BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iaşi
Volumul 69 (73), Numărul 4, 2023
Secţia
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ
DOI:10.2478/bipie-2023-0019

A COMPARISON OF RANKING METHODS USED IN MULTIOBJECTIVE OPTIMIZATION FOR FEATURE SELECTION IN EEG SIGNALS

BY

CORINA CÎMPANU*

“Gheorghe Asachi” Technical University of Iași,
Faculty of Automatic Control and Computer Engineering, Iași, Romania

Received: November 7, 2023
Accepted for publication: December 23, 2024

Abstract. Electroencephalogram recordings provide insightful information concerning the diagnosis and prognosis of human thinking and memory-related processes, aiding researchers and physicians during Brain-Computer Interface systems development. In electroencephalogram memory pattern identification, feature extraction, and feature selection are determining factors for an impartial data description and an accurate classification. The electroencephalogram signals analyzed in this study are collected from sixteen electrodes split into four frequency bands during specific working memory-related tasks on different reasoning scenarios.

Although most genetic algorithm based optimization procedures tackle the minimization of a classifier's error rate and the number of selected features, they are independent of how feature selection procedures are configured, either in single or multi-objective optimization manners, the major problem is multidimensionality and quantity of redundant and noisy electroencephalogram recordings. Since single objective optimization applied separately for two objectives: the minimization of the misclassification rate and the minimization of

*Corresponding author; *e-mail*: corina.cimpanu@academic.tuiasi.ro

the number of selected features bias the final results to a specific objective direction, all these limited explorations ground the use of multi-objective optimization procedures for better and sound results.

Regarding all the multi-objective optimization procedures, the compared Pareto ranking schemes are meant for the selection of parents and survivors in evolutionary multi-objective optimization. Usually, Pareto methods use only the dominance analysis for providing the partial sorting of solutions without considering the specific strength of the conflict between them. The methods compared in this paper assign the ranks by combining the search with the decisional mechanism. The decision is implemented through adaptive grouping schemes meant to guide the search towards the middle of the first Pareto fronts, enabling the progressive rejection of profitless solutions. The population is split into several groups to preserve its diversity, and a supplementary objective is added to control the variety of the most valuable genetic information. Finally, the layout of the available solutions in the objective space is examined based on clustering procedures and by individually ranking of the resulting clusters of solutions to counteract the inherent disadvantages of Pareto methods. All compared ranking schemes demonstrate their effectiveness during the evolutionary selection of features. Furthermore, various classifiers distinctively address the problem at hand, illustrating different decisional mechanisms.

Keywords: classifier, EGG feature selection, genetic algorithm, ranking, multi-objective optimization.

1. Introduction

External devices become controllable by interpreting the brainwave activity through a novel communication pathway defined by the Brain-Computer Interface (BCI) systems (Idowu *et al*., 2021; Wang *et al*., 2023; Zeynali *et al*., 2023). Their fast development in the last few decades attracted researchers' attention due to their inherent advantage of translating/ decoding electrical brain activity into control commands (Idowu *et al*., 2021). Amongst the currently employed measurement techniques of brain activity, direct brain interaction is often measured through electroencephalogram (EEG) signals that remark through non-invasive mechanisms and high temporal resolution (Cabanero-Gomez *et al*., 2021). Patterns associated with user-performed tasks during standardized experiments become measurable through EEG data that can be further used in BCI applications (Zeynali *et al*., 2023). Working memory can hold a limited quantity of information at a specific moment, and that memory pattern is associated with a specific cognitive load level (Cabanero-Gomez *et al*., 2021).

One of the most critical stages in developing a BCI for cognitive load assessment is represented by selecting the most relevant EEG features that offer optimal classifier accuracy and model complexity. Selecting the optimal subset of EEG features has prime importance in reducing computational complexity, reducing overfitting, and eliminating the inconveniences for users during clinical

application (Idowu *et al*., 2021). Filter feature selection algorithms are known to offer limited perspective and impartiality on the importance of the selected subset of features since they work based on preprocessing the EEG data and computing their statistical characteristics. At the same time, wrapper methods become unfeasible whenever the need to evaluate a vast possible combination of feature subsets arises, and embedded feature selection methods employ a particular learning model process to compute feature importance. Different classifier models can be used for identifying a certain working memory load level, although evaluating each model's accuracy is computationally expensive and unreasonable for all existing feature vector alternatives.

Genetic algorithms come in handy for solving the optimization problem that requires exploring a vast collection of candidate solutions, as in the case of finding the EEG subset of optimal features used for classifying cognitive load levels. The related literature pictures both their mono-objective and multi-objective optimization approaches. In (Saibene and Gasparini, 2023), candidate solutions' ranking is achieved with a fitness function based on the accuracy indicator for a linear Support Vector Machine classifier or based on the unsupervised silhouette indicator for k-means, while in (Kocadagli *et al*., 2023), a genetic algorithm that employs the Information Complexity Criteria as a fitness function is used as a relevant EEG feature selection procedure before assessing the preictal phase that precedes an epilepsy seizure.

All the methods compared in this paper assign the ranks by combining the search and the decisional mechanism. The decision is implemented through adaptive grouping schemes meant to guide the search towards the middle of the first Pareto fronts, enabling the progressive rejection of profitless solutions. The population is split into several groups to preserve its diversity, or a supplementary objective is added to control the variety of the most valuable genetic information whenever necessary. Finally, the layout of the available solutions in the objective space is examined based on clustering procedures and by individually ranking of the resulting clusters of solutions to counteract the inherent disadvantages of Pareto methods. All compared ranking schemes demonstrate their effectiveness during the evolutionary selection of features. Furthermore, various classifiers distinctively address the problem at hand, illustrating different decisional mechanisms.

The current study is projected based on the three following sections. The second section details the methodology for acquiring, preprocessing, and organizing the EEG data. It presents the ranking procedures integrated into the multi-objective optimization genetic procedures analyzed as alternatives for EEG-relevant feature selection. Section 3 illustrates the experimental settings and their results in selecting the most relevant features for cognitive load level assessment based on EEG signals. Some conclusions are mentioned in Section 4.

2. Methodology

Electrical brain activity is usually measured in BCIs through EEG flexible and noninvasive sensing devices (Grana and Morais-Quilez, 2023). Their applications spread to clinical practice and psychiatric research areas, such as epilepsy, sleep stages identification, emotion recognition, disability recovery processes, and BCIs. This noninvasive diagnostic tool allows the study of the bioelectrical peculiarities that occur at a functional level in our brains. The wide-available and low-cost EEG data samples reflect the synaptic activity with high time resolution in real-time directly, enabling direct access to neuronal signaling in disease diagnostication.

