	865069,0292	1,075
		T-:	-	-0,5811	-13,0743	-4,7785	65
	BF^*	M :	0,00%	3,16%	17,60%	6239140,2222	24,5556
		Sd:	0,00%	2,00%	9,47%	1741046,5277	13,5565
		T-:	-	3,1736	20,3657	34,9739	140,5385
SRBCT	MBEGA	M :	0,00%	11,29%	32,54%	2253283,4286	5,7143
		Sd:	0,00%	4,43%	5,64%	502583,6223	2,3604
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	11,32%	36,49%	2120388,5	8,5
		Sd:	0,00%	4,96%	6,20%	456188,7659	2,3688
		T-:	-	0,061	3,0256	-1,5557	11,1451
	BF^*	M :	0,00%	18,78%	42,60%	2763555,6667	7,6667
		Sd:	0,00%	11,52%	7,58%	284305,5357	2,8752
		T-:	-	9,6232	9,4849	6,9292	7,7747
9_Tumors	MBEGA	M :	21,11%	75,22%	95,22%	2831195,5	38,6667
		Sd:	2,72%	7,99%	1,66%	514528,9278	17,1075
		T-:	-	-	-	-	-
	AF^*	M :	**17,78%**	79,04%	**93,48%**	**2273256,77**	34,55
		Sd:	7,45%	6,95%	2,26%	486698,7644	17,3789
		T-:	-10,8	2,6213	-7,8089	-17,7342	-1,1634
	BF^*	M :	**16,67%**	76,25%	94,67%	**2542024,87**	39,75
		Sd:	5,04%	6,27%	2,11%	600032,1476	14,0789
		T-:	-9,7473	0,706	-1,4221	-13,9016	0,5589
11_Tumors	MBEGA	M :	3,64%	23,22%	74,25%	10439972,6	49
		Sd:	2,03%	3,59%	0,92%	2127251,9073	2,8284
		T-:	-	-	-	-	-
	AF^*	M :	7,07%	28,45%	**73,03%**	**7719192,88**	**41,88**
		Sd:	4,17%	5,75%	2,72%	2159676,3153	11,5157
		T-:	3,4387	4,0296	-6,9511	-9,959	-2,3749
	BF^*	M :	7,58%	26,13%	73,49%	10918564,1111	**46,88**
		Sd:	2,27%	6,39%	3,91%	1500935,5444	3,9511
		T-:	19,6618	1,8268	-0,8987	5,856	-9,4644

Table C.41: MBEGA *vs* MAs respectively endowed with iterative AF^* and BF^* (2)

Data	Local Search	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	MBEGA	M :	38,40%	61,93%	85,84%	32426149,28	47,42
		Sd:	3,74%	3,48%	1,48%	3598869,18	4,68
		T-:	-	-	-	-	-
	AF^*	M :	38,83%	64,05%	87,90%	**26427243,3**	46,7
		Sd:	3,54%	3,44%	1,48%	2425678,51	4,44
		T-:	0,83	12,37	8,84	-8,09	-0,90
	BF^*	M :	40,55%	**60,92%**	86,78%	**30500610,11**	**43,55**
		Sd:	3,60%	2,70%	1,59%	3420530,24	5,81
		T-:	4,43	-16,94	5,75	-2,54	-4,71
Brain Tumor2	MBEGA	M :	0,00%	51,60%	59,73%	966079,66	10,66
		Sd:	0,00%	15,88%	5,25%	424005,48	7,00
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**40,50%**	**52,90%**	928552,37	11,75
		Sd:	0,00%	11,40%	4,16%	221960,07	9,60
		T-:	-	-6,5238	-5,9131	-2,0107	1,0632
	BF^*	M :	0,96%	**45,50%**	**56,30%**	979041,62	10,5
		Sd:	2,72%	10,85%	4,32%	283275,6847	5,8554
		T-:	8	-3,04	-2,85	0,52	-0,15
Prostate Tumor	MBEGA	M :	0,00%	26,39%	40,11%	2119086,85	27
		Sd:	0,00%	6,07%	7,89%	420355,88	14,73
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	**16,55%**	**15,61%**	**1670356**	6,9
		Sd:	0,00%	2,78%	2,34%	500049,53	1,28
		T-:	-	-26,06	-30,59	-10,1633	-9,3799
	BF^*	M :	0,00%	**26,03%**	**30,54%**	2110164	15,12
		Sd:	0,00%	11,61%	13,69%	429317,4073	13,6532
		T-:	-	-0,82	-6,16	-0,21	-3,94
Lymphoma	MBEGA	M :	0,00%	12,29%	14,81%	1328352	2,85
		Sd:	0,00%	2,56%	2,99%	380575,53	0,37
		T-:	-	-	-	-	-
	AF^*	M :	0,00%	16,06%	18,91%	1282634,1	2,9
		Sd:	0,00%	1,92%	2,52%	203379,02	0,31
		T-:	-	96,27	30,04	-0,84	1,88
	BF^*	M :	0,00%	**10,04%**	14,72%	1444072,57	8,71
		Sd:	0,00%	5,67%	3,53%	325350,20	6,67
		T-:	-	-50,47	-0,20	1,72	31,69

Table C.42: MBEGA *vs* MAs respectively endowed with iterative AF^* and BF^* (3)

Appendix D

Annex IV: Detailed Empirical Results of Chapter IV

Only results of the fourth section of the chapter are reported. Experimental results are organized according to the type LS operator used.

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	IGV	M :	3,60%	36,92%	41,17%	910747,6	2,1
		Sd:	1,26%	1,87%	2,01%	198780,1741	0,3162
		T-:	-	-	-	-	-
	Releif	M :	**1,67%**	39,86%	**40,62%**	**811708,25**	3
		Sd:	2,06%	3,21%	5,08%	174635,3261	1,206
		T-:	-13,3763	20,6611	-1,2457	-3,1985	90
	S.U.	M :	**2,55%**	41,14%	41,25%	**641910,27**	7,27
		Sd:	2,02%	4,48%	4,71%	132212,3444	5,1979
		T-:	-7,6334	7,9446	0,2445	-11,139	10,789
CNS	IGV	M :	0,00%	37,87%	37,87%	257922	2,6
		Sd:	0,00%	4,20%	6,04%	45144,5349	0,5164
		T-:	-	-	-	-	-
	Releif	M :	0,00%	40,27%	38,73%	**213436,65**	2,85
		Sd:	0,00%	4,87%	5,26%	63802,0672	0,7452
		T-:	-	3,692	0,8603	-6,6517	2,1796
	S.U.	M :	0,00%	41,23%	36,46%	**214525,94**	3,21
		Sd:	0,00%	6,77%	5,24%	56784,0274	1,1343
		T-:	-	11,9387	-1,3031	-8,7063	14,7092
Colon	IGV	M :	0,00%	31,74%	33,55%	65576,4	2,6
		Sd:	0,00%	12,16%	6,70%	7768,6605	0,5164
		T-:	-	-	-	-	-
	Releif	M :	0,00%	33,92%	32,94%	**51839,63**	2,6842
		Sd:	0,00%	9,50%	10,17%	16923,5792	0,4776
		T-:	-	9,2094	-0,6075	-16,1292	1,9441
	S.U.	M :	0,00%	30,75%	**30,72%**	**50816,61**	2,72
		Sd:	0,00%	10,01%	10,11%	17702,4633	0,4609
		T-:	-	-6,0048	-3,06	-16,8404	2,8508
Leukemia3C	IGV	M :	0,00%	7,89%	13,61%	277379,2	2,9
		Sd:	0,00%	3,53%	3,40%	48247,0677	0,5676
		T-:	-	-	-	-	-
	Releif	M :	0,00%	9,67%	14,35%	**230605,64**	3,2353
		Sd:	0,00%	3,85%	2,91%	57663,0255	0,7524
		T-:	-	8,7824	1,9978	-4,2941	3,2151
	S.U.	M :	0,00%	8,30%	13,92%	**200753,57**	3,2632
		Sd:	0,00%	3,28%	3,03%	47572,5462	0,8057
		T-:	-	3,9683	0,766	-7,2093	21,2583
Leukemia4C	IGV	M :	5,56%	16,50%	18,56%	273944,8	2,8
		Sd:	0,00%	5,17%	5,55%	79442,4958	0,4216
		T-:	-	-	-	-	-
	Releif	M :	5,56%	**15,11%**	19,52%	**216239,6**	2,9333
		Sd:	0,00%	4,74%	5,97%	48174,6986	0,2582
		T-:	-	-2,0806	3,5931	-17,3263	6,5079
	S.U.	M :	5,56%	16,97%	19,61%	**212549,8**	3
		Sd:	0,00%	5,43%	5,15%	53496,8424	0,6489
		T-:	-	0,6426	5,2936	-15,0689	10

Table D.1: GRASP-FS based on MB operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU	# Attrib
Lung	IGV	M :	5,23%	18,44%	21,54%	496807,4444	6,4444
		Sd:	0,98%	2,02%	1,43%	82916,1851	3,9721
		T-:	-	-	-	-	-
	Releif	M :	**4,36%**	**18,12%**	22,16%	502481,33	8,38
		Sd:	0,84%	2,62%	1,19%	85618,0567	3,0705
		T-:	-7,8095	-3,3745	5,2554	0,2698	7,8597
	S.U.	M :	**4,38%**	18,17%	21,91%	480878,64	8,88
		Sd:	0,86%	2,31%	1,78%	77253,4546	4,0756
		T-:	-7,4032	-2,7768	3,0501	-0,769	8,734
MLL	IGV	M :	0,00%	17,71%	19,65%	531237,625	3,375
		Sd:	0,00%	5,13%	4,42%	75626,1931	0,5175
		T-:	-	-	-	-	-
	Releif	M :	0,00%	17,95%	19,18%	**400824,08**	3,56
		Sd:	0,00%	4,03%	5,39%	100503,2611	0,5069
		T-:	-	1,5618	-1,1857	-9,8286	3,7635
	S.U.	M :	0,00%	19,29%	20,12%	373369,8333	3,6111
		Sd:	0,00%	3,37%	5,53%	71851,595	0,5016
		T-:	-	7,2089	1,1138	-11,9664	4,5745
ovarian	IGV	M :	0,00%	2,16%	4,54%	647397,625	2
		Sd:	0,00%	0,63%	0,47%	65609,9882	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	3,12%	4,79%	**523250,82**	2
		Sd:	0,00%	1,12%	0,56%	93892,0503	0
		T-:	-	21,182	5,2228	-7,7556	-
	S.U.	M :	0,00%	3,02%	4,83%	**519318,46**	2
		Sd:	0,00%	1,05%	0,52%	101626,567	0
		T-:	-	7,2883	5,9254	-7,7638	-
SRBCT	IGV	M :	0,00%	13,11%	34,39%	74685	3,625
		Sd:	0,00%	5,79%	5,61%	14441,5503	0,9161
		T-:	-	-	-	-	-
	Releif	M :	0,00%	14,89%	**32,54%**	**69208**	**3,17**
		Sd:	0,00%	4,13%	7,46%	20701,0933	0,393
		T-:	-	1,5344	-3,1667	-6,1615	-2,6049
	S.U.	M :	0,00%	12,55%	33,40%	**73012,26**	3,47
		Sd:	0,00%	4,36%	6,80%	18360,9135	0,6118
		T-:	-	-0,4823	-1,6738	-2,7089	-0,8713
9_Tumors	IGV	M :	33,33%	85,67%	94,25%	259434,25	4,75
		Sd:	0,00%	4,74%	2,05%	37175,5502	1,5811
		T-:	-	-	-	-	-
	Releif	M :	**31,43%**	**83,67%**	94,24%	**199190,5**	6,5714
		Sd:	4,07%	6,44%	1,62%	43072,1015	4,7993
		T-:	-14	-6,3487	-0,1028	-10,4653	4,3216
	S.U.	M :	**28,10%**	**81,43%**	94,10%	**218275,78**	13,42
		Sd:	5,35%	4,95%	1,65%	52055,9003	8,6355
		T-:	-51,3333	-5,8744	-1,0692	-7,1319	10,1214
11_Tumors	IGV	M :	21,14%	34,64%	69,29%	513053,55	9,8
		Sd:	2,10%	4,99%	1,74%	111159,8372	4,0601
		T-:	-	-	-	-	-
	Releif	M :	**18,02%**	**29,98%**	69,79%	520370,76	13,95
		Sd:	3,11%	4,64%	1,73%	94289,5309	6,1642
		T-:	-24,0642	-24,8452	5,1275	0,9168	12,43
	S.U.	M :	**16,79%**	**28,50%**	69,65%	558506,11	17,18
		Sd:	4,11%	6,83%	1,90%	119276,536	8,3701
		T-:	-40,6351	-31,7976	3,6703	5,6964	28,0227

Table D.2: GRASP-FS based on MB operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	61,22%	74,08%	84,10%	1055734,8571	6,7143
		Sd:	0,90%	1,85%	1,43%	144126,9018	2,4976
		T-:	-	-	-	-	-
	Releif	M :	**59,01%**	**71,26%**	84,75%	1184131,25	9,875
		Sd:	2,60%	3,94%	1,20%	204848,6502	4,3493
		T-:	-41,9518	-17,5103	2,3622	4,8767	8,8887
	S.U.	M :	**56,98%**	**69,02%**	84,75%	1285658,93	13,56
		Sd:	2,68%	4,44%	1,30%	171195,2231	5,8988
		T-:	-55,984	-15,1341	2,5741	9,9768	16,3485
Brain Tumor2	IGV	M :	0,00%	44,70%	56,00%	321544,375	3
		Sd:	0,00%	8,43%	3,28%	67188,1366	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	44,00%	**54,20%**	**280183,68**	2,93
		Sd:	0,00%	9,54%	4,75%	65206,9985	0,25
		T-:	-	-0,8178	-2,5677	-3,3983	-1,0667
	S.U.	M :	0,00%	**43,29%**	54,58%	**293873,83**	3
		Sd:	0,00%	8,94%	4,97%	55972,8705	0
		T-:	-	-2,5758	-1,8011	-2,3054	-
Prostate Tumor	IGV	M :	0,00%	14,36%	17,50%	343600,875	2,5
		Sd:	0,00%	2,99%	2,52%	87981,2744	0,5345
		T-:	-	-	-	-	-
	Releif	M :	0,00%	14,00%	**17,04%**	**288185,23**	2,92
		Sd:	0,00%	2,98%	2,50%	76036,3787	0,2774
		T-:	-	-0,6124	-2,3691	-4,4031	6,7391
	S.U.	M :	0,00%	14,31%	**16,42%**	323304,68	2,93
		Sd:	0,00%	2,45%	2,36%	74430,6706	0,25
		T-:	-	-0,0941	-4,1106	-1,8168	6,9864
Lymphoma	IGV	M :	0,00%	16,06%	17,35%	130114,375	2
		Sd:	0,00%	2,53%	6,95%	23161,9997	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**15,51%**	16,61%	**110298,52**	2,11
		Sd:	0,00%	2,13%	5,15%	27025,8211	0,3321
		T-:	-	-2,6567	-1,0267	-8,1432	2,2667
	S.U.	M :	0,00%	16,26%	17,25%	**92596,47**	2,1765
		Sd:	0,00%	2,21%	4,24%	24579,465	0,393
		T-:	-	1,0108	-0,1314	-22,1042	3,6429
Challenge 2004	IGV	M :	72,83%	96,72%	98,33%	1390955,375	3,875
		Sd:	2,01%	4,25%	1,08%	223913,5908	0,8345
		T-:	-	-	-	-	-
	Releif	M :	**70,65%**	**95,26%**	98,85%	**1365976**	6,25
		Sd:	4,20%	4,49%	1,11%	214262,2442	4,7887
		T-:	-7,1554	-2,5704	5,2063	-4,9306	12,6229
	S.U.	M :	**68,12%**	**93,26%**	98,96%	1400756,5833	13,1667
		Sd:	5,02%	5,17%	0,58%	216638,642	11,2882
		T-:	-9,3858	-4,9807	6,4246	1,5303	12,1611

Table D.3: GRASP-FS based on MB operator (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	IGV	M :	5,33%	42,69%	44,12%	423499,44	2,66
		Sd:	3,46%	8,63%	6,08%	87951,603	0,7071
		T-:	-	-	-	-	-
	Releif	M :	**4,27%**	41,67%	43,97%	1144895,4	6,8
		Sd:	3,53%	5,50%	5,89%	1967813,9708	9,2752
		T-:	-2,7565	-1,4625	-0,4897	17,8946	105,003
	S.U.	M :	**4,00%**	44,48%	47,50%	1583920,75	9,25
		Sd:	3,02%	5,84%	3,99%	2675452,5964	12,4871
		T-:	-2,2941	2,6822	17,7568	9,0512	23,207
CNS	IGV	M :	0,00%	36,83%	41,50%	68493,5	3,25
		Sd:	0,00%	7,41%	2,98%	38010,6904	0,4629
		T-:	-	-	-	-	-
	Releif	M :	0,56%	39,78%	43,06%	78598,25	6,6667
		Sd:	1,92%	5,41%	6,29%	72756,4842	8,0829
		T-:	12	1,6779	6,5879	3,1371	14,166
	S.U.	M :	0,61%	40,36%	**39,88%**	99394,90	8,81
		Sd:	2,01%	8,24%	7,16%	77629,7462	9,185
		T-:	11	1,5415	-8,6316	5,3239	12,4307
Colon	IGV	M :	0,00%	33,47%	34,52%	21787,75	2,25
		Sd:	0,00%	11,28%	8,93%	3609,7527	0,4629
		T-:	-	-	-	-	-
	Releif	M :	0,48%	**30,32%**	**33,50%**	21661,53	3,07
		Sd:	1,73%	9,93%	12,03%	5529,2253	1,0377
		T-:	13	-4,2643	-5,2224	-1,869	9,3397
	S.U.	M :	0,57%	35,25%	36,54%	23020,5455	2,7273
		Sd:	1,88%	7,41%	10,85%	5340,2747	0,9045
		T-:	11	1,9942	13,104	10,0561	6,5265
Leukemia3C	IGV	M :	0,00%	8,77%	14,63%	65137,3333	4,6667
		Sd:	0,00%	7,29%	6,95%	18304,6667	1,2247
		T-:	-	-	-	-	-
	Releif	M :	0,00%	8,33%	14,03%	66645,4167	4,8333
		Sd:	0,00%	6,59%	5,50%	15771,9492	1,4668
		T-:	-	-0,2235	-1,0086	0,713	1,3636
	S.U.	M :	0,00%	10,31%	14,81%	70605,2222	4,1111
		Sd:	0,00%	8,20%	4,54%	16976,6548	0,928
		T-:	-	0,6343	0,3855	2,7064	-7,398
Leukemia4C	IGV	M :	5,56%	21,32%	24,38%	70015,625	4,5
		Sd:	0,00%	4,69%	4,56%	15811,1345	2,0702
		T-:	-	-	-	-	-
	Releif	M :	5,56%	20,83%	26,85%	**61495,16**	6,25
		Sd:	0,00%	3,54%	4,80%	14851,2926	2,9886
		T-:	-	-1,0311	5,4066	-4,5173	8,1588
	S.U.	M :	5,98%	21,67%	28,72%	71081,8462	6,6154
		Sd:	1,54%	3,33%	6,29%	23541,8264	4,2336
		T-:	13	0,7701	11,8379	0,4402	9,4043

Table D.4: GRASP-FS based on MB^+ operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib
Lung	IGV	M :	3,27%	17,34%	22,55%	306600	6
		Sd:	1,70%	5,53%	2,06%	57184,693	2
		T-:	-	-	-	-	-
	Releif	M :	3,10%	18,02%	23,05%	321722,5	6,75
		Sd:	1,77%	3,37%	1,99%	81509,9117	2,8324
		T-:	-0,6081	0,5924	5,3024	11,621	3,2489
	S.U.	M :	3,62%	17,50%	23,15%	358199,2308	9,5385
		Sd:	1,09%	4,18%	1,42%	92485,962	4,1756
		T-:	4,6162	0,1207	4,4243	15,3708	7,364
MLL	IGV	M :	0,00%	12,22%	17,43%	130675	4,125
		Sd:	0,00%	3,93%	4,21%	41884,5966	1,3562
		T-:	-	-	-	-	-
	Releif	M :	0,00%	16,98%	20,25%	224524,3333	10,8889
		Sd:	0,00%	4,86%	8,56%	88324,2561	5,9465
		T-:	-	13,1937	9,8713	6,3196	7,6191
	S.U.	M :	0,00%	13,38%	17,88%	235274,4545	7,6364
		Sd:	0,00%	4,95%	3,49%	219298,3993	5,5186
		T-:	-	2,3287	2,196	6,7713	7,9026
Ovarian	IGV	M :	0,00%	2,70%	4,67%	289944,8889	2,3333
		Sd:	0,00%	0,43%	0,47%	81162,7247	0,5
		T-:	-	-	-	-	-
	Releif	M :	0,00%	2,79%	**4,44%**	263548,2222	2,4444
		Sd:	0,00%	0,42%	0,61%	56475,7135	0,527
		T-:	-	1,9565	-3,2987	-1,663	1,5435
	S.U.	M :	0,00%	2,68%	4,48%	276473,4444	2,4444
		Sd:	0,00%	0,42%	0,70%	87399,4587	0,527
		T-:	-	-0,2432	-3,8831	-0,7425	1,8
SRBCT	IGV	M :	0,00%	15,12%	34,25%	26115,3333	3,2222
		Sd:	0,00%	6,35%	11,48%	2161,6044	0,441
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**11,44%**	**30,55%**	31545,4545	6
		Sd:	0,00%	6,13%	8,17%	12510,638	3,5777
		T-:	-	-2,2233	-4,0916	1,6315	7,6213
	S.U.	M :	0,00%	11,51%	33,85%	28977,1	8,6
		Sd:	0,00%	5,42%	10,22%	5334,4187	8,9343
		T-:	-	-2,3538	-0,4389	6,1149	9,5939
9_Tumors	IGV	M :	25,00%	82,33%	94,08%	91812,125	5,375
		Sd:	6,90%	7,37%	1,31%	27626,1218	1,5059
		T-:	-	-	-	-	-
	Releif	M :	24,67%	82,60%	94,93%	97947,7	5,5
		Sd:	4,50%	6,73%	2,04%	34075,2635	1,9003
		T-:	-0,4273	0,4637	4,3134	0,7697	0,5975
	S.U.	M :	24,00%	84,40%	94,80%	132098,8	7
		Sd:	6,44%	8,59%	1,66%	81073,3564	4,6428
		T-:	-1,4716	3,5196	5,815	7,8248	8
11_Tumors	IGV	M :	24,40%	42,93%	70,99%	366561,7895	5,4737
		Sd:	6,42%	7,92%	3,14%	121331,7967	1,467
		T-:	-	-	-	-	-
	Releif	M :	**20,98%**	42,47%	72,68%	520135,4333	9,0333
		Sd:	5,43%	8,10%	3,28%	363634,9341	7,1847
		T-:	-32,3724	-0,8893	8,8645	18,9016	34,5585
	S.U.	M :	**20,00%**	**39,67%**	72,49%	550855,3	10,75
		Sd:	5,45%	10,70%	3,52%	322024,6297	8,1103
		T-:	-19,7637	-2,9976	6,1379	23,9132	11,2511

Table D.5: GRASP-FS based on MB^+ operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	58,28%	74,77%	85,71%	1018573	7
		Sd:	3,06%	4,02%	1,82%	293844,1716	1,5119
		T-:	-	-	-	-	-
	Releif	M :	**58,04%**	**73,63%**	85,85%	1015016,2308	8,3846
		Sd:	5,01%	4,33%	1,51%	300371,5098	4,9081
		T-:	-6,4417	-3,8006	1,2818	-0,0793	10,7791
	S.U.	M :	**55,45%**	**71,29%**	85,84%	1323763,2	10,4
		Sd:	3,57%	2,79%	1,29%	383480,2002	5,7581
		T-:	-9,4348	-11,1963	1,1094	8,0159	12,5646
Brain Tumor2	IGV	M :	1,92%	49,40%	60,10%	110951,5	4,5
		Sd:	3,56%	10,17%	2,79%	32699,7714	1,6903
		T-:	-	-	-	-	-
	Releif	M :	0,70%	50,25%	59,35%	144247,1818	7
		Sd:	2,32%	13,34%	5,48%	48372,1784	4,1473
		T-:	-1,6904	0,4618	-0,9388	4,5264	11,9975
	S.U.	M :	0,77%	50,96%	58,64%	119221,2	6,7
		Sd:	2,43%	13,52%	5,49%	44836,1484	3,9172
		T-:	-1,591	0,813	-1,8092	0,782	9,6424
Prostate Tumor	IGV	M :	0,00%	22,11%	24,85%	139016,875	4,25
		Sd:	0,00%	8,74%	9,04%	49405,5807	1,669
		T-:	-	-	-	-	-
	Releif	M :	0,70%	**20,57%**	26,99%	182226	5
		Sd:	1,56%	6,59%	8,89%	82938,4914	2,7203
		T-:	11	-3,4522	1,614	7,6517	1,7092
	S.U.	M :	0,48%	22,21%	26,23%	171086,875	5,625
		Sd:	1,36%	8,39%	8,32%	75058,5167	2,5036
		T-:	8	0,1551	0,905	2,0088	3,4437
Lymphoma	IGV	M :	0,00%	18,01%	21,13%	32735	1,7143
		Sd:	0,00%	2,81%	3,20%	4659,5398	0,9512
		T-:	-	-	-	-	-
	Releif	M :	0,53%	**12,89%**	**18,07%**	50436,0909	6,8182
		Sd:	1,77%	6,47%	4,91%	24359,2239	5,0758
		T-:	11	-54,7192	-13,1364	6,2693	40,4197
	S.U.	M :	0,00%	18,48%	22,65%	32802,875	1,875
		Sd:	0,00%	5,01%	2,73%	4218,4935	0,991
		T-:	-	3,6807	10,6213	0,1579	1,0744
Challenge 2004	IGV	M :	73,29%	98,16%	99,30%	2244314,7143	3,5714
		Sd:	4,65%	2,47%	0,62%	1226970,9116	1,3973
		T-:	-	-	-	-	-
	Releif	M :	**70,65%**	**96,89%**	**98,94%**	5096618	10,875
		Sd:	5,57%	2,55%	0,71%	7260376,5843	15,4313
		T-:	-2,1531	-2,5298	-4,2232	8,7595	14,9593
	S.U.	M :	73,04%	97,51%	**98,93%**	5076977,3	8,8
		Sd:	4,00%	3,31%	0,60%	6541440,1996	13,1386
		T-:	-1,9998	-1,1616	-3,2685	7,8319	8,965

