

Low Dimensional EEG Classification for Alzheimer's Disease Recognition

1st Marija Nedeljković
School of Electrical Engineering
University of Belgrade
Belgrade, Serbia
nedeljkovicmajaa@gmail.com

1st Pavle Pađin
School of Electrical Engineering
University of Belgrade
Belgrade, Serbia
pavle.padjin@gmail.com

1st Strahinja Nikolić
School of Electrical Engineering
University of Belgrade
Belgrade, Serbia
superstrale2002@gmail.com

Abstract—Alzheimer's disease is a neurodegenerative disorder that leads to a progressive decline in cognitive function. In this paper, classification methods are used on EEG signals to differentiate between subjects with Alzheimer's disease (AD) and healthy, controlled group subjects (CN). High model interpretability is achieved using Information gain (IG) feature selection technique, reducing the number of features to 5. Accuracy of 77.6% was accomplished with Random Forest classifier, outperforming all non-deep learning methods on the same dataset.

Index Terms—Alzheimer's disease, classification, EEG signal, dimensionality reduction

I. INTRODUCTION

Electroencephalography (EEG) is a widely used technique for measuring the electrical activity of the brain. EEG is used in a broad range of applications such as diagnostics and classification of neural disorders [1], [2]. Aside from EEG, common approaches also include Functional magnetic resonance imaging (fMRI) [3] and Positron emission tomography (PET) [4]. EEG directly measures the neural activity of a person, while fMRI and PET detect secondary metabolic changes in the brain tissue. Furthermore, EEG is the cheapest, non-invasive method, with a time resolution of less than 1ms. This aspect of EEG enables one to analyze high-frequency components of the signals, as well as patterns that occur in the signal.

Alzheimer's disease (AD) is a common neurodegenerative disorder which is characterized by progressive loss of cognitive functions. Research estimates that the number of cases will double every 20 years [5]. Therefore, research about detection and prevention methods is needed. In the rest of this section, three groups of methods for the classification of Alzheimer's disease will be discussed: Brain imaging techniques, Deep learning-based, as well as standard ML techniques for EEG classification.

Due to great matter loss in AD patients, fMRI can provide useful information in identifying the disease [4]. When combined with deep learning methods, it can achieve an accuracy of up to 97% [3]. Besides fMRI, It was shown that PET gives better results when it comes to classifying AD; in this particular paper, an accuracy of 87% (fMRI) and 89% (PET) is obtained [4]. However, effective brain imaging techniques have certain limitations, and the most important of them are the

cost and the availability [6]. Since EEG-based methods have an advantage in this respect, they are considered in further discussion.

Different deep learning techniques have been used for the task of AD classification including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Autoencoders and more recently, Transformers [1], [7]. All of these techniques have shown promising results, with papers achieving results in accuracy as high as 82.0% with CNN and 83.3% with transformers [7], [8].

In addition to promising results, deep learning methods also have their drawbacks. Most noteworthy is reduced interpretability when compared to standard ML techniques, an aspect that is especially important in medical applications.

This study will focus on Logistic Regression, kNN, Random Forest, SVM, and XGBoost, as they proved to perform well on EEG classification tasks [1], [7], [9]–[13], [14]. As there is no agreed-upon database for benchmarking when it comes to AD classification tasks, it is difficult to compare objectively between results in the literature.

This study focuses on the classification between subjects with Alzheimer's disease (AD) and healthy, controlled group subjects (CN). The problem is approached with standard ML techniques and demonstrates the importance of feature selection by surpassing benchmark accuracy reported in [10] while using a set of only 5 features.

II. METHOD

The workflow, as illustrated in Figure 1, includes three main segments: feature extraction, feature analysis and selection, and classification based on machine learning algorithms. Each of these steps was done in Python utilizing the scikit-learn [15] library. Our code is available at: <https://github.com/strajdzsha/alzheimer-classification>.

