SIBERIAN STATE AEROSPACE UNIVERSITY
Krasnoyarsk, Russia




MULTI-OBJECTIVE GENETIC ALGORITHMS AS AN EFFECTIVE TOOL FOR FEATURE SELECTION IN THE SPEECH-BASED EMOTION RECOGNITION PROBLEM

Christina Brester, Olga Semenkina, Maxim Sidorov

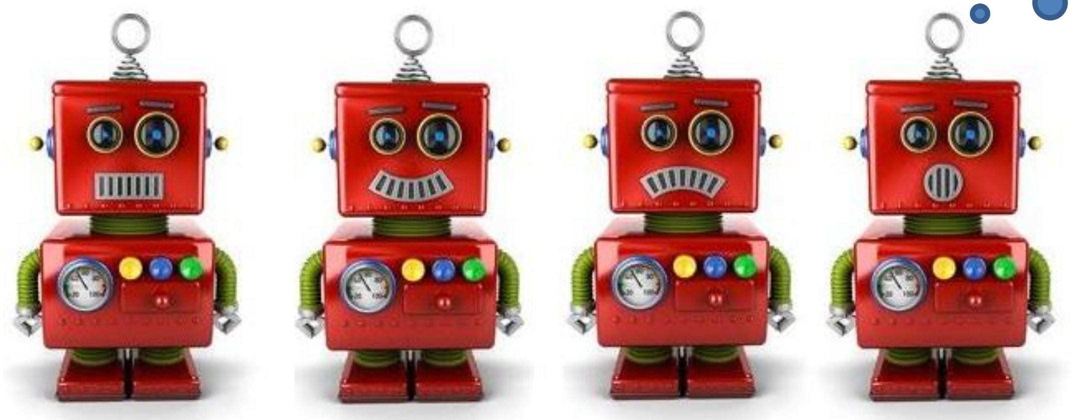



Speech-based Emotion Recognition Problem

- Spoken Dialogue Systems Improvement
- Robotics
- Call-centers quality monitoring

… etc.

Why do we talk about *Feature Selection*?

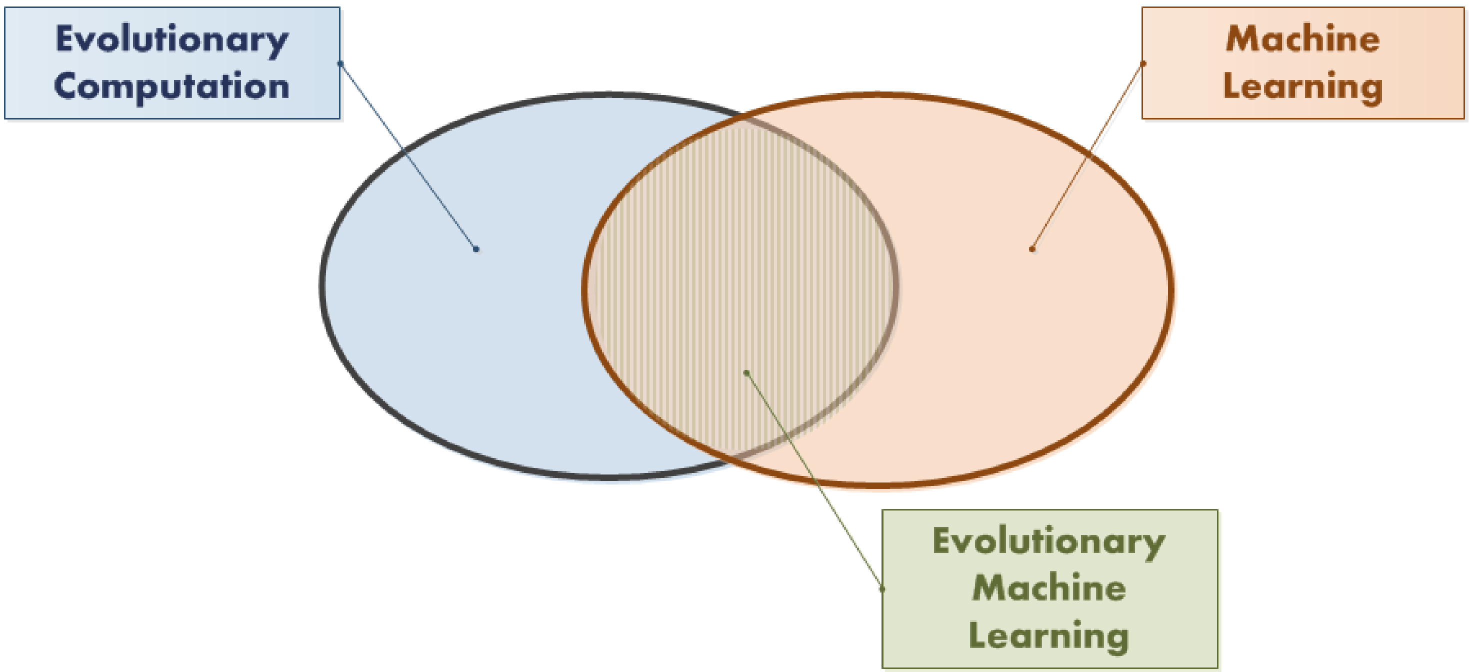


- **The number of features** extracted from the speech signal is **overwhelming**.
- **An optimal feature set** which should be used to represent the speech signals is still an open question.

Outline

- **Motivation**
  - Problem Definition
- **Background**
  - The Evolutionary Computation and Machine Learning Integration
  - Feature Selection: Wrapper or Filter
- **Proposed approach**
  - Two-criteria Filter Feature Approach
  - Multi-objective Genetic Algorithms
  - The Island Model
  - Speech-based Emotion Recognition Problem
  - Corpora Description
- **Results and Discussion**
  - Baseline
  - Heuristic Feature Selection
  - Feature Selection with the Island Model and the ensemble of classifiers
- **Conclusion and Future Plans**

