



RESEARCH ARTICLE - COMPUTER SCIENCE

A Comparison among Different Supervised Parameters-Tuned Machine Learning Algorithms with Application to Epileptic EEG Signal Classification

Sajjad A. Mohammed^{1*}, Sura S. Jasim², Baneen A. Thamer³, Ahmed A. Alabdel Abass⁴

^{1,2,3}Biomedical Engineering, College of Engineering, University of Thi-Qar, Thi-Qar, Iraq

⁴Electrical and Electronics Engineering, College of Engineering, University of Thi-Qar, Thi-Qar, Iraq

* Corresponding author E-mail: engbio1m54@utq.edu.iq

Article Info.	Abstract
<i>Article history:</i> Received 16 September 2024 Accepted 30 September 2024 Publishing 30 September 2025	This study investigates the use of advanced machine learning algorithms to classify different classes of epilepsy EEG signals. Using the Time Series Feature Extraction on basis of Scalable Hypothesis tests (TSFRESH) package, a large set of features was reduced to the most important ones by extracting them from the University of Bonn EEG dataset. Following that, seven distinct machine learning classifiers were trained using these features. Three different classification tasks are used to evaluate the classifiers using cross validation. The evaluation results of each algorithm after testing them on each of the three tasks show that the accuracy of the Support Vector Machine (SVM) reaches 99.60% for binary classification, the accuracy of Histogram-based Gradient Boosting Trees (HGBT) reaches 99.20% for the three-class task, and for the five-class classification, the Random Forest achieves 94.80% accuracy, which is the highest among the others. These results show that there is no dominant algorithm that works for all classification tasks and it is necessary to always run more than one algorithm to get a better assessment.

This is an open-access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>)

The official journal published by the College of Education at Mustansiriya University

Keywords: Machine learning; Feature Extraction; Hyperparameter Tuning; Epilepsy.

1. Introduction

For accurate tracking of brain activity, healthcare professionals rely on measuring the electrical activity through electroencephalography (EEG) [1]. EEG is used for monitoring and identification of harmful neurological disorders such as epilepsy, which is a chronic disease that affects the brain and causes seizures in a periodic way. It might be claimed that while it affects both sexes, men are more affected than women. Since it affects people of various ages, there is no set age for it [2]. Compared to alternative methods used in the evaluation of epilepsy, EEG devices offer portability, cost-effectiveness, and the ability to capture time-domain recordings [3]. However, the analysis of EEG data poses complex challenges due to the complex characteristics of the signal and the large volume of data that requires intensive processing. Due to the nonlinearity of EEG signals, in addition to their non-stationary property, this makes it difficult to obtain accurate and complete information about these biomedical signals and their dynamics. This makes the diagnosis of seizures a difficult challenge for clinicians [4-5]. For this, machine learning (ML) and deep learning (DL) algorithms are applied to EEG data, making it easier and more efficient to decode large, complex, and high-resolution EEG datasets. This is done through a set of tools used to build specific models [6-7]. ML tools consist of algorithms that automatically build these data models from the set of features extracted from the original data (training data). The data models are then used to predict or classify new data (test data) [8]. ML and DL techniques have proven successful in many fields including classification and analysis, making significant contributions to different areas, this is possible due to the use of high computational power to run these models, thus enhancing EEG analysis [9]. The four most widely used types of ML are reinforcement learning, unsupervised learning, semi-supervised learning, and supervised learning. Supervised learning methods are on average of higher accuracy than their unsupervised counterparts. In ML, data is the most crucial component. Numerous fields, including the medical field, may provide this data. Once gathered, this data is examined, converted into the right format for the model, and then entered into various algorithms [10-11]. On the other hand, time series data is a set of sequential measurements recorded over time, where the order of the data points plays a crucial role in the analysis [12]. Differentiating characteristics set time series data apart from other types of data, one important characteristic that shows

how each data point is impacted by its prior observations is temporal dependence, because of this interconnectedness, the data exhibits observable patterns and trends that call for efficient methods of collection and extraction. Inherent characteristics of time series data, such as periodicity, or patterns, can also be seen in EEG signals [13-14]. Comprehensive EEG data analysis requires an understanding of the ability to extract these inherent patterns, this is why researchers and clinicians could benefit from these techniques to help diagnose, monitor, and understand neurological disorders [15-16].