A. Database description

As previously mentioned, all the experiments were performed on the database [10] consisting of 65 subjects. Alzheimer's disease was detected in 36 participants and 29 cases subjects were healthy. All recordings were performed using 19 electrodes placed according to the international 10-20 ratio [16]. The database contained preprocessed signals.

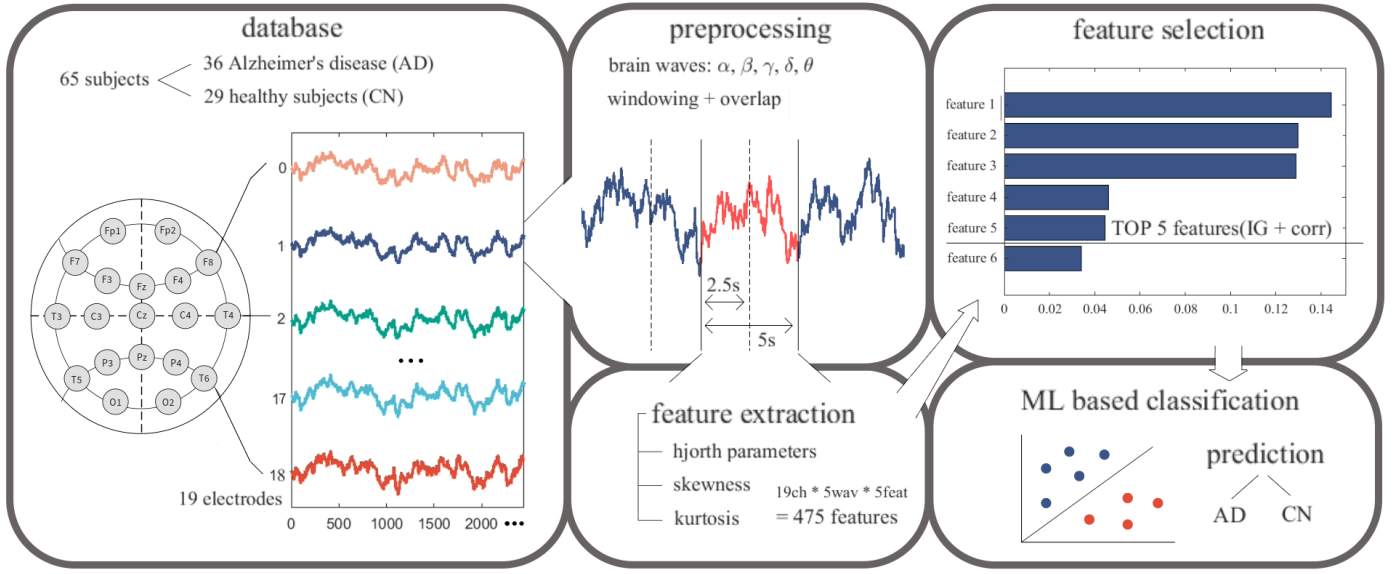


Figure 1. The block diagram of the proposed analysis

All integrated noises caused by the environment were filtered (using a Butterworth bandpass filter (0.5-45Hz)), as well as eye and jaw movements (using Independent Component Analysis, ICA [1]). The degree of the illness was determined by a standardized mini-mental state exam (MMSE) [17].

B. Data preprocessing

Before feature extraction, it is necessary to prepare the data. According to the conclusions drawn from the literature review, subintervals (epochs) of 5 seconds with overlaps of 2.5s were extracted from all signals [9]. The five main types of brain waves that can be recognized are alpha (8-12Hz), beta (13-30Hz), gamma (30-60Hz), delta (0.5-4Hz) and theta (4-8Hz) waves [18]. They provide insight into parts of brain activity and many studies have shown a close relationship between perceptual, cognitive, motor and emotional processes and the type of oscillations in the specific EEG waves [5]. As Alzheimer's disease manifests itself in the form of disruption of the above-mentioned patterns, it is standard practice to perform further analysis over different brain waves [19]. For this reason, all epochs are divided into the mentioned brain waves (using the Butterworth filter) and further feature extraction is performed.