Table D.6: GRASP-FS based on MB^{+} operator (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU(ms)	# Attrib.
Breast	IGV	M :	13,82%	44,66%	49,17%	812317,2727	10,4545
		Sd:	6,54%	7,96%	5,56%	555866,3638	5,6101
		T-:	-	-	-	-	-
	Releif	M :	**7,16%**	**43,55%**	49,41%	1880707,6316	22,1579
		Sd:	5,75%	7,85%	3,53%	749275,5214	7,4853
		T-:	-20,624	-1,0751	0,3979	20,29	14,7258
	S.U.	M :	**6,11%**	44,01%	50,50%	1747183,0526	19,3684
		Sd:	5,39%	8,17%	4,16%	730412,7556	8,3614
		T-:	-34,7421	-0,6065	1,9461	18,2411	11,5868
CNS	IGV	M :	6,11%	44,83%	41,22%	67475	9,4167
		Sd:	6,00%	8,95%	6,29%	21706,1099	5,6962
		T-:	-	-	-	-	-
	Releif	M :	5,26%	43,44%	42,95%	81314,1053	14,5263
		Sd:	5,25%	7,39%	5,91%	27079,5136	7,7199
		T-:	-1,2799	-1,1714	2,7793	6,2713	6,9746
	S.U.	M :	**2,67%**	42,58%	42,71%	88496,7333	14,7333
		Sd:	4,22%	9,40%	6,03%	32360,2494	6,1582
		T-:	-5,4886	-1,6088	2,3947	9,3743	6,7703
Colon	IGV	M :	9,38%	33,48%	38,00%	25228,4	9
		Sd:	5,31%	7,84%	6,78%	6809,4706	4,6904
		T-:	-	-	-	-	-
	Releif	M :	**8,46%**	**26,94%**	**34,50%**	26107,9412	12,4118
		Sd:	5,39%	6,53%	7,36%	4159,6665	9,4344
		T-:	-2,7163	-6,4925	-4,7773	0,88	4,8501
	S.U.	M :	7,07%	26,65%	32,99%	27781,2174	17,3043
		Sd:	5,43%	5,17%	7,92%	5266,3546	10,3108
		T-:	-7,3442	-9,3702	-6,4164	2,9913	15,3763
Leukemia3C	IGV	M :	12,22%	35,17%	28,39%	104000,1	11,3
		Sd:	6,31%	8,98%	15,34%	33294,7675	8,92
		T-:	-	-	-	-	-
	Releif	M :	11,11%	36,48%	31,45%	100385,4444	21
		Sd:	7,13%	11,92%	15,02%	36226,917	10,5663
		T-:	-0,9455	2,2495	3,054	-1,4694	4,4942
	S.U.	M :	**8,47%**	**28,15%**	30,45%	120081,4762	19,8571
		Sd:	7,58%	17,16%	13,71%	46478,4219	10,5085
		T-:	-3,7211	-7,906	2,0456	7,5171	4,1325
Leukemia4C	IGV	M :	14,14%	39,95%	43,89%	94056,9091	14,1818
		Sd:	5,19%	11,28%	9,16%	42217,4803	7,4809
		T-:	-	-	-	-	-
	Releif	M :	**12,09%**	43,43%	45,03%	102183,8235	19,8824
		Sd:	4,49%	10,03%	7,49%	38840,7571	6,818
		T-:	-5,795	1,9361	1,299	1,8655	5,1608
	S.U.	M :	**10,46%**	42,75%	44,25%	93182,2941	17,1765
		Sd:	4,34%	8,16%	6,93%	40620,2916	7,0199
		T-:	-9,995	1,9827	0,5074	-0,2185	3,9294

Table D.7: GRASP-FS based on $IFLS^+$ operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	IGV	M :	7,35%	22,57%	26,17%	484326,9167	10,6667
		Sd:	4,96%	4,57%	2,47%	206322,4853	5,3144
		T-:	-	-	-	-	-
	Releif	M :	**6,10%**	**20,86%**	27,02%	499867,1111	16,6111
		Sd:	3,68%	4,82%	2,80%	158526,8832	5,5321
		T-:	-2,4817	-3,5091	10,5896	0,7894	15,978
	S.U.	M :	**5,39%**	**20,62%**	27,17%	470521,05	15,55
		Sd:	2,19%	5,56%	3,02%	174750,7753	5,2363
		T-:	-4,3572	-4,3593	11,8038	-0,7514	13,4969
MLL	IGV	M :	10,10%	29,44%	33,64%	164992,9091	7,6364
		Sd:	6,95%	11,36%	8,21%	70414,8125	3,2333
		T-:	-	-	-	-	-
	Releif	M :	**4,32%**	**26,08%**	34,04%	333334,7222	22,3889
		Sd:	5,89%	9,37%	6,64%	167997,819	9,8048
		T-:	-14,8876	-3,184	1,6106	13,5189	12,7541
	S.U.	M :	**5,82%**	**25,58%**	33,33%	265968,7143	17,4286
		Sd:	7,13%	8,08%	7,27%	120861,9155	9,3839
		T-:	-15,9637	-3,8212	-1,1611	12,6626	12,1483
Ovarian	IGV	M :	0,57%	3,46%	5,79%	468285,2727	14
		Sd:	0,79%	1,05%	2,36%	100279,8588	6,9714
		T-:	-	-	-	-	-
	Releif	M :	0,64%	**3,00%**	**5,28%**	544126,1364	24,6818
		Sd:	0,92%	1,97%	2,04%	203380,944	11,692
		T-:	1,1984	-20,5641	-5,2562	5,5244	7,608
	S.U.	M :	0,46%	3,70%	**5,31%**	483809,8824	23,7647
		Sd:	0,92%	1,85%	2,33%	161086,3348	8,0973
		T-:	-1,8632	18,9469	-4,8029	1,1826	6,1857
SRBCT	IGV	M :	10,95%	31,56%	50,54%	37839	16,1
		Sd:	6,37%	11,02%	10,46%	10764,0703	9,0609
		T-:	-	-	-	-	-
	Releif	M :	12,96%	**25,01%**	**47,02%**	43744,7778	26,1667
		Sd:	8,92%	12,58%	11,11%	11514,7425	11,5619
		T-:	5,8903	-11,4338	-17,4974	4,1111	7,7789
	S.U.	M :	**8,73%**	**21,76%**	**46,04%**	43961,5556	25,3333
		Sd:	8,21%	13,22%	8,97%	11414,6533	10,9705
		T-:	-2,4647	-16,3818	-6,7717	4,1762	8,0365
9_Tumors	IGV	M :	46,06%	81,94%	93,70%	94663,9091	15,4545
		Sd:	6,96%	7,96%	2,07%	36165,4482	5,6809
		T-:	-	-	-	-	-
	Releif	M :	**41,59%**	81,94%	93,94%	111682,2857	20,381
		Sd:	9,41%	6,50%	1,90%	42924,878	8,4171
		T-:	-18,0324	-0,0047	8,6433	4,5828	19,1177
	S.U.	M :	**38,89%**	**80,52%**	93,78%	115205,4444	22,1111
		Sd:	8,00%	7,33%	2,03%	47660,0852	9,591
		T-:	-16,4641	-2,5695	0,7881	5,8871	14,5389
11_Tumors	IGV	M :	43,39%	59,71%	77,47%	445363,4545	9,3636
		Sd:	4,48%	5,69%	2,65%	155572,5673	3,3845
		T-:	-	-	-	-	-
	Releif	M :	**35,86%**	**48,45%**	78,03%	738540,1111	22,2222
		Sd:	10,47%	13,00%	2,83%	396266,8658	14,0233
		T-:	-7,6122	-12,2371	1,805	10,5071	13,5112
	S.U.	M :	**26,87%**	**40,26%**	**75,55%**	850228,7059	23,7059
		Sd:	11,48%	13,59%	3,92%	389367,9486	10,4568
		T-:	-11,1049	-13,6569	-7,513	11,5216	13,2128

Table D.8: GRASP-FS based on $IFLS^+$ operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	67,01%	78,78%	88,27%	1143892,4	10,1
		Sd:	2,82%	3,51%	2,09%	479642,7669	6,4885
		T-:	-	-	-	-	-
	Releif	M :	**59,94%**	**71,53%**	88,72%	1533576,05	20,95
		Sd:	6,45%	5,30%	1,45%	745159,4704	11,9361
		T-:	-11,8114	-8,9637	6,4944	8,9843	13,3241
	S.U.	M :	**58,28%**	**72,90%**	**87,82%**	1482310,875	17,3125
		Sd:	7,92%	4,75%	1,63%	605066,8498	6,0522
		T-:	-9,6361	-7,3153	-10,2109	7,2397	7,3503
Brain Tumor2	IGV	M :	21,37%	50,13%	62,40%	115419,3333	10,3333
		Sd:	9,25%	10,70%	7,45%	29205,6801	4,5826
		T-:	-	-	-	-	-
	Releif	M :	**14,53%**	54,09%	**60,84%**	170388,1111	13,7778
		Sd:	12,62%	9,40%	4,64%	75049,7691	7,7198
		T-:	-16,0997	7,7161	-2,5	16,1516	21,1358
	S.U.	M :	**11,92%**	54,60%	62,28%	149384,3	11,9
		Sd:	10,73%	8,77%	5,03%	63327,3136	5,2103
		T-:	-16,0307	9,4198	-0,1752	11,3637	7,5575
Prostate Tumor	IGV	M :	1,92%	23,02%	27,18%	119975,7	6,9
		Sd:	2,72%	7,70%	9,95%	42523,2976	3,755
		T-:	-	-	-	-	-
	Releif	M :	2,31%	29,75%	38,93%	229029,7333	23,4
		Sd:	3,18%	7,44%	9,47%	127960,4154	11,831
		T-:	1,5617	13,6172	16,9567	20,2065	16,5708
	S.U.	M :	**0,85%**	28,24%	35,64%	197835,8333	17,7222
		Sd:	1,65%	8,05%	12,04%	100082,7482	11,0337
		T-:	-5,3936	22,3956	13,1219	7,1677	25,259
Lymphoma	IGV	M :	3,21%	12,62%	19,61%	50192	9,2727
		Sd:	3,07%	8,10%	4,45%	6069,1317	8,2109
		T-:	-	-	-	-	-
	Releif	M :	2,94%	13,30%	20,64%	52566,1111	18,3333
		Sd:	4,16%	7,28%	5,82%	13064,5809	9,5116
		T-:	-0,9129	0,7645	3,7252	4,0874	16,0785
	S.U.	M :	2,81%	**11,62%**	20,40%	49883,0435	13,5652
		Sd:	3,00%	6,85%	6,12%	12132,5372	7,6386
		T-:	-1,4254	-2,4936	5,2225	-1,5275	6,7641
Challenge 2004	IGV	M :	85,65%	98,62%	99,11%	1253633,6	8,4
		Sd:	2,93%	2,24%	0,81%	480393,6022	4,2216
		T-:	-	-	-	-	-
	Releif	M :	**84,62%**	98,43%	99,32%	1235476,2308	13,3846
		Sd:	5,78%	1,59%	0,56%	569946,6859	8,5102
		T-:	-2,9318	-0,9481	1,5278	-0,1913	4,3878
	S.U.	M :	**80,56%**	**97,46%**	99,01%	1435311,4118	11,5294
		Sd:	7,70%	3,18%	0,66%	867622,1073	7,6167
		T-:	-15,5509	-5,7034	-0,6895	3,1452	5,0126

Table D.9: GRASP-FS based on $IFLS^+$ operator (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	IGV	M :	3,60%	37,83%	40,65%	274448,85	2,15
		Sd:	1,23%	2,19%	3,18%	58837,5114	0,4894
		T-:	-	-	-	-	-
	Releif	M :	**2,78%**	38,80%	42,65%	**231831,86**	3,2222
		Sd:	1,87%	3,84%	4,53%	64348,0028	2,3067
		T-:	-10,3151	24,2692	5,4411	-20,5419	21,4663
	S.U.	M :	**2,87%**	39,88%	42,62%	**233754,43**	4,4359
		Sd:	1,82%	5,01%	4,24%	60434,7933	4,5526
		T-:	-9,5435	12,1429	14,342	-19,0769	36,3374
CNS	IGV	M :	0,00%	39,40%	40,10%	55841,5	3,05
		Sd:	0,00%	5,65%	4,22%	8892,7351	0,9445
		T-:	-	-	-	-	-
	Releif	M :	0,00%	40,27%	39,62%	**45994,97**	3,25
		Sd:	0,00%	5,23%	5,50%	12321,2464	1,1634
		T-:	-	15,1961	-1,6772	-11,9859	29,3451
	S.U.	M :	0,00%	40,60%	**38,46%**	**46774,44**	3,26
		Sd:	0,00%	5,06%	6,09%	11934,1059	1,1073
		T-:	-	19,1016	-5,1668	-8,9842	28,9513
Colon	IGV	M :	0,00%	30,39%	29,45%	19833,4	2,55
		Sd:	0,00%	10,22%	5,77%	2076,1996	0,5104
		T-:	-	-	-	-	-
	Releif	M :	0,00%	32,10%	**27,31%**	**17568,52**	2,5
		Sd:	0,00%	8,05%	6,21%	2796,1855	0,5061
		T-:	-	14,3924	-7,517	-26,5406	-1,9642
	S.U.	M :	0,00%	33,67%	**28,53%**	**16858,72**	**2,48**
		Sd:	0,00%	8,14%	7,29%	2059,4176	0,5588
		T-:	-	18,8344	-3,022	-38,0486	-2,4025
Leukemia3C	IGV	M :	0,00%	9,47%	14,69%	61589,65	2,6
		Sd:	0,00%	2,85%	2,75%	11642,8988	0,8208
		T-:	-	-	-	-	-
	Releif	M :	0,00%	9,98%	**14,24%**	**50707,05**	2,9167
		Sd:	0,00%	4,14%	3,52%	14518,9582	0,9673
		T-:	-	7,0069	-3,2017	-24,7946	9,7802
	S.U.	M :	0,00%	10,03%	14,43%	**48417,35**	3,325
		Sd:	0,00%	4,69%	3,34%	10838,4666	1,0952
		T-:		7,8277	-1,9452	-36,0486	18,7365
Leukemia4C	IGV	M :	5,56%	20,17%	21,78%	65492,1	2,85
		Sd:	0,00%	5,67%	4,92%	14458,6294	0,6708
		T-:	-	-	-	-	-
	Releif	M :	5,56%	**18,80%**	**20,46%**	**53464,52**	2,80
		Sd:	0,00%	7,11%	4,98%	15136,5831	0,8218
		T-:	-	-7,1673	-6,3855	-31,2005	-0,7203
	S.U.	M :	5,56%	**19,28%**	**20,83%**	**53234,47**	3,11
		Sd:	0,00%	7,17%	5,17%	15149,77	0,9454
		T-:	-	-6,0248	-5,8548	-39,3428	4,1964

Table D.10: GRASP-FS based on MB^* operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	IGV	M :	5,23%	18,01%	20,85%	152316,8889	6,4444
		Sd:	0,95%	1,83%	1,66%	30949,9313	3,3294
		T-:	-	-	-	-	-
	Releif	M :	**4,17%**	**17,56%**	21,09%	**133368,09**	10,90
		Sd:	0,66%	3,50%	1,92%	39518,4522	4,1063
		T-:	-28,6202	-4,5777	7,1858	-12,4947	25,0293
	S.U.	M :	**4,17%**	**16,86%**	21,39%	**132121,5**	11,95
		Sd:	0,66%	2,50%	1,64%	30341,2352	5,43
		T-:	-18,8991	-43,8588	27,7354	-11,7812	28,7668
MLL	IGV	M :	0,00%	23,14%	22,03%	120993,1	3,1
		Sd:	0,00%	6,40%	5,61%	29803,9131	0,4472
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**20,68%**	**19,68%**	**98901,05**	3,51
		Sd:	0,00%	5,94%	6,12%	28423,7364	0,6122
		T-:	-	-10,553	-35,4852	-38,0586	28,0856
	S.U.	M :	0,00%	**20,69%**	**19,97%**	**99334,56**	3,58
		Sd:	0,00%	5,67%	4,78%	30162,6533	0,5906
		T-:	-	-10,5258	-31,4364	-37,3611	43,023
Ovarian	IGV	M :	0,00%	2,47%	4,32%	168058,3	2
		Sd:	0,00%	1,30%	0,54%	28853,9572	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	2,73%	4,39%	**144516,24**	2
		Sd:	0,00%	1,36%	0,74%	36805,6457	0
		T-:	-	3,0451	5,0156	-20,3821	-
	S.U.	M :	0,00%	2,46%	4,30%	**127554,75**	2
		Sd:	0,00%	1,22%	0,60%	31110,4437	0
		T-:	-	-0,0524	-1,2959	-26,8647	-
SRBCT	IGV	M :	0,00%	12,00%	34,32%	26599,8	4
		Sd:	0,00%	4,68%	7,36%	3349,683	0,9177
		T-:	-	-	-	-	-
	Releif	M :	0,00%	12,90%	**33,71%**	**23373,25**	**3,52**78
		Sd:	0,00%	3,58%	7,38%	3678,7417	0,7741
		T-:	-	2,5228	-3,7884	-30,1292	-9,0629
	S.U.	M :	0,00%	11,76%	35,11%	23523,8333	3,8333
		Sd:	0,00%	4,05%	7,56%	2742,335	0,7368
		T-:	-	-0,6617	2,5571	-26,8465	-3,0249
9_Tumors	IGV	M :	33,70%	85,52%	94,74%	65233,9444	4,8889
		Sd:	1,57%	4,71%	1,53%	15986,7671	1,4907
		T-:	-	-	-	-	-
	Releif	M :	33,13%	82,85%	94,00%	54996,8125	4,8125
		Sd:	2,06%	5,58%	1,61%	16307,3275	1,3545
		T-:	-26,8148	-13,8734	-8,3607	-23,4305	-0,6505
	S.U.	M :	29,25%	79,20%	94,30%	57750,9355	11,9677
		Sd:	4,10%	6,11%	1,63%	11918,3827	8,3566
		T-:	-33,4054	-58,0443	-4,6622	-12,2855	25,0562
11_Tumors	IGV	M :	20,80%	31,38%	68,47%	158556,15	10,2
		Sd:	1,33%	5,59%	2,33%	33071,2274	3,2541
		T-:	-	-	-	-	-
	Releif	M :	18,03%	28,57%	69,25%	147523,7667	13,7333
		Sd:	2,73%	5,35%	2,09%	25305,5388	4,3385
		T-:	-38,018	-23,603	8,6325	-8,1777	25,1874
	S.U.	M :	15,48%	25,89%	69,02%	161937,8438	19,0938
		Sd:	3,53%	5,95%	1,95%	26372,1043	8,4942
		T-:	-245,6009	-30,0909	5,9154	2,4466	41,9544

Table D.11: GRASP-FS based on MB^* operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	60,78%	75,25%	84,85%	362638,3	7,85
		Sd:	1,24%	2,35%	1,00%	53673,4209	3,0655
		T-:	-	-	-	-	-
	Releif	M :	**58,69%**	**72,66%**	84,93%	**361446,40**	10,0811
		Sd:	3,04%	3,47%	1,11%	52658,4595	4,5423
		T-:	-35,3175	-30,9618	1,145	-2,6084	14,7839
	S.U.	M :	**57,44%**	**71,89%**	85,04%	362712,0286	12,0857
		Sd:	3,35%	4,06%	0,97%	51520,0092	5,6588
		T-:	-28,9858	-57,1158	2,7404	0,0342	34,4613
Brain Tumor2	IGV	M :	0,00%	43,52%	52,60%	90508,9	3,15
		Sd:	0,00%	10,92%	5,37%	17401,2903	0,4894
		T-:	-	-	-	-	-
	Releif	M :	0,00%	45,28%	**51,20%**	**70930**	**3,08**
		Sd:	0,00%	9,64%	5,42%	16882,3674	0,5466
		T-:	-	5,5922	-6,1148	-59,2994	-2,2847
	S.U.	M :	0,00%	44,84%	**52,06%**	**72602,67**	**3,05**
		Sd:	0,00%	9,90%	5,36%	16092,2334	0,5242
		T-:	-	4,1241	-2,526	-47,7846	-3,257
Prostate Tumor	IGV	M :	0,00%	15,35%	17,90%	93083,95	2,5
		Sd:	0,00%	4,16%	4,76%	14583,301	0,513
		T-:	-	-	-	-	-
	Releif	M :	0,00%	16,11%	18,98%	**74670,30**	2,77
		Sd:	0,00%	5,25%	6,29%	20848,5006	0,5404
		T-:	-	3,3089	4,8016	-30,0223	10,7871
	S.U.	M :	0,00%	15,61%	18,89%	**75678,82**	2,8
		Sd:	0,00%	5,60%	6,04%	22754,811	0,5164
		T-:	-	1,1941	4,6446	-29,4244	11,767
Lymphoma	IGV	M :	0,00%	17,36%	22,97%	32772	1
		Sd:	0,00%	0,53%	1,35%	5783,6399	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	17,86%	23,03%	**28132,77**	1
		Sd:	0,00%	1,76%	2,06%	5168,1466	0
		T-:	-	9,8427	1,0559	-13,7855	-
	S.U.	M :	0,00%	17,79%	23,39%	**28370,74**	1
		Sd:	0,00%	2,24%	2,24%	5850,9964	0
		T-:	-	8,428	11,3599	-11,6266	-
Challenge 2004	IGV	M :	73,67%	96,15%	98,96%	492621,2778	3,3333
		Sd:	1,02%	2,92%	0,76%	90640,7174	0,686
		T-:	-	-	-	-	-
	Releif	M :	**72,20%**	96,01%	98,97%	**454296,78**	4,7105
		Sd:	2,78%	3,45%	0,71%	87680,1889	4,3178
		T-:	-20,9058	-0,9464	0,186	-4,8286	68,7767
	S.U.	M :	**68,58%**	**93,56%**	98,90%	**448516,41**	13,2903
		Sd:	4,59%	5,11%	0,61%	75941,4535	9,8765
		T-:	-29,5053	-15,391	-1,4442	-5,3798	27,3038

Table D.12: GRASP-FS based on MB^* operator (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	IGV	M :	5,60%	37,17%	40,94%	447831,25	2,2
		Sd:	2,01%	3,11%	4,67%	80885,6411	0,4104
		T-:	-	-	-	-	-
	Releif	M :	6,29%	39,48%	44,26%	**316240,14**	3,8
		Sd:	3,11%	4,58%	5,76%	116531,6389	2,1666
		T-:	7,3102	8,9237	22,5202	-22,8587	24,5576
	S.U.	M :	7,65%	41,74%	46,11%	**299546,82**	6,78
		Sd:	3,39%	5,07%	5,48%	98946,9139	9,2193
		T-:	25,2058	20,0177	18,4737	-22,2973	52,542
CNS	IGV	M :	4,33%	37,03%	39,90%	95793,2	2,15
		Sd:	3,26%	3,07%	5,56%	15845,7818	0,3663
		T-:	-	-	-	-	-
	Releif	M :	**2,41%**	38,76%	**38,41%**	**56623,52**	3,13
		Sd:	3,25%	5,50%	7,47%	24210,893	1,3342
		T-:	-14,322	11,0968	-6,922	-36,9296	117,2486
	S.U.	M :	**3,73%**	36,84%	**37,16%**	**65191,5**	2,5882
		Sd:	3,36%	6,40%	8,01%	24920,5472	0,925
		T-:	-3,7977	-0,4786	-7,7432	-26,4058	10,3861
Colon	IGV	M :	1,56%	27,68%	32,10%	31271,15	3,05
		Sd:	2,78%	9,94%	11,61%	3638,7326	1,3169
		T-:	-	-	-	-	-
	Releif	M :	**0,89%**	**26,75%**	28,33%	**22970,94**	4,82
		Sd:	2,22%	8,75%	10,60%	5332,1787	4,449
		T-:	-8,148	-2,6194	-16,299	-36,191	18,4553
	S.U.	M :	**0,74%**	**26,83%**	**29,98%**	**25281,5**	6,5882
		Sd:	2,04%	9,30%	11,24%	6314,0797	6,3728
		T-:	-10,2045	-2,243	-4,9219	-23,4932	11,3881
Leukemia3C	IGV	M :	2,16%	15,99%	16,42%	109086,0556	3,8889
		Sd:	2,79%	7,35%	4,68%	20292,0419	2,3487
		T-:	-	-	-	-	-
	Releif	M :	**0,67%**	15,42%	16,72%	**58935,84**	7,45
		Sd:	1,84%	8,78%	5,58%	20769,2539	6,704
		T-:	-12,2143	-0,6936	2,601	-37,1149	15,0554
	S.U.	M :	**0,79%**	15,57%	16,98%	**73475,68**	7,8
		Sd:	1,97%	8,65%	5,35%	28416,3843	7,2793
		T-:	-11,1898	-0,5112	5,0365	-23,2667	18,2102
Leukemia4C	IGV	M :	6,94%	20,94%	22,61%	105455,6	3,6
		Sd:	2,47%	6,87%	3,61%	12461,8059	1,6351
		T-:	-	-	-	-	-
	Releif	M :	**6,08%**	24,06%	23,39%	**68753,56**	6,65
		Sd:	1,65%	6,61%	5,10%	27721,1139	5,933
		T-:	-4,154	45,9682	13,5552	-34,0015	33,8263
	S.U.	M :	**6,27%**	22,81%	24,62%	**73651,19**	9,74
		Sd:	1,89%	6,80%	6,33%	25869,4954	10,798
		T-:	-3,2061	28,2986	20,6736	-26,0963	62,8251

Table D.13: GRASP-FS based on FLS^{*} operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	IGV	M :	9,70%	23,04%	24,10%	222102,3684	10,1579
		Sd:	1,22%	3,10%	1,90%	35735,1011	7,7836
		T-:	-	-	-	-	-
	Releif	M :	**8,06%**	**19,26%**	**23,96%**	**175091,13**	18,86
		Sd:	1,52%	4,47%	2,26%	42089,1024	9,8069
		T-:	-34,5889	-19,5206	-4,6739	-23,0343	27,46
	S.U.	M :	**7,84%**	**18,80%**	24,33%	**182647,86**	22,33
		Sd:	1,41%	4,03%	2,17%	38023,9199	10,2148
		T-:	-33,9369	-24,5965	9,8807	-15,8227	29,8255
MLL	IGV	M :	0,00%	23,89%	19,47%	182342,7	7,5
		Sd:	0,00%	6,90%	7,58%	22942,388	1,9057
		T-:	-	-	-	-	-
	Releif	M :	0,00%	24,21%	19,09%	**111975,64**	11,16
		Sd:	0,00%	5,67%	5,79%	44731,399	4,0422
		T-:	-	0,719	-0,435	-34,0323	17,0785
	S.U.	M :	0,00%	**22,33%**	20,43%	**119273,57**	15,8
		Sd:	0,00%	7,28%	6,34%	45033,28	8,7843
		T-:	-	-3,4412	1,0013	-34,3062	24,4165
Ovarian	IGV	M :	0,00%	3,79%	4,78%	245353,25	2
		Sd:	0,00%	1,07%	0,73%	39599,099	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**3,33%**	4,83%	**184594,13**	2,94
		Sd:	0,00%	1,08%	0,81%	54910,7285	1,094
		T-:	-	-6,9298	1,891	-29,5406	36
	S.U.	M :	0,00%	**3,41%**	4,86%	**189873,87**	2,87
		Sd:	0,00%	1,15%	0,86%	51576,9358	1,0804
		T-:	-	-5,7922	2,8536	-27,3081	39
SRBCT	IGV	M :	4,52%	25,76%	45,46%	41392,85	4,5
		Sd:	1,88%	10,79%	6,44%	4269,7613	1,9601
		T-:	-	-	-	-	-
	Releif	M :	**1,49%**	**18,93%**	**40,93%**	**34411,06**	9,18
		Sd:	2,24%	8,84%	8,03%	6933,039	4,6934
		T-:	-63,2406	-29,6623	-14,9906	-18,2028	59,1996
	S.U.	M :	**1,34%**	**18,05%**	**39,60%**	**35466**	10,03
		Sd:	2,18%	11,02%	7,43%	5044,8725	5,8059
		T-:	-29,5912	-22,7874	-16,8471	-34,3903	45,7799
9_Tumors	IGV	M :	44,67%	84,30%	93,70%	87804,65	2,9
		Sd:	3,13%	6,15%	2,25%	13144,9434	0,9679
		T-:	-	-	-	-	-
	Releif	M :	**40,86%**	**82,39%**	93,94%	**68438,93**	4,80
		Sd:	4,79%	7,59%	2,03%	17506,2769	5,4432
		T-:	-17,9272	-5,4771	1,0495	-34,8346	21,0266
	S.U.	M :	**39,26%**	**81,65%**	93,68%	**74568,96**	10,70
		Sd:	7,00%	7,24%	1,94%	15787,0501	13,6236
		T-:	-18,5182	-10,0676	-0,0913	-28,0269	27,3463
11_Tumors	IGV	M :	36,36%	52,83%	71,52%	199629,85	5,6
		Sd:	2,33%	6,86%	2,36%	19445,6583	1,9029
		T-:	-	-	-	-	-
	Releif	M :	**32,84%**	**46,40%**	73,01%	**171624,06**	9,41
		Sd:	4,31%	8,82%	2,35%	30238,0167	7,361
		T-:	-31	-13,4915	16,2772	-22,5637	19,8663
	S.U.	M :	**27,14%**	**34,15%**	72,35%	205493,4571	24,9429
		Sd:	6,26%	14,71%	2,28%	48777,7544	16,7032
		T-:	-35	-25,0228	12,0073	3,8717	32,041

