ASTERICS: Projection-based Classification of EEG with Asymmetric Loss Linear Regression and Genetic Algorithm

Krisztian Buza[a,b]

[a]*Department of Artificial Intelligence*, *Eövös Loránd Unviersity (ELTE)*, Budapest, Hungary
[b]*Center for the Study of Complexity*, *Babeş-Bolyai University*, Cluj Napoca, Romania

buza@biointelligence.hu

***Abstract*—Classification of electroencephalograph (EEG) signals is the common theoretical background of various recognition tasks, such as EEG-based diagnosis of diseases, identification of sleep stages and the recognition tasks related to EEG-controlled spelling devices or web browsers. Projection-based classification of EEG is one of the state-of-the-art techniques to solve such tasks. In this paper, we propose (i) to utilize asymmetric loss linear regression for projection-based classification and (ii) to use genetic algorithm to select reference signals. We performed experiments on a publicly available EEG dataset. Our model aimed to classify patients according to a disease (alcoholism). The results show that both proposed techniques may increase accuracy.**

***Index Terms*—Electroencephalography (EEG), Classification, Genetic Algorithm, Dynamic Time Warping, Asymmetric Loss Linear Regression**

I. Introduction

Classification techniques may be used for various EEG-related recognition tasks, such as the recognition of emotions [1], epileptic seizures [2], identification of sleep stages [3] or the diagnosis of autism [4]. Other classification tasks are related to EEG-controlled spelling devices and web browsers for paralysed patients [5], [6].

State-of-the-art solutions of these recognition tasks are based on machine learning. Various models have been used ranging from logistic regression [7] over support vector machines [8], [9] to neural networks [10] and deep learning [11]. Common pre-processing techniques include time-frequency analysis [3], variants of wavelet decomposition [12], [13] and feature extraction [14]. We refer to [15] and [16] for relevant reviews.

We consider the EEG-related recognition tasks as time series classification [17] tasks for which models based on dynamic time warping (DTW) are popular and effective [18]–[20]. As shown in [21], projecting EEG signals using DTW-distances allows fast and accurate classification of EEG signals. On the one hand, the projection-based approach presented in [21] outperformed various techniques such as neural networks, support vector machines, nearest neighbour and hubness-aware classifiers [22]. On the other hand, in contrast with other DTW-based classifiers, such as nearest neighbour or hubness-aware classifiers, only a few DTW-calculations are required in case of projection-based classification of EEG signals.

Compared with other machine learning techniques, projection-based classification offers the following advantages: (1) due to its conceptual simplicity, the model may be visualised and understood by human experts, especially, if the user sets the amount of reference signals to a relatively low value, (2) a reasonable model may be induced even if the amount of training data is moderate which may be a great advantage in case of rare diseases or rare subtypes of diseases, (3) classification of a new signal is computationally cheap.

For the above reasons, we focus on the enhancement of projection-based classification of EEG signals in this paper. In particular, we propose to utilize asymmetric loss linear regression within projection-based classification and to use a genetic algorithm to select reference signals, thus we call the proposed approach *Asymmetric loss linear regression with Genetic Algorithm for Projection-based classification*, or ASTERICS for short.

The reminder of the paper is organised as follows. In Section II we describe the background required to understand our work. This is followed by the details of our approach (Section III). Section IV presents our experimental results while conclusions are drawn in Section V.

II. Background

We begin this section with a short description of the technology used to acquire brain activity data. Simultaneously, we introduce the data used throughout this study. This is followed by the review of projection-based classification.

A. Electroencephalography

Electroencephalography (EEG) is one of the techniques to capture brain activity. In our study, we used a publicly

K. Buza was supported by the project ED_18-1-2019-0030. Project no. ED_18-1-2019-0030 (Application domain specific highly reliable IT solutions subprogramme) has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the Thematic Excellence Programme funding scheme. K. Buza received the "Professor Ferencz Radó" Fellowship of the Babeş-Bolyai University, Cluj Napoca, Romania. We thank for the usage of MTA Cloud (https://cloud.mta.hu/) that significantly helped us achieving the results published in this paper.