C. Feature Extraction

Based on the literature review, 5 characteristic features were selected: skewness, kurtosis and Hjorth parameters (activity, mobility and complexity).

1) *Skewness*: is a measurement of the distortion of symmetrical distribution.

2) *Kurtosis*: is a measure of the tailedness of a distribution.

3) *Hjorth parameters*: They were first introduced in the paper [20], and designed specifically for EEG signals analysis. The potential of Hjorth's parameters in the classification of Alzheimer's disease was demonstrated in [11]. They consist of three parameters: activity, mobility and complexity.

a) *Activity*: represents the signal power.

b) *Mobility*: represents the mean frequency of the standard deviation of the power spectrum.

c) *Complexity*: represents the change in the frequency.

Features are extracted over each of the 5 brain waves of all 19 channels, resulting in $5 \times 5 \times 19 = 475$ different features for each epoch.

D. Feature Selection

Feature reduction is of significant importance for the model interpretability and computational cost of inference. Furthermore, understanding the best features gives us an insight into which one of them is correlated the best with the disruption of a signal when Alzheimer's disease occurs. Due to the large number of features (475), a feature selection technique is applied for their reduction - Information Gain (IG).

E. Classification

After feature extraction and selection, the classification of the AD and CN groups is performed. Algorithms for classification that proved to be the best for this problem, and which were later used in this work, are k-Nearest Neighbors (kNN), Random Forest and Support Vector Machine (SVM) [1], [7], [9]–[13]. Additionally, the XGBoost (eXtreme Gradient Boosting) algorithm was tested, which tended to be the best choice in various classification tasks.

A modified cross-validation method was used, where 5 subjects per group were taken for the test (5 CN + 5 AD), and the rest for the training set. The whole process was repeated for 20 folds and mean metrics were calculated.

III. RESULTS AND DISCUSSION

Two sets of features were considered for experiments. The first set of features consists of all 475 features. The second set of features consists of the top 5 uncorrelated features. This

feature selection procedure is defined as follows: The first feature is selected as the one with the highest information gain parameter. Then, the second feature is selected from the rest as the one with the highest information gain, while having a Pearson correlation with the first feature less than 0.5. The third feature has the highest information gain among the rest, having Pearson correlation with the first and second features less than 0.5, etc.

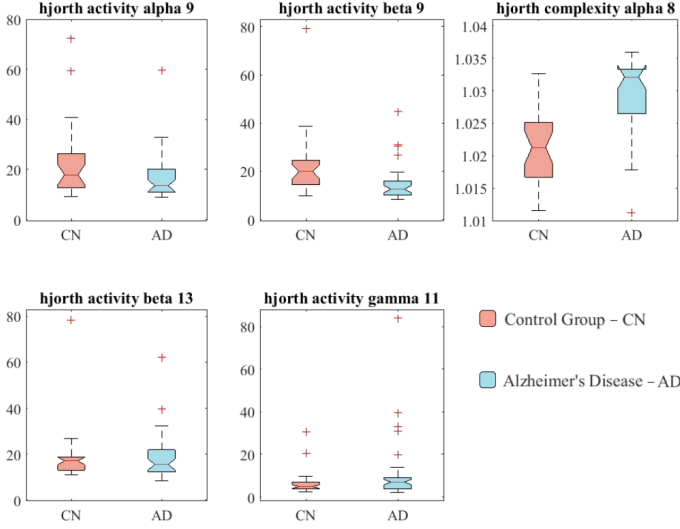


Figure 2. Boxplots for top 5 uncorrelated features. Features are considered uncorrelated if their Pearson correlation is lower than 0.5.