Table D.14: GRASP-FS based on FLS^* operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	67,40%	77,91%	85,36%	385703,4211	2,9474
		Sd:	1,14%	3,17%	1,40%	57277,3769	0,9703
		T-:	-	-	-	-	-
	Releif	M :	**66,19%**	**76,68%**	**85,11%**	439093,8125	5,6875
		Sd:	1,99%	4,45%	1,19%	93054,4854	6,9859
		T-:	-14,9903	-10,5709	-8,016	14,0314	42,4409
	S.U.	M :	**63,28%**	**70,28%**	85,52%	504099,8485	16,4242
		Sd:	3,53%	7,26%	1,12%	125483,3026	13,5417
		T-:	-28,01	-19,6864	10,7771	31,6015	20,5531
Brain Tumor2	IGV	M :	6,33%	47,15%	56,38%	134627,5294	5,1176
		Sd:	3,02%	10,79%	5,21%	17688,4707	2,9556
		T-:	-	-	-	-	-
	Releif	M :	7,69%	**43,39%**	57,02%	**89648,12**	12,81
		Sd:	3,85%	8,26%	6,74%	30865,6082	7,3504
		T-:	17	-5,0411	3,3181	-36,9495	25,7005
	S.U.	M :	9,93%	**43,51%**	57,11%	**89750,77**	15,32
		Sd:	3,55%	10,08%	6,74%	31018,3411	8,7155
		T-:	33,388	-5,0011	2,3359	-73,7017	19,3846
Prostate Tumor	IGV	M :	0,00%	15,08%	19,50%	153742,1667	2,3889
		Sd:	0,00%	2,64%	2,25%	21400,5474	0,6978
		T-:	-	-	-	-	-
	Releif	M :	0,00%	21,03%	26,03%	**89377,25**	8,125
		Sd:	0,00%	6,42%	8,44%	35285,9235	3,9044
		T-:	-	55,3988	22,9672	-45,3362	35,213
	S.U.	M :	0,00%	19,43%	24,89%	**103259,64**	9,67
		Sd:	0,00%	5,88%	9,22%	41837,4035	7,9876
		T-:	-	17,9918	10,9854	-33,4882	52,3421
Lymphoma	IGV	M :	0,00%	15,48%	18,82%	56511,75	2
		Sd:	0,00%	3,23%	2,91%	4879,9782	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**14,72%**	18,77%	**37783,64**	3,74
		Sd:	0,00%	4,14%	2,57%	10214,1012	3,7056
		T-:	-	-3,6198	-0,4648	-36,8341	31
	S.U.	M :	0,00%	15,24%	18,65%	**41001,94**	3,17
		Sd:	0,00%	3,77%	2,77%	12210,9059	2,9153
		T-:	-	-1,0948	-1,8868	-35,0854	35
Challenge 2004	IGV	M :	77,54%	98,15%	98,89%	615434,1667	3,6111
		Sd:	2,24%	2,19%	0,67%	92169,0512	1,2897
		T-:	-	-	-	-	-
	Releif	M :	77,51%	98,19%	**98,79%**	652539,6207	5,6552
		Sd:	2,04%	2,26%	0,62%	253220,6408	8,478
		T-:	-0,1057	0,1169	-2,5765	4,1201	22,7059
	S.U.	M :	**76,96%**	**94,18%**	98,96%	787874,7333	26
		Sd:	2,33%	5,69%	0,76%	331637,639	22,752
		T-:	-2,5718	-10,7666	1,4621	20,9179	27,8568

Table D.15: GRASP-FS based on FLS^{*} operator (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	IGV	M :	3,80%	37,38%	41,77%	337551,35	2,05
		Sd:	0,89%	1,96%	4,09%	100281,0563	0,2236
		T-:	-	-	-	-	-
	Releif	M :	**2,06%**	38,24%	41,63%	**326681,60**	3,03
		Sd:	2,03%	4,53%	6,71%	102526,1601	1,2621
		T-:	-27,5054	8,6575	-1,3886	-3,9435	368,0768
	S.U.	M :	**2,13%**	37,74%	44,06%	389302,3	3,5
		Sd:	2,03%	4,78%	6,17%	104816,4692	2,0299
		T-:	-23,2091	2,9211	13,4041	11,2223	7,9084
CNS	IGV	M :	0,00%	36,93%	37,30%	59832,65	2,7
		Sd:	0,00%	6,43%	4,37%	7881,1242	0,4702
		T-:	-	-	-	-	-
	Releif	M :	0,00%	38,16%	37,39%	**58327,66**	2,66
		Sd:	0,00%	5,30%	5,52%	30766,3276	0,4787
		T-:	-	13,1145	1,1041	-2,1969	-1,8433
	S.U.	M :	0,00%	38,27%	38,69%	92670,6176	2,8235
		Sd:	0,00%	4,64%	5,36%	51462,1563	0,4586
		T-:	-	5,1915	8,6935	18,9348	7,7826
Colon	IGV	M :	0,00%	33,77%	32,68%	22555,1	2,4
		Sd:	0,00%	6,19%	7,35%	2542,7348	0,5026
		T-:	-	-	-	-	-
	Releif	M :	0,00%	32,56%	**31,24%**	**21125,65**	2,57
		Sd:	0,00%	7,56%	6,54%	6987,9131	0,5778
		T-:	-	-1,201	-3,5248	-7,3084	6,8619
	S.U.	M :	0,00%	**31,70%**	**31,11%**	32058,7297	2,6216
		Sd:	0,00%	8,08%	8,57%	14530,3485	0,5452
		T-:	-	-4,0192	-4,6935	18,5128	9,8661
Leukemia3C	IGV	M :	0,00%	10,18%	13,68%	66879,1053	2,6316
		Sd:	0,00%	3,83%	3,17%	12169,2384	0,5973
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**10,02%**	14,17%	**63080,57**	2,57
		Sd:	0,00%	3,14%	3,23%	33547,7179	0,5778
		T-:	-	-4,0037	8,3944	-3,4702	-1,4768
	S.U.	M :	0,00%	10,68%	14,80%	99970,6757	2,8649
		Sd:	0,00%	4,18%	4,70%	51343,6396	0,6734
		T-:	-	6,7526	15,4256	17,9731	6,976
Leukemia4C	IGV	M :	5,56%	19,47%	20,97%	64219,6	2,7
		Sd:	0,00%	5,71%	3,32%	10330,6379	0,4702
		T-:	-	-	-	-	-
	Releif	M :	5,56%	**18,78%**	22,38%	85349,8333	2,8611
		Sd:	0,00%	5,26%	5,97%	48672,0733	0,3507
		T-:	-	-7,0311	12,5566	9,4869	10,4022
	S.U.	M :	5,56%	19,44%	22,84%	104205,6471	2,9118
		Sd:	0,00%	6,22%	6,66%	53816,2393	0,2879
		T-:	-	-0,3204	15,2038	22,159	13,911

Table D.16: GRASP-FS based on $IFLS^*$ operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	IGV	M :	5,37%	19,34%	23,39%	141901,3158	4,6842
		Sd:	0,89%	3,04%	2,15%	38121,2738	2,6885
		T-:	-	-	-	-	-
	Releif	M :	**4,58%**	**19,18%**	**23,13%**	188647,13	6,4
		Sd:	0,94%	3,88%	1,86%	97876,5597	2,7618
		T-:	-9,0277	-19,8085	-17,113	13,9077	5,684
	S.U.	M :	**4,51%**	19,60%	**23,24%**	224988,66	6,3
		Sd:	0,91%	3,98%	2,02%	108798,3735	2,3067
		T-:	-10,9075	1,4096	-2,1178	20,4698	5,7553
MLL	IGV	M :	0,00%	22,44%	24,91%	116891	2,8889
		Sd:	0,00%	6,99%	10,01%	17185,5238	0,3234
		T-:	-	-	-	-	-
	Releif	M :	0,00%	21,32%	23,89%	141650,1795	3
		Sd:	0,00%	6,67%	10,15%	70234,2055	0,3974
		T-:	-	-1,8361	-1,9223	10,2549	2,25
	S.U.	M :	0,00%	**20,65%**	**21,52%**	182483,7419	3,0323
		Sd:	0,00%	7,69%	10,12%	84275,1602	0,4819
		T-:	-	-2,8128	-6,3813	20,974	2,9026
Ovarian	IGV	M :	0,00%	2,51%	4,75%	180974,5789	2
		Sd:	0,00%	1,08%	1,21%	55139,5621	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	2,68%	4,78%	199833,6944	2
		Sd:	0,00%	1,27%	0,76%	90657,8147	0
		T-:	-	6,2137	2,877	6,3832	-
	S.U.	M :	0,00%	2,66%	4,76%	257173,8	2
		Sd:	0,00%	1,00%	0,95%	94861,6352	0
		T-:	-	5,2854	2,2353	15,3364	-
SRBCT	IGV	M :	0,00%	16,20%	37,34%	28581,15	3,55
		Sd:	0,00%	6,26%	7,53%	2907,7653	0,5104
		T-:	-	-	-	-	-
	Releif	M :	0,00%	16,25%	37,80%	29368,84	**3,18**
		Sd:	0,00%	5,93%	8,14%	13122,1128	0,3966
		T-:	-	0,3368	1,6781	2,5225	-15,5911
	S.U.	M :	0,00%	16,95%	**36,30%**	45739,06	**3,22**
		Sd:	0,00%	7,42%	8,63%	25668,1803	0,425
		T-:	-	5,935	-3,3896	21,4842	-13,7082
9_Tumors	IGV	M :	35,79%	87,02%	93,44%	79169,63	3,42
		Sd:	3,30%	4,72%	2,28%	26268,9162	0,607
		T-:	-	-	-	-	-
	Releif	M :	**34,41%**	**86,04%**	93,08%	78292,19	4,1
		Sd:	4,59%	4,00%	1,59%	34127,1851	2,446
		T-:	-6,1562	-6,0497	-2,0023	-0,9949	31,3985
	S.U.	M :	**32,00%**	**86,11%**	93,73%	101581,68	6,08
		Sd:	7,45%	4,27%	1,91%	46758,7523	4,8298
		T-:	-7,1677	-4,5934	1,6258	8,1227	11,1798
11_Tumors	IGV	M :	26,59%	44,46%	70,85%	172928,05	4,2
		Sd:	1,07%	3,86%	1,08%	51691,5702	1,0052
		T-:	-	-	-	-	-
	Releif	M :	**24,49%**	**41,40%**	71,25%	**163977,03**	5,48
		Sd:	1,93%	4,41%	1,57%	53863,8142	1,626
		T-:	-53,9007	-26,8627	3,1337	-3,4423	62,6791
	S.U.	M :	**21,35%**	**38,79%**	70,79%	255331,4643	8,3929
		Sd:	5,11%	4,60%	1,45%	136074,4249	4,81
		T-:	-38,8895	-28,1465	-0,5017	16,6457	34,4849

Table D.17: GRASP-FS based on $IFLS^*$ operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	62,86%	76,49%	84,42%	475281	3,75
		Sd:	1,71%	3,17%	1,33%	134692,2686	0,4443
		T-:	-	-	-	-	-
	Releif	M :	**61,47%**	**74,90%**	84,32%	503156,55	3,88
		Sd:	2,34%	3,27%	1,35%	152152,9631	0,6667
		T-:	-6,2652	-7,8496	-1,0672	5,6532	5,0185
	S.U.	M :	**59,87%**	**73,52%**	84,32%	622786	5,3667
		Sd:	3,04%	4,85%	1,23%	218210,3697	2,7852
		T-:	-13,2253	-15,6568	-0,9143	22,7013	34,2229
Brain Tumor2	IGV	M :	0,00%	41,64%	52,84%	97336,9444	2,9444
		Sd:	0,00%	8,38%	4,21%	14481,7369	0,2357
		T-:	-	-	-	-	-
	Releif	M :	0,00%	44,03%	54,18%	**85503,62**	2,96
		Sd:	0,00%	12,40%	4,76%	17123,2502	0,4211
		T-:	-	3,6419	5,2247	-9,5659	6,3711
	S.U.	M :	0,00%	45,65%	55,39%	138132,6667	3,1
		Sd:	0,00%	10,19%	4,34%	61891,1511	0,712
		T-:	-	6,3913	9,2264	15,1803	34,2422
Prostate Tumor	IGV	M :	0,00%	15,07%	19,03%	100413,8947	2,4211
		Sd:	0,00%	2,77%	4,14%	15312,3316	0,5073
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**14,74%**	18,86%	90318,62	**2,58**
		Sd:	0,00%	3,38%	3,70%	32945,0353	0,5012
		T-:	-	-3,541	-1,1992	-10,902	5,5061
	S.U.	M :	0,00%	15,52%	19,83%	141020,0588	2,5882
		Sd:	0,00%	4,43%	4,80%	69347,0309	0,4996
		T-:	-	2,2655	1,456	18,2532	6,62
Lymphoma	IGV	M :	0,00%	16,88%	15,33%	37983,9	2
		Sd:	0,00%	3,71%	5,25%	4228,2348	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**15,42%**	15,81%	47512,0294	2
		Sd:	0,00%	3,61%	4,89%	27068,6696	0
		T-:	-	-16,2024	1,3585	14,5352	-
	S.U.	M :	0,00%	**15,63%**	17,22%	58277,9459	2
		Sd:	0,00%	3,81%	5,10%	30081,6789	0
		T-:	-	-7,4624	6,027	18,4923	-
Challenge 2004	IGV	M :	73,91%	97,87%	99,30%	705548,7895	2,9474
		Sd:	2,05%	1,36%	0,40%	126948,2368	0,5243
		T-:	-	-	-	-	-
	Releif	M :	**73,41%**	97,83%	**98,96%**	724715,9615	3,1154
		Sd:	1,88%	2,01%	0,71%	250325,5896	0,5159
		T-:	-2,1846	-0,3409	-8,2568	1,4182	32,1166
	S.U.	M :	**70,61%**	**95,98%**	**99,04%**	8202	6,36
		Sd:	5,22%	2,74%	0,64%	251636,9328	5,1952
		T-:	-12,5041	-16,637	-6,5143	6,8689	25,3862

Table D.18: GRASP-FS based on $IFLS^*$ operator (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	IGV	M :	2,40%	36,92%	38,29%	257815,6	2,5
		Sd:	2,07%	2,28%	5,21%	50715,3688	0,8498
		T-:	-	-	-	-	-
	Releif	M :	2,18%	38,43%	41,04%	**240883,81**	5,90
		Sd:	2,04%	4,82%	5,88%	52775,054	6,023
		T-:	-1,2116	7,788	10,4695	-4,1181	24,1149
	S.U.	M :	2,74%	39,89%	41,32%	**236751,73**	8,78
		Sd:	1,91%	6,64%	5,20%	48727,9984	8,3372
		T-:	1,9441	19,9277	11,5558	-5,1155	20,3685
CNS	IGV	M :	0,00%	37,13%	39,40%	62194,4	2,5
		Sd:	0,00%	3,58%	4,67%	14139,3336	0,527
		T-:	-	-	-	-	-
	Releif	M :	0,00%	40,35%	39,79%	**51644,84**	3,31
		Sd:	0,00%	5,57%	5,67%	16140,9987	1,1082
		T-:	-	6,6325	1,0313	-5,7228	5,4436
	S.U.	M :	0,00%	39,14%	39,53%	**48857,88**	3,94
		Sd:	0,00%	6,36%	6,84%	13873,0015	1,7843
		T-:		3,53	0,34	-7,32	10,99
Colon	IGV	M :	0,00%	30,58%	33,35%	20806,4	2,7
		Sd:	0,00%	6,08%	9,59%	2284,3905	0,483
		T-:	-	-	-	-	-
	Releif	M :	0,00%	32,94%	34,12%	**19421,52**	2,82
		Sd:	0,00%	9,05%	10,95%	2764,5003	0,5286
		T-:	-	5,7696	0,7766	-7,7503	1,7456
	S.U.	M :	0,00%	32,39%	32,03%	**18561,35**	2,85
		Sd:	0,00%	8,78%	9,03%	1912,973	0,5871
		T-:	-	2,3705	-1,2257	-11,198	1,8317
Leukemia3C	IGV	M :	0,00%	10,56%	16,17%	60577,3	3
		Sd:	0,00%	4,33%	2,04%	15374,3179	0,4714
		T-:	-	-	-	-	-
	Releif	M :	0,00%	10,40%	**14,69%**	**55522,72**	3,77
		Sd:	0,00%	4,89%	4,01%	14125,9764	0,5483
		T-:	-	-0,4557	-4,2622	-5,8982	63
	S.U.	M :	0,00%	10,15%	**14,35%**	**54006,5**	4,2222
		Sd:	0,00%	5,29%	2,52%	11879,0833	1,896
		T-:	-	-1,1494	-5,4636	-7,3151	99
Leukemia4C	IGV	M :	5,56%	14,94%	19,11%	62169,3	3
		Sd:	0,00%	6,37%	4,34%	11856,383	0,4714
		T-:	-	-	-	-	-
	Releif	M :	5,56%	16,72%	18,78%	**57116,80**	4,04
		Sd:	0,00%	5,36%	3,55%	9855,575	2,0366
		T-:	-	1,5903	-0,3771	-23,5838	11,2683
	S.U.	M :	5,56%	15,61%	19,30%	61390,9474	4,4737
		Sd:	0,00%	5,40%	2,83%	12951,4078	2,3891
		T-:	-	0,5996	0,215	-0,7706	59,1111

Table D.19: GRASP-FS based on AF^* operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	IGV	M :	5,88%	18,46%	22,07%	136556,7778	5,3333
		Sd:	0,00%	1,48%	1,10%	35354,9939	2,2361
		T-:	-	-	-	-	-
	Releif	M :	**4,55%**	**16,73%**	**21,41%**	146559,7727	11,2273
		Sd:	0,93%	3,35%	1,66%	30544,122	5,1449
		T-:	-47,1429	-10,6355	-4,457	2,6287	16,0499
	S.U.	M :	**4,37%**	**15,63%**	**21,12%**	154494,40	13,04
		Sd:	0,84%	2,80%	1,81%	29676,1264	5,2325
		T-:	-74,8	-27,303	-4,0485	4,5219	26,0278
MLL	IGV	M :	0,00%	24,61%	23,50%	112625,1	3
		Sd:	0,00%	5,64%	5,03%	25345,8454	0,4714
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**21,41%**	**17,45%**	**96848**	4,3529
		Sd:	0,00%	4,09%	4,68%	14753,1343	0,9963
		T-:	-	-16,9428	-8,4683	-9,3242	35,5455
	S.U.	M :	0,00%	**21,61%**	**18,11%**	**101250,65**	4,55
		Sd:	0,00%	5,57%	5,64%	21842,5547	1,3945
		T-:	-	-19,7279	-7,3173	-4,479	68,8889
Ovarian	IGV	M :	0,00%	2,33%	4,51%	165405,3	2
		Sd:	0,00%	0,61%	0,44%	32786,7099	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	2,77%	4,59%	**145817,23**	2,80
		Sd:	0,00%	1,31%	0,63%	27112,4048	0,9284
		T-:	-	4,6063	2,3866	-4,0811	14,28
	S.U.	M :	0,00%	2,81%	4,66%	**142868,57**	3,15
		Sd:	0,00%	1,37%	0,65%	25648,0669	1,0679
		T-:	-	4,9116	4,0552	-4,6733	26,125
SRBCT	IGV	M :	0,00%	17,12%	38,83%	25989,1	3,7
		Sd:	0,00%	4,37%	6,43%	2590,7894	0,483
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**12,05%**	**35,80%**	26871,1	5,05
		Sd:	0,00%	5,48%	4,92%	4094,1713	1,4318
		T-:	-	-27,0039	-11,5131	1,9703	13,2339
	S.U.	M :	0,00%	**12,12%**	37,77%	28314,4737	5,4211
		Sd:	0,00%	6,47%	3,92%	4254,7292	1,3464
		T-:	-	-25,1937	-4,6614	5,8736	19,4796
9_Tumors	IGV	M :	34,00%	85,53%	94,00%	66769,3	3,8
		Sd:	2,11%	5,41%	1,91%	14105,8249	0,6325
		T-:	-	-	-	-	-
	Releif	M :	**31,48%**	83,93%	**93,56%**	**62794,88**	8,5
		Sd:	4,46%	5,16%	1,97%	17195,1845	11,9225
		T-:	-3,927	-1,8969	-3,0249	-2,2056	2,1973
	S.U.	M :	**28,57%**	**80,19%**	**93,33%**	67981,57	17,07
		Sd:	4,07%	7,49%	1,63%	16844,3704	13,0647
		T-:	-35,8298	-6,5654	-4,7087	0,9762	8,4723
11_Tumors	IGV	M :	23,41%	36,78%	69,72%	139838,5	6,6
		Sd:	1,53%	5,25%	2,43%	31995,6484	1,1738
		T-:	-	-	-	-	-
	Releif	M :	**20,91%**	**33,79%**	71,13%	151688,4	11,4
		Sd:	3,26%	6,54%	2,81%	35556,7632	6,8626
		T-:	-8,4366	-8,9443	4,9587	94,0632	11,1148
	S.U.	M :	**18,42%**	**29,15%**	70,74%	181290,3158	18,4211
		Sd:	4,54%	6,94%	2,92%	40351,8462	10,4367
		T-:	-16,7457	-18,4001	3,6019	69,9825	21,1763

Table D.20: GRASP-FS based on AF^* operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	62,47%	77,26%	85,18%	335793,7	3,6
		Sd:	2,25%	3,68%	1,14%	63765,7	0,5164
		T-:	-	-	-	-	-
	Releif	M :	**59,16%**	**71,69%**	**84,71%**	378852	7,1111
		Sd:	2,96%	5,69%	1,38%	72287,8555	5,6765
		T-:	-6,9221	-5,9731	-7,0805	8,2791	4,2398
	S.U.	M :	**58,44%**	**70,33%**	**84,59%**	401761,7	8,95
		Sd:	3,34%	5,61%	1,41%	67399,6481	6,3947
		T-:	-9,0016	-9,3451	-8,9035	18,3454	10,6132
Brain Tumor2	IGV	M :	0,00%	42,64%	54,24%	86345,3	3,3
		Sd:	0,00%	10,86%	5,90%	12152,1439	0,483
		T-:	-	-	-	-	-
	Releif	M :	0,00%	47,12%	53,90%	**75116,52**	3,7143
		Sd:	0,00%	10,71%	5,53%	15392,8263	1,1019
		T-:	-	2,9475	-1,2731	-7,2019	5,8096
	S.U.	M :	0,00%	48,44%	54,52%	**76108,8**	4,2
		Sd:	0,00%	10,92%	5,89%	15531,6156	1,2397
		T-:	-	4,3407	1,0842	-6,3827	7,8935
Prostate Tumor	IGV	M :	0,00%	14,94%	20,82%	91083,3	2,4
		Sd:	0,00%	4,35%	4,88%	15086,5811	0,5164
		T-:	-	-	-	-	-
	Releif	M :	0,00%	14,68%	**18,00%**	**79191,09**	3,61
		Sd:	0,00%	4,16%	4,66%	17612,9667	1,7457
		T-:	-	-0,2488	-2,3019	-15,1609	27,7553
	S.U.	M :	0,00%	15,13%	**18,17%**	**81302,90**	3,76
		Sd:	0,00%	4,13%	4,45%	17036,3733	1,5461
		T-:	-	0,1733	-2,172	-18,4402	25,2191
Lymphoma	IGV	M :	0,00%	13,94%	15,70%	34505,1	2
		Sd:	0,00%	5,69%	4,04%	4823,8649	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	15,12%	17,48%	**31975,25**	2,1
		Sd:	0,00%	4,23%	3,56%	4903,648	0,3078
		T-:	-	9,4861	7,5627	-11,5928	2,2222
	S.U.	M :	0,00%	14,68%	16,84%	**33206,91**	2,17
		Sd:	0,00%	4,35%	3,73%	4976,2585	0,3876
		T-:	-	2,7181	4,0572	-6,5982	4,8421
Challenge 2004	IGV	M :	73,04%	97,69%	98,98%	469838,6	3,6
		Sd:	1,83%	2,17%	0,76%	40368,9398	0,9661
		T-:	-	-	-	-	-
	Releif	M :	**72,67%**	**96,57%**	98,92%	477245	6,8571
		Sd:	2,66%	2,35%	0,83%	72493,7797	11,8767
		T-:	-1,5641	-3,085	-0,9436	0,7203	1,1081
	S.U.	M :	**70,33%**	**93,05%**	98,98%	526768,4706	27,4706
		Sd:	3,52%	5,02%	0,66%	75194,0123	22,5447
		T-:	-27,6713	-15,3969	0,039	13,1502	17,9936