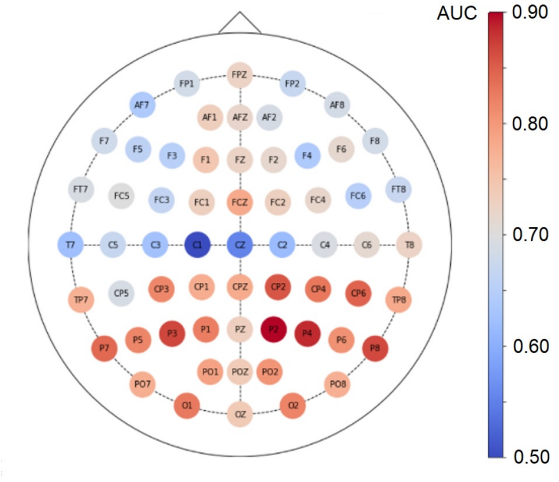

Fig. 1. Performance (area under receiver operator characteristic curve) of nearest neighbour classifier using data from various channels

available EEG dataset[1] from the UCI machine learning repository [23]. In total, this collection contains more than 11,000 EEG signals recorded from 122 persons. Out of the 122 persons, 77 were alcoholic patients and 45 were healthy individuals.

The electrical activity of the brain was captured at 256 Hz for 1 second on 64 channels. For more information about data collection and selection of patients we refer to [23]. In order to filter noise, as a simple preprocessing step, we reduced the length of the signals from 256 to 64 by binning with a window size of four, i.e., we averaged four consecutive values of the signal. Next, we normalised the data: for each time series, we calculated its mean and standard deviation, subsequently, we subtracted the mean from each value and divided it by the standard deviation.

In order to understand the data, we performed an initial experiment with nearest neighbour classifier using DTW as distance measure with a warping window size of 5 time slots (approx. 80 milliseconds). In this experiment, we tried to classify patients using data from a single channel. We performed the experiment for each channel. The results of this analysis (see Fig. 1) show that, compared with other channels, P4 allows for accurate classification. Therefore, we decided to use only this channel subsequently.

B. Projection-based classification of EEG

Projection-based classification has three main steps: (1) selection of reference signals, (2) calculation of DTW distances of each signal to each reference signal, (3) induction of a classifier using the aforementioned distances as features, see also Fig. 2.

In principle, various approaches may be developed within the above framework of projection-based classification. For example, PROCESS [21] performs projection-based classification with random selection of the reference signals (in step 1) and uses logistic regression as classifier (in step 3).

We point out that the selection of reference signals (step 1) and the induction of the classifier (step 3) are essential components of the approach, therefore we focus on these steps in our current study.

When calculating the DTW-distances, we calculated the entire DTW-matrix (i.e., we did not use a warping window). For a detailed description of the calculation of DTW distances we refer to [24].

III. Our Approach

In principle, various optimization techniques and classifiers may be used in the first and third steps of the projection-based approach, i.e., for the selection of appropriate reference signals and to classify the projected signals. We also not that many classifiers, including logistic regression, use an optimization algorithm in the training phase in order to determine appropriate values of their parameters or weights. Just to mention a few of the many techniques, we point out Noisy Extremal Optimization [25], Cuckoo Optimization [26] recent variants of the Grey Wolf Optimizer [27], as well as probabilistic and possibilistic models [28].

In this paper, we propose to use a genetic algorithm for the selection of reference signals in the first step of the projection-based approach. Additionally, we propose to utilize asymmetric loss linear regression (ALLR) in the third step of the projection-based approach. ALLR is a recent classifier which has already been shown to perform surprisingly well for the task of drug-target interaction prediction [29], however, it has not been used in the context of EEG classification before.

This section describes the aforementioned components, the applied genetic algorithm and ALLR.

A. Genetic Algorithm for the Selection of Reference Signals

Genetic algorithms are optimisation techniques motivated by biological evolution. They have been shown to work well in various tasks, such as the detection of Aumann equilibrium [30], Berge and Nash equilibria [31] and feature selection for the classification of brain imaging (fMRI) data [32]. Next, we describe our genetic algorithm for the selection of the appropriate reference signals in order to obtain an accurate classifier.

a) Input: The input is the set of training time-series $T = \{t_i\}_{i=1}^{n}, t_i \in \mathbb{R}^l$, where n is the number of the training time-series.

b) Encoding: Each individual X encodes a set of reference signals. Each individual is a subset of the training time series T. In other words: each training time series t_i is either contained in an individual or not. However, the same training time-series t_i may appear in several individuals.

c) Initialisation: The initial population P contains N individuals, each of them is a random subset of the training time series.