Integration of Evolutionary Computation and Machine Learning

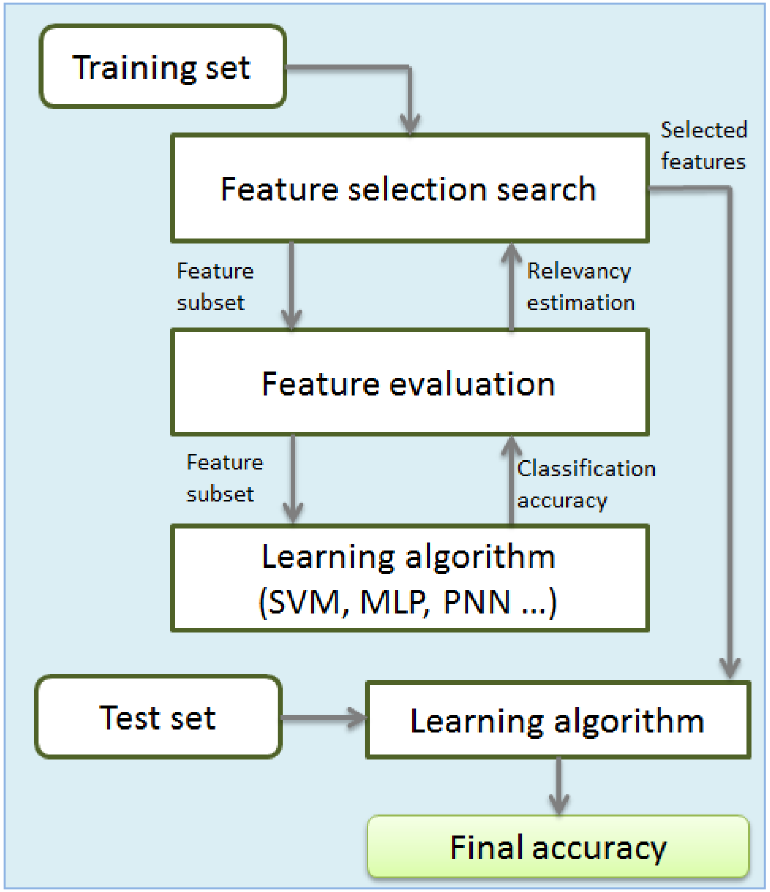


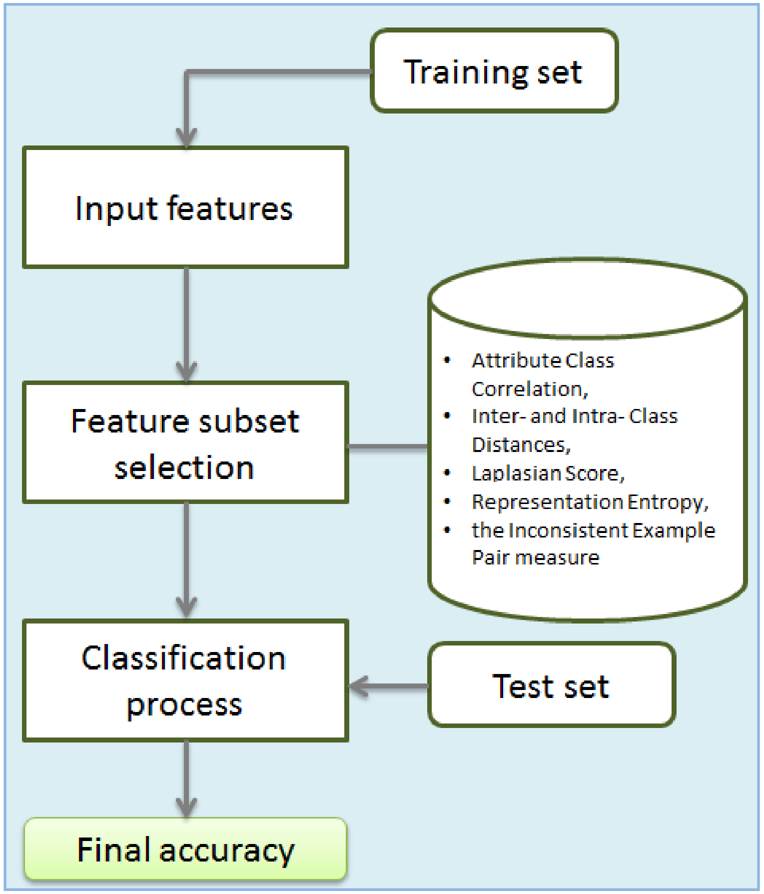
Why evolutionary?

Integration of Evolutionary Computation and Machine Learning

Pros	*Cons*
✓ The classification accuracy of the best **evolutionary** and **non-evolutionary** methods are **comparable;** ✓ Population-based search is easily **parallelized;** ✓ These methods can work in the **dynamic non-stationary environment;** ✓ **Feature selection** and **learning** in one process might be **combined;** ✓ From an optimization perspective, learning problems are typically **large, non-differentiable, noisy, deceptive, multimodal, high-dimensional, and highly constrained**. Evolutionary algorithms are an **effective tool** for such problems.	X Evolutionary methods are generally much **slower** than the non-evolutionary alternatives ➢ *Possible solution: parallelization* X The **performance of evolutionary algorithms varies** significantly for different problems ➢ *Possible solution: cooperative algorithms*

Two main feature selection concepts:
Wrapper vs Filter





Why filter?

Two main feature selection concepts

Wrapper …

✓ involves classification models to evaluate the relevancy of each feature subset: **adjustment to** an applied **classifier**;

X requires **high computational resources.**

vs

Filter …

✓ needs significantly **fewer calculations** therefore it is rather effective in the sense of computational effort;

✓ might be effectively used **in combination with an ensemble** of diverse classifiers (MLP, SVM, Logit)*;

X does **not cooperate** with a **learning algorithm** and so ignores its performance entirely.

*Brester Ch., Semenkin E., Sidorov M. Speech-based emotion recognition: Application of collective decision making concepts // International Conference on Computer Science and Artificial Intelligence (ICCSAI2014), Wuhan, China, 2014.

Two-criteria Filter Approach

- Attribute Class Correlation,
- ***Inter- and Intra- Class Distances,***
- Laplasian Score,
- Representation Entropy,
- the Inconsistent Example Pair measure

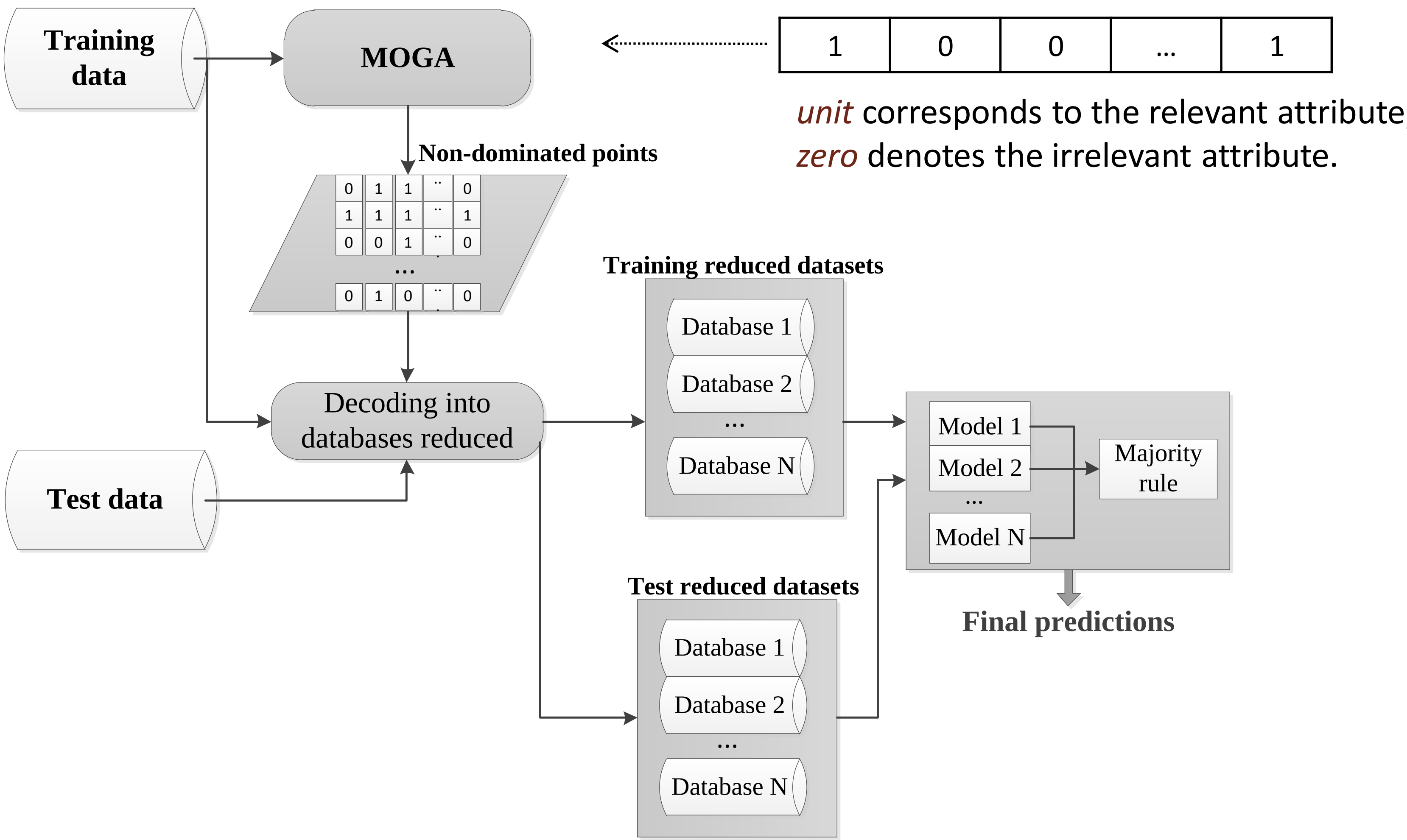
f1 – the Intra-Class Distance (IA),
f2 - the Inter-Class Distance (IE),
f1 → min, f2 → max