Table D.21: GRASP-FS based on AF^* operator (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	IGV	M :	2,95%	37,68%	41,73%	222342,5789	2,3158
		Sd:	1,81%	2,63%	3,03%	61885,4074	0,671
		T-:	-	-	-	-	-
	Releif	M :	**1,71%**	38,51%	42,44%	224994,9048	3,619
		Sd:	2,03%	5,49%	4,92%	55367,8085	2,0366
		T-:	-12,4987	1,3863	1,4279	0,6924	8,0521
	S.U.	M :	**2,40%**	38,57%	42,58%	225513,8	3,76
		Sd:	2,00%	4,69%	3,94%	57809,1515	2,6814
		T-:	-4,9384	1,7459	2,0412	0,9484	11,0531
CNS	IGV	M :	0,00%	38,17%	39,73%	47693,65	2,5
		Sd:	0,00%	2,91%	4,30%	7644,1872	0,513
		T-:	-	-	-	-	-
	Releif	M :	0,00%	39,25%	40,35%	**46407,26**	2,69
		Sd:	0,00%	5,32%	5,21%	8698,8123	0,8221
		T-:	-	11,7687	2,1136	-2,1803	3,1569
	S.U.	M :	0,00%	39,52%	40,48%	**46517,42**	2,96
		Sd:	0,00%	5,75%	4,51%	10174,8784	0,7927
		T-:	-	10,655	4,0333	-3,0374	18,5473
Colon	IGV	M :	0,00%	30,81%	34,19%	18425,45	2,5
		Sd:	0,00%	9,36%	7,50%	2247,3927	0,513
		T-:	-	-	-	-	-
	Releif	M :	0,00%	33,39%	33,34%	**17562,04**	2,56
		Sd:	0,00%	11,10%	8,00%	1872,8557	0,5831
		T-:	-	5,6211	-0,8742	-7,4644	0,9555
	S.U.	M :	0,00%	33,26%	33,62%	**17672,96**	2,66
		Sd:	0,00%	10,20%	8,62%	2049,4833	0,6794
		T-:	-	4,8118	-0,5811	-6,1662	5,9775
Leukemia3C	IGV	M :	0,00%	10,92%	15,47%	50637,1	3,2
		Sd:	0,00%	4,76%	4,56%	11815,236	0,7678
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**10,51%**	15,24%	**49007,82**	3,47
		Sd:	0,00%	5,92%	5,38%	12282,5541	0,9941
		T-:	-	-2,7295	-1,5672	-2,5568	11,2244
	S.U.	M :	0,00%	13,63%	18,38%	**48295,19**	3,80
		Sd:	0,00%	5,64%	5,62%	10523,5108	1,2335
		T-:	-	7,2245	15,0304	-4,6374	12,9474
Leukemia4C	IGV	M :	5,56%	20,31%	21,78%	51218	2,8
		Sd:	0,00%	5,86%	5,01%	13659,6346	0,5231
		T-:	-	-	-	-	-
	Releif	M :	5,56%	18,03%	21,19%	52367,4091	3,3182
		Sd:	0,00%	6,61%	5,02%	11641,2383	0,8937
		T-:	-	-9,6818	-1,8027	1,7274	4,039
	S.U.	M :	5,56%	21,16%	23,96%	**49272,04**	3,39
		Sd:	0,00%	6,40%	5,57%	11307,4877	1,0331
		T-:	-	2,2793	4,4906	-2,7077	12,3285

Table D.22: GRASP-FS based on BF^* operator (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	IGV	M :	5,20%	18,84%	23,01%	102492	7,5
		Sd:	0,96%	3,75%	2,18%	19117,1089	3,5615
		T-:	-	-	-	-	-
	Releif	M :	**4,44%**	19,69%	23,31%	119604	9,5263
		Sd:	0,89%	2,92%	2,21%	31889,6135	4,0328
		T-:	-10,9499	3,272	10,6436	10,1068	7,0728
	S.U.	M :	**4,48%**	18,91%	23,43%	118928,4643	8,9643
		Sd:	0,90%	2,90%	1,95%	28017,1016	3,2487
		T-:	-10,6941	0,2861	17,8207	13,44	5,2834
MLL	IGV	M :	0,00%	18,39%	18,69%	87102,95	3,4
		Sd:	0,00%	7,21%	7,14%	24463,2372	0,5026
		T-:	-	-	-	-	-
	Releif	M :	0,00%	20,07%	19,38%	100580,4	3,52
		Sd:	0,00%	6,27%	8,06%	26346,2383	0,5859
		T-:	-	2,5155	0,7146	9,6852	4,3283
	S.U.	M :	0,00%	19,07%	17,04%	90903,0833	3,7083
		Sd:	0,00%	5,67%	5,51%	23593,9192	0,55
		T-:	-	0,9481	-1,6714	2,7517	8,6484
Ovarian	IGV	M :	0,00%	2,49%	4,77%	116147,15	2
		Sd:	0,00%	1,38%	0,84%	28051,4037	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	2,96%	4,88%	131238,08	2
		Sd:	0,00%	1,19%	0,57%	33942,6977	0
		T-:	-	17,4382	4,8034	10,7333	-
	S.U.	M :	0,00%	2,74%	4,86%	121828,4231	2,0385
		Sd:	0,00%	1,04%	0,68%	25306,2765	0,1961
		T-:	-	12,8805	3,7847	3,4807	26
SRBCT	IGV	M :	0,00%	16,37%	36,71%	21644,45	4
		Sd:	0,00%	5,54%	8,64%	3180,9692	0,9177
		T-:	-	-	-	-	-
	Releif	M :	0,00%	15,63%	37,26%	22892,9565	3,913
		Sd:	0,00%	5,93%	7,58%	3315,9412	0,9493
		T-:	-	-1,2502	1,8043	5,8997	-23
	S.U.	M :	0,00%	**15,30%**	36,57%	23515,6667	4,2963
		Sd:	0,00%	5,67%	5,77%	3263,2916	1,1373
		T-:	-	-4,6675	-0,3901	21,0328	6,1714
9_Tumors	IGV	M :	33,67%	83,83%	94,00%	50068,9	4,25
		Sd:	2,63%	7,10%	2,02%	10269,4184	1,4464
		T-:	-	-	-	-	-
	Releif	M :	**31,75%**	83,78%	94,60%	52632,1429	5,7619
		Sd:	4,17%	7,65%	2,17%	11559,7217	3,5624
		T-:	-7,9218	-0,123	9,0383	2,9997	3,6371
	S.U.	M :	**29,70%**	**82,58%**	93,76%	53861,8636	9,2727
		Sd:	5,72%	6,79%	2,15%	11467,4542	7,3561
		T-:	-8,9998	-4,6146	-1,9904	22,114	7,8611
11_Tumors	IGV	M :	23,07%	38,38%	70,44%	108790,8	6,5
		Sd:	1,85%	3,59%	2,33%	18300,101	1,9057
		T-:	-	-	-	-	-
	Releif	M :	**21,89%**	**36,65%**	70,92%	129190,7407	8,9259
		Sd:	2,76%	5,75%	2,31%	30885,3226	4,7388
		T-:	-8,5558	-4,7474	6,0999	15,5863	4,9509
	S.U.	M :	**19,52%**	**33,00%**	71,31%	139472,3182	14,7273
		Sd:	4,13%	7,84%	2,09%	37912,4314	9,857
		T-:	-13,2118	-9,2304	24,5499	15,2895	11,0662

Table D.23: GRASP-FS based on BF^* operator (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	IGV	M :	61,36%	75,75%	84,44%	267722,7	5,15
		Sd:	1,88%	2,02%	1,06%	41283,9675	1,1821
		T-:	-	-	-	-	-
	Releif	M :	**59,96%**	**72,92%**	84,39%	345046,5	7,2917
		Sd:	2,81%	4,10%	1,24%	85407,1047	5,4572
		T-:	-8,1739	-16,8838	-0,7665	17,9623	4,7001
	S.U.	M :	**58,17%**	**71,78%**	84,55%	366748,4211	10,4211
		Sd:	3,42%	4,86%	1,39%	91200,1688	7,1594
		T-:	-16,699	-10,8348	0,7986	52,2682	11,4364
Brain Tumor2	IGV	M :	0,00%	46,60%	54,92%	68878,8	3,75
		Sd:	0,00%	9,18%	6,95%	12984,9457	0,9105
		T-:	-	-	-	-	-
	Releif	M :	0,31%	**44,80%**	55,84%	77108,92	4,04
		Sd:	1,54%	12,16%	5,49%	19797,366	1,1358
		T-:	25	-2,2484	3,3938	8,2166	7,7263
	S.U.	M :	0,31%	45,12%	55,07%	78924,8	4,24
		Sd:	1,54%	10,00%	4,90%	20275,9427	1,3928
		T-:	25	-1,8787	0,5224	10,569	12,6585
Prostate Tumor	IGV	M :	0,00%	14,86%	18,88%	70317,35	2,4
		Sd:	0,00%	2,05%	2,58%	16408,7966	0,5026
		T-:	-	-	-	-	-
	Releif	M :	0,00%	16,36%	20,11%	71972,5909	2,6818
		Sd:	0,00%	4,41%	5,73%	19048,8628	0,5679
		T-:	-	3,2026	1,7502	1,8778	8,4619
	S.U.	M :	0,00%	14,93%	18,64%	73897,9643	2,75
		Sd:	0,00%	3,57%	4,95%	19228,9211	0,5853
		T-:	-	1,5056	-0,7445	6,3994	6,5072
Lymphoma	IGV	M :	0,00%	17,35%	17,32%	27127,5263	2
		Sd:	0,00%	2,76%	5,02%	3231,0875	0
		T-:	-	-	-	-	-
	Releif	M :	0,00%	**15,71%**	17,58%	26976,4167	2,25
		Sd:	0,00%	4,31%	5,28%	2688,9536	0,4423
		T-:	-	-6,6575	1,1285	-0,7875	24
	S.U.	M :	0,00%	**16,39%**	17,86%	27125,3214	2,25
		Sd:	0,00%	4,22%	5,61%	2876,2523	0,441
		T-:	-	-5,0917	3,3441	-0,0174	28
Challenge 2004	IGV	M :	73,48%	96,96%	99,24%	415043,6	3,2
		Sd:	2,40%	2,31%	0,48%	98875,5844	0,9515
		T-:	-	-	-	-	-
	Releif	M :	72,53%	**96,40%**	**98,99%**	433345,7727	4,2727
		Sd:	3,65%	2,75%	0,81%	88948,0944	3,7184
		T-:	-1,7883	-2,3851	-4,4462	5,3766	2,1926
	S.U.	M :	**70,59%**	**95,22%**	**99,03%**	428864,0588	8,1176
		Sd:	4,99%	3,08%	0,71%	80982,5494	7,0523
		T-:	-9,126	-17,2915	-2,8071	3,7511	12,0881

Table D.24: GRASP-FS based on BF^* operator (3)

Appendix E

Annex V: Detailed Empirical Results of Chapter V

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	BPSO	M :	28,40%	47,75%	49,62%	2050392,8	25,7
		Sd:	4,79%	5,78%	7,18%	886867,5961	17,1338
		T-:	-	-	-	-	-
	PSO	M :	**14,40%**	45,50%	**48,21%**	**577010,2**	118,5
		Sd:	5,06%	7,74%	4,03%	759776,2244	111,1778
		T-:	-18,0259	-2,0769	-2,2492	-6,9165	19,5694
	PSO-FS	M :	**6,46%**	**44,36%**	**46,47%**	**102542,30**	**5,38**
		Sd:	2,60%	4,43%	7,51%	17754,0898	2,256
		T-:	-48,149	-11,2067	-2,9835	-9,3308	-21,7026
CNS	BPSO	M :	26,00%	44,13%	39,80%	125006	28,9
		Sd:	4,92%	10,92%	5,94%	30016,4541	16,5627
		T-:	-	-	-	-	-
	PSO	M :	**24,00%**	42,60%	40,20%	**26223,4**	**25,5**
		Sd:	7,83%	6,07%	2,55%	4920,4688	6,4161
		T-:	-2,3943	-1,7364	1,7786	-24,5769	-7,3395
	PSO-FS	M :	**4,29%**	42,29%	45,00%	**19902,07**	**6,5**
		Sd:	3,31%	7,76%	4,84%	2440,6061	2,8756
		T-:	-27,3252	-1,7422	22,4811	-26,4133	-106,8027
Colon	BPSO	M :	28,12%	26,32%	34,65%	35626,6	21
		Sd:	5,31%	7,91%	6,27%	6058,3094	14,1343
		T-:	-	-	-	-	-
	PSO	M :	**13,54%**	24,95%	**31,94%**	**13023,33**	**16,83**
		Sd:	4,49%	3,77%	5,55%	2203,0788	5,7814
		T-:	-44,9642	-1,5155	-14,0024	-225,4759	-2,5695
	PSO-FS	M :	**3,57%**	**18,43%**	**22,95%**	**12032,78**	**15,92**
		Sd:	3,21%	5,04%	3,80%	931,788	1,8172
		T-:	-67,0097	-7,7737	-54,2854	-181,0441	-3,1696
Leukemia3C	BPSO	M :	23,33%	33,11%	38,56%	218755	28
		Sd:	5,11%	11,56%	13,73%	39573,4139	9,8432
		T-:	-	-	-	-	-
	PSO	M :	**14,44%**	35,61%	39,06%	218427,8	110,6
		Sd:	2,87%	12,84%	13,28%	322195,4758	78,3684
		T-:	-17,8885	1,7142	0,455	-0,0186	14,5123
	PSO-FS	M :	**0,00%**	**5,20%**	**13,61%**	**23559,42**	**10,92**
		Sd:	0,00%	3,93%	4,76%	2869,3525	3,6682
		T-:	-52,5	-21,2839	-42,1438	-43,7802	-28,2237
Leukemia4C	BPSO	M :	21,53%	44,17%	44,31%	20724585,25	379,5
		Sd:	3,56%	10,76%	6,63%	57945089,4663	1012,9607
		T-:	-	-	-	-	-
	PSO	M :	**19,44%**	**22,67%**	32,33%	17430041,5	760,7
		Sd:	2,93%	4,65%	3,24%	16428369,9522	285,7303
		T-:	-7,1586	-10,4648	-14,336	-0,1828	1,2094
	PSO-FS	M :	**5,56%**	**18,57%**	**25,79%**	25305,5	11,35
		Sd:	0,00%	7,01%	6,86%	2575,6222	5,1681
		T-:	-184,0001	-12,4642	-18,9951	-1,1547	-1,175

Table E.1: BPSO compared to PSO and PSO-FS (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	BPSO	M :	14,26%	17,41%	26,84%	1446728,5455	36,8182
		Sd:	1,54%	4,10%	2,02%	184354,5961	9,8875
		T-:	-	-	-	-	-
	PSO	M :	**10,52%**	**8,21%**	**23,62%**	105386371,9091	1101
		Sd:	0,99%	0,65%	1,76%	57521474,0433	337,3319
		T-:	-25,8266	-46,0182	-10,2054	30,125	57,6595
	PSO-FS	M :	**6,05%**	**16,52%**	**23,89%**	**102631,16**	**14,08**
		Sd:	1,55%	4,46%	2,42%	24304,5501	6,543
		T-:	-41,4679	-2,1042	-8,6982	-96,8803	-45,9235
MLL	BPSO	M :	11,11%	29,31%	34,03%	458212,5	31,75
		Sd:	0,00%	7,13%	7,60%	150315,2132	13,2004
		T-:	-	-	-	-	-
	PSO	M :	**10,10%**	**13,94%**	**21,26%**	17698358,8182	756,6364
		Sd:	2,25%	3,27%	5,85%	15757423,9802	292,565
		T-:	-11	-16,0065	-4,5387	4,973	14,4869
	PSO-FS	M :	**0,00%**	**19,15%**	**21,20%**	**43318,53**	**4**
		Sd:	0,00%	4,85%	6,20%	12826,3581	1,1547
		T-:	-	-9,9515	-4,5483	-30,4622	-6,4348
Ovarian	BPSO	M :	7,81%	5,51%	21,32%	4641943,1	28
		Sd:	1,95%	4,10%	8,00%	3164713,1098	12,7454
		T-:	-	-	-	-	-
	PSO	M :	**5,47%**	**1,94%**	**5,79%**	123156899,5	1060,7
		Sd:	1,10%	0,65%	2,42%	65328623,9522	297,2063
		T-:	-30	-16,3954	-31,7583	45,1785	95,8207
	PSO-FS	M :	**0,00%**	**2,37%**	**3,82%**	**77316,38**	**4,30**
		Sd:	0,00%	0,86%	0,75%	16833,6634	2,1364
		T-:	-	-13,9139	-36,1982	-22,348	-56,3231
SRBCT	BPSO	M :	26,98%	19,51%	50,03%	76788,3333	36,7778
		Sd:	5,83%	9,24%	7,60%	8630,6454	7,1725
		T-:	-	-	-	-	-
	PSO	M :	**24,76%**	**5,71%**	**44,34%**	1479923	288,2
		Sd:	3,01%	2,12%	3,78%	879876,815	95,5124
		T-:	-11,0868	-7,2702	-8,5452	30,8607	26,4798
	PSO-FS	M :	**0,00%**	**4,81%**	**31,05%**	**15772**	**10,1429**
		Sd:	0,00%	3,73%	4,06%	2505,4457	2,6561
		T-:	-153	-7,751	-40,8207	-321,6248	-49,8424
9_Tumors	BPSO	M :	64,67%	71,13%	93,60%	35908875,1	747,7
		Sd:	3,22%	13,61%	2,27%	48043880,5583	929,2852
		T-:	-	-	-	-	-
	PSO	M :	**58,67%**	**67,07%**	93,73%	**3319807,6**	**347**
		Sd:	4,22%	6,99%	1,81%	5329301,5329	256,8078
		T-:	-8,4665	-2,8633	0,8839	-9,1163	-5,6213
	PSO-FS	M :	**25,56%**	77,67%	93,78%	**43110**	**26,75**
		Sd:	6,25%	7,40%	1,49%	7192,1514	12,7002
		T-:	-82,207	4,593	1,059	-10,034	-10,1427

Table E.2: BPSO compared to PSO and PSO-FS (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
14_Tumors	BPSO	M :	68,83%	60,26%	89,09%	2832168	48
		Sd:	0	0	0	0	0
		T-:	-	-	-	-	-
	PSO	M :	**66,10%**	**54,13%**	**87,57%**	449823181,3	2070,4
		Sd:	1,78%	19,60%	0,78%	243839378,7599	707,9832
		T-:	-210	-4,7773	-90	56,5711	245,4369
	PSO-FS	M :	**63,35%**	**53,19%**	**86,09%**	5568428,11	246,66
		Sd:	1,81%	5,27%	0,93%	4446302,803	158,4408
		T-:	-13,68	-6,0912	-98,5263	5,3523	9,477
Brain Tumor2	BPSO	M :	30,77%	58,53%	55,47%	191247482,6667	2245,1667
		Sd:	0,00%	13,27%	7,65%	324384557,678	3471,6888
		T-:	-	-	-	-	-
	PSO	M :	30,77%	**46,64%**	55,20%	**2047552**	**366,8**
		Sd:	0,00%	6,54%	7,69%	1818921,7408	141,1034
		T-:	-	-6,6138	-0,4615	-5,9406	-5,1311
	PSO-FS	M :	**6,84%**	**47,91%**	59,56%	**30961,66**	**6,44**
		Sd:	2,56%	9,87%	5,29%	2334,1047	2,9202
		T-:	-252	-6,3249	5,3797	-6,004	-6,119
Prostate Tumor	BPSO	M :	3,85%	-100,00%	29,41%	489730425	7516
		Sd:	0	0	0	0	0
		T-:	-	-	-	-	-
	PSO	M :	**1,54%**	11,61%	36,04%	**31462474,7**	**854,7**
		Sd:	1,99%	3,12%	6,28%	15709491,2227	217,5035
		T-:	-10	296,4583	27,7049	-870,8123	-603,9257
	PSO-FS	M :	**2,56%**	17,82%	**19,56%**	**32435,33**	**7,55**
		Sd:	1,92%	5,91%	4,77%	2194,9423	3,5746
		T-:	-4,5	608,4	-38,3774	-3051429,8918	-19005,75

Table E.3: BPSO compared to PSO and PSO-FS (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Sonar	BPSO	M :	14,50%	26,05%	41,95%	100937,6923	34,5385
		Sd:	3,00%	3,11%	2,97%	33901,6316	11,3477
		T-:	-	-	-	-	-
	PSO	M :	**13,21%**	27,74%	**40,53%**	**48806**	**25,8**
		Sd:	2,04%	2,71%	2,49%	29047,0642	4,9454
		T-:	-2,4668	11,5327	-12,573	-16,4911	-7,6207
	PSO-FS	M :	**12,39%**	28,85%	**37,97%**	**18753,55**	**15,33**
		Sd:	1,70%	3,79%	2,60%	6358,479	6,1237
		T-:	-3,9198	3,6266	-19,6928	-28,8949	-13,8796
Spam Base	BPSO	M :	11,52%	9,00%	16,02%	6317613,7143	47,5714
		Sd:	1,08%	0,55%	0,85%	1862506,8041	3,7358
		T-:	-	-	-	-	-
	PSO	M :	13,67%	9,38%	15,85%	**2932467,15**	**42,38**
		Sd:	1,14%	0,68%	1,44%	1459185,3135	6,3053
		T-:	13,7755	5,3882	-1,5307	-51,3016	-102,89
	PSO-FS	M :	**10,63%**	**8,81%**	**14,88%**	**1362338,77**	**31,33**
		Sd:	1,18%	0,26%	0,97%	307982,4086	4,8218
		T-:	-20,7163	-4,6926	-8,9991	-102,2963	-85,6304
Soybean	BPSO	M :	11,11%	8,43%	58,82%	1621154,2857	31,5714
		Sd:	0,65%	2,17%	0,25%	489379,7182	1,7852
		T-:	-	-	-	-	-
	PSO	M :	14,77%	**6,90%**	59,55%	1685985,625	**28,12**
		Sd:	0,81%	0,40%	1,67%	696435,2705	1,7842
		T-:	85,4787	-16,0896	8,5798	1,1967	-24,8046
	PSO-FS	M :	12,87%	**6,78%**	**58,74%**	**878965,66**	**24,22**
		Sd:	0,97%	0,88%	0,32%	144287,2724	2,9486
		T-:	12,8524	-17,1175	-1,7628	-30,9545	-29,3658
Arrhythmia	BPSO	M :	35,00%	39,84%	43,53%	7026933,4545	181,4545
		Sd:	1,21%	1,52%	1,48%	4575170,306	68,468
		T-:	-	-	-	-	-
	PSO	M :	**34,85%**	41,84%	**42,51%**	**2409928,84**	**118,84**
		Sd:	1,17%	1,38%	1,09%	1804344,2202	36,2097
		T-:	-2,7855	10,1943	-4,2984	-8,922	-6,8885
	PSO-FS	M :	**30,78%**	41,28%	**42,63%**	**202856,44**	**24,44**
		Sd:	2,25%	2,05%	1,59%	54388,1703	7,9861
		T-:	-32,9778	10,7636	-3,9413	-14,5123	-23,221
Secom	BPSO	M :	6,74%	7,19%	8,59%	5484822,5556	26,6667
		Sd:	0,19%	1,19%	0,61%	7644231,3287	55,736
		T-:	-	-	-	-	-
	PSO	M :	**6,43%**	9,32%	**7,87%**	850215,35	34,5
		Sd:	0,25%	1,04%	0,22%	447837,9589	13,2824
		T-:	-8,4237	6,0767	-6,0103	-2,0512	0,4749
	PSO-FS	M :	6,82%	6,61%	10,11%	**54114,71**	2,28
		Sd:	0,14%	0,01%	2,07%	15041,5406	1,2536
		T-:	2,4791	-1,6429	4,7949	-2,4035	-1,4791
semeion	BPSO	M :	16,92%	7,75%	81,11%	36899082,7273	231,0909
		Sd:	0,67%	0,27%	0,16%	10396037,6897	12,6843
		T-:	-	-	-	-	-
	PSO	M :	**15,57%**	8,49%	81,16%	**15039303,91**	**185**
		Sd:	0,94%	0,46%	0,23%	6020993,35	17,1199
		T-:	-8,7289	49,1441	6,6442	-22,0427	-33,8262
	PSO-FS	M :	**15,04%**	9,04%	81,22%	**8725772,66**	**161,77**
		Sd:	0,96%	0,91%	0,15%	2227476,7364	22,3985
		T-:	-17,414	58,6318	10,6652	-30,7131	-72,93

Table E.4: BPSO compared to PSO and PSO-FS (4)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	PSO	M :	14,40%	45,50%	48,21%	577010,2	118,5
		Sd:	5,06%	7,74%	4,03%	759776,2244	111,1778
		T-:	-	-	-	-	-
	PSO(X2)	M :	**6,00%**	44,89%	**45,50%**	**130109,5**	**5,73**
		Sd:	3,05%	6,32%	7,07%	29140,1266	3,8216
		T-:	-13,0312	-0,5658	-12,7053	-10,5298	-24,2472
	PSO(R)	M :	**6,46%**	44,36%	**46,47%**	**102542,30**	**5,38**
		Sd:	2,60%	4,43%	7,51%	17754,0898	2,256
		T-:	-12,1971	-1,0768	-2,0433	-11,1781	-24,3195
	PSO(SU)	M :	**5,54%**	**41,63%**	**45,22%**	**97223,84**	**6,07**
		Sd:	3,07%	5,71%	7,34%	18571,835	4,0919
		T-:	-13,6153	-3,3109	-2,9579	-11,2947	-24,1732
CNS	PSO	M :	24,00%	42,60%	40,20%	26223,4	25,5
		Sd:	7,83%	6,07%	2,55%	4920,4688	6,4161
		T-:	-	-	-	-	-
	PSO(X2)	M :	**5,52%**	48,18%	**38,57%**	**21359,10**	**7,03**
		Sd:	3,59%	8,35%	6,03%	3240,2455	3,3962
		T-:	-45,9817	11,8342	-25,6655	-8,4702	-38,2843
	PSO(R)	M :	**4,29%**	42,29%	**45,00%**	**19902,07**	**6,5**
		Sd:	3,31%	7,76%	4,84%	2440,6061	2,8756
		T-:	-39,1392	-0,527	56,2576	-11,1025	-39,2452
	PSO(SU)	M :	**5,13%**	**39,90%**	**38,21%**	**18925,38**	**6,38**
		Sd:	3,99%	7,83%	9,38%	2081,4884	2,9023
		T-:	-45,241	-4,4166	-24,4169	-11,0689	-25,8775
Colon	PSO	M :	13,54%	24,95%	31,94%	13023,3333	16,8333
		Sd:	4,49%	3,77%	5,55%	2203,0788	5,7814
		T-:	-	-	-	-	-
	PSO(X2)	M :	**1,72%**	**22,47%**	**22,78%**	**10954,17**	**7,93**
		Sd:	2,84%	7,32%	7,84%	1139,6587	3,6246
		T-:	-66,1745	-4,9147	-15,2147	-28,3555	-28,358
	PSO(R)	M :	**3,57%**	**18,43%**	**22,95%**	**12032,78**	**15,92**
		Sd:	3,21%	5,04%	3,80%	931,788	1,8172
		T-:	-47,4552	-12,5473	-31,813	-8,9329	-3,4279
	PSO(SU)	M :	**4,02%**	**23,82%**	**26,45%**	13867,92	21,7857
		Sd:	4,66%	5,03%	5,98%	3442,1121	9,0569
		T-:	-52,4619	-6,0324	-18,4339	8,0522	18,7354
Leukemia3C	PSO	M :	14,44%	35,61%	39,06%	218427,8	110,6
		Sd:	2,87%	12,84%	13,28%	322195,4758	78,3684
		T-:	-	-	-	-	-
	PSO(X2)	M :	**0,00%**	**4,76%**	**15,27%**	**31223,76**	**24,04**
		Sd:	0,00%	5,64%	5,28%	7299,1366	10,8031
		T-:	-65	-46,5161	-24,5602	-11,0189	-15,2622
	PSO(R)	M :	**0,00%**	**5,20%**	**13,61%**	**23559,42**	**10,92**
		Sd:	0,00%	3,93%	4,76%	2869,3525	3,6682
		T-:	-65	-44,5055	-27,3689	-11,471	-17,6082
	PSO(SU)	M :	**0,00%**	**6,84%**	**15,47%**	**22457,92**	**11,84**
		Sd:	0,00%	8,63%	5,33%	3629,9319	7,5922
		T-:	-65	-43,5143	-24,547	-11,5357	-17,41