[1] http://archive.ics.uci.edu/ml/datasets/EEG+Database

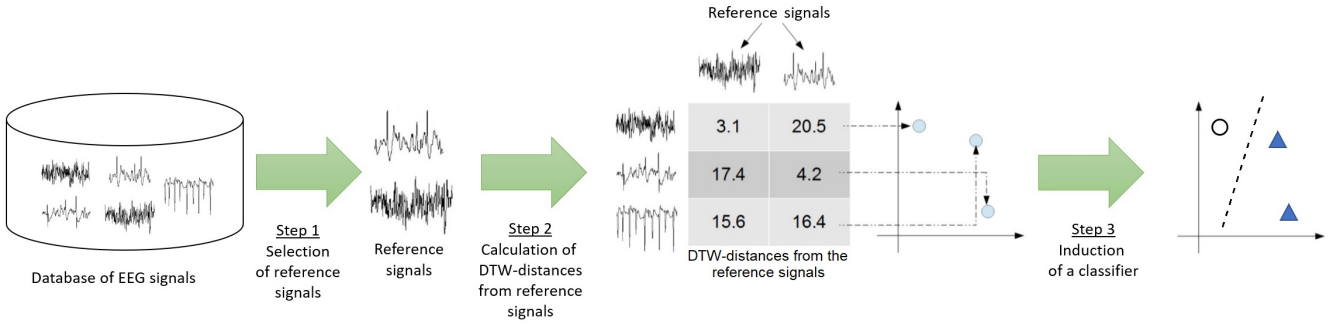

Fig. 2. Projection-based classification of electroencephalograph signals

d) Fitness assignment: In order to calculate the fitness of an individual X, we partition the training data T into two subsets T_1 and T_2. Using the reference signals of X, we project all the instances of T_1 and T_2 into a vector space, i.e., we calculate the DTW-distances of the training time series from the reference signals in X and we use the normalized DTW-distances as features.[2] From now on, we use $T_1^{(p)}$ and $T_2^{(p)}$ to denote the projected data. Using $T_1^{(p)}$, we train a classifier (in particular, asymmetric loss linear regression, which will be described in Section III-B). We evaluate its performance on $T_2^{(p)}$. In particular, we calculate area under receiver operator characteristic curve (AUC) [33] of the classifier on $T_2^{(p)}$ and we use the value of AUC as the fitness of individual X.

e) Offspring: We use recombination (cross-over) and mutation to create the offspring of two randomly selected individuals. We select each reference signal of the offspring one by one. The k-th reference signal of the offspring is either (i) the k-th reference signal of the *first* parent, or (ii) the k-th reference signal of the *second* parent or (iii) a randomly selected signal of the training data. Each of the three above cases has an equal probability of $1/3$.

f) Survival: In each generation, individuals having a fitness value greater than or equal to the median fitness of the population survive.

In our experiments, the size of the population is fixed $N = 10$ throughout all the generations and we run the genetic search for a fixed number of generations $G = 100$.

B. Asymmetric loss linear regression

Given a regression model f_θ where θ is the vector of parameters, f_θ estimates the value of the target y for an instance $x \in \mathbb{R}^m$, i.e., $\hat{y} = f_\theta(x)$. In order to determine the appropriate parameter values θ^*, usually, a loss function $L_D(\theta)$ is minimised:

$$\theta^* = \text{argmin}_\theta \; L_D(\theta). \tag{1}$$

[2]With normalization, we mean that we first calculate the mean and standard deviation of all DTW-distances. Then, from each DTW-distance, we subtract the mean and divide it by the standard deviation and use the resulting values as features.

Note that the actual value of $L_D(\theta)$ depends both on the dataset D and parameters θ. However, once the dataset is fixed, in particular, while the model is being trained using a given training dataset D, the loss can be seen as a function of the parameter vector θ. Therefore, we aim at finding parameters θ^* that minimise the loss. A wide-spread loss function is *mean squared errors*:

$$L_D(\theta) = \frac{1}{|D|} \sum_{(x,y)\in D} (f_\theta(x) - y)^2, \tag{2}$$

where $|D|$ is the number of instances in D.