$$IA = \frac{1}{n}\sum_{r=1}^{k}\sum_{j=1}^{n_r} d(p_j^r, p_r),$$

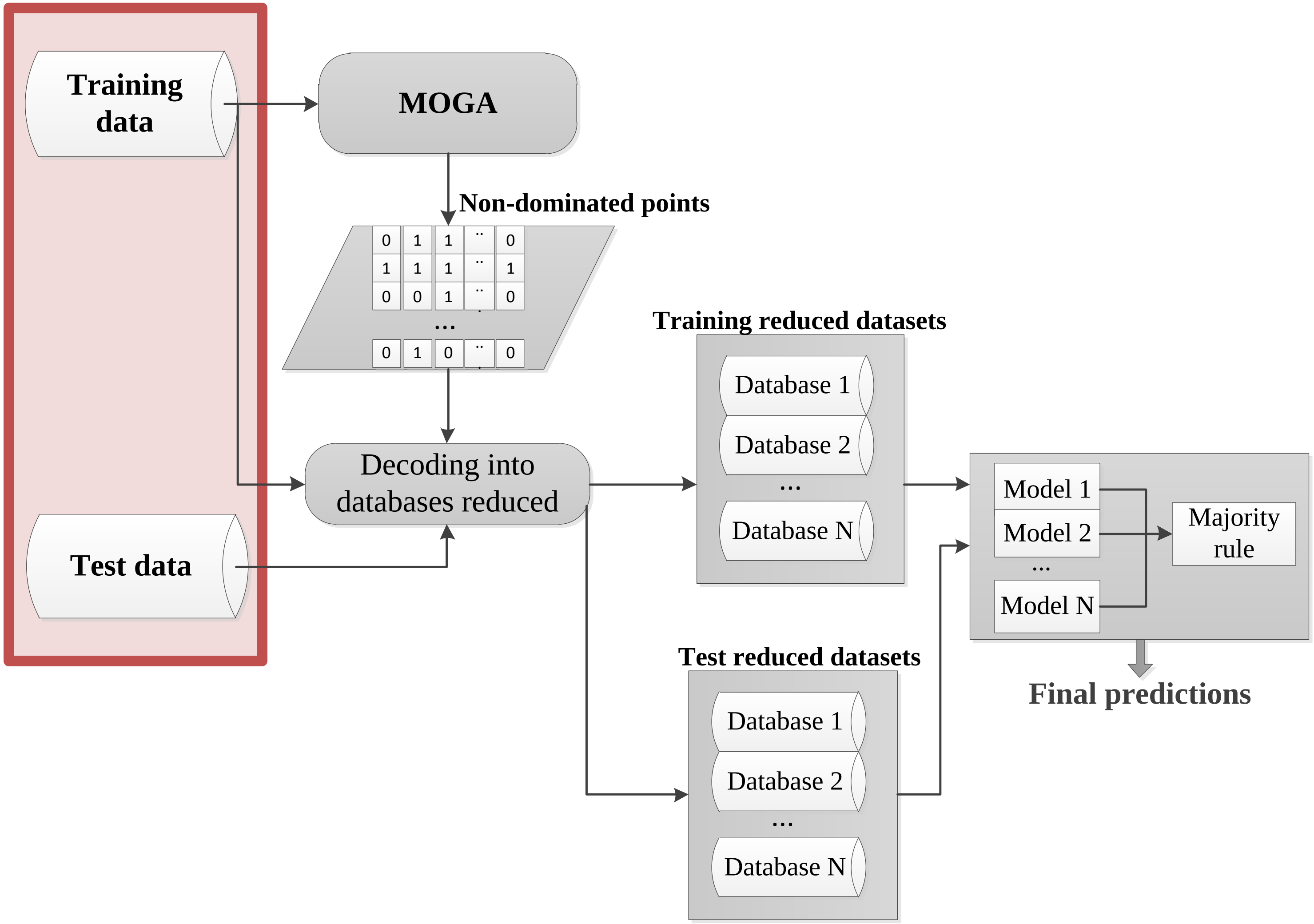
$$IE = \frac{1}{n}\sum_{r=1}^{k} n_r d(p_r, p),$$

where p_j^r is the j-th example from the r-th class,
p is the central example of the data set,
$d(\ldots,\ldots)$ denotes the Euclidian distance,
p_r and n_r represent the central example and the number of examples in the r-th class.

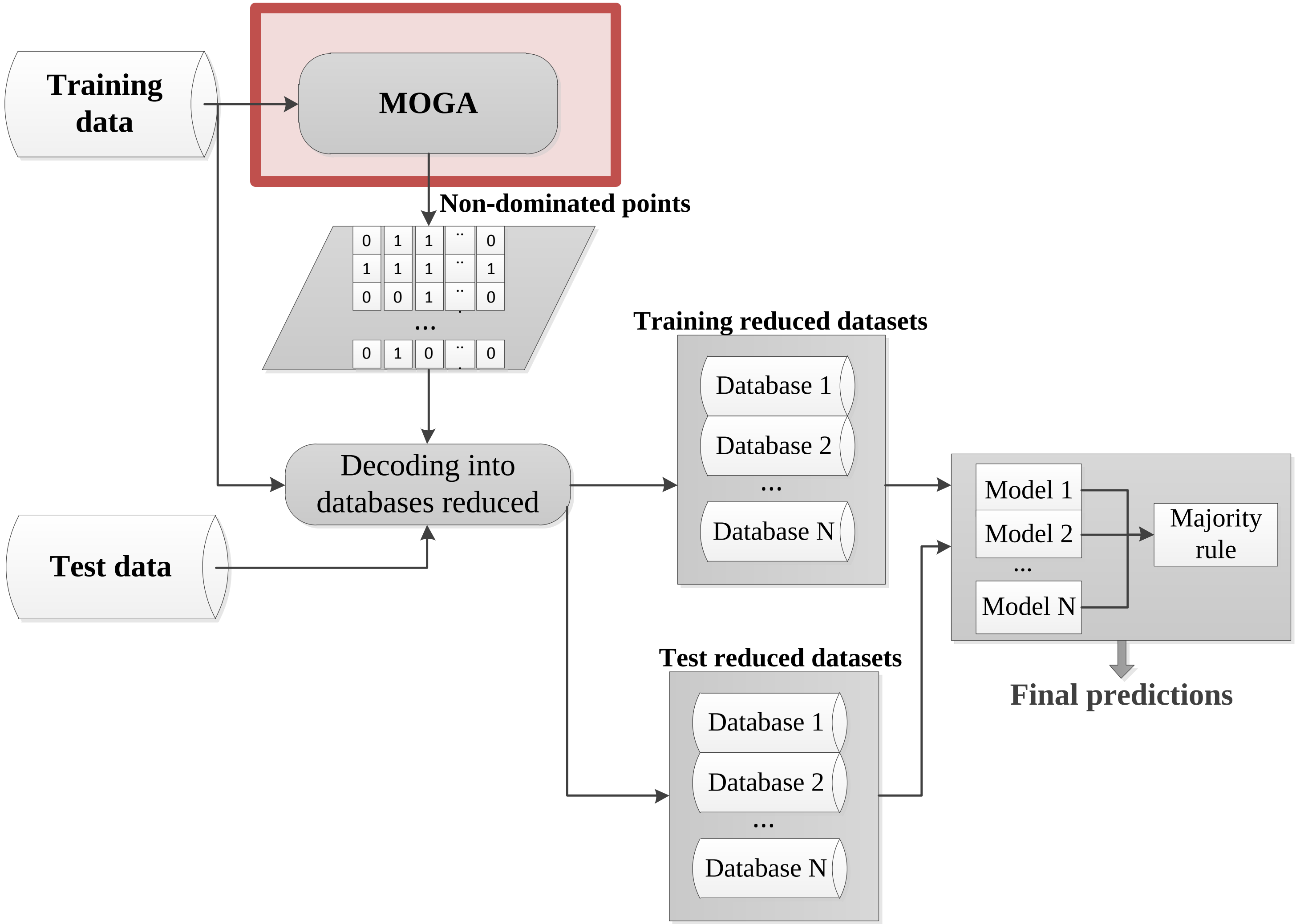
The general scheme of the approach proposed