Table E.5: PSO-FS: filter impact (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU	# Attrib
Leukemia4C	PSO	M :	19,44%	22,67%	32,33%	17430041,5	760,7
		Sd:	2,93%	4,65%	3,24%	16428369,9522	285,7303
		T-:	-	-	-	-	-
	PSO(X2)	M :	**5,56%**	**13,91%**	**18,51%**	**65334,06**	**57,62**
		Sd:	0,00%	4,07%	2,04%	23079,1421	19,6731
		T-:	-50	-50,4243	-75,4763	-9,4091	-20,3563
	PSO(R)	M :	**5,56%**	**18,57%**	**25,79%**	**25305,5**	**11,3571**
		Sd:	0,00%	7,01%	6,86%	2575,6222	5,1681
		T-:	-50	-19,4608	-11,6416	-9,4307	-21,7012
	PSO(SU)	M :	**5,56%**	**20,74%**	**24,54%**	**23764,91**	**11,66**
		Sd:	0,00%	7,24%	7,51%	3886,0856	8,1613
		T-:	-50	-12,1442	-10,1345	-9,4316	-21,6687
Lung	PSO	M :	10,52%	8,21%	23,62%	105386371,9091	1101
		Sd:	0,99%	0,65%	1,76%	57521474,0433	337,3319
		T-:	-	-	-	-	-
	PSO(X2)	M :	**5,88%**	8,92%	**20,56%**	**859882,72**	**134,2**
		Sd:	0,00%	2,14%	1,41%	704337,2295	79,7329
		T-:	-71,5	27,9319	-67,4531	-30,294	-50,6959
	PSO(R)	M :	**6,05%**	16,52%	23,89%	**102631,16**	**14,08**
		Sd:	1,55%	4,46%	2,42%	24304,5501	6,543
		T-:	-27,3959	22,2895	2,1188	-30,5148	-58,8815
	PSO(SU)	M :	**5,23%**	10,23%	23,47%	**3039050,25**	**171,41**
		Sd:	0,97%	5,37%	2,36%	8038525,2874	217,0406
		T-:	-62,4602	4,4283	-0,6342	-25,2937	-17,1251
MLL	PSO	M :	10,10%	13,94%	21,26%	17698358,8182	756,6364
		Sd:	2,25%	3,27%	5,85%	15757423,9802	292,565
		T-:	-	-	-	-	-
	PSO(X2)	M :	**0,00%**	15,10%	**12,68%**	**53434,51**	**6,82**
		Sd:	0,00%	6,06%	5,50%	18741,0901	2,5223
		T-:	-110	5,5962	-408,5647	-5,0898	-15,0409
	PSO(R)	M :	**0,00%**	19,15%	21,20%	**43318,53**	**4**
		Sd:	0,00%	4,85%	6,20%	12826,3581	1,1547
		T-:	-110	11,465	-0,2993	-5,0927	-15,0976
	PSO(SU)	M :	**0,00%**	16,37%	23,25%	**39921,69**	**10,46**
		Sd:	0,00%	4,08%	7,72%	5011,7775	7,1369
		T-:	-110	6,3906	1,9479	-5,0937	-14,9679
Ovarian	PSO	M :	5,47%	1,94%	5,79%	123156899,5	1060,7
		Sd:	1,10%	0,65%	2,42%	65328623,9522	297,2063
		T-:	-	-	-	-	-
	PSO(X2)	M :	**0,00%**	3,13%	**5,41%**	**56532,29**	**2,81**
		Sd:	0,00%	1,52%	1,48%	14053,6897	0,9214
		T-:	-70	33,2185	-4,3433	-47,0693	-98,2248
	PSO(R)	M :	**0,00%**	2,37%	**3,82%**	**77316,38**	**4,30**
		Sd:	0,00%	0,86%	0,75%	16833,6634	2,1364
		T-:	-70	7,0354	-19,8025	-47,0614	-98,0794
	PSO(SU)	M :	**0,00%**	3,28%	**4,92%**	**49072,78**	**4,28**
		Sd:	0,00%	1,06%	1,18%	6420,8494	1,7728
		T-:	-70	20,546	-9,9167	-47,0722	-98,0773

Table E.6: PSO-FS: filter impact (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
SRBCT	PSO	M :	24,76%	5,71%	44,34%	1479923	288,2
		Sd:	3,01%	2,12%	3,78%	879876,815	95,5124
		T-:	-	-	-	-	-
	PSO(X2)	M :	**0,00%**	**2,67%**	**33,14%**	**43836,06**	**45,6**
		Sd:	0,00%	1,80%	3,80%	16531,3243	17,4427
		T-:	-259,9999	-27,5833	-22,1878	-31,5857	-25,5907
	PSO(R)	M :	**0,00%**	**4,81%**	**31,05%**	**15772**	**10,14**
		Sd:	0,00%	3,73%	4,06%	2505,4457	2,6561
		T-:	-259,9999	-10,8209	-27,1427	-32,2029	-29,3303
	PSO(SU)	M :	**0,00%**	6,25%	**33,48%**	**36055,45**	**31,45**
		Sd:	0,00%	2,97%	4,09%	43593,0275	34,1624
		T-:	-259,9999	1,9045	-16,8416	-31,7222	-26,3104
9_Tumors	PSO	M :	58,67%	67,07%	93,73%	3319807,6	347
		Sd:	4,22%	6,99%	1,81%	5329301,5329	256,8078
		T-:	-	-	-	-	-
	PSO(X2)	M :	**29,14%**	66,99%	94,00%	**72648,25**	**46,85**
		Sd:	3,28%	3,92%	1,43%	24721,4531	13,9579
		T-:	-54,5739	-0,6888	2,5	-62,5799	-54,5301
	PSO(R)	M :	**25,56%**	77,67%	93,78%	**43110**	**26,75**
		Sd:	6,25%	7,40%	1,49%	7192,1514	12,7002
		T-:	-61,1684	121,1484	0,2647	-63,148	-56,8255
	PSO(SU)	M :	**28,33%**	73,72%	94,22%	**46030,5**	**29,83**
		Sd:	4,14%	7,82%	1,71%	13055,8242	15,2246
		T-:	-44,8189	13,9355	3,4612	-63,089	-56,0108
11_Tumors	PSO	M :	25,23%	22,53%	72,90%	234685193,8	1872,2
		Sd:	1,29%	24,53%	1,35%	109706988,8342	534,4619
		T-:	-	-	-	-	-
	PSO(X2)	M :	**13,89%**	**12,43%**	**71,32%**	**693765,22**	**126,51**
		Sd:	2,70%	4,47%	1,53%	490053,6127	59,5117
		T-:	-55,3742	-6,4503	-11,1563	-28,8709	-45,7824
	PSO(R)	M :	**17,15%**	**6,44%**	**69,78%**	**10463843,36**	**491**
		Sd:	1,19%	1,17%	1,91%	6882215,4065	127,5429
		T-:	-35,8925	-10,2904	-19,2586	-27,5738	-34,1158
	PSO(SU)	M :	**17,90%**	**7,39%**	**72,56%**	**13698938,87**	**562,87**
		Sd:	1,90%	1,51%	1,78%	7599803,7471	138,1629
		T-:	-22,7684	-9,5876	-2,1125	-26,9806	-28,5542
14_Tumors	PSO	M :	66,10%	54,13%	87,57%	449823181,3	2070,4
		Sd:	1,78%	19,60%	0,78%	243839378,7599	707,9832
		T-:	-	-	-	-	-
	PSO(X2)	M :	**58,04%**	65,46%	**85,57%**	**650084,42**	**48,61**
		Sd:	2,49%	4,00%	1,07%	241460,8851	27,5232
		T-:	-218,2503	8,8287	-21,7398	-56,8472	-244,364
	PSO(R)	M :	**63,35%**	53,19%	**86,09%**	**5568428,11**	**246,66**
		Sd:	1,81%	5,27%	0,93%	4446302,803	158,4408
		T-:	-6,8724	-0,5437	-42,5497	-56,1075	-80,9674
	PSO(SU)	M :	**58,60%**	57,53%	**86,77%**	**1063387,75**	**95**
		Sd:	2,35%	6,10%	1,63%	792778,0725	67,4516
		T-:	-311,3004	2,458	-5,3789	-56,7757	-111,8282

Table E.7: PSO-FS: filter impact (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Brain Tumor2	PSO	M :	30,77%	46,64%	55,20%	2047552	366,8
		Sd:	0,00%	6,54%	7,69%	1818921,7408	141,1034
		T-:	-	-	-	-	-
	PSO(X2)	M :	**7,69%**	48,90%	56,60%	**41058**	**4,12**
		Sd:	4,54%	11,63%	5,20%	8980,4589	2,0283
		T-:	-72	2,903	5,4194	-14,2096	-29,2009
	PSO(R)	M :	**6,84%**	47,91%	59,56%	**30961,66**	**6,44**
		Sd:	2,56%	9,87%	5,29%	2334,1047	2,9202
		T-:	-252	1,4292	8,82	-14,2811	-29,0117
	PSO(SU)	M :	**8,55%**	45,96%	55,29%	**30155,77**	**7**
		Sd:	4,62%	8,10%	4,20%	5244,7227	2,5981
		T-:	-29,25	-0,3947	0,5294	-14,2868	-28,9648
Prostate Tumor	PSO	M :	1,54%	11,61%	36,04%	31462474,7	854,7
		Sd:	1,99%	3,12%	6,28%	15709491,2227	217,5035
		T-:	-	-	-	-	-
	PSO(X2)	M :	1,23%	17,71%	**22,02%**	**39862,44**	**5,6**
		Sd:	1,83%	6,30%	6,02%	7960,1359	2,708
		T-:	-1,2144	15,8462	-55,4992	-59,71	-76,9744
	PSO(R)	M :	2,56%	17,82%	**19,56%**	**32435,33**	**7,55**
		Sd:	1,92%	5,91%	4,77%	2194,9423	3,5746
		T-:	2,7974	14,6767	-46,9629	-59,7241	-76,7544
	PSO(SU)	M :	**0,00%**	22,02%	**25,88%**	**40458,14**	**19,28**
		Sd:	0,00%	9,01%	7,85%	13718,2349	14,2912
		T-:	-6,6667	13,3878	-9,8002	-59,7087	-75,3078
Lymphoma	PSO	M :	5,88%	2,09%	16,31%	476913,5455	193,6364
		Sd:	3,72%	2,07%	2,80%	316990,5845	64,7337
		T-:	-	-	-	-	-
	PSO(X2)	M :	**0,00%**	11,00%	**13,47%**	**17082,11**	**7,62**
		Sd:	0,00%	4,40%	4,81%	3210,0859	4,2621
		T-:	-11	46,5532	-19,0414	-19,6224	-32,6658
	PSO(R)	M :	**0,00%**	13,80%	**17,20%**	**13130,92**	**2,53**
		Sd:	0,00%	6,53%	6,81%	1177,8929	1,3914
		T-:	-11	45,1884	6,4974	-19,7911	-33,5594
	PSO(SU)	M :	**0,00%**	14,59%	**16,06%**	**12316,42**	**2,64**
		Sd:	0,00%	6,69%	6,14%	630,5459	1,0818
		T-:	-11	24,2297	-1,8748	-19,8259	-33,5368

Table E.8: PSO-FS: filter impact (4)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Sonar	PSO	M :	13,21%	27,74%	40,53%	48806	25,8
		Sd:	2,04%	2,71%	2,49%	29047,0642	4,9454
		T-:	-	-	-	-	-
	PSO(X2)	M :	**11,97%**	29,81%	**38,82%**	**17642,70**	**13,33**
		Sd:	1,80%	2,94%	1,92%	3854,8746	3,8431
		T-:	-21,3768	12,7971	-21,0284	-20,026	-49,1522
	PSO(R)	M :	**12,39%**	28,85%	**37,97%**	**18753,55**	**15,33**
		Sd:	1,70%	3,79%	2,60%	6358,479	6,1237
		T-:	-6,7702	1,4109	-12,9701	-17,5639	-12,2663
	PSO(SU)	M :	**12,39%**	28,10%	**38,97%**	**17513,33**	**13,88**
		Sd:	1,70%	2,17%	1,40%	2840,1292	2,3154
		T-:	-2,6267	0,7275	-17,4658	-19,9511	-46,9618
Spam Base	PSO	M :	13,67%	9,38%	15,85%	2932467,1538	42,3846
		Sd:	1,14%	0,68%	1,44%	1459185,3135	6,3053
		T-:	-	-	-	-	-
	PSO(X2)	M :	**10,81%**	10,79%	**14,84%**	**371474,44**	**10,84**
		Sd:	0,66%	1,15%	1,84%	113243,2237	3,0643
		T-:	-18,8552	23,1142	-7,396	-54,7106	-397,6685
	PSO(R)	M :	**10,63%**	**8,81%**	**14,88%**	**1362338,77**	**31,33**
		Sd:	1,18%	0,26%	0,97%	307982,4086	4,8218
		T-:	-20,1807	-9,2225	-14,7035	-32,3276	-58,9296
	PSO(SU)	M :	**10,72%**	10,10%	**14,04%**	**573159,55**	**14,77**
		Sd:	0,71%	1,16%	1,47%	234552,3861	5,7831
		T-:	-19,4845	6,5074	-10,3859	-40,04	-39,8948
Soybean	PSO	M :	14,77%	6,90%	59,55%	1685985,625	28,125
		Sd:	0,81%	0,40%	1,67%	696435,2705	1,7842
		T-:	-	-	-	-	-
	PSO(X2)	M :	**12,45%**	**5,95%**	**58,81%**	**923356,29**	**19,66**
		Sd:	1,14%	0,61%	0,23%	243657,8256	2,3697
		T-:	-39,0108	-147,1236	-9,8302	-15,6669	-59,6934
	PSO(R)	M :	**12,87%**	**6,78%**	**58,74%**	**878965,66**	**24,22**
		Sd:	0,97%	0,88%	0,32%	144287,2724	2,9486
		T-:	-14,589	-6,9534	-10,4588	-16,6105	-13,9203
	PSO(SU)	M :	**13,19%**	**6,35%**	**58,84%**	**877601,88**	**20,77**
		Sd:	0,59%	0,47%	0,26%	135333,8116	2,6352
		T-:	-16,7095	-55,8397	-9,4694	-16,0724	-21,8664

Table E.9: PSO-FS: filter impact (5)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Arrhytmia	PSO	M :	34,85%	41,84%	42,51%	2409928,8462	118,8462
		Sd:	1,17%	1,38%	1,09%	1804344,2202	36,2097
		T-:	-	-	-	-	-
	PSO(X2)	M :	**29,95%**	**38,59%**	43,13%	**330978,38**	**39,46**
		Sd:	1,14%	1,94%	1,31%	100660,2858	11,8566
		T-:	-183,5646	-20,0276	7,2856	-9,6167	-12,8678
	PSO(R)	M :	**30,78%**	**41,28%**	42,63%	**202856,44**	**24,44**
		Sd:	2,25%	2,05%	1,59%	54388,1703	7,9861
		T-:	-33,1456	-3,1734	1,3524	-10,2072	-15,0888
	PSO(SU)	M :	**27,63%**	**36,75%**	42,82%	**164293**	**21,11**
		Sd:	1,06%	1,69%	0,92%	55652,0992	8,0069
		T-:	-211,815	-30,176	1,7123	-10,3879	-15,7562
Secom	PSO	M :	6,43%	9,32%	7,87%	850215,3571	34,5
		Sd:	0,25%	1,04%	0,22%	447837,9589	13,2824
		T-:	-	-	-	-	-
	PSO(X2)	M :	6,44%	**6,77%**	9,86%	**69585,32**	**4,4**
		Sd:	0,20%	0,12%	1,27%	17333,3361	2,0817
		T-:	0,3722	-109,4951	38,1771	-62,1547	-43,9205
	PSO(R)	M :	6,82%	**6,61%**	10,11%	**54114,71**	**2,28**
		Sd:	0,14%	0,01%	2,07%	15041,5406	1,2536
		T-:	15,4952	-118,5639	7,6361	-62,1006	-44,6545
	PSO(SU)	M :	**6,20%**	**6,78%**	10,34%	**66641,37**	**4,37**
		Sd:	0,27%	0,12%	1,80%	21232,9847	2,0659
		T-:	-5,4223	-76,7783	24,1928	-62,0962	-42,5301
semeion	PSO	M :	15,57%	8,49%	81,16%	15039303,9167	185
		Sd:	0,94%	0,46%	0,23%	6020993,35	17,1199
		T-:	-	-	-	-	-
	PSO(X2)	M :	**14,97%**	8,99%	81,17%	**9103508,16**	**157,04**
		Sd:	1,02%	0,75%	0,16%	2681529,3461	19,7978
		T-:	-4,0495	12,9834	1,0845	-14,9188	-18,9787
	PSO(R)	M :	**15,04%**	9,04%	81,22%	**8725772,66**	**161,77**
		Sd:	0,96%	0,91%	0,15%	2227476,7364	22,3985
		T-:	-3,295	25,9422	5,0679	-15,6321	-19,6707
	PSO(SU)	M :	**14,45%**	9,53%	81,13%	**6897631,22**	**147,55**
		Sd:	1,05%	1,27%	0,16%	1832543,704	25,6228
		T-:	-7,4023	17,4509	-1,8221	-20,7458	-34,1324

Table E.10: PSO-FS: filter impact (6)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	PSO	M :	14,80%	40,79%	50,21%	14989102,5	258,7
		Sd:	6,27%	8,29%	3,43%	15819583,1527	248,8641
		T-:	-	-	-	-	-
	PSO-FS	M :	**5,45%**	43,66%	**48,11%**	**6523488,90**	**8,77**
		Sd:	2,32%	5,08%	7,07%	974605,1308	5,5027
		T-:	-32,4833	2,8927	-12,9833	-7,9275	-13,6052
	PSO-FS2	M :	**5,90%**	43,93%	**47,62%**	**185937,05**	**6,27**
		Sd:	2,39%	5,08%	4,76%	54593,0167	3,2737
		T-:	-31,2413	3,1797	-23,9426	-13,8624	-13,7411
CNS	PSO	M :	14,00%	42,67%	39,27%	594637,2	34,6
		Sd:	5,84%	3,96%	5,30%	357285,9256	10,8853
		T-:	-	-	-	-	-
	PSO-FS	M :	**6,11%**	43,83%	39,89%	626050,95	**10,08**
		Sd:	5,53%	7,11%	7,13%	41364,6168	4,7996
		T-:	-9,9522	1,9433	5,4332	11,7969	-68,0987
	PSO-FS2	M :	**4,81%**	44,59%	**38,19%**	**45227,41**	**7,30**
		Sd:	3,36%	6,87%	6,58%	14591,6518	3,6354
		T-:	-12,3939	3,1546	-4,9445	-241,4646	-73,0577
Colon	PSO	M :	9,66%	24,34%	31,55%	98361,7273	20,0909
		Sd:	3,26%	4,79%	4,25%	47964,3065	3,5058
		T-:	-	-	-	-	-
	PSO-FS	M :	**1,30%**	**22,39%**	**26,75%**	**63126,29**	20,33
		Sd:	2,59%	5,00%	3,39%	5767,6176	7,4464
		T-:	-26,5614	-2,9332	-26,4308	-27,1004	1,4205
	PSO-FS2	M :	**3,78%**	**21,85%**	**25,43%**	**19526,95**	**13,41**
		Sd:	3,89%	6,69%	6,78%	6988,6066	6,9906
		T-:	-18,6549	-3,7732	-33,0296	-60,545	-66,9573
Leukemia3C	PSO	M :	16,11%	28,17%	40,44%	6904625,6	311,2
		Sd:	1,76%	11,08%	8,46%	9670346,2715	303,2439
		T-:	-	-	-	-	-
	PSO-FS	M :	**0,00%**	**7,07%**	**12,07%**	**679008,04**	**16,40**
		Sd:	0,00%	7,10%	3,56%	21224,5251	8,5449
		T-:	-32,2222	-13,025	-68,4997	-99,8387	-29,2329
	PSO-FS2	M :	**0,00%**	**5,73%**	**13,23%**	**59425,78**	**14,47**
		Sd:	0,00%	6,67%	6,08%	35213,4072	12,4572
		T-:	-32,2222	-14,0725	-65,7358	-109,7752	-29,4344
Leukemia4C	PSO	M :	17,78%	23,89%	35,44%	39871876,9	935,5
		Sd:	2,34%	3,76%	4,47%	21888598,925	235,0316
		T-:	-	-	-	-	-
	PSO-FS	M :	**5,56%**	**21,24%**	**21,57%**	**658471,13**	**25,68**
		Sd:	0,00%	6,72%	5,27%	72146,103	12,6581
		T-:	-110	-33,5397	-55,2175	-59,0573	-120,4083
	PSO-FS2	M :	**5,56%**	**17,76%**	**23,14%**	**58339,39**	**17,5**
		Sd:	0,00%	6,37%	5,41%	20214,3319	11,3727
		T-:	-110	-25,5185	-50,8836	-59,9612	-121,5848

Table E.11: PSO-FS: Multi-filters (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	PSO	M :	10,78%	8,55%	22,87%	262280152,6	1617,4
		Sd:	1,91%	0,67%	1,89%	119376285,4008	367,9068
		T-:	-	-	-	-	-
	PSO-FS	M :	**4,96%**	9,04%	23,23%	**24327274,11**	**452,76**
		Sd:	1,01%	4,46%	1,80%	35803172,709	494,1124
		T-:	-50,43	0,6854	1,3282	-36,8833	-32,8044
	PSO-FS2	M :	**5,83%**	8,79%	**20,74%**	**3277129,83**	**178,51**
		Sd:	0,32%	3,32%	1,71%	3779166,1201	126,0449
		T-:	-50,5351	2,3275	-8,5794	-41,0104	-58,012
MLL	PSO	M :	11,11%	20,89%	23,11%	110196512,3	1817,1
		Sd:	0,00%	15,21%	5,13%	63219803,2856	657,5251
		T-:	-	-	-	-	-
	PSO-FS	M :	**0,00%**	**14,66%**	23,02%	**2244662,38**	**58,44**
		Sd:	0,00%	6,17%	4,39%	109637,0996	44,4075
		T-:	-	-7,0578	-0,2632	-639,8892	-47,1393
	PSO-FS2	M :	**0,00%**	**14,72%**	**14,87%**	**99694,33**	**7,30**
		Sd:	0,00%	5,36%	7,02%	68273,9982	3,7252
		T-:	-	-7,0806	-49,554	-653,0878	-48,6372
Ovarian	PSO	M :	6,64%	11,83%	4,54%	356989523,125	2078,375
		Sd:	0,72%	13,13%	1,89%	117505367,7324	468,4003
		T-:	-	-	-	-	-
	PSO-FS	M :	**0,00%**	**2,80%**	4,54%	**1369084,58**	**4,52**
		Sd:	0,00%	0,40%	1,04%	22981,9931	2,3216
		T-:	-45,3333	-3,0221	-0,0305	-17,7388	-18,3399
	PSO-FS2	M :	**0,00%**	**3,05%**	4,37%	**172204,92**	**3,46**
		Sd:	0,00%	1,32%	1,08%	113087,782	0,9772
		T-:	-45,3333	-2,9384	-0,9042	-17,7985	-18,3494
SRBCT	PSO	M :	17,99%	5,75%	43,47%	2702994	346,3333
		Sd:	5,20%	6,05%	2,93%	1675240,1434	165,2188
		T-:	-	-	-	-	-
	PSO-FS	M :	**0,00%**	**4,79%**	**33,08%**	**120370,12**	**37,56**
		Sd:	0,00%	4,82%	6,47%	82681,815	36,2031
		T-:	-153	-4,5602	-44,6688	-12,5258	-15,909
	PSO-FS2	M :	**0,00%**	5,39%	**31,18%**	**29495,14**	**14,12**
		Sd:	0,00%	3,54%	4,74%	8994,8415	7,5802
		T-:	-153	-1,9213	-50,7726	-12,9666	-17,1164
9_Tumors	PSO	M :	55,56%	71,48%	94,74%	874830	149,8889
		Sd:	5,77%	10,81%	1,35%	489574,27	116,9406
		T-:	-	-	-	-	-
	PSO-FS	M :	**30,53%**	69,82%	**94,28%**	**476369,36**	**50**
		Sd:	4,05%	5,92%	2,18%	51782,9044	18,3242
		T-:	-20,0047	-1,6984	-3,2697	-8,7633	-7,688
	PSO-FS2	M :	**27,48%**	73,07%	**94,00%**	**89345,97**	**31,97**
		Sd:	4,27%	7,43%	2,02%	25118,2624	14,2364
		T-:	-22,5909	1,6544	-5,2587	-17,2751	-9,0785

Table E.12: PSO-FS: Multi-filters (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
11_Tumors	PSO	M :	23,74%	45,72%	72,57%	326287385,5556	2537,8889
		Sd:	2,00%	35,77%	1,47%	179105755,9801	735,6314
		T-:	-	-	-	-	-
	PSO-FS	M :	**18,06%**	**8,01%**	**71,67%**	**29800745**	**748,44**
		Sd:	2,40%	4,90%	1,86%	11605883,9798	241,8688
		T-:	-40,4495	-17,7607	-6,9181	-44,554	-49,5458
	PSO-FS2	M :	**15,45%**	**8,45%**	**71,77%**	**5459088,55**	**271,77**
		Sd:	2,20%	3,20%	2,09%	5178874,5277	134,7704
		T-:	-58,8473	-17,5536	-5,7951	-48,537	-64,7175
14_Tumors	PSO	M :	66,75%	63,12%	87,23%	552404446,2	2526,9
		Sd:	1,10%	17,04%	1,15%	165292469,6776	463,5459
		T-:	0	0	0	0	0
	PSO-FS	M :	**60,82%**	**54,10%**	**86,16%**	**3911628,94**	**141,16**
		Sd:	2,00%	3,34%	0,84%	872627,1664	36,7011
		T-:	-59,9583	-5,902	-4,9113	-52,5189	-154,2766
	PSO-FS2	M :	**58,96%**	62,38%	**85,69%**	**1727936,35**	**74,92**
		Sd:	2,03%	6,66%	1,04%	1237781,8377	52,115
		T-:	-145,5213	-0,4872	-7,0921	-52,7287	-161,1389
Brain Tumor2	PSO	M :	30,77%	45,22%	54,25%	25252029,0952	938
		Sd:	0,00%	3,58%	6,68%	15765058,7788	313,9256
		T-:	-	-	-	-	-
	PSO-FS	M :	**3,37%**	52,50%	58,15%	**1272232**	**11,62**
		Sd:	3,94%	9,35%	4,51%	208896,0895	4,3493
		T-:	-130,2857	57,841	8,0214	-69,6755	-175,0994
	PSO-FS2	M :	**9,57%**	48,55%	57,46%	**72471,34**	**4,85**
		Sd:	5,65%	13,25%	6,31%	22716,5443	2,4654
		T-:	-149,4516	5,2163	66,7848	-73,1902	-176,5398
Prostate Tumor	PSO	M :	1,28%	30,82%	32,75%	181622434,1667	2175,5
		Sd:	1,89%	16,90%	5,28%	70355673,1626	593,58
		T-:	0	0	0	0	0
	PSO-FS	M :	**0,00%**	**11,45%**	**23,75%**	**1646577,81**	**92,25**
		Sd:	0,00%	2,71%	4,37%	348646,3169	71,9949
		T-:	-12	-23,1453	-39,3629	-707,4502	-204,1891
	PSO-FS2	M :	1,25%	**18,43%**	**22,39%**	**77045,60**	**6,02**
		Sd:	1,82%	5,26%	6,62%	23842,5382	3,1281
		T-:	-0,2692	-14,8792	-29,3406	-727,1205	-403,5891
Lymphoma	PSO	M :	6,47%	1,58%	17,45%	7289321,2	454,8
		Sd:	4,34%	1,72%	2,25%	3538625,7463	78,2387
		T-:	-	-	-	-	-
	PSO-FS	M :	**0,00%**	15,62%	18,52%	**224784,5**	**3,22**
		Sd:	0,00%	4,80%	5,66%	3034,4406	1,896
		T-:	-12,2222	80,5281	3,158	-32,248	-1593,91
	PSO-FS2	M :	**0,00%**	12,51%	**14,04%**	**27030,34**	**3,56**
		Sd:	0,00%	5,78%	5,34%	7993,7106	1,8848
		T-:	-12,2222	43,2493	-13,0834	-33,1507	-1596,8734