Widely studied by BCIs in educational applications or medical diagnosis, the cognitive load level is determined by self-reporting task difficulty level, by performing dual tasks, or by monitoring physiological measures through EEG data (Cabanero-Gomez *et al*., 2021). The EEG data used in this study was recorded on a reference group including five undergraduate students who share similar learning capabilities (Cîmpanu *et al*., 2017; Cîmpanu *et al*., 2021; Cîmpanu, 2023) under some standard working memory load scenarios associated with n-back memory tests (Kane *et al*., 2007), where each student has to visually match objects in sequences that were presented 2, or 3 times ago. Some fully relaxed neutral sessions were added to the set as a baseline comparison. Each EEG recording experiment had stages for training, trials, and breaks of a similar span (20 minutes each). The consistency and consequence of the recorded data were maintained by discarding 30% of the initial and final EEG recording in each session associated with reaching a stable concentration level and task' habituation, respectively. The characteristics of brain functions and connectivity reflected by the final data set targeted only the successful trials (over 85% success rate).

The EasyCap headset with passive wet electrodes from Brain Products, connected to a V-Amp 16 channels amplifier was used to acquire data at a sample rate of 512 Hz from 14 scalp locations organized according to the 10-20 International System (from FP1, FP2, F3, F4, F7, F8, P3, P4, P7, P8, C3, C4, O1, and O2) (Cîmpanu, 2023). The OpenVibe application is used for both EEG data acquisition and pre-processing stages. The EEG signal is band-filtered according to the frequency ranges from Table 1 from the 0.5 to 100 Hz interval to obtain the frequency bands of interest, namely alpha, beta, gamma, and theta. The powerline interferences (50 Hz) are removed through Notch filtering. Body artifacts such as eye blinking and muscle movements are tackled with an online Savitzky-Golay zero-order filter for smoothing. Finally, the Fast Fourier Transform is applied to commute the EEG data into frequency spectral patterns (Cîmpanu, 2023).

The EEG dataset $D = [X, Y]$ used for experiments gathers the signal envelopes achieved on four of the five bandwidths, removing delta waves since they are associated with sleep stages or with brain reorganization based on

experience in awake subjects (in infancy or in traumatic brain injuries, learning problems, incapacity to think, severe ADHD, or dementia) as stated in Table 1. The data collected from 14 electrodes on each alpha, beta, gamma, and theta frequency bands translates into data samples described through a total of $M = 56$ features. The class labels in Y associated with the corresponding data sample in X mark out one of the $c = 3$ types of working memory load patterns as n-back-2, n-back-3, or resting state. The division of D into training and testing data is done by maintaining the same sample-to-class distribution in Y while randomly separating 5% of the data for training and keeping the unseen 95% for testing and demonstrating the generalization capability. The resulting sample per class distribution on each division is presented in Table 2.

Table 1
Brain Waves Characteristics

Brain wave	**Frequency**	**Brain state description**
Gamma	>35 Hz	Concentration
Beta	12-35 Hz	Anxiety, active, attention, relaxed
Alpha	8-12 Hz	Very relaxed, passive attention
Theta	4-8 Hz	Deeply relaxed, focused
Delta	0.5-4 Hz	Sleep, traumatic brain injuries, learning problems, incapacity to think, severe ADHD, or dementia

Table 2
Sample-to-class distribution in training and testing datasets

Class	**Train**	**Test**
n-back-2	41.22%	41.04%
n-back-3	21.86%	23.23%
rest	36.92%	35.73%

The genetic multi-objective procedures work on a population of $NIND$ binary encoded, fixed-length arrays of M genes randomly generated as candidate solutions for EEG assessment relevant predictors' selection structured as in the Eq. (1) definition. Each gene from T, t_i is defined in $\{0, 1\}$, where 0 indicates rejecting the corresponding predictor, while 1 marks out choosing the corresponding predictor for EEG classification.

$$T = [\, t_1, t_2, \dots, t_M] \tag{1}$$

The proposed genetic multi-objective procedures for EEG-relevant feature selection aim to simultaneously optimize the accuracy of a classifier model (O_{ERR}), as well as the model's complexity through the number of selected predictors (O_{NSF}). The optimization procedures are dispatched given the following objective functions from Eq. (2) and Eq. (3):

$$O_{ERR}(x) = \frac{N - C}{N} \tag{2}$$

$$O_{NSF}(x) = \sum_{i=1}^{M} t_i \tag{3}$$

Algorithm 1: Genetic Multi-Objective Optimization Procedure for the EEG Relevant Feature Selection

1: Load the randomly generated population of NIND binary encoded individuals

2: Load the evaluation for the initial population. Evaluation is solved by training different multi-class classifiers on 5% samples and predicting their response on the remainder of data samples

3: For GEN generations execute

4: Assign fitness values to each individual via *ranking*

5: Select NIND/2 individuals for recombination via universal stochastic sampling

6: Generate NIND/2 offspring via discrete crossover operator according to the $PC = 0.7$ probability

7: Apply small variation to the genetic material of the generated offspring according to uniform mutation with $PM = 0.1$ probability

8: Compute fitness values via *ranking* for the offspring population

9: Insert the best NIND/4 offspring into the population replacing the worst existing candidate solutions

10: Select the final result from the center of the non-dominated Pareto front

A randomly generated population P of $NIND$ binary individuals representing feature selection alternatives evolved over GEN generations. The same initial population $P(0)$ and an initial preliminary evaluation of each candidate solution inside $P(0)$ are used to maintain experimental compatibility and a fair comparison between methods. The best-adapted individuals are selected as parents for the next generation of children via stochastic universal sampling based on ranking. Offspring are then created through discrete recombination and uniform mutation, and the best half of them is inserted in the population to replace the worst existing solutions. Fitness values are assigned to each individual via three ranking procedures. The Nondominated Sorting Genetic Algorithm (NSGA-II) (Deb *et al*., 2002), the Multi-Objective Optimization that incorporates interactions with a Decision Block (MOO-DB) (Ferariu and Cîmpanu, 2013), and the Multi-Objective Optimization with Adaptive Ranking (MOO-AR) (Cîmpanu and Ferariu, 2013). The solution located closest to 0, as described in the normalized objective space defined by the O_{ERR} and O_{NSF}, is considered the result of the genetic procedure. More details about the general flow of the genetic procedure are given in Algorithm 1.

The genetic multi-objective optimization procedures proposed in Algorithm 1 involve an infinite set of Pareto optimal solutions, each revealing the possible compromise between O_{ERR} and O_{NSF}. Moreover, if the objectives are conflicting, their linear aggregation into a single objective (Cîmpanu, 2023) does not prove to be effective for exploring the search space, motivating the need to use multi-objective optimization techniques that offer an insightful perspective about the landscape of the search space, by generating solutions that are located closer to the Pareto optimal set, or distributed along it.

Working memory load assessment based on EEG data deals with selecting the most adequate subset of relevant features and choosing the most suitable classifier model to obtain better results. Multi-objective optimizations involve an infinite set of Pareto optimal solutions that reveal the compromise between O_{ERR} and O_{NSF}.