The contribution of this paper is testing different ML algorithms in epilepsy signal classification. By dividing the problem into three sub-problems, it has been possible to look deeper into each ML algorithm and its efficiency. Specifically, the first sub-problem is a multi-class classification task for 5 classes, the second one is a multi-class classification task for 3 classes, normal, inter-ictal, and seizure. The last one is a binary classification task to classify between seizure and no seizure classes. The paper is organized as follows. In the second part, Sec.2 the review of similar studies on EEG classification is provided. The third section provides the proposed system model and describes the data and methods used. In the fourth section, the results are presented and discussed. Finally, the conclusion of the paper is in Sec. 5.

2. Related Works

This section presents a review of previous studies on EEG classification. EEG and epilepsy classification have been studied previously, and different traditional machine learning algorithms have been applied to different tasks.

Savadkoobi et al. [17] extracted EEG features in three domains: time, frequency, and time-frequency, they also applied feature selection for classification, comparing both support vector machine (SVM) and K-Nearest Neighbors (KNN) algorithms where SVM's accuracy was higher. AlShorman et al. [18] used SVM and Naive Bayes and achieved an accuracy ranging from 90% to 98.21% for detecting psychological stress from EEG. Bashivan et al. [19] investigated robust representations of EEG using ML and DL techniques, the results for test errors were 12.39% and 8.89% for both single and multi frame classifications respectively. Raghu et al. [20] aimed to classify seven types of seizures and non-seizure EEG patterns with pretrained Convolutional Neural Networks (CNNs) architecture and SVM, their method achieved an accuracy of 88.30%. Xu et al. [21] researched to develop a model of recognizing seizures automatically using epileptic signals with a One-Dimensional Convolutional Neural Network-Long Short-Term Memory model (1D CNN-LSTM). It has achieved high accuracy in binary and multi-class tasks related to seizures. Varalakshmi et al. [22] applied the global feature extraction using the Tunable Q Factor Wavelet Transformation (TQWT) for EEG decomposition, EEG signals were classified using a simple Artificial Neural Network (ANN).

Hussain et al. [23] proposed a Multichannel Vision Transformer (MViT) architecture to classify preictal and interictal EEG activities with the use of Continuous Wavelet Transform (CWT), this method achieved 99.8% accuracy, on the CHB-MIT Dataset. Potter et al. [24] introduced a transformer-based unsupervised learning approach for seizure identification in EEG data, the achieved results were an AUC of 0.93 ± 0.005 on the CHB-MIT dataset. Yan et al. [25] highlighted the design of a model to predict epileptic seizures for patients with epilepsy via Three-tower Transformer, achieving a sensitivity result of 96.01% on the CHB-MIT data. Qi et al. [26] described an automatic epilepsy detection method by using feature extraction and CNNs with an attention mechanism for classification. Achieved 98.89% for triple and 99.70% for binary classification accuracies respectively in Bonn and Bern-Barcelona datasets. Moreover, Lebal et al. [27] proposed and evaluated Epilepsy-Net. Both CNN, Recurrent Neural Network (RNN), and attention mechanism were used, achieving high accuracy on Bonn EEG dataset. Lastly, Pontes et al. [28] proposed a seizure prediction algorithm incorporating a data-batch selection method using logistic regression (Seizure-batch Regression), with dynamic integration of classifiers (Dynamic Weighted Ensemble), a combination of univariate linear features and SVM classifiers were used. Considering a group of 37 patients with Temporal Lobe Epilepsy from the Epilepsiae database. The best-performing approach was the Backwards-Landmark Window, which achieved results of 0.75 ± 0.33 for sensitivity and 1.03 ± 1.00 for false positive rate per hour.

3. System Model

This section briefly describes the data set used in the study, the algorithms, the metrics used to assess the efficiency of each algorithm, and the feature selection approaches. The system model is shown in Figure 2 below.

3.1 Dataset Description

The Bonn University EEG dataset is one of the resources that is widely used in epileptic seizure classification tasks, it is collected by the Department of Epileptology at the University of Bonn, Germany, and the recording of the signals is performed using the international 10-20 system, they were sampled at 173.61 Hz sampling rate and filtered with a 0.53-40 Hz bandpass filter. The dataset is organized into five classes, each class contains 100 single-channel EEG records, and each record is 23.6 seconds long with 4097 samples in each. Each class's records show a distinct state of brain activity: class A records from healthy volunteers with their eyes open, class B records from healthy volunteers who have their eyes closed, class C and D feature interictal signals from patients with epilepsy during intervals without seizures, and class E contains ictal signals recorded during a seizure. Table 1 shows an overview of the dataset used [29].