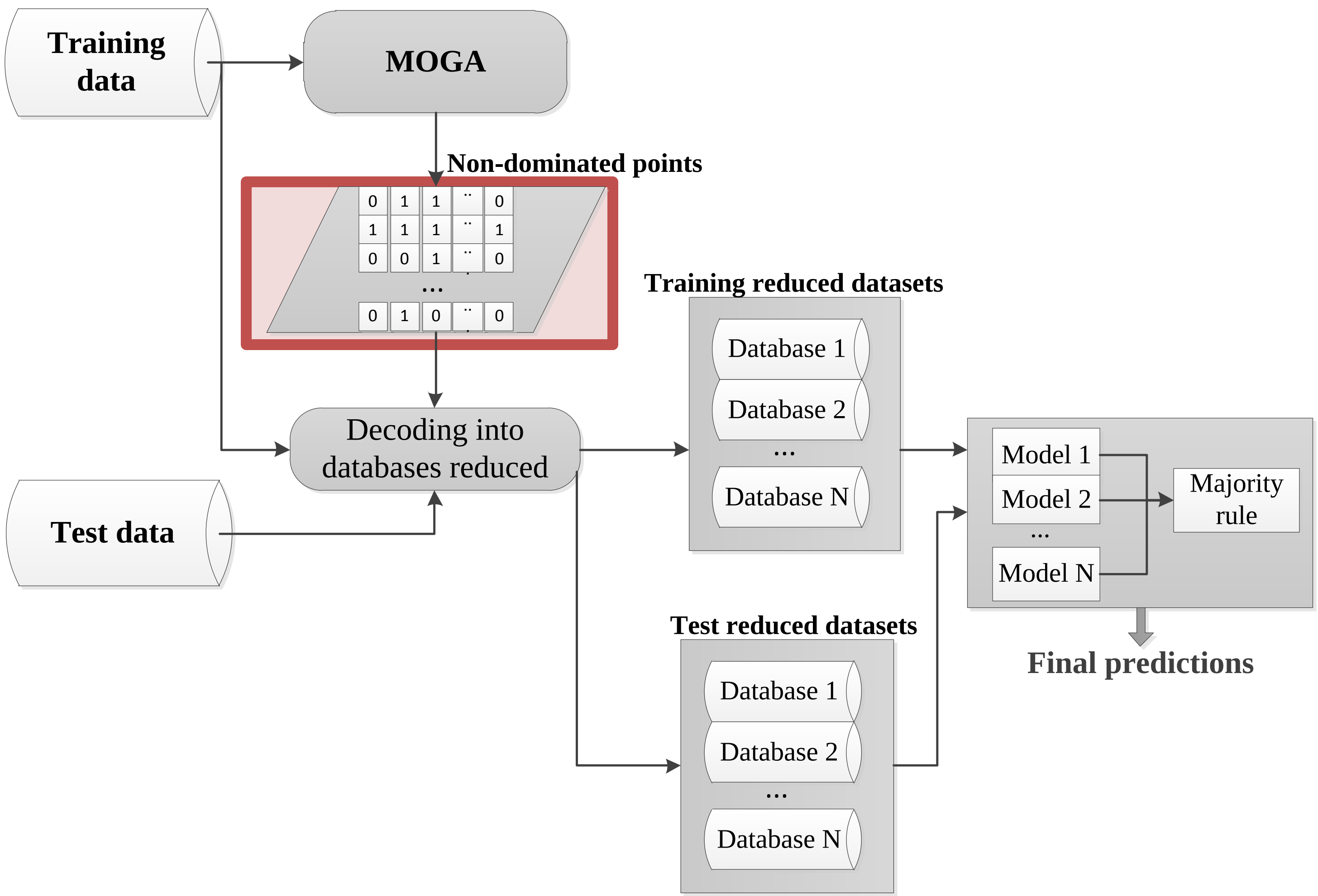
The general scheme of the approach proposed



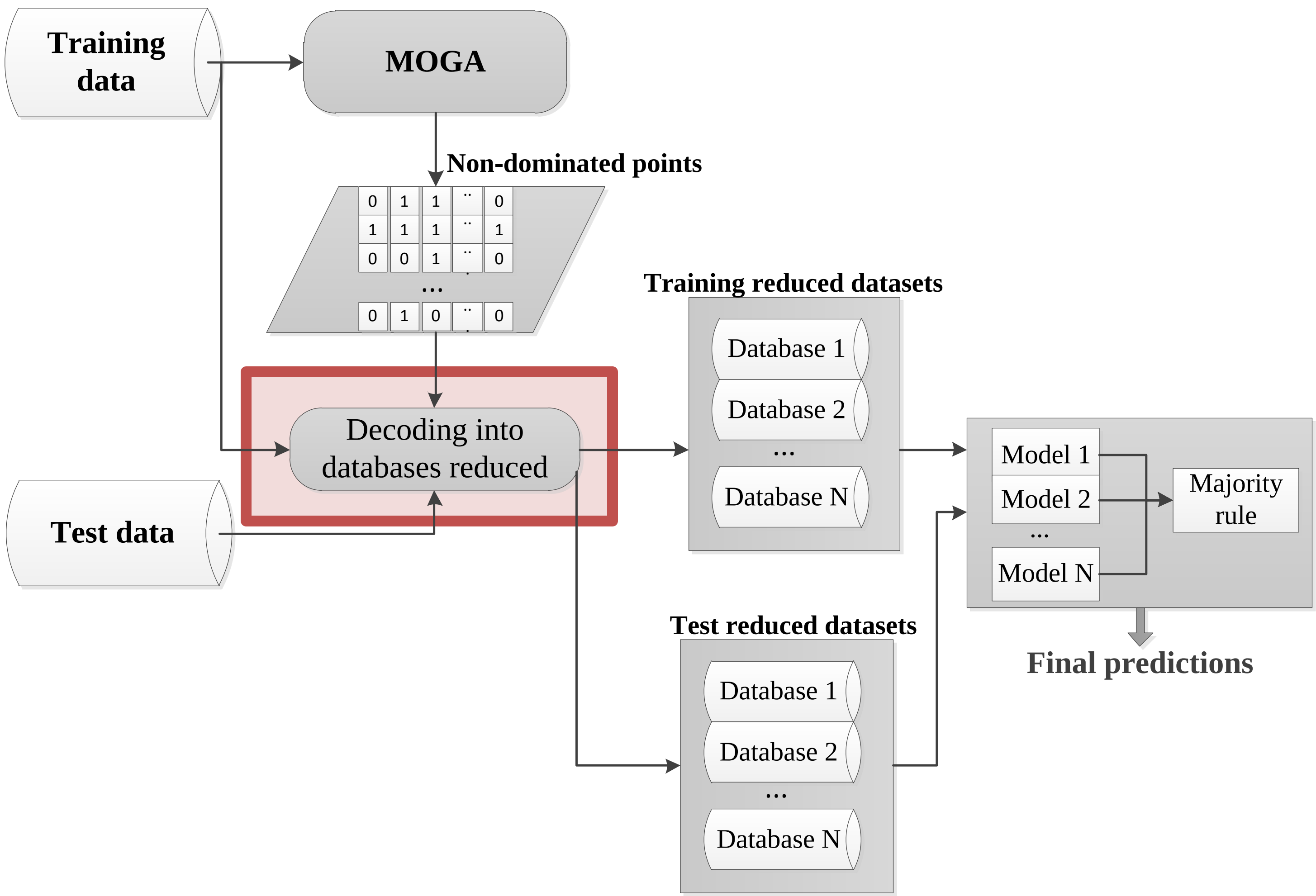
The general scheme of the approach proposed