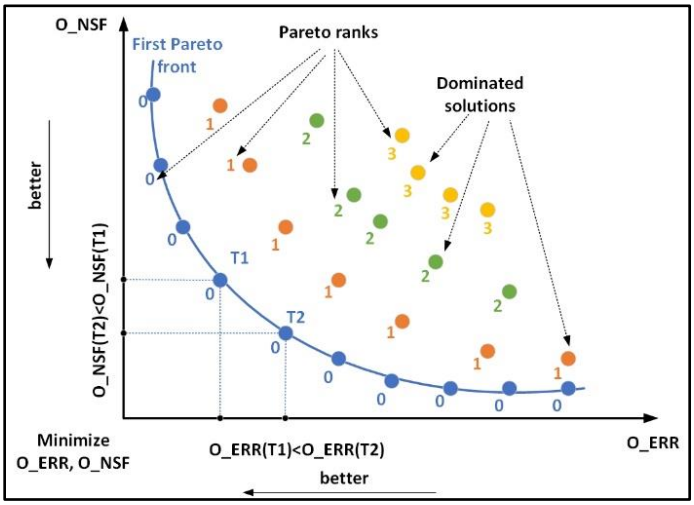

Fig. 1 – A graphical representation of Deb's ranking scheme.

Pareto ranking techniques are recommended for solving multi-objective optimization problems with few objectives. Ranks are computed after dominance analysis, keeping in mind to maintain the diversity of the population of candidate solutions. An individual encrypting T_1 is labeled as dominant of an individual encrypting T_2 if and only if each one of them surpasses the other in at least one objective direction as presented in Eq. (4). Deb's ranking scheme (Deb, 2001) assumes assigning the first rank to the nondominated candidate solutions inside the population, then iteratively excluding the already ranked ones and assigning the second rank to the new group of nondominated solutions, and so on. NSGA-II based ranking (Deb *et al*., 2002) will be used as a baseline comparison, as well as per solutions clusters, as detailed in the following section.

$$T_1 \prec T_2 \Leftrightarrow \Big(\big(O_{ERR}(T_1) < O_{ERR}(T_2)\big) AND \big(O_{NSF}(T_1) \leq O_{NSF}(T_2)\big)\Big) OR\big(\big(O_{ERR}(T_1) \leq O_{ERR}(T_2)\big) AND \big(O_{NSF}(T_1) < O_{NSF}(T_2)\big)\big) \quad (4)$$

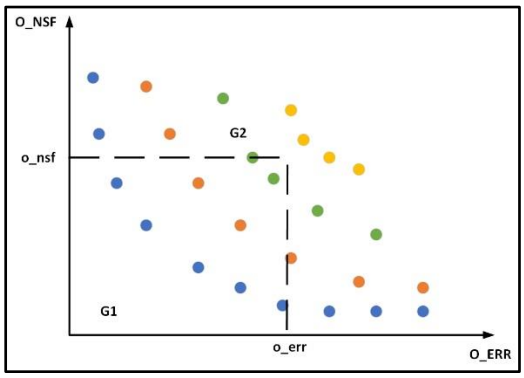

Fig. 2 – Interaction with the Decision Block mechanism in the [MOO-DB] ranking scheme. Setting clusters' limits.

Let us consider $P(k)$ as the populations of solutions obtained at generation k. As was already discussed, an essential aspect of the whole process is guiding the search procedure to explore the middle of the first Pareto fronts. Therefore, in order to avoid solutions located at the extremities of the search space, the population of candidate solutions is clustered into two groups: G_1 containing the higher-ranked preferred individuals, and G_2 containing the remainder of solutions as detailed in Eq. (7) and Eq. (8). The first clusters are delimitated by using the borders in Eq. (5) and Eq. (6). The clusters' borders are adaptively updated according to the layout of solutions in the objective space. Moreover, if the count of preferred candidate solutions becomes either too small (less than 10% of the number of individuals inside the population) or too large (more than 90% of the number of individuals inside the population), the clustering borders will be increased or correspondingly decreased. At each two consecutive generations, the layout of the G_1 and G_2 clusters are compared for detecting fast convergence or for direct diversity control. If G_1 records a considerable shrinking, or the mean distance to zero computed to the origin of the objective space in Eq. (12) registers a fast decrease as well, fast convergence is detected. In this particular situation, all individuals in the G_1 cluster will be ranked considering a supplementary objective function O_{DCN} from Eq. (13) purposely designed to directly controlling the diversity of the genetic material. Detailed conditions and settings used for describing risking convergence speed are detailed in Eq. (9) – (11). A detailed description of the ranking procedure employed by the multi-

objective optimization procedure that incorporates the interaction with the decision block [MOO-DB] is given in Algorithm 2.

Algorithm 2: Ranking procedure incorporated into [MOO-DB]

1: Split the population of candidate solutions into two clusters according to Eq. (5)-(8).
2: Adjust the clusters' borders if the size of G_1 becomes too small or too large as noted in (Fig. 2).
3: Verify condition in Eq. (9), (10), or (11) and activate the O_DCN objective function in Eq. (13) if at least one of them is true.
4: If risky convergence is detected at step 3:
5: Apply NSGA for ranking candidate solutions in G_1 subject to Eq. (2), (3), and (13)
6: Else
7: Apply NSGA for ranking candidate solutions in G_1 subject to Eq. (2), and (3)
8: Apply NSGA for ranking candidate solutions in G_2 subject to Eq. (2), and (3). Assign them worst ranks than those assigned to individuals in G_1.
9: Linearly map the associated ranks into fitness values.

$$\mathrm{o_{err}} = \sum_{i=1}^{|P(k)|} \frac{O_{ERR}(T_i)}{|P(k)|} \tag{5}$$

$$o_{nsf} = \sum_{i=1}^{|P(k)|} \frac{O_{NSF}(T_i)}{|P(k)|} \tag{6}$$

$$G_1(k) = \{T_i \in P(k) | O_{ERR}(T_i) < o_{err},\ O_{NSF}(T_i) < o_{nsf}\} \tag{7}$$

$$G_2(k) = P(k) \backslash G_1(k) \tag{8}$$

$$|G_1(k)| < 0.8|G_1(k-1)| \tag{9}$$

$$\underset{T_i \in G_1(k)}{mean}\ dcn\left(T_i|_{G_1(k)}\right) < 0.85\ \underset{T_i \in G_1(k-1)}{mean}\ dcn\left(T_i|_{G_1(k-1)}\right) \tag{10}$$

$$\underset{T_i \in P\ (k)}{mean}\ d0_k(T_i) < 0.8\ \underset{T_i \in P\ (k-1)}{mean}\ d0_{k-1}(T_i) \tag{11}$$

$$d0_k(T_i) = \sqrt{\left(\frac{O_{ERR}(T_i)}{\sum_{T_i \in P(k)} O_{ERR}(T_i)}\right)^2 + \left(\frac{O_{NSF}(T_i)}{\sum_{T_i \in P(k)} O_{NSF}(T_i)}\right)^2} \tag{12}$$

$$O_{DCN}(T) = dcn(T|_{G(k)}) \tag{13}$$

The idea behind the Adaptive Ranking (AR) mechanism integrated into [MOO-AR] is to efficiently exploit the information from the current population to check for strongly conflicting objectives to activate diversity enhancement. The weak nadir point wn is defined intuitively as the opposite corner of the ideal point in the hypercube that delimitates the first-order Pareto front. Methodically, at each generation, wn is searched according to Eq. (14), (15), and (16) where the nondominated solutions $\boldsymbol{x}^*_{ERR}$ and $\boldsymbol{x}^*_{NSF}$ indicate the optimum solutions obtained for the individual minimization of the two objective functions in the current population. The weak nadir point is graphically illustrated in Fig. 3a.