Boxplots for the top 5 uncorrelated features can be seen in figure 2. A decrease in Hjorth activity alpha and beta in the AD group compared to CN and an increase in Hjorth activity gamma are noticeable. Both of these observations have been discussed in literature before [21].

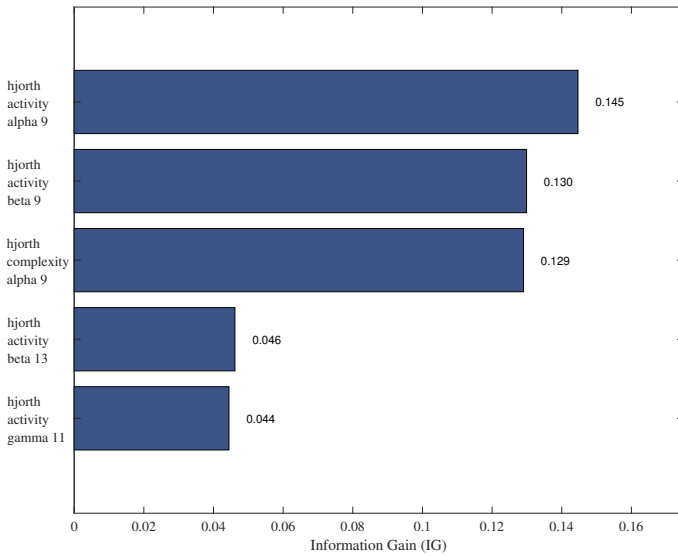


Figure 3. Information gain for 5 uncorrelated features. Features are considered uncorrelated if their Pearson correlation is lower than 0.5.

Figure 3 shows exact values of information gain for the top 5 uncorrelated features. Hjorth parameters are the most informative out of all features; the top 100 features based on information gain are all Hjorth parameters.

Tables I and II show different metrics for two sets of features described above using modified cross validation. The best classification accuracy of 77.6% was achieved with a random forest algorithm on the second set of top 5 features. Furthermore, all algorithms show increased accuracy on the set of 5 features.

Table I
ACCURACY ON ALL 475 FEATURES, FOR DIFFERENT CLASSIFIERS

Model	Logistic Regression	kNN	Random Forest	SVM	XGBoost
Accuracy [%]	75.8 ±4.8	67.6 ±4.5	73.7 ±7.3	73.3 ±7.1	73.7 ±6.9
Sensitivity [%]	75.9 ±5.4	65.3 ±4.0	68.8 ±7.1	71.4 ±8.0	71.6 ±6.5
Specificity [%]	78.2 ±7.3	78.4 ±8.6	86.4 ±8.3	78.6 ±8.5	77.1 ±8.9
F1 [%]	76.7 ±4.9	70.9 ±4.4	76.3 ±6.1	74.4 ±6.2	74.0 ±6.5

Table II
ACCURACY ON 5 FEATURES, FOR DIFFERENT CLASSIFIERS

Model	Logistic Regression	kNN	Random Forest	SVM	XGBoost
Accuracy [%]	76.5 ±7.1	74.4 ±5.8	77.6 ±7.2	74.3 ±6.4	73.8 ±6.9
Sensitivity [%]	72.3 ±8.5	73.6 ±7.1	74.6 ±7.6	69.2 ±7.2	71.6 ±7.9
Specificity [%]	84.6 ±8.5	77.6 ±7.7	84.5 ±9.6	89.3 ±7.2	80.0 ±8.9
F1 [%]	77.5 ±7.8	75.2 ±5.6	78.8 ±6.9	77.9 ±5.2	74.6 ±7.1

This has several important implications. First, it was shown that training on a whole set of features, as often done in literature [1], [7], [9]–[11], [13], can decrease accuracy. Second, reducing the feature set decreases the computational cost of training and inference. Third, a reduced set of features is more interpretable, which is an important aspect in all machine learning applications, especially those closely related to medicine.