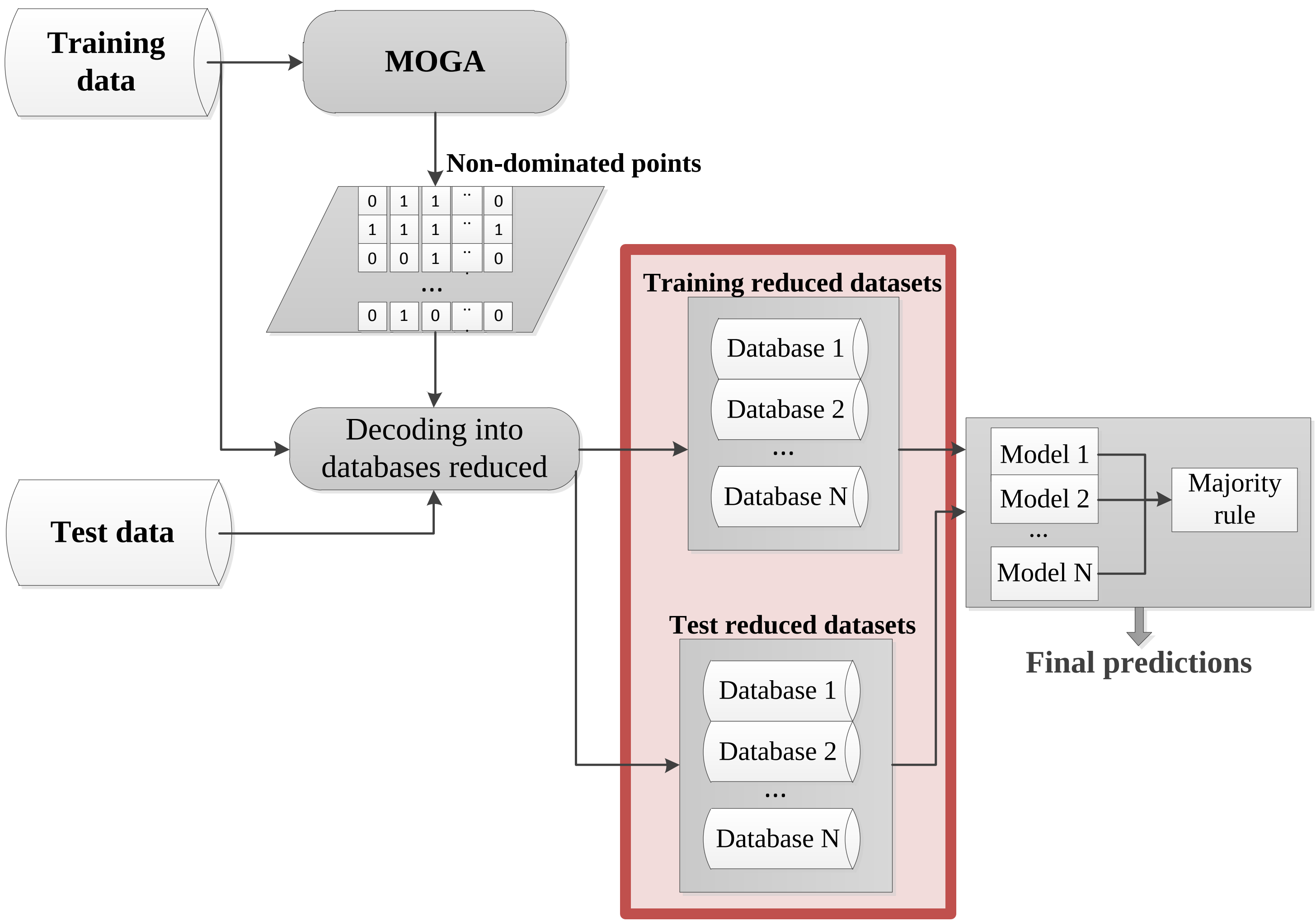
The general scheme of the approach proposed



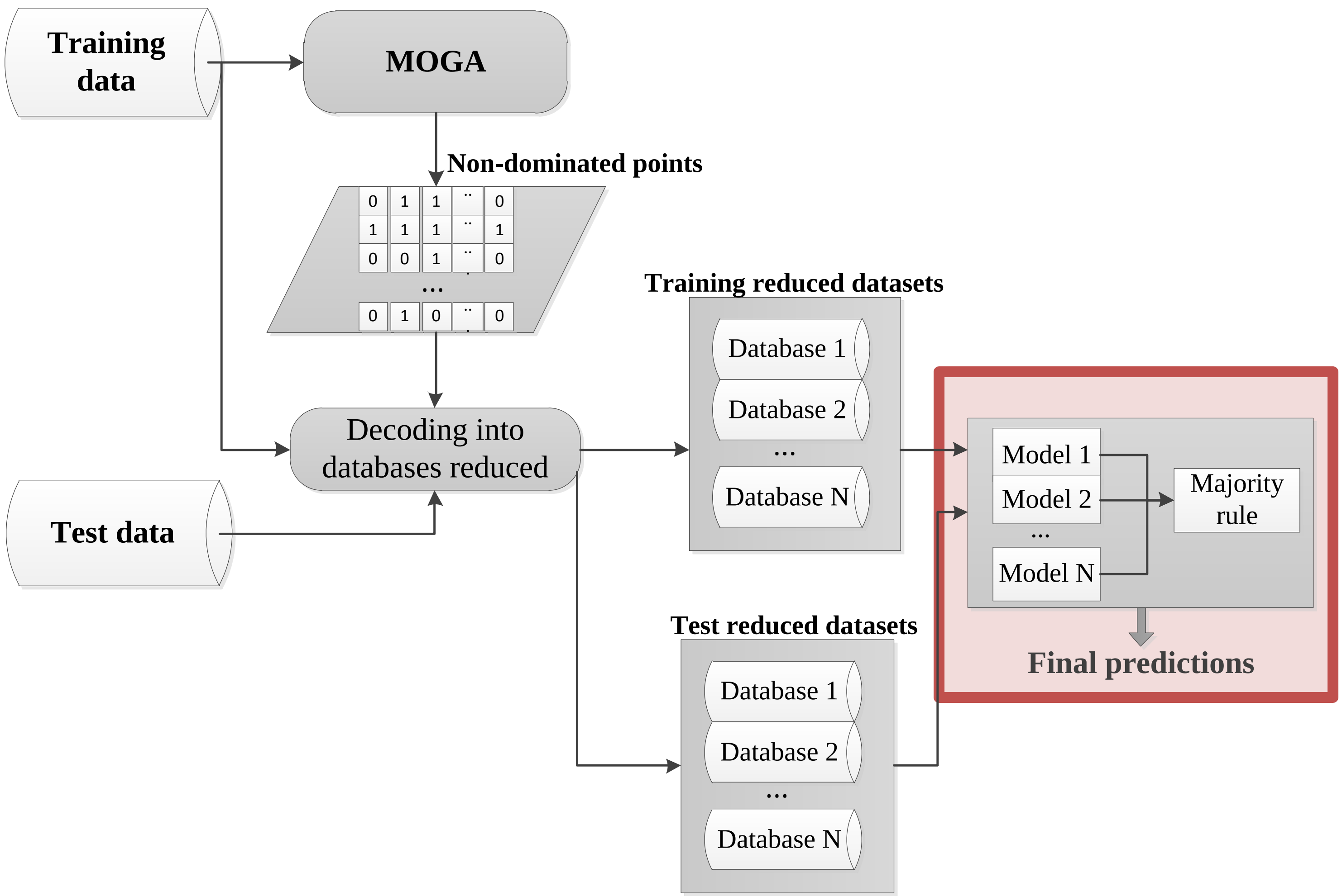
The general scheme of the approach proposed



The general scheme of the approach proposed



The general scheme of the approach proposed



Multi-Objective Genetic Algorithms

- *Generate the **initial population***
- *Evaluate **criteria** values*
- *While (stop-criterion!=true), do:*

{

*- Estimate **fitness-values**;*

*- Choose the most appropriate individuals with the mating **selection** operator based on their fitness-values;*

*- Produce new candidate solutions with **recombination**;*

*- Modify the obtained individuals with **mutation**;*

*-Evaluate **criteria** values for new individuals. Compose the new population (**environmental selection**);*