$$\boldsymbol{wn} = [wn_{ERR}, wn_{NSF}] \tag{14}$$

$$wn_{ERR} = \max\{O_{ERR}(\boldsymbol{x}^*_{ERR}), O_{ERR}(\boldsymbol{x}^*_{NSF})\} \tag{15}$$

$$wn_{NSF} = \max\{O_{NSF}(\boldsymbol{x}^*_{NSF}), O_{NSF}(\boldsymbol{x}^*_{ERR})\} \tag{16}$$

Let us suppose the population of candidate solutions contains enough top-ranked solutions - the virtual solution dominates at least 30% of the solutions in the population as pictured in Fig. 3b. In this particular case, the search is guided to the middle of the first Pareto fronts by clustering the individuals into two groups, where the first one contains individuals that are better than average. If the first cluster contains too many solutions, maximizing the distance to the closest neighbour inside G_1 becomes imperative. The clustering rules are presented in Eq. (17) and (18), where k denotes the current generation and $\overline{O_j}$ denotes the mean value of the objective function. If the first fronts in G_1 contain too many solutions, the O_{DCN} maximization is required for all individuals located in the first cluster.

$$G_1(k) = \{\boldsymbol{x} \in P(k) | O_j(\boldsymbol{x}) < \max_{j \in \{ERR,NSF\}}\{\overline{O_j}, wn_j\}\} \tag{17}$$

$$G_2(k) = P(k) \backslash G_1(k) \tag{18}$$

On the contrary when the virtual solution dominates less than 30% of the solutions in the population, the algorithm ensures the exploration of the best solutions since they are already few, and eliminating a part of them will damage the whole process. In this case, the population gets split into three clusters: G_1 containing the whole first Pareto front and the middle of other first Pareto fronts, G_2 encouraging the solutions located in their neighborhood, and G_3 with the remainder of individuals. Since maintaining the diversity of individuals inside G_1 is crucial, maximizing O_{DCN} is always required. The clusters' limits are detailed in Eq. (19), (20), and (21).

In both cases, NSGA I (Deb, 2001) is separately applied within each group of candidate solutions. Since the O_{DCN} is evaluated only for the best

solutions in P, the solitary solutions in the first cluster are strongly protected comparative to the case of employing classical crowding.

$$G_1(k) = \{\boldsymbol{x} \in P(k) | O_j(\boldsymbol{x}) < \min_{j \in \{ERR,NSF\}} \{\overline{O}_j, wn_j\}\} \tag{19}$$

$$G_2(k) = \{\boldsymbol{x} \in P(k) \backslash G_1(k) | O_j(\boldsymbol{x}) < \max_{j \in \{ERR,NSF\}} \{\overline{O}_j, wn_j\}\} \tag{20}$$

$$G_3(k) = P(k) \backslash (G_1(k) \cup G_2(k)) \tag{21}$$

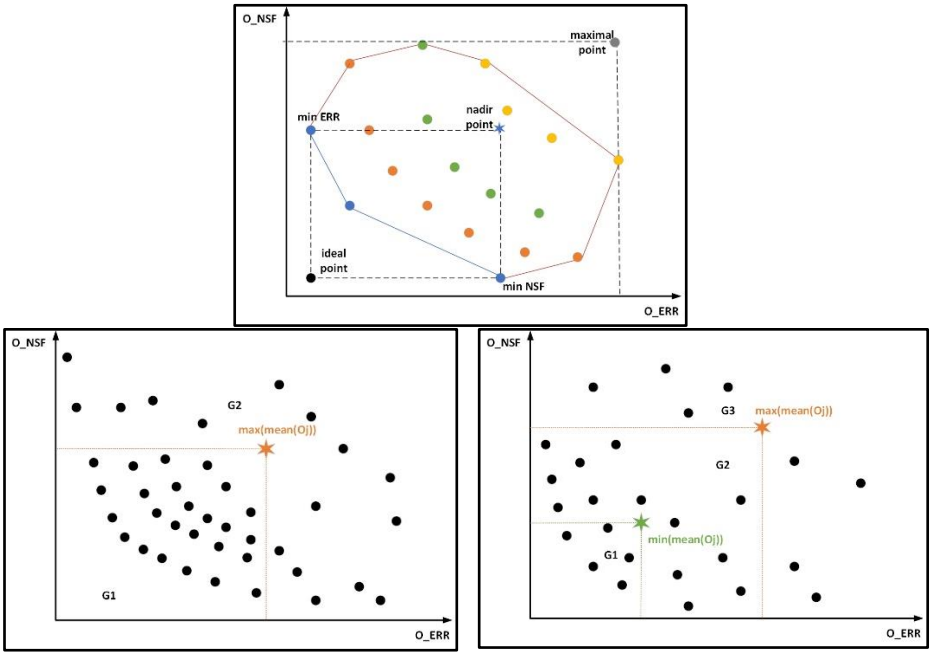

Fig. 3 – Insights of the AR scheme on individuals' clustering: a) defining the weak nadir point; b) the virtual solution dominates at least 30% of the solutions in P(k); c) the virtual solution dominates less than 30% of the solutions in P(k).

3. Experimental Analysis

The relevant EEG data features selection and the data classification-related experiments use a labeled EEG data collection that gathers samples that describe three distinct cognitive load levels, namely n-back-2, n-back-3, and a baseline resting state. Data partitioning between training and testing data ensures randomly selecting the information while maintaining the same sample-to-class ratio from the original dataset. Multiple classifiers from both lazy and eager

learners categories are evaluated in terms of misclassification rate (ERR), as well as computational complexity through the number of predictors (NSF).