While the sum of squared errors is popular, we argue that in case of classification tasks related to EEG, it is not fully consistent with the underlying medical reality. While patients are seen as being affected by a disease or not, there may be important differences between two healthy patients, as well as between two patient suffering from a disease (see e.g. the severity of a disease). Consequently, considering the both classes corresponding to the presence ($y = +1$) and absence of a disease ($y = -1$) in case of a particular patient (described by the instance x, we should not penalise a model that predicts a score that is higher than $+1$, if the patient is indeed affected by the disease. Similarly, in case of a healthy patient ($y = -1$), we do not want to penalise a model that predicts a score that is lower than -1.

Therefore, we propose an asymmetric loss function. First, we define the error of the model f_θ for a single prediction $f_\theta(x)$, for instance x with label y as

$$\text{err}(f_\theta, x, y) = \begin{cases} 0 \text{ if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 \text{ if } f_\theta(x) < -1 \text{ and } y = -1 \\ \big(f_\theta(x) - y\big)^2 \quad \text{otherwise.} \end{cases} \tag{3}$$

We define *mean asymmetric loss (MAL)* as the mean of the above errors for all instances of the dataset D:

$$MAL_D(\theta) = \frac{1}{|D|} \sum_{(x,y)\in D} \text{err}(f_\theta, x, y). \tag{4}$$

The above loss can be minimised with various optimisation techniques ranging from gradient-based methods to more advanced approaches, see e.g. [25]. For simplicity, we decided

Algorithm Asymmetric loss linear regression (ALLR)

Require: Training data D, number of epochs e, learning rate η, standard deviation σ

Ensure: Weights $w_0, w_1, \dots w_k$

Initialise weights $w_0, w_1, \dots w_k$ from standard normal distribution with zero mean and standard deviation σ
for epoch in $1 \dots e$ **do**
  for each $(x, y) \in D$ in random order **do**
    $\hat{y} \leftarrow w_0 + \sum_{i=1}^{k} w_i x_i$
    if $\big(\hat{y} > 1$ **and** $y = 1\big)$ **or** $\big(\hat{y} < -1$ **and** $y = -1\big)$ **then**
      continue
    $w_0 \leftarrow w_0 - \eta\, 2(\hat{y} - y)$
    for i in $1 \dots k$ **do**
      $w_i \leftarrow w_i - \eta\, 2(\hat{y} - y) x_i$
  end for
end for
return $w_0, w_1, \dots, w_k$

Fig. 3. **Pseudocode of asymmetric loss linear regression (ALLR)**

to use gradient descent. The partial derivative $\frac{\partial MAL_D(\theta)}{\partial \theta}$ of $MAL_D(\theta)$ is:

$$\frac{\partial MAL_D(\theta)}{\partial \theta} = \frac{1}{|D|} \sum_{(x,y) \in D} \frac{\partial \text{err}(f_\theta, x, y)}{\partial \theta}, \tag{5}$$

where

$$\frac{\partial \text{err}(f_\theta, x, y)}{\partial \theta} = \begin{cases} 0 & \text{if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 & \text{if } f_\theta(x) < -1 \text{ and } y = -1 \\ 2\big(f_\theta(x) - y\big) \frac{\partial f_\theta(x)}{\partial \theta} & \text{otherwise.} \end{cases} \tag{6}$$

In case of linear regression where $x = (x_1, ..., x_k)$, $\theta = \{w_0, w_1, ...w_k\}$, and the model is $f_\theta(x) = w_0 + \sum_{i=1}^{k} w_i x_i$, the partial derivatives of $\text{err}(f_\theta, x, y)$ according to w_i, $1 \leq i \leq k$, are

$$\frac{\partial \text{err}(f_\theta, x, y)}{\partial w_i} = \begin{cases} 0 & \text{if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 & \text{if } f_\theta(x) < -1 \text{ and } y = -1 \\ 2\big(f_\theta(x) - y\big) x_i & \text{otherwise,} \end{cases} \tag{7}$$

while the partial derivative according to w_0 is

$$\frac{\partial \text{err}(f_\theta, x, y)}{\partial w_0} = \begin{cases} 0 \text{ if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 \text{ if } f_\theta(x) < -1 \text{ and } y = -1 \\ 2\big(f_\theta(x) - y\big) & \text{otherwise.} \end{cases} \tag{8}$$

We propose to use stochastic gradient descent to optimise MAL_D. The pseudocode of the resulting asymmetric loss linear regression (ALLR) is shown in Fig. 3.