Table 1. Overview of Bonn University EEG Dataset

Class	Description	State
A	Healthy volunteers with eyes open	Normal
B	Healthy volunteers with eyes closed	Normal
C	Epileptic patients, interictal (opposite hemisphere)	Interictal
D	Epileptic patients, interictal (epileptogenic zone)	Interictal
E	Epileptic patients, ictal (during seizures)	Seizure (Ictal)

3.2 Feature Extraction and Selection

Feature extraction in this study is performed using the Time Series Feature Extraction on basis of Scalable Hypothesis tests (TSFRESH) package in Python 3.12.0, this package makes feature extraction faster by automatic calculation of features using 63 different feature calculation methods, capturing various statistical properties of each EEG signal as well as time-domain, and frequency-domain characteristics. Feature selection is performed using the feature selection method in the same package to select the most relevant features for more efficient computation [30]. The package implements standard APIs of time series and machine learning libraries (e.g. pandas and scikit-learn) and it is designed for both exploratory analyses as well as straightforward integration into operational data science applications. All the features are standardized using a standard scalar before they are fed to the classifiers so that they have a mean of zero and standard deviation of one, the standard scalar is expressed as:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

where z is the standardized value, x is the original feature, μ is the mean of the feature values and σ is the standard deviation of the feature values.

This process is crucial, especially for algorithms that rely on distance calculation like SVM and are sensitive to feature scales, the standardization will reduce the bias towards features that have large numeric ranges.

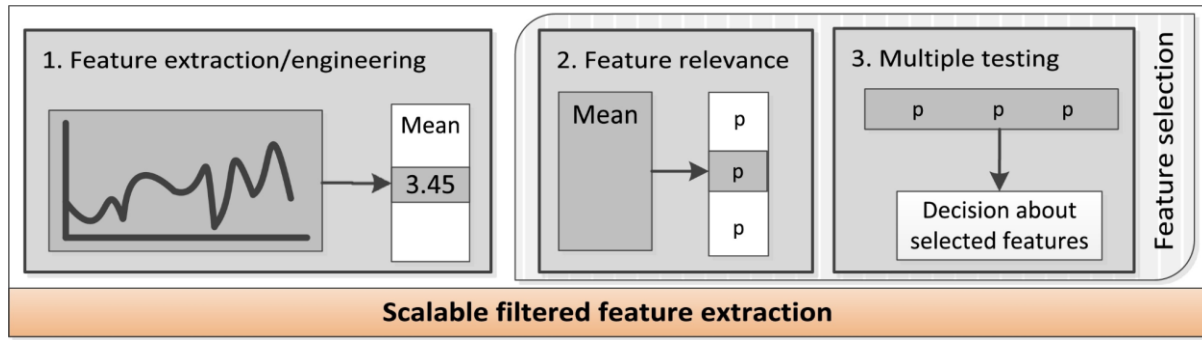


Figure 1. The three steps of the TSFRESH algorithm are feature extraction (1.), calculation of p -values (2.) and a multiple testing procedure (3.), image from [41].

3.3 Classification methods

We assess the performance of a range of ML classifiers on the EEG dataset for epilepsy classification. Logistic regression estimates the likelihood that a given input belongs to a specific class by fitting a logistic function to the input features, it is useful in situations when there is a linear relationship between the target and the features [31]. The non-parametric, instance-based learning technique KNN classifies a data point according to the majority class among its k -nearest neighbors in the feature space [32]; however, because of its computational cost, its performance may suffer with larger datasets, so we employ KNN with Bagging (Bootstrap Aggregating) to enhance the performance of KNN [33]. Another non-parametric model is a decision tree, which is used independently to evaluate their performance, but in order to reduce overfitting, the ensemble learning technique, random forest is used. It uses a bootstrap sample of the data to train each decision tree and only selects a random subset of features for splitting at each node. This increases accuracy and reduces overfitting [34]. An additional ensemble technique also used is Histogram-based Gradient Boosting Trees (HGBT) which builds models sequentially with each new model correcting the errors of its predecessor [35]. Lastly, to find the hyperplane that separates the classes in the feature space with the application of kernel functions, we use SVM, which can handle both linear and non-linear classification by maximizing the margin between the nearest points of different classes [36].

