

Enhancing EEG-based brain-computer interface systems through efficient machine learning classification techniques

Ferdi Ahmed Yassine^{1,2}, Ghazli Abdelkader¹

¹Department of Mathematics and Computer Sciences, Faculty of Exact Sciences, Tahri Mohamed University, Bechar, Algeria

²Laboratory of LTIT, Tahri Mohamed University, Bechar, Algeria

Article Info

Article history:

Received Jan 21, 2024

Revised Feb 24, 2024

Accepted Mar 10, 2024

Keywords:

Attention

Brain-computer interface

Classification

Electroencephalography

Focus

Machine learning

ABSTRACT

Advances in the fields of neuroscience and computer science have greatly enhanced the human brain's ability to communicate and interact with the surrounding environment. In addition, recent steps in machine learning (ML) have increased the use of electroencephalography (EEG)-based BCIs for artificial intelligence (AI) applications. The prevailing challenge in recording EEG sensor data is that the captured signals are mixed with noise, which makes their effective use difficult. Therefore, strengthening the classification stage becomes extremely important and plays a major role in addressing this problem. In this study, we chose five most widely used classification models that obtained the best results in this field and tested them on two open-source databases. We also focused on improving the hyperparameters of each algorithm to obtain best results. Our results indicate excellent results on the first dataset and acceptable for most models on the second, while RF showed superior performance on both with an accuracy of 100% on the first dataset and 86.47% on the second. This was achieved with the lowest training costs, and better performance compared to previous works we evaluated that used the same databases. These results provide valuable insights and advance the development of brain-computer interface (BCI) technology and design.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Ferdi Ahmed Yassine

Department of Mathematics and Computer Sciences, Faculty of Exact Sciences, Tahri Mohamed University

Kenadsa Street, Bechar, Algeria

Email: ferdi.ahmed@univ-bechar.dz

1. INTRODUCTION

A brain-computer interface (BCI) is a technology that captures and interprets an individual's brain signals to perform a desired action. Among the various techniques employed in BCI applications, one of the most widely used is electroencephalography (EEG) [1], [2]. BCI provides a unique opportunity to develop innovative forms of communication technology controlled by the brain, offering significant advantages to individuals with motor impairments [3]. Brain-computer interfaces can be employed to create a range of applications such as brain-controlled prosthetic limbs, adaptive chairs, speech systems, emotion detection, states of focus and attention and more. For instance, interfacing a humanoid robot with this communication system opens up numerous possibilities for replicating human movements, both in terms of physical appearance and the range of motions it can achieve [4], [5]. Several methods can be used to obtain brain signals, including electrocorticography (ECoG), near-infrared spectroscopy (NIRS) and electroencephalography (EEG). The detection and analysis of EEG is referred to as electroencephalography where electroencephalogram (electro=electrical, encephalo=brain, gram=record). An EEG captures the electrical signals generated by brain cells. These signals, also known as local field potentials, are recorded using

electrodes either placed on the scalp or inserted directly into the cortex, referred to as an electrocorticogram. The monitoring of EEG can occur in various contexts, such as in response to stimuli (event-related potential or ERP) or in the absence of any specific stimulus, termed spontaneous EEG [6]. EEG has been a fundamental technique in clinical neurology for many years. Bioelectric potentials are created by the electrochemical activity of excitable cells found in neural, muscular, or glandular tissue [7]. The first observations of bioelectric potentials in the brains of rabbits and monkeys date back to the 1870s, thanks to the work of English physiologist Richard Caton. Meanwhile, the human EEG was first identified in 1924 by German psychiatrist Hans Berger, who even believed he had experienced mental telepathy with his sister during a serious accident hundreds of kilometers away [8]. These voltages are generated by the brain's neuronal activity in response to various external circumstances, events, or stimuli. The examination of EEG rhythms allows for the assessment of shifts in neural activities for clinical diagnosis. EEGs typically exhibit frequencies ranging from 0.5 to 40 Hz and amplitudes between 10 and 200 V [9]. Five distinct EEG rhythms have been identified: delta (0.5-4) Hz, theta (4-8) Hz, alpha (8-13) Hz, beta (13-30) Hz and gamma (over 30 Hz), as shown in Table 1.