When comparing results from this paper to other relevant works, two main things need to be considered: the database and validation method that was used. The database used in this paper was also used in previous papers [7], [9], [10]. Two validation methods are most frequent in literature: leave-one-person-out and k-fold cross-validation. When it comes to the validity of these methods, it is essential that data is split in such a manner that no epoch from patients in the test set is contained in the train set. Only then the obtained results can be considered realistic, as the model was validated on segments from patients it was never trained on.

Given these two remarks, the most relevant paper for our work is [10] which used the same database as in this paper, and used leave-one-person-out as validation method. Benchmark accuracy, which was achieved using 95 features, in [10] is 77.0%, which was surpassed in this paper by 0.6% using fewer features. Similar work was done earlier by the same authors and the reported accuracy was 78.5% [9], however, the database used in that paper was a subset of database [10], as it consisted of only 20 subjects (10 CN + 10 AD). The best accuracy overall on the database used in this paper was achieved using novel transformer-based architecture and is equal to 83.3% [7]. The best accuracy among all non-deep learning methods on this database was achieved in this paper.

Few other papers met our remarks and are eligible for comparison, as they used similar-sized databases and validation methods [11], [12], [13]. T. Araújo et al. also focused on reducing the feature set, achieving 78.9% on 5 features using decision trees [12]. Similar to our work, M. S. Safi et al. made extensive use of Hjorth parameters and achieved an accuracy of 81.0% using SVM [11]. S. J. Ruiz-Gómez et al. reported an accuracy of 78.43% on a larger dataset (37 CN + 37 AD) using MLP on a set of features that is, in good part, different from one discussed in this paper [13].

Limitations of the approach used in this paper should also be discussed. One of them lies in the high noisiness of EEG data, as well as the limited number of patients in the database. Additionally, higher interpretability comes with a performance decrease in accuracy, when compared to deep learning methods [7].

CONCLUSION

This paper addresses the challenge of classifying subjects into two distinct groups: the control group and individuals diagnosed with Alzheimer's disease. The classification metrics obtained through the utilization of all features surpass those reported in [10] on the same database. This study highlights the efficacy of a feature selection approach that leverages information gain (IG) to enhance classification results. The highest accuracy achieved, $(77.6 \pm 7.2)\%$, was obtained with Random Forest, utilizing the subset of five features with the highest IG values and cross-correlation coefficients lower than 0.5. The primary contribution of this research lies in its capacity to improve classification accuracy while significantly reducing dimensionality.

ACKNOWLEDGEMENTS

We would like to express our gratitude for the guidance, dedication, and support our professor Milica Janković provided throughout our academic journey.