3.4 Cross Validation (CV)

We can use previously unseen data to evaluate our models by dividing our dataset into train and test sets. However, it is possible that the split used for testing does not represent the true performance of our models because of data variability. For that, we will use K-Fold CV with shuffling to generate randomly shuffled pairs of data, and stratification to maintain class distribution, we train and test the model on the complete dataset and average the results to get a more accurate model evaluation [37-38]. We use 5 splits for the five-class classification task and 10 splits for both binary and the three-class classification tasks.

3.5 Model Selection and Hyperparameters Tuning

We select the best-performing model based on the CV results, to further enhance the performance of the selected model, the optimal hyperparameters should be chosen. Grid search [39], also known as exhaustive grid search, is a popular method for determining the optimal set of hyperparameters. It is done by defining the range of potential values for the hyperparameters and then evaluating every combination that can be found. Furthermore, it is designed to conduct hyperparameter tuning in a systematic way by going through each of the sets of hyperparameter values automatically during the

model training process [40]. The fact that the hyperparameter settings are independent is one benefit of grid search. It is hence appropriate for parallel processing.

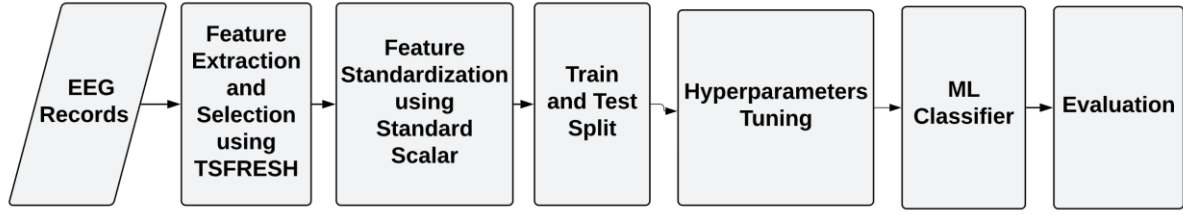


Figure 2. Overview of the Proposed System Model

3.6 Evaluation Metrics

To understand the performance of the classification models the following evaluation metrics are used:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall (Sensitivity)} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (5)$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents False Negatives. These metrics collectively provide a comprehensive view of a model's strengths and weaknesses in classification tasks.

4. Simulation Results

This section presents a comprehensive evaluation of performance metrics, including accuracy, precision, recall, and F1-score, for the previously mentioned classification methods across three distinct tasks: binary, five-class, and three-class classification. A comparative analysis was conducted among Logistic Regression, KNN, KNN with bagging, SVM, HGBT, Random Forest, and Decision Trees.

4.1 Experimental Setup

The dataset is combined using three different ways, specifically, three classification tasks to comprehensively evaluate the model performance as follows: binary classification (four non-seizure classes vs. seizure class), multi-class classification for three classes (normal, interictal, and seizure), and multi-class classification task for five classes.

All experiments were conducted using the 1.51 version of scikit-learn [37] library in Python.

4.2 Feature Extraction and Selection Results

A total of 777 features were computed from each EEG recording using the TSFRESH library for extracting features. In order to improve the efficiency and performance of the model, feature selection was used; the three-class, five-class, and binary classification tasks produced 383, 343, and 281 features, respectively. Because of the decrease in dimensionality, it is possible to estimate the significance of feature engineering in EEG-based classification by considering the value of distinct feature subsets in differentiating between numbers of classes.

4.3 Classification Results

The classification results for the binary classification task are summarized in Table 2, while the results for the three-class and five-classes classification tasks are summarized in Table 3 and Table 4 respectively.

Table 2. Evaluation Results for Binary Classification Task

Model	Accuracy (%)	F1-Score (%)	Precision (%)	Recall (%)	Specificity (%)
SVM	99.60	99.60	99.63	99.60	99.90
Logistic Regression	99.20	99.20	99.25	99.20	99.90
KNN	98.20	98.16	98.26	98.20	98.40
KNN with Bagging	98.40	98.35	98.44	98.40	98.35
Decision Tree	97.00	96.96	97.10	97.00	92.60
Random Forest	96.60	96.59	96.62	96.60	98.20
HGBT	98.00	97.98	98.07	98.00	95.10

Table 3. Evaluation Results for Three-Class Classification Task

Model	Accuracy (%)	F1-Score (%)	Precision (%)	Recall (%)	Specificity (%)
SVM	98.40	98.39	98.47	98.40	99.01
Logistic Regression	98.40	98.37	98.48	98.40	98.93
KNN	96.60	96.55	96.77	96.60	98.06
KNN with Bagging	96.80	96.76	96.96	96.80	98.30
Decision Tree	95.40	95.39	95.63	95.40	97.23
Random Forest	97.60	97.58	97.72	97.60	98.30
HGBT	99.00	98.99	99.03	99.00	99.46