Nowadays, many researchers are exploring the integration of deep learning techniques in the world of brain-computer interfaces. However, when it comes to real-world applications, using deep learning requires complex calculations and a deeper understanding of tuning various parameters, including architectural setup and hyperparameters [10]. Therefore, machine learning techniques have become widely used because their results are superior to deep learning techniques. In this study, five machine learning techniques were applied, which are considered common, most widely used, and obtain the best results in this field. support vector machine (SVM), linear discriminant analysis (LDA), k-nearest neighbor (KNN), decision tree (DT) and random forest (RF) were selected for use in this experiment. By comparing the accuracy rates, these five methods are evaluated and compared, with the aim of determining the most effective classifier. Two open-source databases are used. The first contains electroencephalography (EEG) data to detect the state of mental attention, and the second contains brainwave data for the state of concentration for students as shown in Figure 1. In contrast, in most current works, the accuracy is very low, which makes this data not recommended or its use is limited. This is what causes very slow progress in this field. The new results in this work are better than the current works in terms of the accuracy obtained, thanks to the methods of data processing and improving the machine learning algorithms used in this study. The remainder of the paper is structured as follows: the "Previous works" part highlights earlier research in this field, and the "Material and methods" section describes the data and machine learning methods utilized in this study for categorization. After then, the numerical evaluation, a thorough description of the techniques and instruments employed in the inquiry, and the stages of the applied aspect are explained in the "Experimental setup" section The EEG classification, discussion and conclusion are presented in the "Results and Discussion" section. In the "Conclusion" section, we offer a summary at the conclusion.

Table 1. Classification of frequency bands

Brain wave	Frequency (Hz)	Amplitude	Brain states
delta (Δ)	0.5-4	higher	deep sleep, deepest meditation
theta (Θ)	4-8	high	drowsiness, dreaming, deeply relaxed
alpha (α)	8-12	medium	very relaxed, alert, positive attention
beta (β)	13-35	low	active, attentive, judgment, relaxed
gamma (γ)	>35	lower	concentration, integrated thoughts

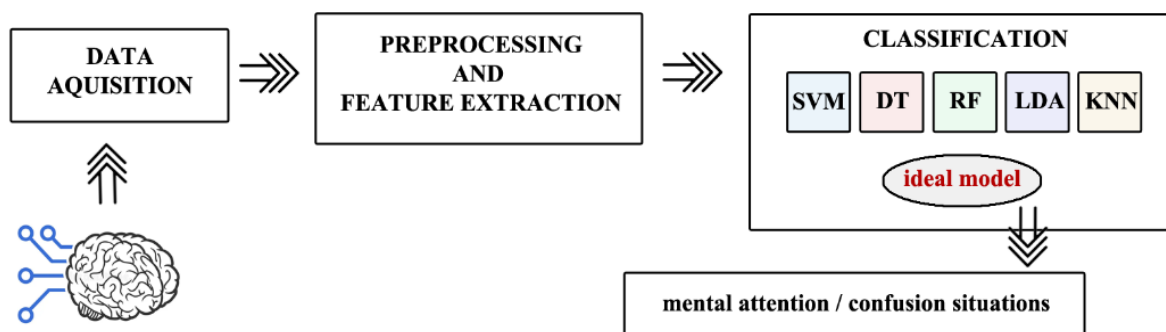


Figure 1. Flowchart of using ML techniques to classify EEG-based mental attention and confusion situations