A preliminary series of experiments from (Cîmpanu, 2023b) was used for deciding the detailed configurations of the classifier models that were further applied in the multi-objective optimization procedures: an Random Forest (RF) classifier configured with 75 trees in E#1-E#3, an Adaptive Boosting (AB) classifier configured with 50 trees in E#4-E#6, a multiclass Support Vector Machine (SVM) classifier with Radial Basis Function (RBF) kernel function and one vs. one comparison mechanism (E#7-E#9) and two k Nearest Neigbours (kNN) classifier models with k=5, kd-tree search and cityblock distance (E#10-E#12), and exhaustive search and Spearman distance (E#13-E#15). For each set of experiments associated with a particular classifier model, three ranking methods are applied for associating fitness values to each candidate solution in a population of $NIND = 40/20$ parents/children that have evolved over $MAXGEN = 10$ generations. NSGA-II is used as a baseline comparison for the DB and AR methods, as detailed in Table 3.

Table 3
Experimental configurations for MOO's

ETC	Classifier	Model Settings	NIND	MAX GEN	MOO Method
E#1	RF	noTrees=75	40	10	NSGA-II
E#2					DB
E#3					AR
E#4	AB	noTrees=50			NSGA-II
E#5					DB
E#6					AR
E#7	SVM	kernFcn = RBF			NSGA-II
E#8					DB
E#9					AR
E#10	kNN	k=5, search=kdtree, distance=cityblock			NSGA-II
E#11					DB
E#12					AR
E#13		k=5, search=exhaustive, distance=spearman			NSGA-II
E#14					DB
E#15					AR

The middle part of the experimental section is dedicated to illustrating the evolution of the clusters' size during the [MOO-DB] and [MOO-AR] procedures in Table 4 and Table 5, respectively. The decision block rule integrated into the fitness assignment procedure guides the search towards the center of the first Pareto front, avoiding declaring Pareto-optimal, the solutions characterized by enormous misclassification rates yet reduced predictor complexity, or vice versa. The DB-preferred individuals in cluster G_1 are assigned

with higher fitness values. Clustering borders are adaptively set with respect to the average content of the population, yet the first cluster contains a small number of individuals. The [MOO-DB] works with two or three objectives, enhances the diversity of the genetic material if necessary, considers adaptive limits when ranking solutions, and reduces computational costs by applying NSGA in each cluster. Nevertheless, the number of preferred individuals in G_1 remains small.

Table 4

E#11: kNN, k=5, search=kdtree, distance=cityblock, MOO-DB, NIND = 40, MAXGEN = 10

Gen	Maximize O_{DCN}	Clustering borders		Cluster G_1	Cluster G_2
1	yes	0.943094	0.940140	4	36
2	yes	0.946054	0.939943	4	36
3	no	0.947207	0.969290	4	36
4	yes	0.943959	0.971792	2	38
5	no	0.965927	0.985039	3	37
6	yes	0.965207	0.929792	1	38
7	yes	0.956427	0.888124	1	39
8	no	0.929944	0.968056	3	37
9	yes	0.973148	0.970052	5	35
10	yes	0.930119	0.971739	2	38

As presented in Table 5, the current generation's population $P(k)$ contains a lot of top-ranked solutions when the virtual solution corresponding to the virtual nadir point dominates at least 30% solutions ($Method = 2$) in $P(k)$ for $k = \{1, 2, 3, 5, 6, 7, 8, 9, 10\}$ while ranking the parent population and for $k = \{3, 5, 9, 10\}$ while ranking the offspring population. On the contrary, if the best solutions in $P(k)$ are already too few, and their elimination damages the effectiveness of the search procedure, the virtual solution associated with the weak nadir point dominates less than 30% of solutions ($Method = 1$). For this situation, the complete first-order Pareto front is located in G_1. Their neighbours are also encouraged, and the maximization of O_{DCN} is always employed inside G_1. As indicated in Table 5, $Method\ 1$ is used in the fourth generation while ranking the parent populations, and at generations $k = \{1, 2, 4, 6, 7, 8\}$ while ranking the children populations.

The last part of the experimental section presents the subsets of EEG-relevant features obtained after applying the multi-objective optimization procedures under the simultaneous optimization of the classifier error rate and the number of selected features. The multi-objective optimization procedures in (E#1-E#15) are compared in terms of the classifier's error rate (ERR), number of selected features (NSF), and number of selected electrodes (NSE) based on the results centralized in Table 6. The number of signal sources in each brain region

– frontal (F), parietal (P), central (C), and occipital (O) is centralized in the middle section of the same table. The feature selection in each frequency range is presented in the last section of Table 6.

Table 5
E#3: RF, noTrees = 75, MOO-AR, NIND = 40, MAXGEN = 10
Method = 1(Nadir), 2(MeanandDCN)

Gen	Offspring population ranking						Parent population ranking					
	Method	Ideal objectives		Clusters			Method	Ideal objectives		Clusters		
1	1	0.7286	15	0	5	15	2	0.6861	20	22	18	-
2	1	0.7108	21	1	7	12	2	0.6861	20	22	18	-
3	2	0.7477	22	13	7	-	2	0.6861	21	22	18	-
4	1	0.6190	24	1	6	13	1	0.6190	21	1	15	24
5	2	0.6861	18	18	2	-	2	0.6190	18	24	16	-
6	1	0.6697	23	1	6	13	2	0.6190	18	26	14	-
7	1	0.5834	20	3	1	16	2	0.6190	18	31	9	-
8	1	0.5286	23	0	8	12	2	0.5286	20	17	26	-
9	2	0.5971	19	10	10	-	2	0.5286	19	24	16	-
10	2	0.5779	23	16	4	-	2	0.5286	19	29	11	-

Table 6
Feature Selection Results – MOO-GA

ETC	ERR[%]	NSF	NSE	FS in each Brain Region				FS in each Frequency Band			
				F	P	C	O	α	β	γ	θ
E#1	0.7587	23	14	8	5	2	8	7	4	6	6
E#2	0.8614	20	14	8	6	4	2	7	3	3	7
E#3	0.6190	28	14	12	7	5	4	9	4	8	7
E#4	0.4930	21	13	6	5	5	5	4	4	5	8
E#5	0.4122	21	13	7	6	5	3	5	3	5	8
E#6	0.3862	28	14	10	6	5	6	7	5	6	9
E#7	12.0926	21	12	10	6	3	2	5	3	6	7
E#8	12.6856	23	14	10	7	4	2	7	5	5	6
E#9	10.6697	26	12	9	8	5	4	8	5	6	7
E#10	4.8069	17	12	9	7	0	1	8	3	3	3
E#11	4.4933	18	13	5	7	3	3	7	3	2	6
E#12	3.6511	21	14	7	6	3	5	8	1	1	11
E#13	2.4158	23	14	9	7	5	2	4	2	7	10
E#14	2.3185	20	13	8	5	3	4	5	1	5	9
E#15	2.4884	23	13	10	5	3	4	5	4	8	6

Brain regions do not emit simultaneously the same brain wave frequency, and an EEG signal recorded from multiple electrodes translates into distinctive brain wave characteristics. Since large amounts of subject-related, particularly patterned EEG data originate from a single EEG recording, its interpretation becomes difficult (Abhang *et al*., 2016). Sensitive to various creativity-related demands, the alpha waves increase in the frontal and right-parietal brain regions.