Table 4. Evaluation Results for Five-Class Classification Task

Model	Accuracy (%)	F1-Score (%)	Precision (%)	Recall (%)	Specificity (%)
SVM	91.60	91.43	92.13	91.60	97.90
Logistic Regression	92.00	91.87	92.54	92.00	98.00
KNN	86.00	85.62	86.32	86.00	96.50
KNN with Bagging	87.60	87.46	88.67	87.60	96.50
Decision Tree	83.20	83.11	84.26	83.20	95.90
Random Forest	93.60	93.57	94.12	93.60	98.30
HGBT	93.20	93.08	93.64	93.20	98.30

From the results, Random Forest showed as the best-performing model for the five classes in a multi classification task. It achieved an accuracy of 93.60%, which means that 93.60% of its predictions were correct. Furthermore, it had an F1-score of 93.57% which signifies a balance between precision and recall. Specifically, recall measures how effectively the model could identify all true positives, whereas precision is the number of positive predictions that were truly positive. produced scores for precision and recall of 94.12% and 93.60%, respectively.

Similarly, in binary classification, SVM had the best scores with an accuracy of 99.60% and other metrics as well. On the other hand, for multi-class three-class classification tasks, HGBT showed very high performance across all metrics, it achieved almost the perfect accuracy of 99%. The F1-score of 98.99% reflects a strong balance between precision 99.03% and recall 99%, indicating the model's ability to correctly identify positive cases and minimizing false negatives.

It can be noticed that HGBT demonstrated the most consistent and superior performance across all classification tasks. For more specificity, in binary classification, SVM and Logistic Regression showed excellent results, with SVM slightly outperforming the Logistic Regression. Conversely, in multi-class classification of the three-class, HGBT achieved perfect scores, while SVM and Logistic Regression maintained strong performance. In multi-class classification of the five-class, Logistic Regression performed slightly better than SVM, and both significantly outperformed other models. However, KNN and KNN with Bagging performed well in binary classification but lagged in multi-class classification, on the other hand, Decision Trees and Random Forest are generally underperformed compared to other models, especially in multi-class tasks. But HGBT demonstrates exceptional adaptability to different classification complexities, consistently achieving top results.

4.4 Hyperparameters Tuning Results

The best hyperparameters for each best-performing model in each task are identified through grid search with cross-validation, these parameters will further improve the performance of the models. For the Random Forest model, the best hyperparameters were: 1000 estimators, minimum samples split of 5, a minimum samples leaf of 2, max features set to 'sqrt', a max depth of 100, and no bootstrap, resulting in improved classification metrics. The HGBT model achieved its best performance with an l2_regularization of 0.0, a learning rate of 0.1, a max depth of 5, and 100 iterations. The SVM model

showed enhanced results on binary classification by setting the hyperparameters of C to 0.1, and gamma to scale while using a linear kernel.

Table 5 presents the evaluation results for each split as well as the average results for SVM classifier in the binary classification task. While the results for HGBT are presented in Table 6 show a slight increase in the performance of the classifier after hyperparameter optimization. Table 7 shows improved results attained by random forest for each split and the overall average results for the five-class classification task.

Table 5. Evaluation Results for Binary Classification Tasks for SVM

Split	Accuracy	F1-score	Precision	Recall	Specificity
1	100%	100%	100%	100%	100%
2	100%	100%	100%	100%	100%
3	100%	100%	100%	100%	100%
4	100%	100%	100%	100%	100%
5	100%	100%	100%	100%	100%
6	98%	98.03%	98.18%	98%	99.5%
7	100%	100%	100%	100%	100%
8	98%	98.03%	98.18%	98%	99.5%
9	100%	100%	100%	100%	100%
10	100%	100%	100%	100%	100%
Average	99.60%	99.60%	99.63%	99.60%	99.90%

Table 6. Evaluation Results for Three-Class Classification Task for HGBT

Split	Accuracy	F1-score	Precision	Recall	Specificity
1	100%	100%	100%	100%	100%
2	98%	97.99%	98.09%	98%	98.66%
3	98%	97.97%	98.09%	98%	98.66%
4	98%	97.99%	98.09%	98%	98.66%
5	100%	100%	100%	100%	100%
6	100%	100%	100%	100%	100%
7	100%	100%	100%	100%	100%
8	98%	97.97%	98.09%	98%	98.66%
9	100%	100%	100%	100%	100%
10	100%	100%	100%	100%	100%
Average	99.20%	99.19%	99.23%	99.20%	99.46%