2. PREVIOUS WORKS

Rani *et al.* [11] claim that when using the same physiological data, support vector machines (SVM) perform the best, featuring an 85.81% classification accuracy, closely followed by a Bayesian network at 74.03%, a regression tree at 83.5%, and a k-nearest neighbor (K-NN) at 75.16%. By utilizing informative information, Bayesian network and K-NN algorithms can perform better. When it comes to physiological signal databases obtained from ten to hundreds of users, SVM exhibits 25% and 33.3% accuracy for three and four emotion categories, respectively. Using Marquardt forward propagation, K-NN, and discriminant function analysis, Nasoz *et al.* [12] were able to discriminate between six emotions with a classification accuracy ranging from 71% to 83%. Conati [13] proposed that probabilistic models may be created using a process that takes into account the user's personality, numerous body expressions, and the setting of the interaction. Artificial neural network (ANN) has been used to assess mental fatigue, and the average classification accuracies for the baseline, low task difficulty, and high difficult task states, respectively, were 85%, 82%, and 86% [14]. Fisher created an emotion-recognizer based on SVMs that had accuracy rates for three, four, and five emotion categories of 78.4%, 61.8%, and 41.7%, respectively [15]. K-NN is one of the most popular strategies for categorizing EEG data linked to certain affective/emotional states, according to a thorough survey conducted by Rani *et al.* [11]. When analyzing EEG data to identify emotion sickness, Yu *et al.* [16] discovered that K-NN was the most successful classifier. K-NN is said to be very effective for classifying EEG data by Bhattacharyya *et al.* because it can handle discriminant analysis of challenging probability densities [17]. In the medical industry, RT is frequently used to classify data like EEG, say Wilson and Russell [14]. Additionally, Brown *et al.* indicate that RT is frequently used to categorize EEG data [18]. Macas *et al.* [19] classified many emotional states using BN successfully. Rani *et al.* [11] fully endorse SVM in their study and suggest using it to correctly identify EEG data. Chen and Hou [20] also lend credence to this assertion. Yu *et al.* [16] and Huang *et al.* [21] experiments show that SVM can categorize EEG data effectively and with promising results. Because ANN can handle noisy data effectively, Chen and Hou [20] suggest that it is a useful approach for classifying EEG data. These five techniques (decision tree (DT), random forest (RF), neural network (MLP-ANN), K-nearest neighbor (KNN) and support vector machine (SVM)) have been found to be used in most experimental experiments.

3. MATERIAL AND METHODS

3.1. Datasets

3.1.1. EEG data for mental attention state detection (Dataset 1)

An original dataset gathered in [22] consisted of 25 hours of EEG recordings from 5 participants engaged in 34 trials. These recordings were utilized to monitor attention states (focused, unfocused, and drowsy) through passive EEG BCI. During a low-intensity control task, participants operated a computer-simulated train using Microsoft Train Simulator. Throughout the trials, the experiment supervisor closely observed participants, ensuring there were no significant disturbances like movement or speech, and recorded the sessions on video. Each participant took part in 7 trials, with a maximum of one trial per day. The initial 2 trials served as habituation, while the subsequent 5 trials were designated for data collection. EEG data was captured utilizing a modified EMOTIV Epoc EEG headset with its classic wet electrodes. This portable EEG acquisition device offered 12 channels of real-time EEG data at a sampling rate of 128 Hz, a voltage resolution of 0.51 V, and a bandwidth of 0.2-43 Hz. The device was connected to a computer via a wireless Bluetooth link. Electrode positions followed the standard 10-20 system: C3, Cz, C4, F3, Fz, F4, T3, T4, T5, T6, and Pz. Data extraction was accomplished through a customized Matlab script developed based on the eeglogger.m sample program. Each MATLAB file contained the data object acquired from the EMOTIV device during a single experiment.

3.1.2. Confused student EEG brainwave data (Dataset 2)

This dataset [23] was generated through a series of exercises involving 10 university students who watched massive open online course (MOOC) videos. The selected videos aimed to be comprehensible to college students, focusing on subjects like basic algebra and geometry. However, intentionally confusing content was also prepared, featuring topics such as quantum mechanics and stem cell research. In total, 20 videos were prepared, with 10 falling into each category. Each video had a duration of approximately 2 minutes, deliberately ending in the middle of a topic to heighten confusion. The students wore a single-channel wireless MindSet device, which measured activity in their frontal lobes. This device recorded the voltage between an electrode on the forehead and two electrodes (one acting as the ground and another as the reference) attached to the ears. After viewing each video, students rated their confusion level on a scale from 1 to 7. These ratings were further normalized to indicate whether the students were confused or not. This self-labeled confusion was used in conjunction with a pre-specified confusion label. Throughout the data collection process, each student watched ten video clips, resulting in a total of 100 data points. Although

there are over 12,000 rows in the dataset, considering each video clip as a single data point, there are over 120 rows sampled every 0.5 seconds within each data point. Notably, EEG data was collected only during the middle 1-minute segment of each 2-minute video, with the first and last 30 seconds removed. The average values of the highest frequency signals were reported over each 0.5-second interval. More details shown in Table 2.