It varies as a function of creativity-related task demands and indicates the presence of specific memory processes as it increases during creative cognition, reflecting internal processing requisitions (Fink and Benedek, 2014). Alpha waves play an essential role in cognition; their lower amplitude waves increase selective attention among relevant neural populations (Eggermont, 2021). In healthy, awake adults, the alpha waves occur during a relaxed - resting state, when the subjects keep their eyes closed, vanish when the subject is concentrated on a particular mental task, and are more prominent over the occipital region (Moini and Piran, 2020). Alpha waves, prominent in the frontal lobe, point out momentary memory storage (Baars and Gage, 2013). Beta waves become prominent in frontocentral scalp areas while the subject manifests attention to mental tasks or stimuli. The transition between relaxed with eyes closed to beginning mental activity with eyes open translates into decreased and attenuated alpha waves that are replaced by beta waves all over the scalp (Moini and Piran, 2020). Theta waves remark themselves by facilitating the encoding of temporary episodic memory into long-term episodic memory (Baars and Gage, 2013). Also, alpha-theta synchronization or desynchronization is somehow related to cognitive processes (Baars and Gage, 2013).

A detailed feature selection in each frequency band is listed in Table 7. The large number of selected features from the alpha frequency band from all experimental test case results is sustained by higher brain functions that rely on metal representations (Stevens and Zabelina, 2019). Theta waves validate reaching distinct cognitive load levels and illustrate experiment-related engagement through the junction between the emotional and motivational aspects of the whole process (Hipp *et al*., 2012). The fastest brain activity, the Gamma waves, is in charge of cognitive functions, memory processes, and data representations (Hipp *et al*., 2012). The state of concentration and problem-solving periods are associated with focused attention and reflective thoughts from the Beta EEG features in the low – mid-range frequency interval when the subject is observant and concentrated or in the higher frequency range when the subject manifests increased energy level and task-related performance (Cîmpanu, 2023b). The frontal lobe-activated EEG signals are associated with conscious thoughts related to processing information in short-term memory and decision-making while problem-solving (Li *et al*., 2023). Conversely, the parietal lobe-activated EEG signals demonstrate that the subject integrates information from various senses to determine spatial orientation. Lastly, the occipital lobe-activated EEG signals are related to offering interpretation of incoming visual data.

Table 7

EEG Detailed Feature Selection

ETC	Feature Selection
E#1	Alpha band from electrodes: FP1 P7 O1 O2 FP2 F4 P8 Beta band from electrodes: F7 F3 O1 O2 Gamma band from electrodes: F3 C3 O1 O2 P4 P8 Theta band from electrodes: FP1 P7 O1 O2 F8 C4
E#2	Alpha band from electrodes: F7 C3 P3 O1 F8 F4 P4 Beta band from electrodes: P7 O1 P4 Gamma band from electrodes: F7 C3 F8 Theta band from electrodes: FP1 F7 C3 P7 F8 C4 P8
E#3	Alpha band from electrodes: FP1 C3 P7 O1 FP2 F8 F4 C4 P4 Beta band from electrodes: F7 F3 O2 P4 Gamma band from electrodes: C3 P7 O2 FP2 F8 C4 P4 P8 Theta band from electrodes: FP1 F7 C3 P3 O1 F8 F4
E#4	Alpha band from electrodes: O1 O2 F8 C4 Beta band from electrodes: F3 O1 FP2 F4 Gamma band from electrodes: C3 O2 C4 P4 P8 Theta band from electrodes: C3 P7 O1 FP2 F8 C4 P4 P8
E#5	Alpha band from electrodes: FP1 C3 O1 F8 P4 Beta band from electrodes: P7 O1 FP2 Gamma band from electrodes: F3 C3 P7 C4 P4 Theta band from electrodes: FP1 C3 P3 O1 FP2 F8 C4 P8
E#6	Alpha band from electrodes: FP1 F7 P7 O1 O2 F8 C4 Beta band from electrodes: F7 F3 O1 FP2 F4 Gamma band from electrodes: F7 C3 O1 O2 C4 P4 Theta band from electrodes: FP1 C3 P3 P7 O1 FP2 C4 P4 P8
E#7	Alpha band from electrodes: FP1 P7 O1 F4 C4 Beta band from electrodes: F7 F3 P4 Gamma band from electrodes: FP1 F7 C3 O1 P4 P8 Theta band from electrodes: FP1 C3 FP2 F8 F4 P4 P8
E#8	Alpha band from electrodes: F3 C3 P7 O1 FP2 F4 C4 Beta band from electrodes: F3 P3 FP2 F4 P4 Gamma band from electrodes: P3 FP2 F8 C4 P4 Theta band from electrodes: F3 C3 P7 O2 FP2 P4
E#9	Alpha band from electrodes: FP1 F7 F3 C3 P3 O1 F4 P8 Beta band from electrodes: F3 C3 O1 F4 P4 Gamma band from electrodes: F3 C3 P3 O1 F4 P4 Theta band from electrodes: C3 P3 P7 O2 F4 C4 P8
E#10	Alpha band from electrodes: FP1 F3 P3 P7 O1 FP2 F4 P8 Beta band from electrodes: FP1 P3 P8 Gamma band from electrodes: F7 FP2 P4 Theta band from electrodes: FP1 F3 P3
E#11	Alpha band from electrodes: FP1 P3 P7 O1 O2 P4 P8 Beta band from electrodes: FP1 C3 O1 Gamma band from electrodes: F4 P4 Theta band from electrodes: C3 P7 F8 F4 C4 P8
E#12	Alpha band from electrodes: F3 C3 P3 O1 O2 F4 P4 P8 Beta band from electrodes: O1 Gamma band from electrodes: FP1 Theta band from electrodes: F7 C3 P7 O1 O2 FP2 F8 F4 C4 P4 P8

E#13	Alpha band from electrodes: C3 F8 F4 P4 Beta band from electrodes: C4 P4 Gamma band from electrodes: P3 P7 O1 FP2 F8 C4 P8 Theta band from electrodes: FP1 F7 F3 C3 P7 O2 FP2 F4 C4 P4
E#14	Alpha band from electrodes: F3 O1 FP2 F4 C4 Beta band from electrodes: O1 Gamma band from electrodes: FP1 P3 O2 P4 P8 Theta band from electrodes: FP1 F7 C3 O1 F8 F4 C4 P4 P8
E#15	Alpha band from electrodes: F7 C3 P7 O1 F8 Beta band from electrodes: F3 O1 F8 C4 Gamma band from electrodes: FP1 F3 O1 O2 F8 C4 P4 P8 Theta band from electrodes: FP1 O1 F8 F4 P4 P8