IV. Experimental Evaluation

The goal of our experiments is two-fold. On the one hand, as our approach is based on PROCESS [21], we compare it with PROCESS. On the other hand, we will show that both proposed components (i.e., genetic algorithm for the selection of reference signals, and asymmetric loss linear regression) contribute to the accuracy of ASTERICS.

In our experiments, we aimed to simulate scenarios in which EEG is used to asses the presence or the severity of a disease. In particular, our models tried to recognize whether a patient is affected by alcoholism or not. We took into account which signals originate from the same person: in order to calculate the likelihood of a person being alcoholic, we average the predictions over all the signals originating from that person.

TABLE I
AUC $\pm$ its standard deviation in case of our approach, ASTERICS, and PROCESS

	AUC
ASTERICS	**0.854 $\pm$ 0.161**
PROCESS	0.775 $\pm$ 0.176

To measure the performance of our approach, we used *area under receiver-operator characteristic curve* (AUC) [33].

We performed 10-fold cross-validation to evaluate our approach, i.e., first, we partitioned the data into 10 subsets and used one of these subsets as test data while the remaining 9 subsets were used as training data. Training and testing the model was repeated 10 times, in each round, a different subset was used as test set. We calculated the average of these AUC values in order to assess the quality of the model.

While partitioning the data for cross-validation, we paid attention that all the signals belonging to the same person were assigned to the same partition, and therefore each person either appeared in the training data or in the test data, but not in both. On the one hand, this allowed to simulate the real-world scenario in which the recognition system is applied to new patients. On the other hand, EEG signals are known to be characteristic to individuals, see e.g. person identification systems using EEG [34], therefore, if the same person would appear in both the train and test data, this could lead to overoptimistic results.

In order to assess if the observed differences are statistically significant we used t-test with a significance threshold of $\alpha = 0.05$.

Throughout the experiments, we set the number of reference signals to $r = 5$.

A. Comparison with PROCESS

As our approach, ASTERICS, is based on PROCESS [21], first we compare ASTERICS and PROCESS. Tab. I shows the results. As one can see, ASTERICS outperforms PROCESS. The difference is statistically significant.

B. Contribution of genetic selection

In order to assess the contribution of the genetic algorithm for the selection of reference signals, we run ASTERICS with and without genetic selection. In the later case, randomly selected reference signals are used. The results are summarized in Tab. II. As one can see, the genetic algorithm indeed has a remarkable contribution to the classification performance.

C. Contribution of the classifier

In order to assess the contribution of the classifier, instead of asymmetric loss linear regression, we run ASTERICS with standard linear regression and logistic regression. The results

TABLE II
AUC $\pm$ ITS STANDARD DEVIATION IN CASE OF OUR APPROACH, ASTERICS, WITH AND WITHOUT GENETIC ALGORITHM FOR THE SELECTION OF REFERENCE SIGNALS

	AUC
ASTERICS with genetic selection	**0.854 $\pm$ 0.161**
ASTERICS with random selection	0.813 $\pm$ 0.169

TABLE III
AUC $\pm$ ITS STANDARD DEVIATION IN CASE OF OUR APPROACH, ASTERICS, WITH ASYMMETRIC LOSS LINEAR REGRESSION (ALLR), STANDARD LINEAR REGRESSION AND LOGISTIC REGRESSION

	AUC
ASTERICS with ALLR	**0.854 $\pm$ 0.161**
ASTERICS with linear regression	0.837 $\pm$ 0.169
ASTERICS with logistic regression	0.833 $\pm$ 0.171

are summarized in Tab. III. As one can see, asymmetric loss linear regression contributes to the classification performance.