}

Multi-Objective Genetic Algorithms

MOGA	Fitness Assignment	Diversity Preservation	Elitism
NSGA-II	Pareto-dominance (*niching mechanism*) and diversity estimation (*crowding distance*)	Crowding distance	Combination of the previous population and the offspring
PICEA-g	Pareto-dominance (*with generating goal vectors*)	Nearest neighbour technique	The archive set and combination of the previous population and the offspring
SPEA2	Pareto-dominance (*niching mechanism*) and density estimation (*the distance to the k-th nearest neighbour in the objective space*)	Nearest neighbour technique	The archive set

Multi-Objective Genetic Algorithms

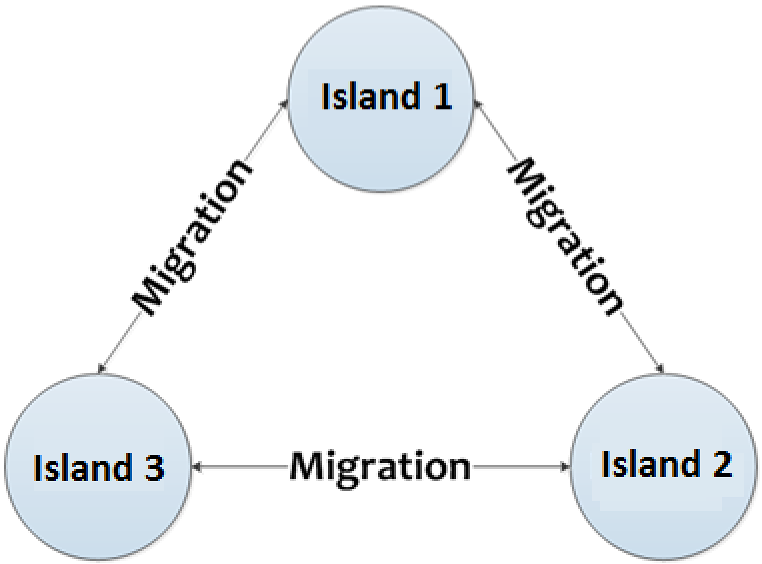
Designing a MOGA, researchers are faced with some issues:

- fitness assignment strategies,
- diversity preservation techniques,
- ways of elitism implementation.

Tasks:

1. To investigate the **effectiveness of MOGAs**, which are based **on various heuristic mechanisms**, from the perspective of the feature selection procedure;
2. To implement the **cooperation of MOGAs** and observe its effectiveness.

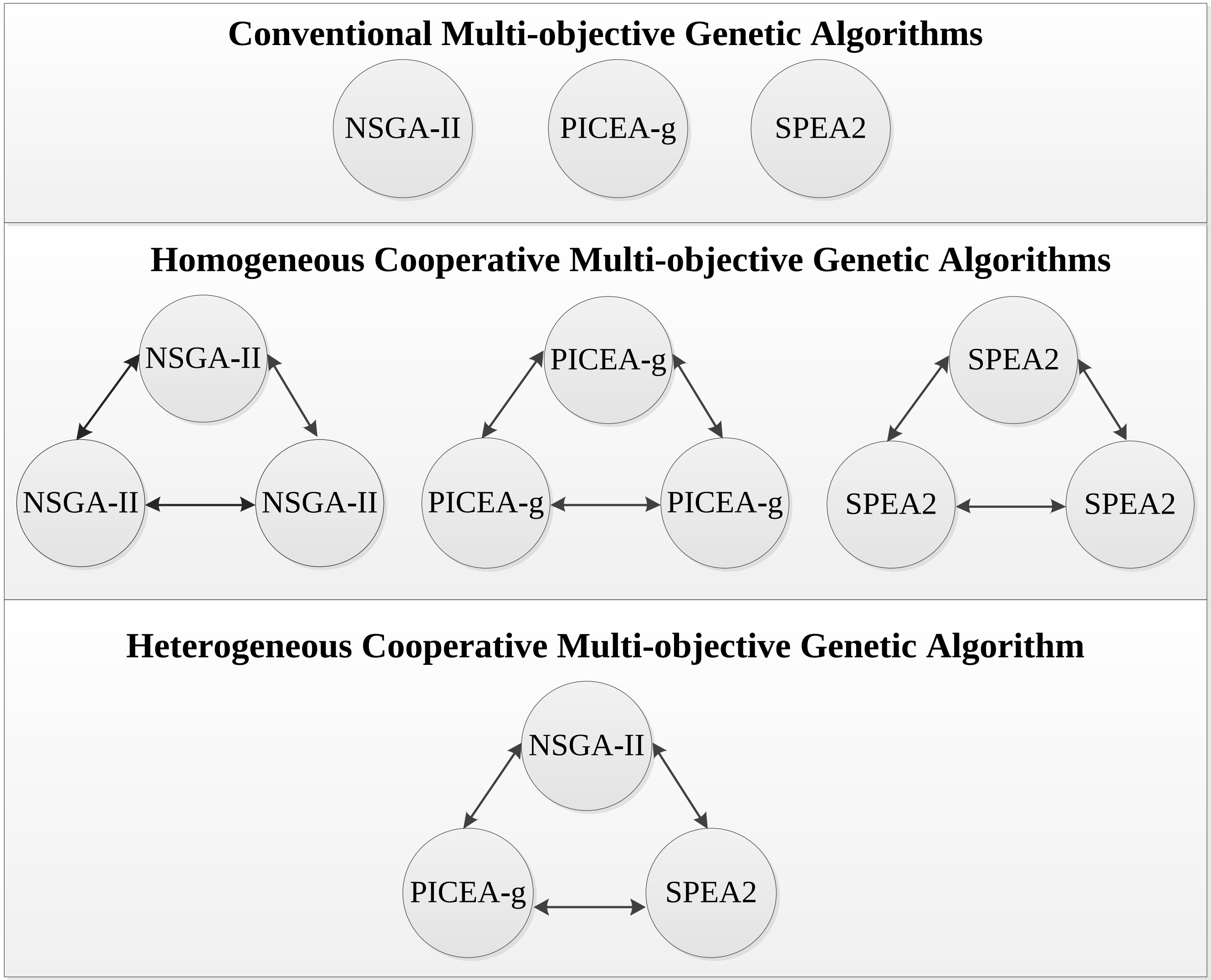
Cooperative Modification



Island model …

- ✓ is based on parallel work of islands;
- ✓ has an ability to preserve genetic diversity;
- ✓ could be applied to separable problems.

The three categories of algorithms used



Speech-based Emotion Recognition Problem

List of extracted features

- General features: Power, Mean, Root mean square, Jitter, Shimmer
- Mel-frequency cepstral coefficients (MFCCs):12 MFCCs
- Formants: 5 Formants
- Pitch, Intensity and harmonicity based features: Mean, Minimum, Maximum, Range, Deviation
- Etc.

Voice

Voice conversion into the digital form

Extraction of numerical characteristics

Classification of sound signals

The **emotion** is detected

Sample

$x_{1,1}$	$x_{1,2}$	…	$x_{1,m}$	y_1
$x_{2,1}$	$x_{2,2}$	…	$x_{2,m}$	y_2
$x_{3,1}$	$x_{3,2}$	…	$x_{3,m}$	y_3
…	…	…	…	…
$x_{n,1}$	$x_{n,2}$	…	$x_{n,m}$	y_n

$\bar{x}_i$ – independent variable,
y_i – dependent variable, $i=\overline{1,n}$,
$y_i \in C$, where $C=\{c_1, c_2, \dots, c_r\}$ – finite set,
r – the number of classes.

New examples

$x_{1,1}$	$x_{1,2}$	…	$x_{1,m}$	?
…	…	…	…	…
$x_{l,1}$	$x_{l,2}$	…	$x_{l,m}$	?

Goal:
To classify new objects based on the sample (supervised learning).

Corpora description

Database	Language	Full length (min.)	Number of emotions	File level duration		Notes
				Mean (sec.)	Std. (sec.)	
EMO-DB	German	24.7	7	2.7	1.02	Acted
SAVEE	English	30.7	7	3.8	1.07	Acted
UUDB	Japanese	113.4	4	1.4	1.7	Non-acted

Experiments conducted

Common for all experiments:

- 6-fold cross-validation procedure
- Conventional classifiers (WEKA):
  - *Support Vector Machine – SMO;*
  - *Multilayer Perceptron – MLP;*
  - *Linear Logistic Regression – Logit*.
- The *F-score* metric was evaluated.