Table 7. Evaluation Results for Five-Class Classification Task for Random Forest

Split	Accuracy	F1-score	Precision	Recall	Specificity
1	98%	97.99%	98.18%	98%	99.5%
2	92%	91.88%	92.14%	92%	97.5%
3	92%	91.97%	92.32%	92%	98%
4	96%	95.99%	96.18%	96%	99%
5	96%	96%	96%	96%	99%
Average	94.80%	94.77%	94.96%	94.80%	98.60%

The results showed that SVM consistently achieved the highest performance in binary classification, with very high accuracy, and other metrics as well. It also maintained strong performance in the three-class classification and five-class classification. Logistic Regression results were comparable to SVM but slightly better than it in the five-class classification. HGBT excelled in the three-class classification task with excellent scores around 99% after hyperparameters tuning. Random Forest displayed significant performance in the five-class classification at 94.8% with optimized hyperparameters, but it was less competitive in binary and three-class classification tasks. For example, binary classifications by KNN and KNN with Bagging produced approximately 98.20% and 98.40% respectively but performed poorly on the three-class classification task, the classifiers having about 3 percent less accuracy of 96.60%, while at the same time falling significantly at approximately 86% accuracy with the five-class classification, however, Decision Tree has been found to always have low scores across all tests especially within the three-class classification.

The application of these algorithms to epilepsy diagnosis and monitoring could significantly enhance the accuracy and efficiency of identifying epileptic events. The provided analysis offers a comprehensive approach by comparing the performance of various ML algorithms on three different kinds of classifications, Binary, three-class, and five-class classification, thus allowing us to easily examine their strengths in terms of speed as well as how well they perform and it's also assessed using accuracy, precision, recall and F1-score for comprehensive evaluation regarding model performance, considering both right predictions plus covering the balance between true positives against false positives/negatives, also clear results are provided enabling understanding about the performance of algorithms during various tasks. Moreover, the high accuracy of these models can easily enhance the automated monitoring systems that continuously analyse EEG data in real time, providing immediate alerts for potential epileptic events.

5. Conclusion

The aim of this research was to classify Electroencephalography (EEG) signals into different categories related to epilepsy, which is a chronic brain disease that causes periodic seizures. This is done by using Machine Learning (ML) applied to the extracted features from EEG data of Bonn EEG dataset, the features were identified using the Time Series Feature Extraction on basis of Scalable Hypothesis tests (TSFRESH) package and then fed into seven machine learning classifiers, which were evaluated through three classification tasks, binary, three-class, and five-class classification tasks to identify the best-performing model in each task. The results showed that Support Vector Machine (SVM) was the top-performing model in binary classification with an accuracy of 99.6% and a specificity of 99.90%. For three-class classification, Histogram-based Gradient Boosting Trees (HGBT) achieved an accuracy of 99.2% and a specificity of 99.46%. Finally, the random forest model achieved 94.80% accuracy and also a specificity of 98.30% in the five-class classification. These results make it clear how important these methods are for using EEG data to diagnose neurological conditions. In order to improve the diagnosis and treatment of epilepsy and other neurological

illnesses, future research should concentrate on the improvement of these techniques and their speed, in addition to lowering their false positive rate, to classify negative cases more robustly, and the use of more comprehensive feature extraction.