Table E.13: PSO-FS: Multi-filters (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Sonar	PSO	M :	12,24%	27,19%	40,35%	31605,2727	23,5455
		Sd:	2,76%	2,43%	2,24%	6516,7455	4,298
		T-:	-	-	-	-	-
	PSO-FS	M :	11,97%	28,13%	**38,91%**	**16108,33**	**13,27**
		Sd:	2,04%	2,58%	2,11%	3661,0476	2,3962
		T-:	-1,2355	5,2451	-24,6444	-11,9608	-13,1816
	PSO-FS2	M :	12,11%	27,67%	**39,30%**	32962,75	**14,93**
		Sd:	1,93%	3,19%	1,95%	13987,2921	4,337
		T-:	-1,1359	6,177	-10,2563	1,0507	-11,1887
Spam Base	PSO	M :	14,06%	9,71%	15,62%	2003230	38,1
		Sd:	1,01%	1,13%	2,01%	450056,7096	4,9766
		T-:	-	-	-	-	-
	PSO-FS	M :	**9,83%**	**9,55%**	**13,75%**	**513861,77**	**13,94**
		Sd:	0,94%	0,76%	1,98%	145717,9986	3,7803
		T-:	-36,0939	-2,9991	-9,4577	-47,8746	-72,5492
	PSO-FS2	M :	**10,58%**	10,04%	**14,14%**	**907429,20**	**15,05**
		Sd:	0,77%	1,39%	2,00%	429717,5333	7,3804
		T-:	-28,8566	5,7659	-13,7928	-34,4666	-79,2002
Soybean	PSO	M :	12,76%	7,23%	60,49%	1326505,0909	27,3636
		Sd:	1,92%	0,54%	2,31%	145369,2342	2,1106
		T-:	-	-	-	-	-
	PSO-FS	M :	12,93%	**6,30%**	**58,98%**	**887488,5**	**21,11**
		Sd:	0,44%	0,48%	1,09%	120433,3917	2,1113
		T-:	1,1227	-12,9032	-5,7104	-21,0433	-24,0603
	PSO-FS2	M :	13,06%	**6,20%**	**58,91%**	1526667,09	**20,71**
		Sd:	0,83%	0,74%	0,76%	470670,6078	3,218
		T-:	1,9955	-14,6107	-5,9657	11,7062	-43,6658
Arrhythmia	PSO	M :	33,95%	41,09%	42,72%	1756141,0909	109,3636
		Sd:	1,07%	2,30%	0,55%	761999,0311	17,7329
		T-:	-	-	-	-	-
	PSO-FS	M :	**26,55%**	**36,78%**	42,84%	**182711**	**23,88**
		Sd:	1,46%	1,95%	0,88%	48903,0314	6,8246
		T-:	-252,9999	-16,0584	3,9012	-18,7317	-62,2445
	PSO-FS2	M :	**28,14%**	**37,89%**	43,18%	**354455,17**	**26,77**
		Sd:	1,64%	2,36%	1,25%	113644,6349	9,1329
		T-:	-117,5489	-11,9218	9,3242	-16,6906	-60,8355
Secom	PSO	M :	6,41%	9,34%	7,75%	681233,5	35,5
		Sd:	0,24%	1,13%	0,16%	366131,4789	14,5239
		T-:	-	-	-	-	-
	PSO-FS	M :	**6,06%**	**6,78%**	10,71%	**75904,55**	**4,77**
		Sd:	0,32%	0,11%	0,85%	9476,9187	1,166
		T-:	-9,6912	-85,9831	124,3101	-35,2093	-32,3365
	PSO-FS2	M :	6,36%	**6,84%**	9,67%	**138875,61**	**4,76**
		Sd:	0,28%	0,23%	1,55%	57824,3409	2,9061
		T-:	-1,6572	-85,9674	78,8573	-31,3938	-32,3476
semeion	PSO	M :	14,76%	9,15%	81,17%	9704920,1818	158,2727
		Sd:	1,49%	0,27%	0,16%	1828334,9384	6,4667
		T-:	-	-	-	-	-
	PSO-FS	M :	**13,37%**	9,21%	81,16%	**6984514,77**	**149,27**
		Sd:	0,92%	0,48%	0,18%	914462,7034	8,6691
		T-:	-20,8352	1,5122	-0,784	-24,1223	-9,2728
	PSO-FS2	M :	**14,45%**	**8,95%**	**81,13%**	16115367,55	160,075
		Sd:	0,95%	0,79%	0,18%	5351789,3788	18,4007
		T-:	-7,2709	-5,6025	-2,2559	35,2408	1,8238

Table E.14: PSO-FS: Multi-filters (4)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	CLPSO	M :	14,80%	40,79%	50,21%	14989102,5	258,7
		Sd:	6,27%	8,29%	3,43%	15819583,1527	248,8641
		T-:	-	-	-	-	-
	PSO(R)	M :	**6,46%**	44,36%	46,47%	**102542,30**	**5,38**
		Sd:	2,60%	4,43%	7,51%	17754,0898	2,256
		T-:	-27,4308	3,6661	-4,3698	-13,9405	-13,7894
	CLPSO(R)	M :	**5,70%**	44,24%	**45,37%**	**6966383,54**	**8,57**
		Sd:	2,65%	4,78%	5,38%	1086915,3481	3,9924
		T-:	-31,5457	3,585	-37,3314	-7,5113	-13,6155
CNS	CLPSO	M :	14,00%	42,67%	39,27%	594637,2	34,6
		Sd:	5,84%	3,96%	5,30%	357285,9256	10,8853
		T-:	-	-	-	-	-
	PSO(R)	M :	**4,29%**	42,29%	45,00%	**19902,07**	**6,5**
		Sd:	3,31%	7,76%	4,84%	2440,6061	2,8756
		T-:	-12,2244	-0,452	56,0055	-253,268	-69,9257
	CLPSO(R)	M :	**1,62%**	**40,26%**	43,25%	622270,06	**7,84**
		Sd:	2,90%	7,29%	6,22%	54263,7199	3,3552
		T-:	-16,5309	-3,9594	15,8843	10,5845	-74,121
Colon	CLPSO	M :	9,66%	24,34%	31,55%	98361,7273	20,0909
		Sd:	3,26%	4,79%	4,25%	47964,3065	3,5058
		T-:	-	-	-	-	-
	PSO(R)	M :	**3,57%**	**18,43%**	**22,95%**	**12032,78**	**15,92**
		Sd:	3,21%	5,04%	3,80%	931,788	1,8172
		T-:	-16,7143	-7,1927	-30,9611	-66,2385	-41,9148
	CLPSO(R)	M :	**3,57%**	**19,32%**	**23,98%**	**62126,51**	**14,91**
		Sd:	4,09%	4,66%	4,23%	3412,6605	5,1983
		T-:	-19,07	-7,585	-37,4967	-27,8814	-34,632
Leukemia3C	CLPSO	M :	16,11%	28,17%	40,44%	6904625,6	311,2
		Sd:	1,76%	11,08%	8,46%	9670346,2715	303,2439
		T-:	-	-	-	-	-
	PSO(R)	M :	**0,00%**	**5,20%**	**13,61%**	**23559,42**	**10,92**
		Sd:	0,00%	3,93%	4,76%	2869,3525	3,6682
		T-:	-32,2222	-14,3212	-70,1638	-110,3503	-29,788
	CLPSO(R)	M :	**0,00%**	**6,40%**	**10,59%**	**669026,93**	**10,36**
		Sd:	0,00%	6,53%	5,32%	50593,7262	4,4639
		T-:	-32,2222	-13,5849	-77,0924	-99,7463	-29,8423
Leukemia4C	CLPSO	M :	17,78%	23,89%	35,44%	39871876,9	935,5
		Sd:	2,34%	3,76%	4,47%	21888598,925	235,0316
		T-:	-	-	-	-	-
	PSO(R)	M :	**5,56%**	**18,57%**	**25,79%**	**25305,5**	**11,35**
		Sd:	0,00%	7,01%	6,86%	2575,6222	5,1681
		T-:	-110	-34,932	-17,3851	-60,0109	-122,4008
	CLPSO(R)	M :	**5,56%**	**17,22%**	**23,99%**	**703333,02**	**19,85**
		Sd:	0,00%	5,50%	4,15%	14850,1825	8,8013
		T-:	-110	-67,4772	-68,3489	-58,9898	-121,1695

Table E.15: CLPSO based Relief *vs* random CLPSO and PSO-FS (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	CLPSO	M :	10,78%	8,55%	22,87%	262280152,6	1617,4
		Sd:	1,91%	0,67%	1,89%	119376285,4008	367,9068
		T-:	-	-	-	-	-
	PSO(R)	M :	**6,05%**	16,52%	23,89%	**102631,16**	**14,08**
		Sd:	1,55%	4,46%	2,42%	24304,5501	6,543
		T-:	-26,4702	20,947	3,6805	-41,5174	-65,5923
	CLPSO(R)	M :	**7,14%**	10,85%	23,18%	**2947791,46**	**126,53**
		Sd:	2,15%	4,00%	1,51%	1650031,1773	105,5443
		T-:	-26,669	29,4946	1,1902	-41,0666	-60,508
MLL	CLPSO	M :	11,11%	20,89%	23,11%	110196512,3	1817,1
		Sd:	0,00%	15,21%	5,13%	63219803,2856	657,5251
		T-:	-	-	-	-	-
	PSO(R)	M :	**0,00%**	19,15%	**21,20%**	**43318,53**	**4**
		Sd:	0,00%	4,85%	6,20%	12826,3581	1,1547
		T-:	-	-1,8232	-7,2688	-653,3314	-48,7261
	CLPSO(R)	M :	**0,00%**	**14,50%**	**18,43%**	**2096728,14**	**10,75**
		Sd:	0,00%	6,71%	6,05%	57640,0522	4,758
		T-:	-	-7,0107	-19,2878	-641,2373	-48,5439
Ovarian	CLPSO	M :	6,64%	11,83%	4,54%	356989523,125	2078,375
		Sd:	0,72%	13,13%	1,89%	117505367,7324	468,4003
		T-:	-	-	-	-	-
	PSO(R)	M :	**0,00%**	**2,37%**	**3,82%**	**77316,38**	**4,30**
		Sd:	0,00%	0,86%	0,75%	16833,6634	2,1364
		T-:	-45,3333	-3,1662	-3,6666	-17,8032	-18,3419
	CLPSO(R)	M :	**0,00%**	**2,62%**	4,22%	**1398506,66**	**4,73**
		Sd:	0,00%	0,77%	0,66%	26316,2297	2,1324
		T-:	-45,3333	-3,0813	-1,6908	-17,7373	-18,3381
SRBCT	CLPSO	M :	17,99%	5,75%	43,47%	2702994	346,3333
		Sd:	5,20%	6,05%	2,93%	1675240,1434	165,2188
		T-:	-	-	-	-	-
	PSO(R)	M :	**0,00%**	**4,81%**	**31,05%**	**15772**	**10,14**
		Sd:	0,00%	3,73%	4,06%	2505,4457	2,6561
		T-:	-153	-6,2195	-52,3143	-13,0332	-17,3227
	CLPSO(R)	M :	**0,00%**	6,44%	**30,11%**	**86627,9**	**13,86**
		Sd:	0,00%	5,08%	4,80%	11505,6794	6,8266
		T-:	-153	3,0972	-57,2256	-12,6895	-17,1309
9_Tumors	CLPSO	M :	55,56%	71,48%	94,74%	874830	149,8889
		Sd:	5,77%	10,81%	1,35%	489574,27	116,9406
		T-:	-	-	-	-	-
	PSO(R)	M :	**25,56%**	77,67%	**93,78%**	**43110**	**26,75**
		Sd:	6,25%	7,40%	1,49%	7192,1514	12,7002
		T-:	-24,2319	6,5096	-5,0486	-18,2908	-9,439
	CLPSO(R)	M :	**23,45%**	74,92%	**93,38%**	**414520,10**	**27,06**
		Sd:	4,22%	7,34%	1,82%	74761,4618	13,9282
		T-:	-25,8871	3,4027	-9,5835	-9,9879	-9,4562

Table E.16: CLPSO based Relief *vs* random CLPSO and PSO-FS (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
11_Tumors	CLPSO	M :	23,74%	45,72%	72,57%	326287385,5556	2537,8889
		Sd:	2,00%	35,77%	1,47%	179105755,9801	735,6314
		T-:	-	-	-	-	-
	PSO(R)	M :	**17,15%**	**6,44%**	**69,78%**	**10463843,36**	**491**
		Sd:	1,19%	1,17%	1,91%	6882215,4065	127,5429
		T-:	-39,0264	-18,5018	-14,5808	-47,5434	-54,476
	CLPSO(R)	M :	**16,67%**	**5,47%**	**70,78%**	**23737858,92**	**678,25**
		Sd:	1,54%	1,10%	1,59%	10805476,9711	187,8824
		T-:	-46,7952	-18,9621	-12,7511	-45,7721	-52,9894
14_Tumors	CLPSO	M :	66,75%	63,12%	87,23%	552404446,2	2526,9
		Sd:	1,10%	17,04%	1,15%	165292469,6776	463,5459
		T-:	-	-	-	-	-
	PSO(R)	M :	**63,35%**	**53,19%**	**86,09%**	**5568428,11**	**246,66**
		Sd:	1,81%	5,27%	0,93%	4446302,803	158,4408
		T-:	-8,4255	-5,192	-5,2248	-52,2984	-88,0811
	CLPSO(R)	M :	**62,24%**	**52,10%**	**86,27%**	**5710696,96**	**211,78**
		Sd:	1,76%	3,51%	0,78%	2395489,664	86,9714
		T-:	-86,6098	-7,2285	-4,4344	-52,3451	-144,588
Brain Tumor2	CLPSO	M :	30,77%	45,22%	54,25%	25252029,0952	938
		Sd:	0,00%	3,58%	6,68%	15765058,7788	313,9256
		T-:	-	-	-	-	-
	PSO(R)	M :	**6,84%**	47,91%	59,56%	**30961,66**	**6,44**
		Sd:	2,56%	9,87%	5,29%	2334,1047	2,9202
		T-:	-252	6,0485	10,7035	-73,3109	-176,1591
	CLPSO(R)	M :	**4,56%**	47,67%	57,04%	**1260191,77**	**8,88**
		Sd:	3,85%	11,23%	4,63%	184846,2031	4,9485
		T-:	-155,25	3,7947	20,6417	-69,7335	-175,7126
Prostate Tumor	CLPSO	M :	1,28%	30,82%	32,75%	181622434,1667	2175,5
		Sd:	1,89%	16,90%	5,28%	70355673,1626	593,58
		T-:	-	-	-	-	-
	PSO(R)	M :	2,56%	**17,82%**	**19,56%**	**32435,33**	**7,55**
		Sd:	1,92%	5,91%	4,77%	2194,9423	3,5746
		T-:	4,2135	-15,2372	-39,5688	-727,299	-402,2534
	CLPSO(R)	M :	2,71%	**17,44%**	**21,31%**	**1472073,37**	**6,51**
		Sd:	1,79%	4,88%	6,86%	77578,6907	2,8739
		T-:	12,4007	-14,7031	-14,4074	-721,5327	-403,4129
Sonar	CLPSO	M :	12,24%	27,19%	40,35%	31605,2727	23,5455
		Sd:	2,76%	2,43%	2,24%	6516,7455	4,298
		T-:	-	-	-	-	-
	PSO(R)	M :	12,39%	28,85%	**37,97%**	**18753,55**	**15,33**
		Sd:	1,70%	3,79%	2,60%	6358,479	6,1237
		T-:	0,9552	2,1486	-12,4528	-8,7739	-7,3315
	CLPSO(R)	M :	**9,55%**	**26,35%**	**37,81%**	**20885,41**	**17,72**
		Sd:	2,09%	2,58%	1,70%	4214,0676	3,9087
		T-:	-20,9464	-6,3798	-23,0093	-8,3731	-7,4935

Table E.17: CLPSO based Relief *vs* random CLPSO and PSO-FS (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Spam Base	CLPSO	M :	14,06%	9,71%	15,62%	2003230	38,1
		Sd:	1,01%	1,13%	2,01%	450056,7096	4,9766
		T-:	-	-	-	-	-
	PSO(R)	M :	**10,63%**	**8,81%**	**14,88%**	**1362338,77**	**31,33**
		Sd:	1,18%	0,26%	0,97%	307982,4086	4,8218
		T-:	-29,1607	-16,6232	-6,2857	-19,3588	-19,6658
	CLPSO(R)	M :	**9,74%**	**9,05%**	**14,83%**	**1123008,58**	**27,62**
		Sd:	0,83%	0,72%	1,01%	276360,7724	4,5858
		T-:	-34,7951	-7,4198	-8,0067	-26,2295	-28,393
Soybean	CLPSO	M :	12,76%	7,23%	60,49%	1326505,0909	27,3636
		Sd:	1,92%	0,54%	2,31%	145369,2342	2,1106
		T-:	-	-	-	-	-
	PSO(R)	M :	12,87%	**6,78%**	**58,74%**	**878965,66**	**24,22**
		Sd:	0,97%	0,88%	0,32%	144287,2724	2,9486
		T-:	0,5361	-6,2396	-6,6067	-28,2392	-11,3701
	CLPSO(R)	M :	12,78%	**7,00%**	**58,77%**	**891868,68**	**24,13**
		Sd:	0,73%	0,86%	0,24%	151398,1124	2,5735
		T-:	0,1646	-2,9393	-6,5071	-27,4318	-17,0657
Arrhythmia	CLPSO	M :	33,95%	41,09%	42,72%	1756141,0909	109,3636
		Sd:	1,07%	2,30%	0,55%	761999,0311	17,7329
		T-:	-	-	-	-	-
	PSO(R)	M :	**30,78%**	41,28%	**42,63%**	**202856,44**	**24,44**
		Sd:	2,25%	2,05%	1,59%	54388,1703	7,9861
		T-:	-25,6557	0,6612	-1,8014	-18,4492	-49,886
	CLPSO(R)	M :	**30,18%**	41,88%	**42,59%**	**213292,21**	**28,28**
		Sd:	1,82%	1,72%	1,13%	69584,6381	8,4847
		T-:	-127,9048	2,7692	-2,9726	-18,3685	-59,6686
Secom	CLPSO	M :	6,41%	9,34%	7,75%	681233,5	35,5
		Sd:	0,24%	1,13%	0,16%	366131,4789	14,5239
		T-:	-	-	-	-	-
	PSO(R)	M :	6,82%	**6,61%**	10,11%	**54114,71**	**2,28**
		Sd:	0,14%	0,01%	2,07%	15041,5406	1,2536
		T-:	12,8042	-93,8449	8,0118	-36,0378	-33,8556
	CLPSO(R)	M :	6,56%	**6,71%**	8,73%	**99209,81**	5,88
		Sd:	0,30%	0,15%	0,92%	27196,7619	2,9264
		T-:	4,7972	-90,0489	66,0185	-33,5695	-30,0366
semeion	CLPSO	M :	14,76%	9,15%	81,17%	9704920,1818	158,2727
		Sd:	1,49%	0,27%	0,16%	1828334,9384	6,4667
		T-:	-	-	-	-	-
	PSO(R)	M :	15,04%	**9,04%**	81,22%	**8725772,66**	**161,77**
		Sd:	0,96%	0,91%	0,15%	2227476,7364	22,3985
		T-:	2,9028	-3,0197	2,3641	-6,4442	3,3538
	CLPSO(R)	M :	**13,88%**	9,17%	**81,11%**	**7770737,82**	**154,58**
		Sd:	0,75%	0,53%	0,17%	1358144,7059	10,37
		T-:	-20,2347	0,8489	-3,643	-17,2727	-3,9319
Lymphoma	CLPSO	M :	6,47%	1,58%	17,45%	7289321,2	454,8
		Sd:	4,34%	1,72%	2,25%	3538625,7463	78,2387
		T-:	0-	-	-	-	-
	PSO(R)	M :	**0,00%**	13,80%	17,20%	**13130,92**	**2,53**
		Sd:	0,00%	6,53%	6,81%	1177,8929	1,3914
		T-:	-12,2222	51,79	-0,9484	-33,2141	-1602,3901
	CLPSO(R)	M :	**0,00%**	11,43%	17,66%	**224054,58**	**3,48**
		Sd:	0,00%	6,80%	5,51%	5009,3233	2,2139
		T-:	-12,2222	58,4754	0,725	-32,2513	-1539,5655

Table E.18: CLPSO based Relief *vs* random CLPSO and PSO-FS (4)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	CLPSO	M :	**14,80%**	40,79%	**50,21%**	**14989102,5**	**258,7**
		Sd:	6,27%	8,29%	3,43%	15819583,1527	248,8641
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**3,64%**	42,34%	**44,57%**	**6628826,73**	**18,62**
		Sd:	2,81%	5,90%	6,43%	1130932,801	14,3088
		T-:	-37,655	1,5745	-40,3869	-7,8288	-13,0675
	EPSO-FS(R)	M :	**5,70%**	44,24%	**45,37%**	**6966383,54**	**8,57**
		Sd:	2,65%	4,78%	5,38%	1086915,3481	3,9924
		T-:	-31,5457	3,585	-37,3314	-7,5113	-13,6155
	EPSO-FS(SU)	M :	**5,45%**	**43,66%**	**48,11%**	**6523488,90**	**8,77**
		Sd:	2,32%	5,08%	7,07%	974605,1308	5,5027
		T-:	-32,4833	2,8927	-12,9833	-7,9275	-13,6052
CNS	CLPSO	M :	14,00%	42,67%	39,27%	594637,2	34,6
		Sd:	5,84%	3,96%	5,30%	357285,9256	10,8853
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**2,78%**	48,92%	**35,86%**	**569435,52**	**10,75**
		Sd:	3,32%	7,34%	6,00%	148728,5232	4,5265
		T-:	-15,2556	10,2581	-42,8431	-10,9086	-66,1877
	EPSO-FS(R)	M :	**1,62%**	**40,26%**	43,25%	622270,06	**7,84**
		Sd:	2,90%	7,29%	6,22%	54263,7199	3,3552
		T-:	-16,5309	-3,9594	15,8843	10,5845	-74,121
	EPSO-FS(SU)	M :	**6,11%**	43,83%	39,89%	626050,95	**10,08**
		Sd:	5,53%	7,11%	7,13%	41364,6168	4,7996
		T-:	-9,9522	1,9433	5,4332	11,7969	-68,0987
Colon	CLPSO	M :	9,66%	24,34%	31,55%	98361,7273	20,0909
		Sd:	3,26%	4,79%	4,25%	47964,3065	3,5058
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**1,11%**	23,40%	**23,94%**	**54274,86**	**11,75**
		Sd:	2,42%	5,33%	7,31%	11215,3871	6,1465
		T-:	-27,4944	-1,4217	-39,1617	-33,8839	-80,9595
	EPSO-FS(R)	M :	**3,57%**	**19,32%**	**23,98%**	**62126,51**	**14,91**
		Sd:	4,09%	4,66%	4,23%	3412,6605	5,1983
		T-:	-19,07	-7,585	-37,4967	-27,8814	-34,632
	EPSO-FS(SU)	M :	**1,30%**	**22,39%**	**26,75%**	**63126,29**	**20,33**
		Sd:	2,59%	5,00%	3,39%	5767,6176	7,4464
		T-:	-26,5614	-2,9332	-26,4308	-27,1004	1,4205
Leukemia3C	CLPSO	M :	16,11%	28,17%	40,44%	6904625,6	311,2
		Sd:	1,76%	11,08%	8,46%	9670346,2715	303,2439
		T-:	0-	-	-	-	-
	EPSOFS(X2)	M :	**0,00%**	**1,01%**	**15,40%**	**724600,39**	**55,60**
		Sd:	0,00%	2,31%	4,86%	65116,357	35,4398
		T-:	-32,2222	-17,0276	-64,3678	-99,0895	-25,34
	EPSO-FS(R)	M :	**0,00%**	**6,40%**	**10,59%**	**669026,93**	**10,36**
		Sd:	0,00%	6,53%	5,32%	50593,7262	4,4639
		T-:	-32,2222	-13,5849	-77,0924	-99,7463	-29,8423
	EPSO-FS(SU)	M :	**0,00%**	**7,07%**	**12,07%**	**679008,04**	**16,40**
		Sd:	0,00%	7,10%	3,56%	21224,5251	8,5449
		T-:	-32,2222	-13,025	-68,4997	-99,8387	-29,2329

Table E.19: CLPSO-FS: filter comparison (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Leukemia4C	CLPSO	M :	17,78%	23,89%	35,44%	39871876,9	935,5
		Sd:	2,34%	3,76%	4,47%	21888598,925	235,0316
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**5,56%**	**11,93%**	**19,60%**	**815816,91**	**94,73**
		Sd:	0,00%	3,11%	1,98%	84481,1971	23,5397
		T-:	-110	-214,325	-101,3925	-58,8203	-111,354
	EPSO-FS(R)	M :	**5,56%**	**17,22%**	**23,99%**	**703333,02**	**19,85**
		Sd:	0,00%	5,50%	4,15%	14850,1825	8,8013
		T-:	-110	-67,4772	-68,3489	-58,9898	-121,1695
	EPSO-FS(SU)	M :	**5,56%**	**21,24%**	**21,57%**	**658471,13**	**25,68**
		Sd:	0,00%	6,72%	5,27%	72146,103	12,6581
		T-:	-110	-33,5397	-55,2175	-59,0573	-120,4083
Lung	CLPSO	M :	10,78%	8,55%	22,87%	262280152,6	1617,4
		Sd:	1,91%	0,67%	1,89%	119376285,4008	367,9068
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**5,75%**	**8,43%**	**19,96%**	**6668015,95**	**256,56**
		Sd:	0,50%	2,36%	1,86%	4613655,1388	204,39
		T-:	-51,339	-1,4431	-11,7113	-40,4777	-55,4047
	EPSO-FS(R)	M :	**7,14%**	10,85%	23,18%	**2947791,46**	**126,53**
		Sd:	2,15%	4,00%	1,51%	1650031,1773	105,5443
		T-:	-26,669	29,4946	1,1902	-41,0666	-60,508
	EPSO-FS(SU)	M :	**4,96%**	9,04%	23,23%	**24327274,11**	**452,76**
		Sd:	1,01%	4,46%	1,80%	35803172,709	494,1124
		T-:	-50,43	0,6854	1,3282	-36,8833	-32,8044
MLL	CLPSO	M :	11,11%	20,89%	23,11%	110196512,3	1817,1
		Sd:	0,00%	15,21%	5,13%	63219803,2856	657,5251
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**0,00%**	**14,18%**	**11,80%**	**1954399,34**	**10,34**
		Sd:	0,00%	4,35%	5,20%	66545,9714	4,8016
		T-:	-	-7,7212	-74,0138	-642,0846	-48,5558
	EPSO-FS(R)	M :	**0,00%**	**14,50%**	**18,43%**	**2096728,14**	**10,75**
		Sd:	0,00%	6,71%	6,05%	57640,0522	4,758
		T-:	-	-7,0107	-19,2878	-641,2373	-48,5439
	EPSO-FS(SU)	M :	**0,00%**	**14,66%**	23,02%	**2244662,38**	**58,44**
		Sd:	0,00%	6,17%	4,39%	109637,0996	44,4075
		T-:	-	-7,0578	-0,2632	-639,8892	-47,1393
Ovarian	CLPSO	M :	6,64%	11,83%	4,54%	356989523,125	2078,375
		Sd:	0,72%	13,13%	1,89%	117505367,7324	468,4003
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**0,00%**	**2,97%**	4,40%	**3070593,82**	**3,93**
		Sd:	0,00%	1,48%	0,94%	525183,1468	1,5433
		T-:	-45,3333	-2,9654	-0,7777	-17,6539	-18,3452
	EPSO-FS(R)	M :	**0,00%**	**2,62%**	4,22%	**1398506,66**	**4,73**
		Sd:	0,00%	0,77%	0,66%	26316,2297	2,1324
		T-:	-45,3333	-3,0813	-1,6908	-17,7373	-18,3381
	EPSO-FS(SU)	M :	**0,00%**	**2,80%**	4,54%	**1369084,58**	**4,52**
		Sd:	0,00%	0,40%	1,04%	22981,9931	2,3216
		T-:	-45,3333	-3,0221	-0,0305	-17,7388	-18,3399