V. Conclusion and Future Work

In this paper, we focused on projection-based classification of EEG. We proposed a genetic algorithm for the selection of reference signals and asymmetric loss linear regression as the final classifier. As we discussed, the assumptions underlying our classifier are more consistent with medical reality compared with that of conventional machine learning techniques.

We evaluated our approach on a publicly available real-world EEG dataset and demonstrated that our approach, ASTERICS, outperforms PROCESS and both proposed components contribute its accuracy.

In our future work, we will study the effect of the number of reference signals in more detail, as well as the parameters of the genetic algorithm (such as population size or the number of generations). It would also be interesting to apply ASTERICS to other time series classification tasks.

Acknowledgements

The author thanks for the fruitful discussion with Noémi Gaskó, Mihai Suciu and Rodica Ioana Lung (Babeş-Bolyai University, Cluj Napoca, Romania).

References

[1] V. Kurbalija, M. Ivanović, M. Radovanović, Z. Geler, W. Dai, and W. Zhao, "Emotion perception and recognition: an exploration of cultural differences and similarities," *Cognitive Systems Research*, vol. 52, pp. 103–116, 2018.

[2] R. B. Pachori and S. Patidar, "Epileptic seizure classification in eeg signals using second-order difference plot of intrinsic mode functions," *Computer methods and programs in biomedicine*, vol. 113, no. 2, pp. 494–502, 2014.

[3] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, and H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single eeg channel and random forest classifier," *Computer methods and programs in biomedicine*, vol. 108, no. 1, pp. 10–19, 2012.

[4] E. Grossi, C. Olivieri, and M. Buscema, "Diagnosis of autism through eeg processed by advanced computational algorithms: a pilot study," *Computer methods and programs in biomedicine*, vol. 142, pp. 73–79, 2017.

[5] C. Neuper, G. R. Müller-Putz, R. Scherer, and G. Pfurtscheller, "Motor imagery and eeg-based control of spelling devices and neuroprostheses," *Progress in brain research*, vol. 159, pp. 393–409, 2006.

[6] J. L. Sirvent, J. M. Azorín, E. Iáñez, A. Úbeda, and E. Fernández, "P300-based brain-computer interface for internet browsing," in *Trends in practical applications of agents and multiagent systems*. Springer, 2010, pp. 615–622.

[7] H. Hung and C.-C. Wang, "Matrix variate logistic regression model with application to eeg data," *Biostatistics*, vol. 14, no. 1, pp. 189–202, 2012.

[8] M. Sharma, P. Sharma, R. B. Pachori, and U. R. Acharya, "Dual-tree complex wavelet transform-based features for automated alcoholism identification," *International Journal of Fuzzy Systems*, vol. 20, no. 4, pp. 1297–1308, 2018.

[9] M. Sharma, D. Deb, and U. R. Acharya, "A novel three-band orthogonal wavelet filter bank method for an automated identification of alcoholic eeg signals," *Applied Intelligence*, vol. 48, no. 5, pp. 1368–1378, 2018.

[10] S. Ghosh-Dastidar and H. Adeli, "Improved spiking neural networks for eeg classification and epilepsy and seizure detection," *Integrated Computer-Aided Engineering*, vol. 14, no. 3, pp. 187–212, 2007.

[11] S. Kim, J. Kim, and H.-W. Chun, "Wave2vec: Vectorizing electroencephalography bio-signal for prediction of brain disease," *International journal of environmental research and public health*, vol. 15, no. 8, p. 1750, 2018.

[12] N. Ahmadi, Y. Pei, and M. Pechenizkiy, "Detection of alcoholism based on eeg signals and functional brain network features extraction," in *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2017, pp. 179–184.

[13] S. Patidar, R. B. Pachori, A. Upadhyay, and U. R. Acharya, "An integrated alcoholic index using tunable-q wavelet transform based features extracted from eeg signals for diagnosis of alcoholism," *Applied Soft Computing*, vol. 50, pp. 71–78, 2017.

[14] S. Fattah, K. Fatima, and C. Shahnaz, "An approach for classifying alcoholic and non-alcoholic persons based on time domain features extracted from eeg signals," in *2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*. IEEE, 2015, pp. 479–482.