Table 2. Description of databases

	Dataset 1	Dataset 2
Type	Mental states	Confusion situations
Subject	5 subjects	10 students
Time	5 hours for each subject	40 min for each subject
Equipment	EMOTIV EPOC + 14 channels	Mindset NeuroSky
Channel	14 electrode channels	3 electrode channels
Frequency	128 Hz	512 Hz
Classes	3 classes (focused/unfocused/drowsy)	2 classes (attention/mediation)
Files types	MATLAB file	CSV file
Size	584 MB	144

3.2. Machine learning

3.2.1. Linear discriminant analysis

Linear discriminant analysis (LDA) is proficient in resolving both binary and multiclass classification challenges. It employs a linear classifier to allocate variables to specific classes. Due to its linear nature, LDA inherits stringent assumptions regarding linearity and normality [24]. It shares similarities with linear regression and k-means clustering, yet it also differs significantly from both methods. Unlike k-means clustering, which is an unsupervised classification technique, LDA is supervised. This means LDA's classification is trained on known data, making it a supervised learning method. LDA aims to find the best linear function that can effectively separate data points into specific categories or groups. To achieve this, LDA maximizes the variances between group means, ensuring that the differences between the means of different classes are as large as possible. Simultaneously, LDA minimizes the variances within each group, making the data points within the same class as similar to each other as possible. The main goal of LDA is to strike a balance between enhancing inter-class variation and lowering intra-class variance [25].

3.2.2. Support vector machine (SVM)

SVM is a fundamental and crucial technique for classifying various data points. It categorizes these points, also known as support vectors, by creating a hyperplane using the kernel function. There are different types of kernel functions, including radial, radial-integral, polynomial, and linear kernels. A hyperplane is a plane that passes through the centers of the data points and is responsible for properly separating classes within the given dataset. It ensures the largest margin within the area bounded by the hyperplane. The support vectors, belonging to subgroups +1 and -1, are the closest to the dividing hyperplane and the edge of the slab. By correctly identifying support vectors, the margin can be maximized using appropriate methods [26], [27].

3.2.3. K-nearest neighbors (K-NN)

K-NN is a straightforward supervised machine learning algorithm used for both classification and regression tasks. It operates by referencing a database of data points categorized into different classes and attempts to classify a new sample data point provided to it. K-NN is considered non-parametric as it doesn't assume anything about the underlying data distribution. K-NN offers several advantages: it is user-friendly, cost-effective to construct, and adaptable for classes with diverse communication patterns. It can be highly effective in certain scenarios where other methods might fail. However, there are drawbacks to KNN as well. Classifying unclassified records can be expensive, requiring the calculation of distances to the k-nearest neighbors. As the method becomes more computationally intensive with larger training sets, accuracy might decrease, especially when dealing with numerous distracting or irrelevant elements. Additionally, KNN is a slow learner as it computes distances across k-neighbors. It retains all the training data without generalizing, making it less efficient in handling large datasets due to costly calculations. Moreover, higher dimensional data can lead to reduced accuracy in defining regions, making it necessary to carefully consider the choice of features when applying the K-NN algorithm [28].

3.2.4. Decision tree (DT)

DT are hierarchical structures used for classifying instances based on their feature values. Each node in a decision tree represents a feature of the instance to be classified, and each branch corresponds to a possible value the node can have. The classification process starts at the root node, where instances are sorted based on their feature values [29]. In the context of data mining and machine learning, DT learning involves using a decision tree as a predictive model. It maps observations about an item to conclusions about the item's target value. These tree models are also known as classification trees or regression trees [30]. Decision tree classifiers often utilize post-pruning techniques to enhance their performance. These techniques involve evaluating decision trees by using a validation set and removing nodes, assigning them the most common class of the training instances they are associated with [29].

3.2.5. Random forest (RF)

The RF algorithm utilizes the collective strength of multiple DT to make decisions in the field of machine learning [31]. It comprises a set of n decision trees, each generating distinct results for a given input. In this context, the model's output is determined by the majority of outcomes from these n decision trees. Random forest serves as a notable illustration of ensemble learning [32]. It has the capability to address classification and regression (CART) challenges by employing the bootstrap clustering technique, commonly known as bagging [33].