References

- [1] D. L. Sherman and N. V. Thakor, "EEG Signal Processing: theory and applications," in Springer eBooks, 2020, pp. 97–129. doi: 10.1007/978-3-030-43395-6_3.
- [2] E. Beghi, "The Epidemiology of Epilepsy," *Neuroepidemiology*, vol. 54, no. 2, pp. 185–191, Dec. 2019, doi: 10.1159/000503831.
- [3] P. L. Nunez and R. Srinivasan, *Electric fields of the brain*. 2006. doi: 10.1093/acprof:oso/9780195050387.001.0001.
- [4] S. Chambon, M. N. Galtier, P. J. Arnal, G. Wainrib, and A. Gramfort, "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 4, pp. 758–769, Apr. 2018, doi: 10.1109/tnsre.2018.2813138.
- [5] N. Aaysha, M. B. Qureshi, M. Afzaal, M. S. Qureshi, and M. Fayaz, "Machine learning-based EEG signals classification model for epileptic seizure detection," *Multimedia Tools and Applications*, vol. 80, no. 12, pp. 17849–17877, Feb. 2021, doi: 10.1007/s11042-021-10597-6.
- [6] H. Zhang et al., "The applied principles of EEG analysis methods in neuroscience and clinical neurology," *Military Medical Research*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40779-023-00502-7.
- [7] R. Shoorangiz, S. J. Weddell, and R. D. Jones, "EEG-Based Machine Learning: Theory and Applications," in Springer eBooks, 2021, pp. 1–39. doi: 10.1007/978-981-15-2848-4_70-1.
- [8] J. R. De Miras, A. J. Ibáñez-Molina, M. F. Soriano, and S. Iglesias-Parro, "Schizophrenia classification using machine learning on resting state EEG signal," *Biomedical Signal Processing and Control*, vol. 79, p. 104233, Jan. 2023, doi: 10.1016/j.bspc.2022.104233.
- [9] X. Zhang, X. Xu, J. Ren, Y. Zhang, and H. Wang, "A lightweight dual-stage CNN framework for automatic sleep stage classification using single-channel EEG signal," in 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Los Angeles, CA, USA, 2022, pp. 128–133, doi: 10.1109/BIBM55620.2022.10031507.
- [10] S. Tufail, H. Riggs, M. Tariq, and A. I. Sarwat, "Advancements and Challenges in Machine Learning: A comprehensive review of models, libraries, applications, and algorithms," *Electronics*, vol. 12, no. 8, p. 1789, Apr. 2023, doi: 10.3390/electronics12081789.
- [11] M.-P. Hosseini, A. Hosseini, and K. Ahi, "A review on Machine learning for EEG signal Processing in bioengineering," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 204–218, Jan. 2021, doi: 10.1109/rbme.2020.2969915.
- [12] W. Zheng and J. Hu, "Multivariate Time Series Prediction based on Temporal Change Information learning Method," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 10, pp. 7034–7048, Oct. 2023, doi: 10.1109/tnnls.2021.3137178.
- [13] X. Song and F. Xiao, "Combining time-series evidence: A complex network model based on a visibility graph and belief entropy," *Applied Intelligence*, vol. 52, no. 9, pp. 10706–10715, Jan. 2022, doi: 10.1007/s10489-021-02956-5.
- [14] A. Zigler, S. Straw, I. Tokuda, E. Bronson, and T. Riede, "Critical calls: Circadian and seasonal periodicity in vocal activity in a breeding colony of Panamanian golden frogs (*Atelopus zeteki*)," *PLoS ONE*, vol. 18, no. 8, p. e0286582, Aug. 2023, doi: 10.1371/journal.pone.0286582.
- [15] D. Bethge, P. Hallgarten, O. Ozdenizci, R. Mikut, A. Schmidt, and T. Grosse-Puppenthal, "Exploiting Multiple EEG Data Domains with Adversarial Learning," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Jul. 2022, doi: 10.1109/embc48229.2022.9871743.