Experiments conducted

Experiment 1:

Conventional classifiers (SMO, MLP, Logit) and their ensemble without Feature Selection -> **Baseline**

Experiment 2: The two-criterion filter approach with conventional MOGAs

NSGA-II, PICEA-g, and SPEA2 were used as optimizers in combination with SMO, MLP, Logit classifiers and their ensemble.

- All algorithms were provided with the same amount of resources (90 generations and 150 individuals in populations).
- For each MOGA the following settings were defined:

  - binary tournament selection,

  - uniform recombination,

  - the mutation probability $p_m=1/n$, where n is the length of the chromosome.

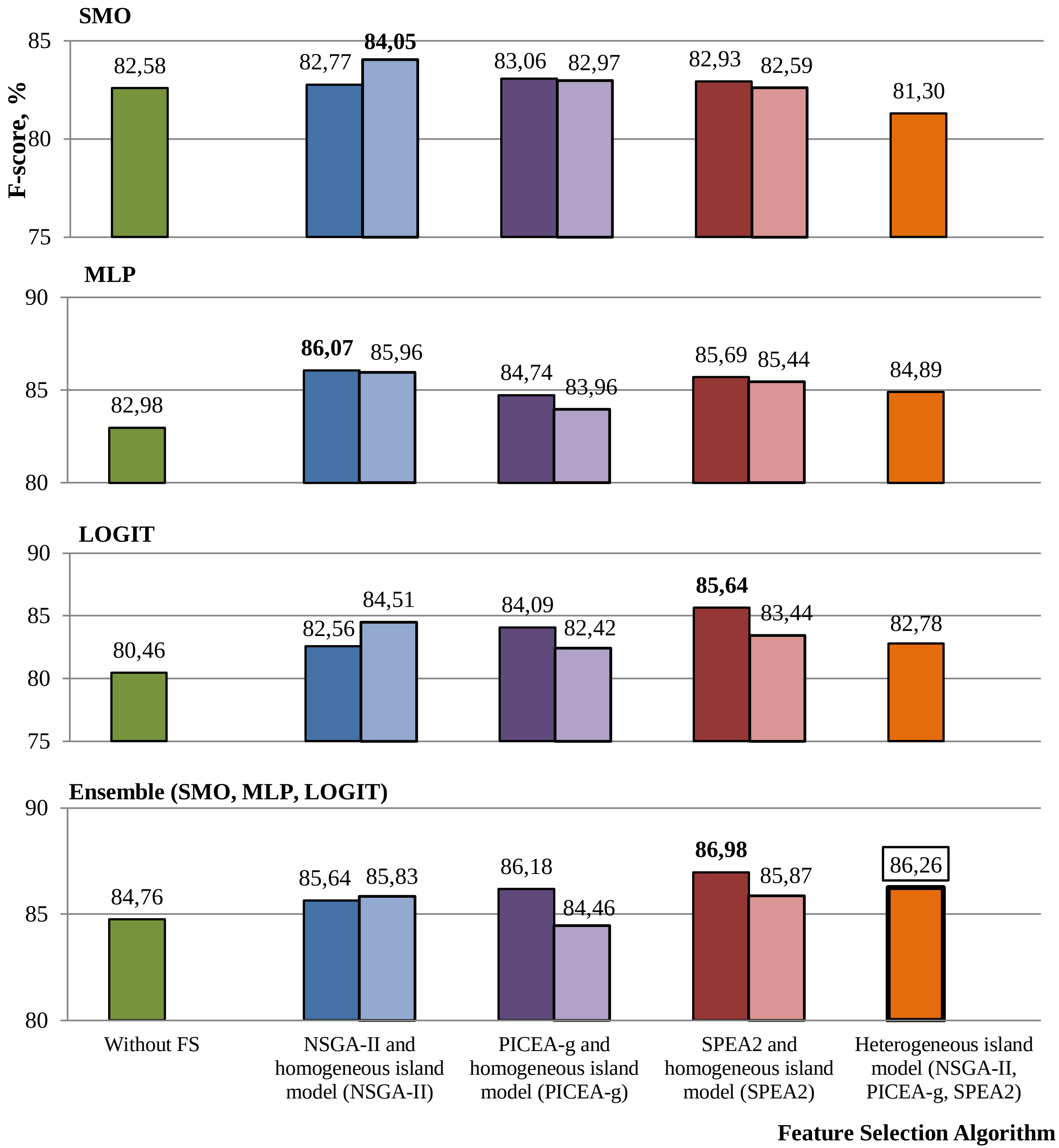
Experiments conducted

Experiment 3: The two-criterion filter approach with the cooperative MOGAs

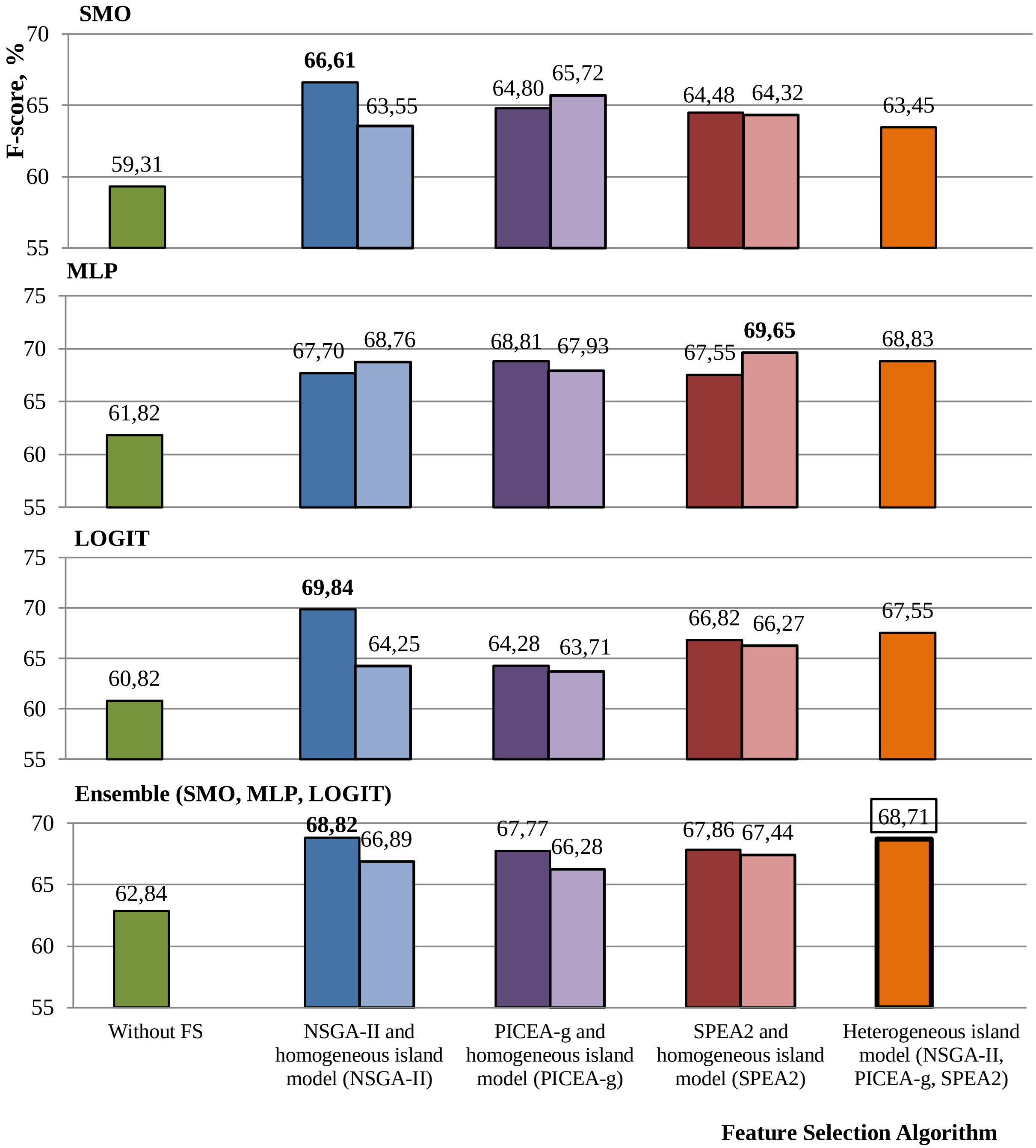
We applied three **homogeneous** algorithms (NSGA-II – NSGA-II – NSGA-II; PICEA-g – PICEA-g – PICEA-g; SPEA2 – SPEA2 – SPEA2) and the **heterogeneous** one (NSGA-II – PICEA-g – SPEA2)

- All islands had an equal amount of resources (90 generations and 150/3 = 50 individuals in populations), the migration size was equal to 10 (in total each island got 20 points from two others), and the migration interval was equal to 10 generations.
- In the final set of non-dominated points we had 30 binary strings (as well as in Experiment 2).