Table E.20: CLPSO-FS: filter comparison (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
SRBCT	CLPSO	M :	17,99%	5,75%	43,47%	2702994	346,3333
		Sd:	5,20%	6,05%	2,93%	1675240,1434	165,2188
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**0,00%**	**2,08%**	**31,65%**	**180284,80**	**67,31**
		Sd:	0,00%	1,75%	3,24%	161670,1203	24,9897
		T-:	-153	-24,1984	-47,1112	-12,2353	-14,3762
	EPSO-FS(R)	M :	**0,00%**	6,44%	**30,11%**	**86627,9**	**13,86**
		Sd:	0,00%	5,08%	4,80%	11505,6794	6,8266
		T-:	-153	3,0972	-57,2256	-12,6895	-17,1309
	EPSO-FS(SU)	M :	**0,00%**	**4,79%**	**33,08%**	**120370,12**	**37,56**
		Sd:	0,00%	4,82%	6,47%	82681,815	36,2031
		T-:	-153	-4,5602	-44,6688	-12,5258	-15,909
9_Tumors	CLPSO	M :	55,56%	71,48%	94,74%	874830	149,8889
		Sd:	5,77%	10,81%	1,35%	489574,27	116,9406
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**26,80%**	**66,38%**	**94,26%**	**463140,65**	**52**
		Sd:	3,72%	4,48%	1,69%	97428,8365	11,4546
		T-:	-23,2897	-5,3874	-3,4405	-9,0541	-7,5359
	EPSO-FS(R)	M :	**23,45%**	74,92%	**93,38%**	**414520,10**	**27,06**
		Sd:	4,22%	7,34%	1,82%	74761,4618	13,9282
		T-:	-25,8871	3,4027	-9,5835	-9,9879	-9,4562
	EPSO-FS(SU)	M :	**30,53%**	69,82%	**94,28%**	**476369,36**	**50**
		Sd:	4,05%	5,92%	2,18%	51782,9044	18,3242
		T-:	-20,0047	-1,6984	-3,2697	-8,7633	-7,688
11_Tumors	CLPSO	M :	23,74%	45,72%	72,57%	326287385,5556	2537,8889
		Sd:	2,00%	35,77%	1,47%	179105755,9801	735,6314
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**12,71%**	**9,18%**	**71,25%**	**3590808,84**	**209,86**
		Sd:	2,10%	2,61%	1,43%	1716002,0386	94,2897
		T-:	-76,83	-17,2107	-11,8087	-48,8202	-66,4889
	EPSO-FS(R)	M :	**16,67%**	**5,47%**	**70,78%**	**23737858,92**	**678,25**
		Sd:	1,54%	1,10%	1,59%	10805476,9711	187,8824
		T-:	-46,7952	-18,9621	-12,7511	-45,7721	-52,9894
	EPSO-FS(SU)	M :	**18,06%**	**8,01%**	**71,67%**	**29800745**	**748,44**
		Sd:	2,40%	4,90%	1,86%	11605883,9798	241,8688
		T-:	-40,4495	-17,7607	-6,9181	-44,554	-49,5458
14_Tumors	CLPSO	M :	66,75%	63,12%	87,23%	552404446,2	2526,9
		Sd:	1,10%	17,04%	1,15%	165292469,6776	463,5459
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**55,02%**	63,24%	**85,21%**	**2826602,47**	**70**
		Sd:	1,98%	4,29%	0,89%	571025,0777	31,3146
		T-:	-212,4484	0,0835	-9,3301	-52,6235	-161,7265
	EPSO-FS(R)	M :	**62,24%**	**52,10%**	**86,27%**	**5710696,96**	**211,78**
		Sd:	1,76%	3,51%	0,78%	2395489,664	86,9714
		T-:	-86,6098	-7,2285	-4,4344	-52,3451	-144,588
	EPSO-FS(SU)	M :	**60,82%**	**54,10%**	**86,16%**	**3911628,94**	**141,16**
		Sd:	2,00%	3,34%	0,84%	872627,1664	36,7011
		T-:	-59,9583	-5,902	-4,9113	-52,5189	-154,2766

Table E.21: CLPSO-FS: filter comparison (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Brain Tumor2	CLPSO	M :	30,77%	45,22%	54,25%	25252029,0952	938
		Sd:	0,00%	3,58%	6,68%	15765058,7788	313,9256
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**6,09%**	46,67%	57,03%	**1197634,77**	**5,02**
		Sd:	4,19%	13,39%	5,41%	168267,6695	1,5227
		T-:	-739,2	8,3428	43,536	-69,9194	-176,5083
	EPSO-FS(R)	M :	**4,56%**	47,67%	57,04%	**1260191,77**	**8,88**
		Sd:	3,85%	11,23%	4,63%	184846,2031	4,9485
		T-:	-155,25	3,7947	20,6417	-69,7335	-175,7126
	EPSO-FS(SU)	M :	**3,37%**	52,50%	58,15%	**1272232**	**11,62**
		Sd:	3,94%	9,35%	4,51%	208896,0895	4,3493
		T-:	-130,2857	57,841	8,0214	-69,6755	-175,0994
Prostate Tumor	CLPSO	M :	1,28%	30,82%	32,75%	181622434,1667	2175,5
		Sd:	1,89%	16,90%	5,28%	70355673,1626	593,58
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**0,96%**	**19,65%**	**22,15%**	**1354122,22**	**6,34**
		Sd:	1,68%	7,70%	7,23%	114596,1402	2,736
		T-:	-2,9391	-13,4212	-48,6207	-721,9745	-403,5242
	EPSO-FS(R)	M :	2,71%	**17,44%**	**21,31%**	**1472073,37**	**6,51**
		Sd:	1,79%	4,88%	6,86%	77578,6907	2,8739
		T-:	12,4007	-14,7031	-14,4074	-721,5327	-403,4129
	EPSO-FS(SU)	M :	**0,00%**	**11,45%**	**23,75%**	**1646577,81**	**92,25**
		Sd:	0,00%	2,71%	4,37%	348646,3169	71,9949
		T-:	-12	-23,1453	-39,3629	-707,4502	-204,1891
Sonar	CLPSO	M :	12,24%	27,19%	40,35%	31605,2727	23,5455
		Sd:	2,76%	2,43%	2,24%	6516,7455	4,298
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	12,37%	28,80%	**38,83%**	**14856,78**	**11,80**
		Sd:	1,65%	2,70%	1,75%	3284,5232	2,1768
		T-:	0,9073	20,9004	-25,5272	-13,0834	-15,2467
	EPSO-FS(R)	M :	**9,55%**	**26,35%**	**37,81%**	**20885,41**	**17,72**
		Sd:	2,09%	2,58%	1,70%	4214,0676	3,9087
		T-:	-20,9464	-6,3798	-23,0093	-8,3731	-7,4935
	EPSO-FS(SU)	M :	11,97%	28,13%	**38,91%**	**16108,33**	**13,27**
		Sd:	2,04%	2,58%	2,11%	3661,0476	2,3962
		T-:	-1,2355	5,2451	-24,6444	-11,9608	-13,1816
Spam Base	CLPSO	M :	14,06%	9,71%	15,62%	2003230	38,1
		Sd:	1,01%	1,13%	2,01%	450056,7096	4,9766
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**10,26%**	10,76%	**14,38%**	**383114,19**	**11,23**
		Sd:	0,62%	1,20%	1,98%	129138,1385	3,5288
		T-:	-32,209	20,3082	-12,7292	-53,3566	-92,4734
	EPSO-FS(R)	M :	**9,74%**	**9,05%**	**14,83%**	**1123008,58**	**27,62**
		Sd:	0,83%	0,72%	1,01%	276360,7724	4,5858
		T-:	-34,7951	-7,4198	-8,0067	-26,2295	-28,393
	EPSO-FS(SU)	M :	**9,83%**	**9,55%**	**13,75%**	**513861,77**	**13,94**
		Sd:	0,94%	0,76%	1,98%	145717,9986	3,7803
		T-:	-36,0939	-2,9991	-9,4577	-47,8746	-72,5492

Table E.22: CLPSO-FS: filter comparison (4)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Soybean	CLPSO	M :	12,76%	7,23%	60,49%	1326505,0909	27,3636
		Sd:	1,92%	0,54%	2,31%	145369,2342	2,1106
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**11,82%**	**5,70%**	**58,94%**	**860082,72**	**19,36**
		Sd:	1,24%	0,57%	0,30%	214200,5806	2,1308
		T-:	-6,253	-21,462	-5,8614	-29,1557	-59,8214
	EPSO-FS(R)	M :	12,78%	**7,00%**	**58,77%**	**891868,68**	**24,13**
		Sd:	0,73%	0,86%	0,24%	151398,1124	2,5735
		T-:	0,1646	-2,9393	-6,5071	-27,4318	-17,0657
	EPSO-FS(SU)	M :	12,93%	**6,30%**	**58,98%**	**887488,5**	**21,11**
		Sd:	0,44%	0,48%	1,09%	120433,3917	2,1113
		T-:	1,1227	-12,9032	-5,7104	-21,0433	-24,0603
Arrhytmia	CLPSO	M :	33,95%	41,09%	42,72%	1756141,0909	109,3636
		Sd:	1,07%	2,30%	0,55%	761999,0311	17,7329
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**28,90%**	**37,17%**	42,86%	**319936,81**	**34,45**
		Sd:	0,85%	1,68%	0,78%	95337,2984	5,151
		T-:	-168,0096	-14,6188	3,9273	-17,1025	-56,0473
	EPSO-FS(R)	M :	**30,18%**	41,88%	**42,59%**	**213292,21**	**28,28**
		Sd:	1,82%	1,72%	1,13%	69584,6381	8,4847
		T-:	-127,9048	2,7692	-2,9726	-18,3685	-59,6686
	EPSO-FS(SU)	M :	**26,55%**	**36,78%**	42,84%	**182711**	**23,88**
		Sd:	1,46%	1,95%	0,88%	48903,0314	6,8246
		T-:	-252,9999	-16,0584	3,9012	-18,7317	-62,2445
Secom	CLPSO	M :	6,41%	9,34%	7,75%	681233,5	35,5
		Sd:	0,24%	1,13%	0,16%	366131,4789	14,5239
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**6,25%**	**6,73%**	10,08%	62414,43	3,52
		Sd:	0,26%	0,09%	1,22%	8436,1375	0,8488
		T-:	-5,4257	-89,8007	75,8461	-35,9929	-33,6581
	EPSO-FS(R)	M :	6,56%	**6,71%**	8,73%	**99209,81**	**5,88**
		Sd:	0,30%	0,15%	0,92%	27196,7619	2,9264
		T-:	4,7972	-90,0489	66,0185	-33,5695	-30,0366
	EPSO-FS(SU)	M :	**6,06%**	**6,78%**	10,71%	**75904,55**	**4,77**
		Sd:	0,32%	0,11%	0,85%	9476,9187	1,166
		T-:	-9,6912	-85,9831	124,3101	-35,2093	-32,3365
semeion	CLPSO	M :	14,76%	9,15%	81,17%	9704920,1818	158,2727
		Sd:	1,49%	0,27%	0,16%	1828334,9384	6,4667
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**13,52%**	**9,07%**	81,17%	**7413861,91**	**151,04**
		Sd:	0,94%	0,61%	0,18%	1172657,8542	8,9193
		T-:	-28,2096	-2,5261	0,0056	-19,9386	-7,707
	EPSO-FS(R)	M :	**13,88%**	9,17%	**81,11%**	**7770737,82**	**154,58**
		Sd:	0,75%	0,53%	0,17%	1358144,7059	10,37
		T-:	-20,2347	0,8489	-3,643	-17,2727	-3,9319
	EPSO-FS(SU)	M :	**13,37%**	9,21%	81,16%	**6984514,77**	**149,27**
		Sd:	0,92%	0,48%	0,18%	914462,7034	8,6691
		T-:	-20,8352	1,5122	-0,784	-24,1223	-9,2728
Lymphoma	CLPSO	M :	6,47%	1,58%	17,45%	7289321,2	454,8
		Sd:	4,34%	1,72%	2,25%	3538625,7463	78,2387
		T-:	-	-	-	-	-
	EPSOFS(X2)	M :	**0,00%**	7,51%	**12,83%**	**227372,60**	**15,32**
		Sd:	0,00%	5,19%	5,20%	20685,1587	9,7981
		T-:	-12,2222	31,774	-14,4047	-32,2362	-1487,8666
	EPSO-FS(R)	M :	**0,00%**	11,43%	17,66%	**224054,58**	**3,48**
		Sd:	0,00%	6,80%	5,51%	5009,3233	2,2139
		T-:	-12,2222	58,4754	0,725	-32,2513	-1539,5655
	EPSO-FS(SU)	M :	**0,00%**	15,62%	18,52%	**224784,5**	**3,22**
		Sd:	0,00%	4,80%	5,66%	3034,4406	1,896
		T-:	-12,2222	80,5281	3,158	-32,248	-1593,91

Table E.23: CLPSO-FS: filter comparison (5)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Breast	CLPSO(R)	M :	5,70%	44,24%	45,37%	6966383,5455	8,5758
		Sd:	2,65%	4,78%	5,38%	1086915,3481	3,9924
		T-:	-	-	-	-	-
	PSO-FS 2	M :	5,90%	43,93%	47,62%	**185937,05**	**6,27**
		Sd:	2,39%	5,08%	4,76%	54593,0167	3,2737
		T-:	2,3248	-1,4696	27,4643	-308,4938	-17,0064
	EPSO-FS2	M :	**4,77%**	**41,71%**	47,39%	13109555,69	**7,57**
		Sd:	2,53%	5,45%	5,12%	3774419,5844	3,0617
		T-:	-12,2391	-53,3626	6,8654	26,0292	-6,8975
CNS	CLPSO(R)	M :	14,00%	42,67%	39,27%	594637,2	34,6
		Sd:	5,84%	3,96%	5,30%	357285,9256	10,8853
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**4,81%**	44,59%	**38,19%**	**45227,41**	**7,30**
		Sd:	3,36%	6,87%	6,58%	14591,6518	3,6354
		T-:	-12,3939	3,1546	-4,9445	-241,4646	-73,0577
	EPSO-FS2	M :	**2,05%**	46,85%	**38,21%**	1171319,88	**8,92**
		Sd:	3,14%	7,59%	6,06%	308822,593	4,0785
		T-:	-16,2002	6,9654	-7,5338	88,5421	-65,7805
Colon	CLPSO(R)	M :	9,66%	24,34%	31,55%	98361,7273	20,0909
		Sd:	3,26%	4,79%	4,25%	47964,3065	3,5058
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**3,78%**	**21,85%**	**25,43%**	**19526,95**	**13,41**
		Sd:	3,89%	6,69%	6,78%	6988,6066	6,9906
		T-:	-18,6549	-3,7732	-33,0296	-60,545	-66,9573
	EPSO-FS2	M :	**2,50%**	**22,48%**	**28,90%**	116675,52	**11,52**
		Sd:	4,03%	6,48%	5,75%	38090,1747	5,0918
		T-:	-20,7926	-2,3964	-5,941	14,0047	-84,8479
Leukemia3C	CLPSO(R)	M :	16,11%	28,17%	40,44%	6904625,6	311,2
		Sd:	1,76%	11,08%	8,46%	9670346,2715	303,2439
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**0,00%**	**5,73%**	**13,23%**	**59425,78**	**14,47**
		Sd:	0,00%	6,67%	6,08%	35213,4072	12,4572
		T-:	-32,2222	-14,0725	-65,7358	-109,7752	-29,4344
	EPSO-FS2	M :	**0,00%**	**5,43%**	**14,77%**	**1496311,95**	**18,22**
		Sd:	0,00%	4,56%	5,11%	395832,9724	9,5963
		T-:	-32,2222	-14,2198	-63,9055	-85,8225	-29,0595
Leukemia4C	CLPSO(R)	M :	17,78%	23,89%	35,44%	39871876,9	935,5
		Sd:	2,34%	3,76%	4,47%	21888598,925	235,0316
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**5,56%**	**17,76%**	**23,14%**	**58339,39**	**17,5**
		Sd:	0,00%	6,37%	5,41%	20214,3319	11,3727
		T-:	-110	-25,5185	-50,8836	-59,9612	-121,5848
	EPSO-FS2	M :	**5,56%**	**16,44%**	**21,26%**	**1462794,40**	**25,72**
		Sd:	0,00%	6,19%	4,82%	203925,3168	17,7071
		T-:	-110	-30,5292	-74,9121	-57,8458	-117,1067

Table E.24: CLPSO: multi-filters assessement (1)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Lung	CLPSO(R)	M :	10,78%	8,55%	22,87%	262280152,6	1617,4
		Sd:	1,91%	0,67%	1,89%	119376285,4008	367,9068
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**5,83%**	8,79%	**20,74%**	**3277129,83**	**178,51**
		Sd:	0,32%	3,32%	1,71%	3779166,1201	126,0449
		T-:	-50,5351	2,3275	-8,5794	-41,0104	-58,012
	EPSO-FS2	M :	**5,32%**	9,84%	**21,41%**	**7364130,85**	**163,19**
		Sd:	0,91%	3,22%	1,99%	4928803,359	157,6355
		T-:	-46,0658	6,5544	-5,7759	-40,3547	-58,1286
MLL	CLPSO(R)	M :	11,11%	20,89%	23,11%	110196512,3	1817,1
		Sd:	0,00%	15,21%	5,13%	63219803,2856	657,5251
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**0,00%**	**14,72%**	**14,87%**	**99694,33**	**7,30**
		Sd:	0,00%	5,36%	7,02%	68273,9982	3,7252
		T-:	-	-7,0806	-49,554	-653,0878	-48,6372
	EPSO-FS2	M :	**0,00%**	**17,64%**	**15,33%**	**4050809**	**8,8**
		Sd:	0,00%	5,64%	6,80%	995981,8346	4,7784
		T-:	-	-3,7429	-53,6182	-628,9388	-48,5969
Ovarian	CLPSO(R)	M :	6,64%	11,83%	4,54%	356989523,125	2078,375
		Sd:	0,72%	13,13%	1,89%	117505367,7324	468,4003
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**0,00%**	**3,05%**	4,37%	**172204,92**	**3,46**
		Sd:	0,00%	1,32%	1,08%	113087,782	0,9772
		T-:	-45,3333	-2,9384	-0,9042	-17,7985	-18,3494
	EPSO-FS2	M :	**0,00%**	**2,48%**	4,31%	**4986064,82**	**4,43**
		Sd:	0,00%	0,83%	0,75%	2067667,7646	2,6939
		T-:	-45,3333	-3,1304	-1,2422	-17,5582	-18,3407
SRBCT	CLPSO(R)	M :	17,99%	5,75%	43,47%	2702994	346,3333
		Sd:	5,20%	6,05%	2,93%	1675240,1434	165,2188
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**0,00%**	5,39%	**31,18%**	**29495,14**	**14,12**
		Sd:	0,00%	3,54%	4,74%	8994,8415	7,5802
		T-:	-153	-1,9213	-50,7726	-12,9666	-17,1164
	EPSO-FS2	M :	**0,00%**	**3,59%**	**31,91%**	**221740,09**	**28,63**
		Sd:	0,00%	3,80%	4,13%	183338,0398	20,3997
		T-:	-153	-9,6952	-51,3855	-12,0342	-16,3679
9_Tumors	CLPSO(R)	M :	55,56%	71,48%	94,74%	874830	149,8889
		Sd:	5,77%	10,81%	1,35%	489574,27	116,9406
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**27,48%**	73,07%	**94,00%**	**89345,97**	**31,97**
		Sd:	4,27%	7,43%	2,02%	25118,2624	14,2364
		T-:	-22,5909	1,6544	-5,2587	-17,2751	-9,0785
	EPSO-FS2	M :	**26,67%**	71,06%	94,47%	1134513,79	**33,54**
		Sd:	4,82%	7,72%	1,95%	523201,2959	15,3962
		T-:	-22,8293	-0,45	-1,859	5,7087	-8,9041

Table E.25: CLPSO: multi-filters assessement (2)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
11_Tumors	CLPSO(R)	M :	23,74%	45,72%	72,57%	326287385,5556	2537,8889
		Sd:	2,00%	35,77%	1,47%	179105755,9801	735,6314
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**15,45%**	**8,45%**	**71,77%**	**5459088,55**	**271,77**
		Sd:	2,20%	3,20%	2,09%	5178874,5277	134,7704
		T-:	-58,8473	-17,5536	-5,7951	-48,537	-64,7175
	EPSO-FS2	M :	**14,05%**	**7,19%**	**72,05%**	**13122508,77**	**373,63**
		Sd:	2,49%	1,67%	2,54%	7933093,8697	144,6232
		T-:	-59,149	-18,1486	-4,5721	-47,3622	-61,8091
14_Tumors	CLPSO(R)	M :	66,75%	63,12%	87,23%	552404446,2	2526,9
		Sd:	1,10%	17,04%	1,15%	165292469,6776	463,5459
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**58,96%**	62,38%	**85,69%**	**1727936,35**	**74,92**
		Sd:	2,03%	6,66%	1,04%	1237781,8377	52,115
		T-:	-145,5213	-0,4872	-7,0921	-52,7287	-161,1389
	EPSO-FS2	M :	**57,14%**	**59,17%**	**85,66%**	**6003182,16**	**84,5**
		Sd:	2,51%	5,09%	0,95%	2256850,5579	40,5688
		T-:	-80,0553	-2,5979	-7,1596	-52,316	-160,7717
Brain Tumor2	CLPSO(R)	M :	30,77%	45,22%	54,25%	25252029,0952	938
		Sd:	0,00%	3,58%	6,68%	15765058,7788	313,9256
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**9,57%**	48,55%	57,46%	**72471,34**	**4,85**
		Sd:	5,65%	13,25%	6,31%	22716,5443	2,4654
		T-:	-149,4516	5,2163	66,7848	-73,1902	-176,5398
	EPSO-FS2	M :	**5,02%**	44,70%	56,94%	**2345996,91**	**4,69**
		Sd:	4,41%	13,72%	5,81%	584160,9526	2,0546
		T-:	-221,375	-0,5016	10,6126	-66,5521	-176,5705
Prostate Tumor	CLPSO(R)	M :	1,28%	30,82%	32,75%	181622434,1667	2175,5
		Sd:	1,89%	16,90%	5,28%	70355673,1626	593,58
		T-:	-	-	-	-	-
	PSO-FS 2	M :	1,25%	**18,43%**	**22,39%**	**77045,60**	**6,02**
		Sd:	1,82%	5,26%	6,62%	23842,5382	3,1281
		T-:	-0,2692	-14,8792	-29,3406	-727,1205	-403,5891
	EPSO-FS2	M :	1,75%	**17,13%**	**20,25%**	**2521959,63**	**6,45**
		Sd:	1,96%	4,03%	6,31%	827152,3457	3,0194
		T-:	3,5013	-15,8998	-42,0877	-717,0533	-403,539
Sonar	CLPSO(R)	M :	12,24%	27,19%	40,35%	31605,2727	23,5455
		Sd:	2,76%	2,43%	2,24%	6516,7455	4,298
		T-:	-	-	-	-	-
	PSO-FS 2	M :	12,11%	27,67%	**39,30%**	32962,75	**14,93**
		Sd:	1,93%	3,19%	1,95%	13987,2921	4,337
		T-:	-1,1359	6,177	-10,2563	1,0507	-11,1887
	EPSO-FS2	M :	**11,45%**	28,06%	**38,71%**	29799,36	**13,09**
		Sd:	2,10%	3,13%	2,57%	10395,1351	3,1306
		T-:	-3,7358	6,0333	-10,2131	-1,4078	-13,0246

Table E.26: CLPSO: multi-filters assessement (3)

Data	construction stage	Measure	Fitness	Validation1	Validation2	CPU (ms)	# Attrib.
Spam Base	CLPSO(R)	M :	14,06%	9,71%	15,62%	2003230	38,1
		Sd:	1,01%	1,13%	2,01%	450056,7096	4,9766
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**10,58%**	10,04%	**14,14%**	**907429,20**	**15,05**
		Sd:	0,77%	1,39%	2,00%	429717,5333	7,3804
		T-:	-28,8566	5,7659	-13,7928	-34,4666	-79,2002
	EPSO-FS2	M :	**9,53%**	**9,54%**	**13,68%**	**1416780,73**	**18,08**
		Sd:	0,70%	0,71%	1,20%	630673,4262	5,6882
		T-:	-36,1952	-2,4729	-19,7909	-15,5243	-47,2971
Soybean	CLPSO(R)	M :	12,76%	7,23%	60,49%	1326505,0909	27,3636
		Sd:	1,92%	0,54%	2,31%	145369,2342	2,1106
		T-:	-	-	-	-	-
	PSO-FS 2	M :	13,06%	**6,20%**	**58,91%**	1526667,09	**20,71**
		Sd:	0,83%	0,74%	0,76%	470670,6078	3,218
		T-:	1,9955	-14,6107	-5,9657	11,7062	-43,6658
	EPSO-FS2	M :	12,63%	**6,25%**	58,79%	1654804,5	**21,77**
		Sd:	0,76%	0,55%	0,23%	333589,0991	2,5622
		T-:	-0,8848	-13,8041	-6,4145	16,2701	-44,9442
Arrhythmia	CLPSO(R)	M :	33,95%	41,09%	42,72%	1756141,0909	109,3636
		Sd:	1,07%	2,30%	0,55%	761999,0311	17,7329
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**28,14%**	**37,89%**	43,18%	**354455,17**	**26,77**
		Sd:	1,64%	2,36%	1,25%	113644,6349	9,1329
		T-:	-117,5489	-11,9218	9,3242	-16,6906	-60,8355
	EPSO-FS2	M :	**26,74%**	**37,75%**	42,83%	**364101,39**	**26,91**
		Sd:	1,19%	2,54%	0,92%	110878,833	7,5553
		T-:	-96,6794	-12,043	2,3814	-16,5755	-61,9657
Secom	CLPSO(R)	M :	6,41%	9,34%	7,75%	681233,5	35,5
		Sd:	0,24%	1,13%	0,16%	366131,4789	14,5239
		T-:	-	-	-	-	-
	PSO-FS 2	M :	6,36%	**6,84%**	9,67%	**138875,61**	**4,76**
		Sd:	0,28%	0,23%	1,55%	57824,3409	2,9061
		T-:	-1,6572	-85,9674	78,8573	-31,3938	-32,3476
	EPSO-FS2	M :	**6,21%**	**6,71%**	10,11%	**130686,47**	**3,69**
		Sd:	0,35%	0,06%	1,43%	39824,3147	1,0196
		T-:	-6,6724	-90,4542	94,7166	-32,0233	-33,475
semeion	CLPSO(R)	M :	14,76%	9,15%	81,17%	9704920,1818	158,2727
		Sd:	1,49%	0,27%	0,16%	1828334,9384	6,4667
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**14,45%**	**8,95%**	**81,13%**	16115367,55	160,07
		Sd:	0,95%	0,79%	0,18%	5351789,3788	18,4007
		T-:	-7,2709	-5,6025	-2,2559	35,2408	1,8238
	EPSO-FS2	M :	**13,44%**	9,14%	81,23%	15351833,6522	**153,13**
		Sd:	0,58%	0,55%	0,16%	4786029,9937	10,1323
		T-:	-30,3947	-0,2555	2,874	40,7089	-4,8889
Lymphoma	CLPSO(R)	M :	6,47%	1,58%	17,45%	7289321,2	454,8
		Sd:	4,34%	1,72%	2,25%	3538625,7463	78,2387
		T-:	-	-	-	-	-
	PSO-FS 2	M :	**0,00%**	**12,51%**	14,04%	**27030,34**	**3,56**
		Sd:	0,00%	5,78%	5,34%	7993,7106	1,8848
		T-:	-12,2222	43,2493	-13,0834	-33,1507	-1596,8734
	EPSO-FS2	M :	**0,00%**	**10,04%**	14,84%	**437518,43**	**7,43**
		Sd:	0,00%	6,83%	4,99%	119953,4412	6,4443
		T-:	-12,2222	23,2788	-5,1937	-31,2769	-1130,2255

Table E.27: CLPSO: multi-filters assessement (4)

Bibliography

[1] *Kent Ridge Bio-medical Dataset*, 2002.