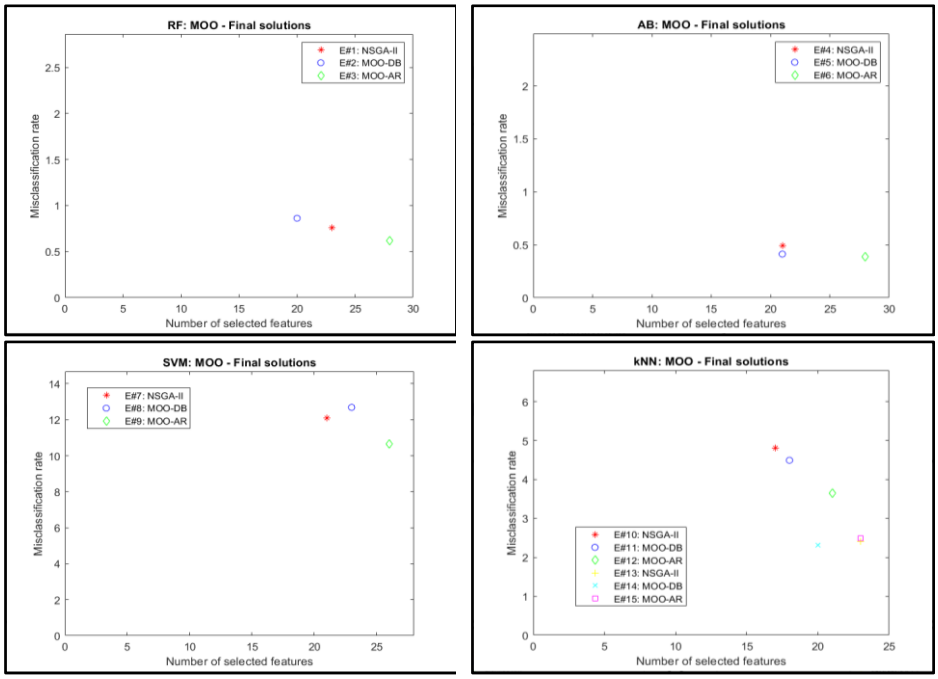

Fig. 4 – Final solutions' layout in the objective space.

The final solutions' layout in the objective space defined by ERR and NSF is pictured in Fig. 4 for each classifier model separately since each model imprints their noise in evaluation: a) RF, b) AB, c) SVM, and d) kNN. Compared to the separate optimization of the misclassification rate or number of selected features presented in (Cîmpanu, 2023b), the multi-objective optimizations in (E#1-E#15) achieve superior results. Encouraged by the mono-objective optimization based on the linear aggregation of the two objectives in (Cîmpanu, 2023b), the multi-objective genetic procedures analyzed in this study can streamline the search to the middle of the first Pareto fronts, considering both

error rate and predictor complexity simultaneously while maintaining the diversity of the genetic material by maximizing the distance to the closest neighbor as a supplementary objective.

4. Conclusions

This comparative analysis surveys the impact of three ranking methods on exploring the search space defined by a classifier's misclassification rate, namely AB, RF, SVM, and kNN, and by the number of selected features during the genetic multi-objective optimization procedure for choosing the most relevant EEG features for cognitive load assessment.

Ranks are associated with the candidate solutions by combining the search with the feedback given by the decisional mechanism. The decision block is implemented through two adaptive grouping schemes, [MOO-DB] and [MOO-AR], that aim to guide the search to the middle of the first Pareto fronts, ensuring the progressive elimination of the less desired candidate solutions. The method used as baseline comparison is [MOO-NSGA]. The diversity of the genetic material is maintained through the maximization of a supplementary objective function in the form of a distance to the closest neighbor.

All compared ranking schemes demonstrate their effectiveness during the evolutionary selection of features. Furthermore, various classifiers distinctively address the problem at hand, illustrating different decisional mechanisms. The final results indicate that the cognitive load levels can be assessed precisely using a limited selection of the connected electrodes and waveforms.

Acknowledgements. This paper is financially supported by the Project "Network of excellence in applied research and innovation for doctoral and postdoctoral programs / InoHubDoc", project co-funded by the European Social Fund financing agreement no. POCU/993/6/13/153437.

This paper is published as work of Post-doctoral advanced research program affiliated to IOSUD-TUIASI "Performance and excellence in post-doctoral research – 2022".

REFERENCES

Abhang P.A., Gawali B.W., Mehrotra S.C., *Technological Basics of EEG Recording and Operation of Apparatus*, in Introduction to EEG- and Speech-Based Emotion Recognition, Eds. Abhang P.A., Gawali B.W., Mehrotra S.C., **2,** 19-50, Academic Press, Elsevier, 2016.

Baars B.J., Gage N.M., *The brain is conscious*, Fundamentals of Cognitive Neuroscience - A Beginner's Guide, **8**, 211-252, Academic Press, Elsevier, 2013.

Cabanero-Gomez L., Hervas R., Gonzalez I., Villarreal V., *Studying the generalizability of cognitive load measured with EEG*, Biomedical Signal Processing and Control, **70**(103032), Elsevier, 2021.

Cîmpanu C., Ferariu L., *Genetic Multiobjective Fitness Assignment Scheme Applied to Robot Path Planning*, Proceedings of the 10th International Conference on Electronics Computer and Computation, ICECO 2013, Ankara, Turkey, 196-199, IEEE.

Cîmpanu C., Ferariu L., Ungureanu F., Dumitriu T., *Genetic Feature Selection for Large EEG Data with Commutation between Multiple Classifiers*, Proceedings of the IEEE 21st International Conference on System Theory, Control and Computing, ICSTCC 2017, Sinaia, Romania, 618-623, IEEE Xplore.

Cîmpanu C., Ferariu L., Dumitriu T., *Feature Selection via Genetic Multiobjective Optimization with Fuzzy Rejection Mechanisms*, Proceedings of the IEEE 17th International Conference on Intelligent Computer Communication and Processing, ICCP 2021, Cluj, Romania, 437-444, IEEE.

Cîmpanu C., *EEG Multi-Objective Feature Selection using a Genetic Procedure with Hybrid Mutation Operator*, Proceedings of the IEEE 24th International Conference on Control Systems and Computer Science, CSCS 2023, Bucharest, Romania, 128-135, IEEE.

Cîmpanu C., *On the impact of a classification model in EEG feature selection for cognitive load assessment*, Buletinul Institutului Politehnic din Iași, published by „Gheorghe Asachi" Technical University of Iași, Automatic Control and Computer Science Section, Iași, România, 2023.

Deb K., *Multiobjective Optimization Using Evolutionary Algorithms*, WileyandSons, 2001.

Deb K., Pratap A., Agarwal S., Meyarivan T., *A fast elitist multiobjective genetic algorithm: NSGA-II*, IEEE Transactions on Evolutionary Computation, **6**(2), 182-197, IEEE, 2002.