EMO-DB

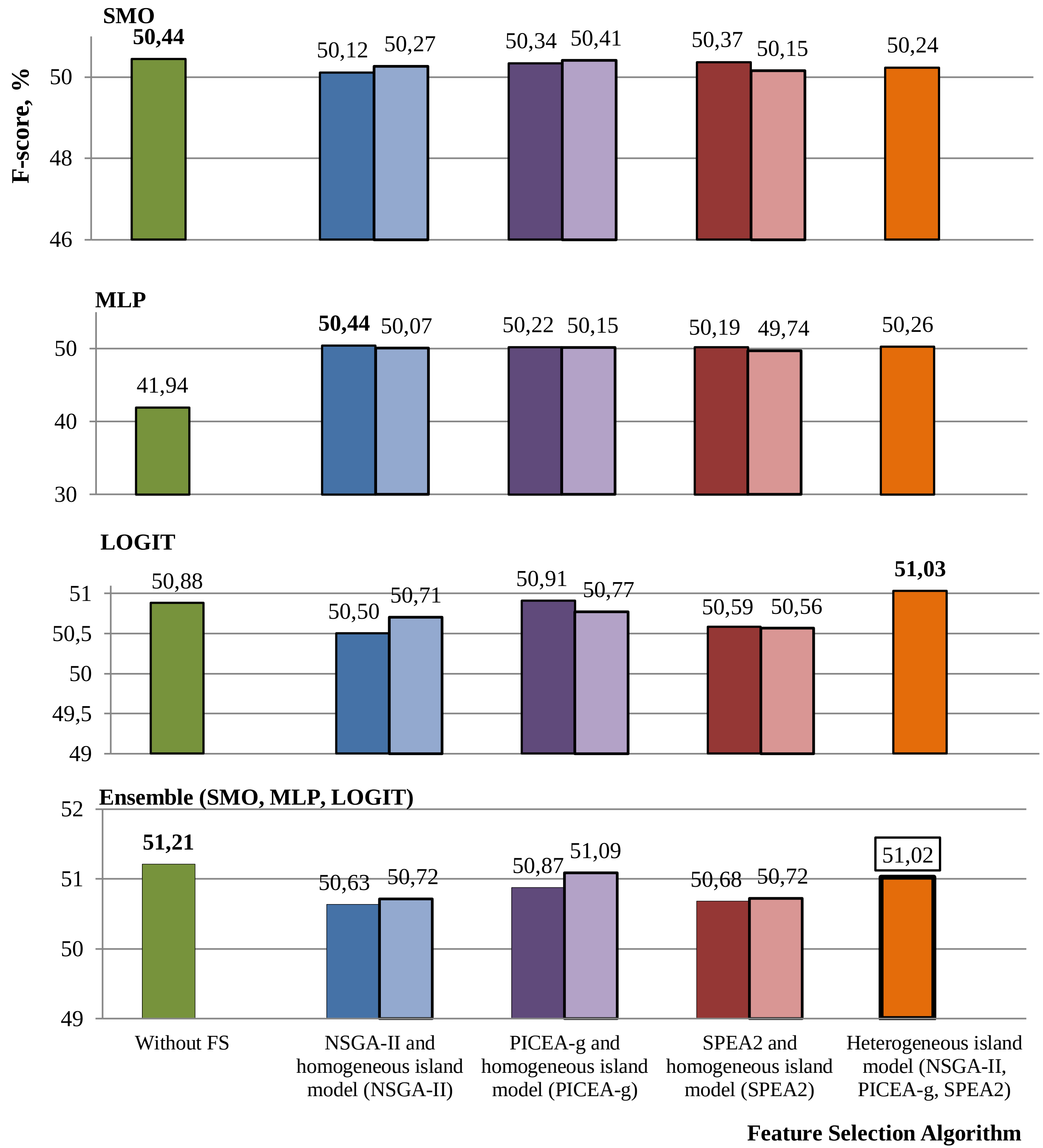

SMO
F-score, %
85
80
75
82,58
82,77
84,05
83,06
82,97
82,93
82,59
81,30
MLP
90
85
80
82,98
86,07
85,96
84,74
83,96
85,69
85,44
84,89
LOGIT
90
85
80
75
80,46
82,56
84,51
84,09
82,42
85,64
83,44
82,78
Ensemble (SMO, MLP, LOGIT)
90
85
80
84,76
85,64
85,83
86,18
84,46
86,98
85,87
86,26
Without FS
NSGA-II and homogeneous island model (NSGA-II)
PICEA-g and homogeneous island model (PICEA-g)
SPEA2 and homogeneous island model (SPEA2)
Heterogeneous island model (NSGA-II, PICEA-g, SPEA2)
Feature Selection Algorithm

SAVEE



SMO
F-score, %
70
65
60
55
59,31
66,61
63,55
64,80
65,72
64,48
64,32
63,45
MLP
75
70
65
60
55
61,82
67,70
68,76
68,81
67,93
67,55
69,65
68,83
LOGIT
75
70
65
60
55
60,82
69,84
64,25
64,28
63,71
66,82
66,27
67,55
Ensemble (SMO, MLP, LOGIT)
70
65
60
55
62,84
68,82
66,89
67,77
66,28
67,86
67,44
68,71
Without FS
NSGA-II and homogeneous island model (NSGA-II)
PICEA-g and homogeneous island model (PICEA-g)
SPEA2 and homogeneous island model (SPEA2)
Heterogeneous island model (NSGA-II, PICEA-g, SPEA2)
Feature Selection Algorithm

UUDB



SMO
F-score, %
50,44
50,12
50,27
50,34
50,41
50,37
50,15
50,24
MLP
50,44
50,07
50,22
50,15
50,19
49,74
50,26
41,94
LOGIT
50,88
50,50
50,71
50,91
50,77
50,59
50,56
51,03
Ensemble (SMO, MLP, LOGIT)
51,21
50,63
50,72
50,87
51,09
50,68
50,72
51,02
Without FS
NSGA-II and homogeneous island model (NSGA-II)
PICEA-g and homogeneous island model (PICEA-g)
SPEA2 and homogeneous island model (SPEA2)
Heterogeneous island model (NSGA-II, PICEA-g, SPEA2)
Feature Selection Algorithm

Conclusions and Future Plans

- The proposed feature selection technique is **an effective alternative to the conventional PCA** because in most cases the application of any MOGA leads to an improvement in the classifier performance and a significant dimensionality reduction.

- It is reasonable to use **the heterogeneous MOGA** and **the ensemble of classifiers** to **eliminate the choice** of the most effective heuristic algorithm and the best model without detriment to the classification quality.

- In comparison with conventional MOGAs, **homogeneous modifications** are often **preferable** only in the sense of **time costs**, whereas **the heterogeneous one** shows **high F-score values**, especially in combination with the ensemble of classifiers.

- Finally, the promising results prove that the proposed algorithmic scheme might be applied to solve **some other problems** related to the speech-based recognition of human qualities such as gender or speaker identification.

Thanks a lot!