[15] U. R. Acharya, S. Bhat, H. Adeli, A. Adeli *et al.*, "Computer-aided diagnosis of alcoholism-related eeg signals," *Epilepsy & Behavior*, vol. 41, pp. 257–263, 2014.

[16] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, p. 031005, 2018.

[17] A. Nanopoulos, R. Alcock, and Y. Manolopoulos, "Feature-based classification of time-series data," *International Journal of Computer Research*, vol. 10, no. 3, pp. 49–61, 2001.

[18] X. Xi, E. Keogh, C. Shelton, L. Wei, and C. A. Ratanamahatana, "Fast time series classification using numerosity reduction," in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 1033–1040.

[19] M. Radovanović, A. Nanopoulos, and M. Ivanović, "Time-series classification in many intrinsic dimensions," in *Proceedings of the 2010 SIAM International Conference on Data Mining*. SIAM, 2010, pp. 677–688.

[20] H. A. Dau, D. F. Silva, F. Petitjean, G. Forestier, A. Bagnall, A. Mueen, and E. Keogh, "Optimizing dynamic time warping's window width for time series data mining applications," *Data mining and knowledge discovery*, vol. 32, no. 4, pp. 1074–1120, 2018.

[21] K. Buza, J. Koller, and K. Marussy, "Process: projection-based classification of electroencephalograph signals," in *International Conference on Artificial Intelligence and Soft Computing*. Springer, 2015, pp. 91–100.

[22] K. A. Buza and J. Koller, "Classification of electroencephalograph data: A hubness-aware approach," *Acta Polytechnica Hungarica*, vol. 13, no. 2, pp. 27–46, 2016.

[23] X. L. Zhang, H. Begleiter, B. Porjesz, W. Wang, and A. Litke, "Event related potentials during object recognition tasks," *Brain Research Bulletin*, vol. 38, no. 6, pp. 531–538, 1995.

[24] K. Buza, "Time series classification and its applications," in *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*. ACM, 2018, p. 4.

[25] M. Suciu, R. I. Lung, and N. Gaskó, "Noisy extremal optimization," *Soft Computing*, vol. 21, no. 5, pp. 1253–1270, 2017.

[26] A. Goli, A. Aazami, and A. Jabbarzadeh, "Accelerated cuckoo optimization algorithm for capacitated vehicle routing problem in competitive

conditions," *International Journal of Artificial Intelligence*, vol. 16, no. 1, pp. 88–112, 2018.

[27] R.-C. Roman, R.-E. Precup, C.-A. Bojan-Dragos, and A.-I. Szedlak-Stinean, "Combined model-free adaptive control with fuzzy component by virtual reference feedback tuning for tower crane systems," *Procedia Computer Science*, vol. 162, pp. 267–274, 2019.

[28] C. Carlsson and R. Fullér, "Probabilistic versus possibilistic risk assessment models for optimal service level agreements in grid computing," *Information Systems and e-Business Management*, vol. 11, no. 1, pp. 13–28, 2013.

[29] K. Buza, L. Peška, and J. Koller, "Modified linear regression predicts drug-target interactions accurately," *PloS one*, vol. 15, no. 4, p. e0230726, 2020.

[30] D. Dumitrescu, R. I. Lung, N. Gaskó, and T. M. Dan, "Evolutionary detection of aumann equilibrium," in *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, 2010, pp. 827–828.

[31] N. Gaskó, D. Dumitrescu, and R. I. Lung, "Evolutionary detection of berge and nash equilibria," in *Nature Inspired Cooperative Strategies for Optimization (NICSO 2011)*. Springer, 2011, pp. 149–158.

[32] A. Szenkovits, R. Meszlényi, K. Buza, N. Gaskó, R. I. Lung, and M. Suciu, "Feature selection with a genetic algorithm for classification of brain imaging data," in *Advances in feature selection for data and pattern recognition*. Springer, 2018, pp. 185–202.

[33] V. Bewick, L. Cheek, and J. Ball, "Statistics review 13: receiver operating characteristic curves," *Critical care*, vol. 8, no. 6, p. 508, 2004.

[34] M. Poulos, M. Rangoussi, N. Alexandris, and A. Evangelou, "Person identification from the eeg using nonlinear signal classification," *Methods of information in Medicine*, vol. 41, no. 01, pp. 64–75, 2002.