[2] *GEMS: Gene Expression Model Selector*, 2005.

[3] T. Abeel, T. Helleputte, Y. Van de Peer, P. Dupont, and Y. Saeys. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. *Bioinformatics*, 26(3):392–398, 2010.

[4] A. Abraham, C. Grosan, and V. Ramos. *Swarm Intelligence in Data Mining*, volume 34 of *Studies in Computational Intelligence*. Springer, 2006.

[5] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri. Text feature selection using ant colony optimization. *Expert Systems with Applications*, 36(3):6843 – 6853, 2009.

[6] R. M. Aiex, S. Binato, and M. G. C. Resende. Parallel GRASP with path-relinking for job shop scheduling. *Parallel Computing*, 29:393–430, April 2003.

[7] A. Al-Ani. Ant colony optimization for feature subset selection. In *World Enformatika Conference, (WEC'05), Istanbul, Turkey*, pages 35–38, February 2005.

[8] A. Al-Ani. A dependency-based search strategy for feature selection. *Expert Systems with Applications*, 36(10):12392 – 12398, 2009.

[9] H. Almuallim and T. G. Dietterich. Learning with many irrelevant features. In *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI-91)*, volume 2, pages 547–552, Anaheim, California, 1991. AAAI Press.

[10] M. Argüello, J. Bard, and G. Yu. A GRASP for aircraft routing in response to groundings and delays. *Journal of Combinatorial Optimization*, 1:211–228, 1997.

[11] R. Armañanzas, I. Inza, R. Santana, Y. Saeys, J. L. Flores, J. A. Lozano, Y. V. de Peer, R. Blanco, V. Robles, C. Bielza, and P. Larrañaga. A review of estimation of distribution algorithms in bioinformatics. *BioData Mining*, 1, 2008.

[12] I. Babaoglu, O. Findik, and E. lker. A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine. *Expert Systems with Applications*, 37(4):3177 – 3183, 2010.

[13] L. Backstrom and R. Caruana. C2FS: An algorithm for feature selection in cascade neural networks. In *Proceedings of the International Joint Conference on Neural Networks, (IJCNN'06), part of the IEEE World Congress on Computational Intelligence, (WCCI'06), Vancouver, BC, Canada*, pages 4748–4753, July 2006.

[14] J. Bala, K. D. Jong, J. Huang, H. Vafaie, and H. Wechsler. Using learning to facilitate the evolution of features for recognizing visual concepts. *Evolutionary Computation*, 4(3):297–311, 1996.

[15] R. Battiti. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks*, 5:537–550, 1994.

[16] C. Blake and C. Merz. *UCI repository of machine learning databases*, 1998.

[17] R. Blanco, P. Larrañaga, I. Inza, and B. Sierra. Gene selection for cancer classification using wrapper approaches. *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*, 18(8):1373–1390, 2004.

[18] M. Boudia, M. A. O. Louly, and C. Prins. A reactive GRASP and path relinking for a combined production-distribution problem. *Computers and Operations Research*, 34(11):3402–3419, 2007.

[19] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[20] C. Carreto and B. Baker. An improved GRASP interactive approach for the vehicle routing problem with backhauls. In *Essays and Surveys in Metaheuristics*, pages 185–199. Kluwer Academic Publishers, 2001.

[21] L. Chuang, H. Chang, C. Tu, and C. Yang. Improved binary PSO for feature selection using gene expression data. *Computational Biology and Chemistry*, 32(1):29–38, 2008.

[22] M. Clerc and J. Kennedy. The particle swarm - explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6(1):58–73, 2002.

[23] M. Dash. Feature selection via set cover. In *KDEX'97: Proceedings of the 1997 IEEE Knowledge and Data Engineering Exchange Workshop*, page 165, Washington, DC, USA, 1997. IEEE Computer Society.

[24] M. Dash and H. Liu. Consistency-based search in feature selection. *Artificial Intelligence*, 151(1-2):155–176, 2003.

[25] Del, G. K. Venayagamoorthy, S. Mohagheghi, J. C. Hernandez, and R. G. Harley. Particle swarm optimization: Basic concepts, variants and applications in power systems. *IEEE Transactions on Evolutionary Computation*, 12(2):171–195, 2008.

[26] H. Delmaire, J. A. Díaz, E. Fernández, and M. Ortega. Reactive GRASP and Tabu search based heuristics for the single source capacitated plant location problem. *Information Systems and Operational Research (INFOR)*, 37:194–225, 1999.

[27] C. Ding and H. Peng. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology*, 3(2):185–205, April 2005.

[28] J. Doak. An evaluation of feature selection methods and their application to computer security. Technical report, University of California, Department of Computer Science, 1992.

[29] W. Duch. Filter methods. In *Feature extraction, foundations and applications*, pages 89–118. Physica Verlag, Springer, 2006.

[30] C. Emmanouilidis, A. Hunter, and J. MacIntyre. A multiobjective evolutionary setting for feature selection and a commonality-based crossover operator. In *Proceedings of the Congress on Evolutionary Computation (CEC'00)*, pages 309–316, California, USA, 2000. IEEE Press.

[31] A. P. Engelbrecht. *Computational Intelligence: An Introduction*. John Wiley & Sons, InterEditions, 2007.

[32] M. A. Esseghir. Effective wrapper-filter hybridization through GRASP schemata. *International Journal of Machine Learning Research*, 10(45-54), 2010.

[33] M. A. Esseghir, G. Goncalves, and Y. Slimani. Distributed feature selection: benchmarking collaboration protocol. In *Colloque sur l'optimisation et les systèmes d'information (COSI'09), Annaba*, May 2009.

[34] M. A. Esseghir, G. Goncalves, and Y. Slimani. Memetic feature selection: Benchmarking hybridization schemata. In *Proceedings of the 5th International Conference on Hybrid Artificial Intelligence Systems (HAIS'10), San Sebastián, Spain, Part I*, pages 351–358, June 2010.

[35] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada. Normalized mutual information feature selection. *IEEE Transactions on Neural Networks*, 20(2):189–201, 2009.

[36] T. A. Feo and M. G. C. Resende. A probabilistic heuristic for a computationally difficult set covering problem. *Operations Research Letters*, 8(2):67 – 71, 1989.

[37] T. A. Feo and M. G. C. Resende. Greedy randomized adaptive search procedures. *Journal of Global Optimization*, 6:109–133, 1995.

[38] P. Festa and M. Resende. GRASP: basic components and enhancements. *Telecommunication Systems*, 46:253–271, 2011.

[39] P. Festa and M. G. C. Resende. An annotated bibliography of grasp part i: Algorithms. *International Transactions in Operational Research*, 16(1):1–24, 2009.

[40] H. A. Firpi and E. Goodman. Swarmed feature selection. In *AIPR '04: Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop*, pages 112–118, Washington, DC, USA, 2004. IEEE Computer Society.

[41] A. A. Freitas. *Data Mining and Knowledge Discovery with Evolutionary Algorithms*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2002.

[42] J. García-Nieto, E. Alba, L. Jourdan, and E. Talbi. Sensitivity and specificity based multiobjective approach for feature selection: Application to cancer diagnosis. *Information Processing Letters*, 109(16):887–896, 2009.

[43] I. A. Gheyas and L. S. Smith. Feature subset selection in large dimensionality domains. *Pattern Recognition*, 43(1):5–13, 2010.

[44] F. Glover. Tabu Search - Part I. *INFORMS Journal on Computing*, 1(3):190–206, 1989.

[45] F. Glover. Tabu search and adaptive memory programing advances, applications and challenges. In *Interfaces in Computer Science and Operations Research*, pages 1–75. Kluwer, 1996.

[46] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.

[47] F. C. Gomes, P. Pardalos, C. S. Oliveira, and M. G. C. Resende. Reactive GRASP with path relinking for channel assignment in mobile phone networks. In *Proceedings of the 5th International Workshop on Discrete Algorithms and methods for mobile computing and communications*, DIALM '01, pages 60–67, New York, NY, USA, 2001. ACM.

[48] P. M. Granitto, F. Biasioli, C. Furlanello, and F. Gasperi. Efficient feature selection for PTR-MS fingerprinting of agroindustrial products. In *ICANN'08: Proceedings of the 18th International Conference on Artificial Neural Networks*, pages 42–51, 2008.

[49] P. M. Granitto, C. Furlanello, F. Biasioli, and F. Gasperi. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. *Chemometrics and Intelligent Laboratory Systems*, 83(2):83–90, 2006.

[50] P. M. Granitto, P. F. Verdes, and H. A. Ceccatto. Neural network ensembles: evaluation of aggregation algorithms. *Artificial Intelligence*, 163(2):139–162, 2005.

[51] A. Guillén, H. Pomares, J. González, I. Rojas, O. Valenzuela, and B. Prieto. Parallel multiobjective memetic rbfnns design and feature selection for function approximation problems. *Neurocomputing*, 72:3541–3555, October 2009.

[52] M. Gutlein, E. Frank, M. Hall, and A. Karwath. Large-scale attribute selection using wrappers. In *CIDM'09: Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2009, part of the IEEE Symposium Series on Computational Intelligence 2009, Nashville, TN, USA,*, pages 332–339, March 2009.

[53] I. Guyon and A. Elisseff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, March 2003.

[54] I. Guyon, S. Gunn, M. Nikravesh, and L. Zadeh. *Feature Extraction, Foundations and Applications.* Series Studies in Fuzziness and Soft Computing. Springer, 2006.

[55] D. Hadjidj and H. Drias. GRASP and guided local search for the examination timetabling problem. *International Journal of Artificial Intelligence and Soft Computing*, 2:103–114, April 2010.

[56] M. A. Hall. *Correlation-based feature selection for machine learning*. PhD thesis, Department of Computer Science, University of Waikato, Hamilton, New Zealand, 1998.

[57] M. A. Hall and G. Holmes. Benchmarking attribute selection techniques for discrete class data mining. *IEEE Transactions on Knowledge and Data Engineering*, 15(3), June 2003.

[58] W. E. Hart, N. Krasnogor, and J. Smith. *Recent Advances in Memetic Algorithms*. Studies in Fuzziness and Soft Computing. Springer, 2004.

[59] S. J. Hong. Use of contextual information for feature ranking and discretization. *IEEE Transactions on Knowledge and Data Engineering*, 9:718–730, 1997.

[60] H. H. Hoos and T. Stützle. *Stochastic Local Search: Foundations and Applications*. Elsevier / Morgan Kaufmann, 2004.

[61] C. Huang and C. Wang. A GA-based feature selection and parameters optimization for support vector machines. *Expert Systems with Applications*, 31(2):231–240, August 2006.

[62] J. Huang, Y. Cai, and X. Xu. A hybrid genetic algorithm for feature selection wrapper based on mutual information. *Pattern Recognition Letters*, 28(13):1825–1844, October 2007.

[63] F. Hussein. Genetic algorithms for feature selection and weighting, a review and study. In *ICDAR '01: Proceedings of the Sixth International Conference on Document Analysis and Recognition*, pages 1240–1244, Washington, DC, USA, 2001. IEEE Computer Society.

[64] I. Inza, M. Merino, P. Larraaga, J. Quiroga, B. Sierra, and M. Girala. Feature subset selection by genetic algorithms and estimation of distribution algorithms.

a case study in the survival of cirrhotic patients treated with tips. *Artificial Intelligence in Medicine*, 23:187–205, October 2001.

[65] A. Jain and D. Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:153–158, 1997.

[66] A. Jakulin and I. Bratko. Analyzing attribute dependencies. In N. Lavrac, D. Gamberger, L. Todorovski, and H. Blockeel, editors, *Knowledge Discovery in Databases: PKDD 2003*, volume 2838 of *Lecture Notes in Computer Science*, pages 229–240. Springer Berlin / Heidelberg, 2003.

[67] A. Jakulin and I. Bratko. Testing the significance of attribute interactions. In *In Proc. of 21st International Conference on Machine Learning (ICML'04)*, pages 409–416. ACM Press, 2004.

[68] G. John, R. Kohavi, and K. Pfleger. Irrelevant features and the subset selection problem. In *International Conference on Machine Learning (ICML'94),New Brunswick, NJ, USA*, pages 121–129, 1994.

[69] H. R. Kanan, K. Faez, and S. M. Taheri. Feature selection using ant colony optimization (ACO): A new method and comparative study in the application of face recognition system. In *Industrial Conference on Data Mining, Leipzig, Germany*, pages 63–76, 2007.

[70] S. S. Kannan and N. Ramaraj. A novel hybrid feature selection via symmetrical uncertainty ranking based local memetic search algorithm. *Knowledge-Based Systems*, 23(6):580–585, 2010.

[71] J. Kennedy. Population structure and particle swarm performance. In *In: Proceedings of the Congress on Evolutionary Computation (CEC'02)*, pages 1671–1676. IEEE Press, 2002.

[72] J. Kennedy and R. C. Eberhart. *Swarm intelligence.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2001.

[73] R. N. Khushaba, A. Al-Ani, and A. Al-Jumaily. Feature subset selection using differential evolution. In *ICONIP: 15th International Conference on Neural Information Processing, Auckland, New Zealand*, pages 103–110, 2008.

[74] R. N. Khushaba, A. Al-Ani, A. Al-Sukker, and A. Al-Jumaily. A combined ant colony and differential evolution feature selection algorithm. In *Ant Colony Optimization and Swarm Intelligence, 6th International Conference, (ANTS'08), Brussels, Belgium*, pages 1–12, September 2008.

[75] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[76] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97:273–324, 1997.

[77] D. Koller and M. Sahami. Toward optimal feature selection. pages 284–292. Morgan Kaufmann, 1996.

[78] I. Kononenko. Estimating attributes: Analysis and extensions of RELIEF. In F. Bergadano and L. De Raedt, editors, *Machine Learning: ECML-94*, volume 784 of *Lecture Notes in Computer Science*, pages 171–182. Springer, 1994.

[79] M. Kudo and J. Sklansky. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, 33(1):25–41, 2000.

[80] M. Laguna and R. Martí. GRASP and path relinking for 2-layer straight line crossing minimization. *INFORMS Journal on Computing*, 11:44–52, 1999.

[81] P. Larraanaga and J. A. Lozano. *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.

[82] G. V. Lashkia and L. Anthony. Relevant, irredundant feature selection and noisy example elimination. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 34(2):888–897, 2004.

[83] H. Lei and V. Govindaraju. A comparative study on the consistency of features in on-line signature verification. *Pattern Recognition Letters*, 26(15):2483–2489, 2005.

[84] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. *IEEE Trans. Evolutionary Computation*, 10(3):281–295, 2006.

[85] S. Lin, Z. Lee, S. Chen, and T. Tseng. Parameter determination of support vector machine and feature selection using simulated annealing approach. *Applied Soft Computing*, 8(4):1505–1512, 2008. Soft Computing for Dynamic Data Mining.

[86] S. Lin, T. Tseng, S. Chou, and S. Chen. A simulated-annealing-based approach for simultaneous parameter optimization and feature selection of back-propagation networks. *Expert Systems with Applications*, 34(2):1491–1499, 2008.

[87] H. Liu, E. R. Dougherty, J. G. Dy, K. Torkkola, E. Tuv, H. Peng, C. Ding, F. Long, M. Berens, L. Parsons, Z. Zhao, L. Yu, and G. Forman. Evolving feature selection. *IEEE Intelligent Systems*, 20(6):64–76, 2005.

[88] H. Liu, L. Liu, and H. Zhang. Feature selection using mutual information: An experimental study. In *PRICAI '08: Proceedings of the 10th Pacific Rim International Conference on Artificial Intelligence*, pages 235–246, Berlin, Heidelberg, 2008. Springer-Verlag.

[89] H. Liu, L. Liu, and H. Zhang. Boosting feature selection using information metric for classification. *Neurocomputing*, 73:295–303, December 2009.

[90] H. Liu and H. Motoda. *Computational methods of feature selection*. Chapman and Hall/CRC Editions, 2008.

[91] H. Liu, H. Motoda, and M. Dash. A monotonic measure for optimal feature selection. In *In Proceedings of European Conference on Machine Learning*, pages 101–106, 1998.

[92] H. Liu and R. Setiono. Chi2: Feature selection and discretization of numeric attributes. In *International Conference on Tools with Artificial Intelligence (ICTAI '95)*, pages 388–391, 1995.

[93] H. Liu and L. Yu. Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, 17(4):491–502, 2005.

[94] F. G. López, M. G. Torres, B. M. Batista, J. A. M. Pérez, and J. M. Moreno-Vega. Solving feature subset selection problem by a parallel scatter search. *European Journal of Operational Research*, 2:477–489, march 2006.

[95] J. A. Lozano, P. Larra naga, I. Inza, and E. Bengoetxea. *Towards a New Evolutionary Computation: Advances on Estimation of Distribution Algorithms (Studies in Fuzziness and Soft Computing)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[96] P. E. Lutu and A. P. Engelbrecht. A decision rule-based method for feature selection in predictive data mining. *Expert Systems with Applications*, 37(1):602–609, 2010.

[97] R. Meiri and J. Zahavi. Using simulated annealing to optimize the feature selection problem in marketing applications. *European Journal of Operational Research*, 171(3):842–858, 2006.

[98] R. Mendes, J. Kennedy, and J. Neves. The fully informed particle swarm: Simpler, maybe better. *IEEE Transactions on Evolutionary Computation*, 8(3):204–210, 2004.

[99] A. Moser and M. N. Murty. On the scalability of genetic algorithms to very large-scale feature selection. In *EvoWorkshops, Edinburgh, Scotland, UK*, pages 77–86, April 2000.

[100] S. Nakariyakul and D. P. Casasent. Adaptive branch and bound algorithm for selecting optimal features. *Pattern Recognition Letters*, 28(12):1415–1427, 2007.

[101] S. Nakariyakul and D. P. Casasent. An improvement on floating search algorithms for feature subset selection. *Pattern Recognition*, 42(9):1932–1940, 2009.

[102] S. Nemati, M. E. Basiri, N. Ghasem-Aghaee, and M. H. Aghdam. A novel ACO-GA hybrid algorithm for feature selection in protein function prediction. *Expert Systems with Applications*, 36(10):12086–12094, 2009.

[103] I. Oh, J. Lee, and B. Moon. Hybrid genetic algorithms for feature selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11):1424–1437, 2004.

[104] C. H. Ooi, M. Chetty, and S. W. Teng. Differential prioritization in feature selection and classifier aggregation for multiclass microarray datasets. *Data Mining and Knowledge Discovery*, 14(3):329–366, 2007.

[105] K. E. Parsopoulos and M. N. Vrahatis. Unified particle swarm optimization in dynamic environments. In *EvoWorkshops*, pages 590–599, 2005.

[106] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27:1226–1238, 2005.

[107] O. M. Pérez, M. Hidalgo-Conde, F. J. Marín, and O. Trelles. Weighting and feature selection on gene-expression data by the use of genetic algorithms. In *IWANN '03: Proceedings of the 7th International Work-Conference on Artificial and Natural Neural Networks*, pages 377–384, Berlin, Heidelberg, 2003. Springer-Verlag.

[108] M. Prais and C. C. Ribeiro. Reactive GRASP: An application to a matrix decomposition problem in TDMA traffic assignment. *INFORMS Journal on Computing*, 12:164–176, July 2000.

[109] P. Pudil, J. Novovicová, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, 1994.

[110] J. R. Quinlan. Induction of decision trees. In B. G. Buchanan and D. C. Wilkins, editors, *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Expert Systems*, pages 349–361. Kaufmann, San Mateo, CA, 1993.

[111] M. G. C. Resende and C. C. Ribeiro. GRASP with path-relinking: Recent advances and applications. In *Metaheuristics: Progress as Real Problem Solvers*, pages 29–63. Springer, 2005.

[112] M. Robnik-Sikonja and I. Kononenko. An adaptation of Relief for attribute estimation in regression. In *ICML '97: Proceedings of the Fourteenth International Conference on Machine Learning*, pages 296–304, 1997.

[113] M. Robnik-Sikonja and I. Kononenko. Theoretical and empirical analysis of ReliefF and RReliefF. *Machine Learning*, 53(1-2):23–69, 2003.

[114] R. Ruiz, J. C. Riquelme, and J. S. Aguilar-Ruiz. Incremental wrapper-based gene selection from microarray data for cancer classification. *Pattern Recogn.*, 39(12):2383–2392, 2006.

[115] Y. Saeys, T. Abeel, and Y. Peer. Robust feature selection using ensemble feature selection techniques. In *Proceedings of the European conference on Machine Learning and Knowledge Discovery in Databases - Part II*, ECML PKDD '08, pages 313–325, Berlin, Heidelberg, 2008. Springer-Verlag.

[116] Y. Saeys, S. Degroeve, D. Aeyels, Y. Peer, and P. Rouzé. Fast feature selection using a simple estimation of distribution algorithm: a case study on splice site prediction. In *Proceedings of the European Conference on Computational Biology (ECCB'03)*, pages 179–188, 2003.

[117] Y. Saeys, I. Inza, and P. Larrañaga. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, 2007.

[118] Y. Shi and R. Eberhart. A modified particle swarm optimizer. In *IEEE World Congress on Computational Intelligence, Anchorage, Alaska, USA*, pages 69–73, 1998.

[119] W. Siedlecki and J. Sklansky. Handbook of pattern recognition and computer vision. pages 88–107, 1993.

[120] P. Somol, P. Pudil, J. Novovičová, and P. Paclík. Adaptive floating search methods in feature selection. *Pattern Recognition Letters*, 20(11-13):1157–1163, 1999.

[121] R. Storn and K. Price. Differential Evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[122] Y. Sun and J. Li. Iterative RELIEF for feature weighting. In *ICML '06: Proceedings of the 23rd International Conference on Machine Learning*, pages 913–920, New York, NY, USA, 2006. ACM.

[123] M. A. Tahir, A. Bouridane, and F. Kurugollu. Simultaneous feature selection and feature weighting using hybrid Tabu Search/K-nearest neighbor classifier. *Pattern Recognition Letters*, 28(4):438–446, 2007.

[124] E.-G. Talbi, L. Jourdan, J. García-Nieto, and E. Alba. Comparison of population based metaheuristics for feature selection: Application to microarray data classification. In *AICCSA '08 Proceedings of the 2008 IEEE/ACS International Conference on Computer Systems and Applications*, pages 45–52, 2008.

[125] K. C. Tan, E. J. Teoh, Q. Yu, and K. C. Goh. A hybrid evolutionary algorithm for attribute selection in data mining. *Expert Systems with Applications*, 36(4):8616–8630, 2009.

[126] A. Trabelsi and M. A. Esseghir. New evolutionary bankruptcy forecasting model based on genetic algorithms and neural networks. In *the 17th IEEE International Conference on Tools with Artificial Intelligence*, pages 241–245, Hong Kong, 2005.

[127] O. Uncu and U. Turksen. A novel feature selection approach: Combining feature wrappers and filters. *Information Sciences*, 177(2):449–466, January 2007.

[128] Y. Wang, L. Li, J. Ni, and S. Huang. Feature selection using Tabu search with long-term memories and probabilistic neural networks. *Pattern Recognition Letters*, 30(7):661–670, 2009.

[129] WEKA. Machine learning project - university of waikato. URL http://www.cs.waikato.ac.nz/˜ml/weka.

[130] L. Yu and H. Liu. Feature selection for high-dimensional data: A fast correlation-based filter solution. pages 856–863. Twentieth International Conference on Machine Leaning (ICML-03).

[131] L. Yu and H. Liu. Efficient feature selection via analysis of relevance and redundancy. *The Journal of Machine Learning Research*, 5:1205–1224, 2004.

[132] S. Yu. *Feature Selection and Classifier Ensembles: A Study on Hyperspectral Remote Sensing Data*. PhD thesis, The University of Antwerp, 2003.

[133] S. C. Yusta. Different metaheuristic strategies to solve the feature selection problem. *Pattern Recognition Letters*, 30(5):525 – 534, 2009.

[134] H. Zhang and G. Sun. Feature selection using Tabu search method. *Pattern Recognition*, 35(3):701–711, 2002.

[135] Z. Zhao and H. Liu. On interacting features in subset selection. *Encyclopedia of Data Warehousing and Mining*, pages 1079–1084, September 2008.

[136] H. Zhu, L. Jiao, and J. Pan. Multi-population genetic algorithm for feature selection. In *Advances in Natural Computation*, volume 4222 of *Lecture Notes in Computer Science*, pages 480–487. Springer Berlin / Heidelberg, 2006.

[137] Z. Zhu, S. Jia, and Z. Ji. Towards a memetic feature selection paradigm. *IEEE Computational Intelligence Magazine*, 5:41–53, May 2010.

[138] Z. Zhu and Y. Ong. Memetic algorithms for feature selection on microarray data. In *ISNN'07: Proceedings of the 4th international symposium on Neural Networks: Advances in Neural Networks*, pages 1327–1335, Berlin, Heidelberg, 2007. Springer-Verlag.

[139] Z. Zhu, Y. Ong, and M. Dash. Markov Blanket-embedded genetic algorithm for gene selection. *Pattern Recogn.*, 40:3236–3248, November 2007.

[140] Z. Zhu, Y. Ong, and M. Dash. Wrapper-Filter feature selection algorithm using a memetic framework. *IEEE Transactions on Systems, Man, and Cybernetics*, 37(1):70–76, 2007.