- [16] P. Mathur and V. K. Chakka, "Graph signal processing based Cross-Subject mental task classification using Multi-Channel EEG signals," *IEEE Sensors Journal*, vol. 22, no. 8, pp. 7971–7978, Apr. 2022, doi: 10.1109/jsen.2022.3156152.
- [17] M. Savadkoohi, T. Oladunni, and L. Thompson, "A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal," *Journal of Applied Biomedicine*, vol. 40, no. 3, pp. 1328–1341, Jul. 2020, doi: 10.1016/j.bbe.2020.07.004.
- [18] O. AlShorman et al., "Frontal lobe real-time EEG analysis using machine learning techniques for mental stress detection," *Journal of Integrative Neuroscience*, vol. 21, no. 1, p. 020, Jan. 2022, doi: 10.31083/j.jin2101020.
- [19] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks," *arXiv (Cornell University)*, Jan. 2015, doi: 10.48550/arxiv.1511.06448.
- [20] S. Raghu, N. Sriraam, Y. Temel, S. V. Rao, and P. L. Kubben, "EEG based multi-class seizure type classification using convolutional neural network and transfer learning," *Neural Networks*, vol. 124, pp. 202–212, Apr. 2020, doi: 10.1016/j.neunet.2020.01.017.
- [21] G. Xu, T. Ren, Y. Chen, and W. Che, "A One-Dimensional CNN-LSTM model for epileptic seizure recognition using EEG signal analysis," *Frontiers in Neuroscience*, vol. 14, Dec. 2020, doi: 10.3389/fnins.2020.578126.
- [22] V. P, R. V, C. S. C, and A. V. R, "EEG Signal based epileptic seizure Forecasting using deep learning models," *Ieee*, Mar. 2021, doi: 10.1109/wispnet51692.2021.9419411.
- [23] R. Hussein, S. Lee, and R. Ward, "Multi-Channel Vision Transformer for Epileptic Seizure Prediction," *Biomedicines*, vol. 10, no. 7, p. 1551, Jun. 2022, doi: 10.3390/biomedicines10071551.
- [24] İ. Y. Potter, G. Zerveas, C. Eickhoff, and D. Duncan, "Unsupervised Multivariate Time-Series Transformers for Seizure Identification on EEG," *Ieee*, Dec. 2022, doi: 10.1109/icmla55696.2022.00208.
- [25] J. Yan, J. Li, H. Xu, Y. Yu, and T. Xu, "Seizure prediction based on transformer using scalp electroencephalogram," *Applied Sciences*, vol. 12, no. 9, p. 4158, Apr. 2022, doi: 10.3390/app12094158.
- [26] Q. Xin, S. Hu, S. Liu, L. Zhao, and Y.-D. Zhang, "An Attention-Based Wavelet Convolution Neural Network for Epilepsy EEG classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 957–966, Jan. 2022, doi: 10.1109/tnsre.2022.3166181.
- [27] A. Lebal, A. Moussaoui, and A. Rezgoui, "Epilepsy-Net: attention-based 1D-inception network model for epilepsy detection using one-channel and multi-channel EEG signals," *Multimedia Tools and Applications*, vol. 82, no. 11, pp. 17391–17413, Oct. 2022, doi: 10.1007/s11042-022-13947-0.
- [28] E. D. Pontes, M. Pinto, F. Lopes, and C. Teixeira, "Concept-drifts adaptation for machine learning EEG epilepsy seizure prediction," *Scientific Reports*, vol. 14, no. 1, Apr. 2024, doi: 10.1038/s41598-024-57744-1.
- [29] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review. E, Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 64, no. 6, Nov. 2001, doi: 10.1103/physreve.64.061907.
- [30] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package)," *Neurocomputing*, vol. 307, pp. 72–77, Sep. 2018, doi: 10.1016/j.neucom.2018.03.067.
- [31] T. Hastie, R. Tibshirani, and J. Friedman, "Linear methods for classification," in *Springer series in statistics*, 2008, pp. 101–137. doi: 10.1007/978-0-387-84858-7_4.

- [32] E. Fix and J. L. Hodges, “Discriminatory analysis. Nonparametric discrimination: Consistency properties,” *International Statistical Review*, vol. 57, no. 3, p. 238, Dec. 1989, doi: 10.2307/1403797.
- [33] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, no. 2, pp. 123–140, Aug. 1996, doi: 10.1007/bf00058655.
- [34] T. Hastie, R. Tibshirani, and J. Friedman, “Random forests,” in *Springer series in statistics*, 2008, pp. 587–604. doi: 10.1007/978-0-387-84858-7_15.
- [35] “Multiclass classification problems,” in *The MIT Press eBooks*, 2012, pp. 303–340. doi: 10.7551/mitpress/8291.003.0016.
- [36] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/bf00994018.
- [37] T. Hastie, J. Friedman, and R. Tibshirani, *The elements of statistical learning*. 2001. doi: 10.1007/978-0-387-21606-5.
- [38] PedregosaFabian et al., “SciKit-Learn: Machine Learning in Python,” *Journal of Machine Learning Research*, Nov. 2011, doi: 10.5555/1953048.2078195.
- [39] X. Wang, G. Gong, N. Li, and S. Qiu, “Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization,” *Frontiers in Human Neuroscience*, vol. 13, Feb. 2019, doi: 10.3389/fnhum.2019.00052.
- [40] J. Brownlee, “How to Grid Search Hyperparameters for Deep Learning Models in Python with Keras. <https://machinelearningmastery.com/grid-search-hyperparameters-deep-learning-models-python-keras/> (accessed July. 20, 2024).
- [41] M. Christ, A. W. Kempa-Liehr, and M. Feindt, “Distributed and parallel time series feature extraction for industrial big data applications,” *arXiv.org*, Oct. 25, 2016. <https://arxiv.org/abs/1610.07717>