Eggermont J.J., *The alpha and delta rhythms and their interaction with other brain rhythms*, Brain Oscillations, Synchrony, and Plasticity - Basic Principles and Application to Auditory-Related Disorders, **3**, 43-58, Academic Press, Elsevier, 2021.

Ferariu L., Cîmpanu C., *Multi-Objective Genetic Algorithm with Clustering-based Ranking and Direct Control of Diversity*, Proceedings of the 17th International Conference on System Theory, Control and Computing, ICSTCC 2013, Sinaia, Romania, 213-218, IEEE Xplore.

Fink A., Benedek M., *EEG alpha power and creative ideation*, Neuroscience and Biobehavioral Reviews, **44**(100), 111-123, Elsevier, 2014.

Grana M., Morais-Quilez I., *A review of Graph Neural Networks for Electroencephalography data analysis*, Neurocomputing, 562, 126901, 1-11, Elsevier, 2023.

Hipp J.F., Hawellek D.J., Corbetta M., Siegel M., Engel A.K., *Large-scale cortical correlation structure of spontaneous oscillatory activity*, Nature, **15**(6), 884-892, Nature Neuroscience, 2012.

Idowu O.P., Adelopo O., Ilesanmi A.E., Li X., Samuel O.W., Fang P., Li G., *Neuro-evolutionary approach for optimal selection of EEG channels in motor imagery based BCI application,* Biomedical Signal Processing and Control, **68,** 102621, 1-16, Elsevier, 2021.

Kane M.J., Conway A.R.A., Miura T.K., Colflesh G.J.H., *Working memory, attention control, and the n-back task: A question of construct validity*, Journal of

Experimental Psychology: Learning, Memory, and Cognition, **33**(3), 615–622, American Psychological Association, 2007.

Kocadagli O., Ozer E., Batista A.G., *Preictal phase detection on EEG signals using hybridized machine learning classifiers with a novel feature selection procedure based Gas and ICOMP*, Expert Systems With Applications, 212, 118825, 1-11, Elsevier, 2023.

Li C.T., Chen C.S., Cheng C.M., Chen C.P., Chen J.P., Chen M.H., Bai Y.M., Tsai S.J., *Prediction of antidepressant responses to non-invasive brain stimulation using frontal electroencephalogram signals: Cross-dataset comparisons and validation*, Journal of Affective Disorders, 1-10, Elsevier, 2023.

Moini J., Piran P., *Cerebral Cortex*, in Functional and Clinical Neuroanatomy - A Guide for Health Care Professionals, **6**, 177-240, Academic Press, Elsevier, 2020.

Saibene A., Gasparini F., *Genetic algorithm for feature selection of EEG heterogeneous data*, Expert Systems with Applications, 217, 119488, 1-28, Elsevier, 2023.

Stevens C.E., Zabelina D.L., *Creativity comes in waves: an EEG-focused exploration of the creative brain*, Current Opinion in Behavioral Sciences, **27**, 154-162, Elsevier, 2019.

Wang Y., Fang Z., Sun X., Lin X., Niu L., Ma W., *An adaptive driver fatigue classification framework using EEG and attention-based hybrid neural network with individual feature subset*, Biomedical Signal Processing and Control, **85**, 105045, Elsevier, 2023.

Zeynali M., Seyedarabi H., Afrouzian R., *Classification of EEG signals using Transformer based deep learning and ensemble methods*, Biomedical Signal Processing and Control, **86**, 105130, Elsevier, 2023.

ANALIZĂ COMPARATIVĂ A UNOR METODE DE GRUPARE PE RANGURI FOLOSITE DE PROCEDURA DE OPTIMIZARE MULTIOBIECTIV PENTRU SELECȚIA DE TRĂSĂTURI DIN SEMNALE EEG

(Rezumat)

Înregistrările de tip electroencefalogramă (ElectroEncephaloGram - EEG) oferă informații ilustrative despre diagnoza și predicția gândirii umane, dar și a proceselor aferente memoriei, sprijinind cercetătorii și medicii în dezvoltarea sistemelor de tip interfață creier-calculator. Pentru identificare tiparelor de gândire pe baza datelor EEG, extragerea și selecția de trăsături sunt factori determinanți pentru descrierea imparțială a datelor și pentru obținerea de acuratețe ridicată la clasificare. Semnalele EEG analizate în acest studiu sunt colectate în timpul unor activități de gândire standard pentru încărcarea memoriei de lucru de la 16 electrozi și mai apoi divizate pe 4 benzi de frecvență.

Deși majoritatea procedurilor de optimizare bazate pe folosirea algoritmilor genetici (Genetic Algorithm – GA) abordează minimizarea erorii de clasificare și a numărului de trăsături selectate, ele sunt independente de configurarea procedurii de selecție de trăsături, fie ea mono-obiective (Single Objective Optimization – SOO) sau multi-obiectiv (Multi-Objective Optimization – MOO), dar sunt clar afectate de caracterul multi-dimensional, redundant, sau zgomotos al datelor EEG înregistrate. În

plus, distribuția soluțiilor în spațiul obiectiv pentru o serie de populații de soluții candidat generate aleator motivează folosirea abordărilor SOO și MOO. Limitările abordărilor de tip SOO aplicate separat pentru fiecare din cele două obiective: minimizarea erorii de clasificare, respectiv minimizarea numărului de trăsături selectate fundamentează utilizarea abordărilor de tip MOO pentru a obține rezultate superioare și sigure.

În ceea ce privește procedurile de optimizare multiobiectiv, schemele de grupare pe ranguri Pareto care sunt analizate sunt menite să sprijine selecția părinților și a supraviețuitorilor în cadrul buclei evolutive. În general, metodele de grupare pe ranguri Pareto utilizează analiza dominanței pentru a genera ordonarea parțială a soluțiilor, fără a ține cont în mod special de caracterul lor conflictual. Metodele comparate în această lucrare asociază valorile rangurilor combinând procesul de căutare cu mecanismul decizional. Mecanismul decizional este implementat cu ajutorul unor scheme de grupare adaptive, menite să ghideze căutarea către centrul primelor fronturi Pareto, activând eliminarea progresivă a soluțiilor neprofitabile. Populația este divizată în câteva grupuri, cu scopul de a i se menține diversitatea, iar un obiectiv suplimentar este activat ori de câte ori este necesară controlarea diversității materialului genetic valoros. Distribuția soluțiilor în spațiul obiectiv este examinată pe baza unor proceduri de grupare nesupervizată și prin gruparea pe ranguri Pareto în interiorul clusterelor rezultate pentru a contracara dezavantajul inerent al metodelor Pareto. Toate schemele de grupare pe ranguri analizate în cadrul acestei lucrări își demonstrează eficiența în selecția evolutivă a trăsăturilor relevante din semnale EEG. Suplimentar, o serie de modele de clasificatori abordează problema ilustrând diferite mecanisme decizionale.