REFERENCES

- [1] M. Saeidi, W. Karwowski, F. V. Farahani, K. Fiok, R. Taiar, P. A. Hancock, and A. M. Aljuaid, "Neural Decoding of EEG Signals with Machine Learning: A Systematic Review," *Brain Sciences*, vol. 11, p. 1525, 11 2021.
- [2] J. C. Henry, "Electroencephalography: Basic principles, clinical applications, and Related Fields, fifth edition," *Neurology*, vol. 67, p. 2092, 12 2006.
- [3] S. Sarraf and G. Tofighi, "Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks," *CoRR*, vol. abs/1603.08631, 2016.
- [4] J. Samper-González, N. Burgos, S. Bottani, S. Fontanella, P. Lu, A. Marcoux, A. Routier, J. Guillon, M. Bacci, J. Wen, A. Bertrand, H. Bertin, M. O. Habert, S. Durrleman, T. Evgeniou, and O. Colliot, "Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data," *NeuroImage*, vol. 183, pp. 504–521, 12 2018.
- [5] L. Chu, "Alzheimer's disease: early diagnosis and treatment," *Hong Kong Medical Journal*, vol. 18, no. 3, p. 228, 2012.
- [6] D. N. Lenkov, A. B. Volnova, A. R. Pope, and V. Tsytsarev, "Advantages and limitations of brain imaging methods in the research of absence epilepsy in humans and animal models," *Journal of neuroscience methods*, vol. 212, no. 2, pp. 195–202, 2013.
- [7] A. Miltiadous, E. Gionanidis, K. D. Tzamourta, N. Giannakeas, and A. T. Tzallas, "Dice-net: A novel convolution-transformer architecture for alzheimer detection in eeg signals," *IEEE Access*, 2023.
- [8] F. C. Morabito, M. Campolo, C. Ieracitano, J. M. Ebadi, L. Bonanno, A. Bramanti, S. Desalvo, N. Mammone, and P. Bramanti, "Deep convolutional neural networks for classification of mild cognitive impaired and alzheimer's disease patients from scalp eeg recordings," in *2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)*, pp. 1–6, IEEE, 2016.
- [9] A. Miltiadous, K. D. Tzamourta, N. Giannakeas, M. G. Tsiouras, T. Afrantou, P. Ioannidis, and A. T. Tzallas, "Alzheimer's Disease and Frontotemporal Dementia: A Robust Classification Method of EEG Signals and a Comparison of Validation Methods," *Diagnostics*, vol. 11, p. 1437, 8 2021.
- [10] A. Miltiadous, K. D. Tzamourta, T. Afrantou, P. Ioannidis, N. Grigoriadis, D. G. Tsalikakis, P. Angelidis, M. G. Tsiouras, E. Glavas, N. Giannakeas, *et al.*, "A dataset of scalp eeg recordings of alzheimer's disease, frontotemporal dementia and healthy subjects from routine eeg," *Data*, vol. 8, no. 6, p. 95, 2023.
- [11] M. S. Safi and S. M. M. Safi, "Early detection of alzheimer's disease from eeg signals using hjorth parameters," *Biomedical Signal Processing and Control*, vol. 65, p. 102338, 2021.
- [12] T. Araújo, J. P. Teixeira, and P. M. Rodrigues, "Smart-data-driven system for alzheimer disease detection through electroencephalographic signals," *Bioengineering*, vol. 9, p. 141, 03 2022.
- [13] S. Ruiz-Gómez, C. Gómez, J. Poza, G. Gutiérrez-Tobal, M. Tola-Arribas, M. Cano, and R. Hornero, "Automated multiclass classification of spontaneous eeg activity in alzheimer's disease and mild cognitive impairment," *Entropy*, vol. 20, p. 35, 01 2018.
- [14] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou, *et al.*, "Xgboost: extreme gradient boosting," *R package version 0.4-2*, vol. 1, no. 4, pp. 1–4, 2015.
- [15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [16] R. W. Homan, "The 10-20 electrode system and cerebral location," *American Journal of EEG Technology*, vol. 28, no. 4, pp. 269–279, 1988.
- [17] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "mini-mental state": a practical method for grading the cognitive state of patients for the clinician," *Journal of psychiatric research*, vol. 12, no. 3, pp. 189–198, 1975.
- [18] F. L. da Silva, "Neural mechanisms underlying brain waves: from neural membranes to networks," *Electroencephalography and clinical neurophysiology*, vol. 79, no. 2, pp. 81–93, 1991.
- [19] J. S. Kumar and P. T. V. Bhuvaneshwari, "Analysis of Electroencephalography (EEG) Signals and its Categorization—A Study," *Procedia Engineering*, vol. 38, pp. 2525–2536, 1 2012.
- [20] B. Hjorth, "Eeg analysis based on time domain properties," *Electroencephalography and clinical neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.
- [21] D. V. Moretti, C. Babiloni, G. Binetti, E. Cassetta, G. Dal Forno, F. Ferrer, R. Ferri, B. Lanuzza, C. Miniussi, F. Nobili, *et al.*, "Individual analysis of eeg frequency and band power in mild alzheimer's disease," *Clinical Neurophysiology*, vol. 115, no. 2, pp. 299–308, 2004.