4. EXPERIMENTAL SETUP

To showcase our work, we utilized the Python environment (version 3.11.3) on a HP EliteBook laptop equipped with an 8th generation Intel Core i7 processor, 16 GB of RAM, and running Windows 11 (64-bit). In the initial stages of both preprocessing and feature extraction, we employed power spectral density (PSD), a method used in signal processing and physics to describe the distribution of power over different frequencies in a signal. It provides information about how the power of a signal is distributed across its frequency components. PSD is particularly useful in analyzing signals that vary over time. The PSD is typically calculated for a continuous signal or a discrete signal. For a continuous signal, the PSD $S(f)$ is defined as the Fourier transform of the autocorrelation function $R(t)$ of the signal.

$$S(f) = \int_{-\infty}^{\infty} R(t)e^{-j2\pi ft} dt \quad (1)$$

Here, f represents frequency, $R(t)$ is the autocorrelation function, and t is the time lag. For a discrete signal, the PSD can be estimated using methods such as the periodogram, which is a tool for estimating the spectral density of a signal. The data underwent normalization, a process aimed at reorganizing it to facilitate the application of machine learning algorithms. Data normalization is a technique used in data preprocessing to scale and standardize the features of a dataset. The goal of normalization is to bring the values of different features into a similar range, preventing some features from dominating others in machine learning algorithms that are sensitive to the scale of the input features. This normalization served the dual purpose of eliminating repetitive and disorganized data while ensuring uniformity across all records and fields. In addition to PSD and normalization, the SMOTE technique (synthetic minority oversampling technique) is also an added value in this work. It is a common technique in machine learning to address class imbalance by creating artificial samples for the minority class, and it had a major role and magical effect. In performing algorithms with data and improving results. In Python we use the "imbalanced learning" library, commonly referred to as "non-learning", to implement SMOTE.

Our work was carried out on each database in a separate program, and these steps are common to them: Initially, in our Python program, we imported data files, utilizing the MATLAB format for the first dataset and csv format for the second. Subsequently, we amalgamated all elements from these files into a single table. We proceeded with table level partitioning, defining inputs and outputs, and subsequently fed the data into our machine learning algorithms. During the classification phase for the first database, the classifier was trained to categorize values as 0 (indicating drowsiness), 1 (representing unfocused attention), or 2 (indicating focused attention). For the second database, the classifier discerned values as 0 (Attention ≤ 50) or 1 (Attention > 50). In all experiments, we split the datasets into training sets (80%) and test sets (20%), Table 3 shows the number of cases used for training and testing. The classification process involved a range of classifiers, including LDA, SVM, KNN, DT and RF. Unlike many studies that rely on default algorithm parameters, we fine-tuned and modified these settings multiple times to enhance the performance of each algorithm. Finally, the selection of the most suitable machine learning algorithm was guided by the accuracy measure, with the algorithm demonstrating the highest accuracy being chosen. After selection, we have a ready-made model with an optimal algorithm. Once we've selected and finetuned the optimal model,

we've taken the extra step to export it into Python (.py) format for computer usage. Additionally, we've saved the model in widely used formats such as joblib (.sav) and pickle (.pkl), making it ready for integration into any Android application. This step is driven by the fact that mobile phones are the most ubiquitous communication devices in history [34], and mobile networks enjoy global coverage and are currently the most widely used network type [35]. Moreover, the vast majority of people now possess a mobile phone equipped with internet or network connectivity [36]. By taking this step, we've ensured that our work is highly versatile and can be utilized on a wide range of devices.

Table 3. Distribution of datasets

	Instances	Value
Data 1	Training	16,236 (80%)
	Testing	4,060 (20%)
Data 2	Training	598 (80%)
	Testing	200 (20%)

5. RESULTS AND DISCUSSION

5.1. Accuracy

In our results evaluation, we utilized the accuracy metric, a fundamental measure for evaluating classification models. Accuracy represents the proportion of correct predictions made by the model. Formally, accuracy is calculated (2).

$$Accuracy = \frac{(Total\ Number\ of\ Predictions)}{(Number\ of\ Correct\ Predictions)} \quad (2)$$

In the context of classification, accuracy can also be expressed in terms of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) as (3).

$$Accuracy = \frac{(TN+TP)}{(TP+FP+TN+FN)} \quad (3)$$

Where, TP is instances where the model correctly predicts the positive class, TN is instances where the model correctly predicts the negative class, FP is instances where the model incorrectly predicts the positive class and FN is instances where the model incorrectly predicts the negative class. Accuracy provides a clear and intuitive measure of the model's overall correctness in its predictions. The classification results for all classifiers for each of the two databases used are shown in Table 4.

For the first, the highest accuracy was obtained for RF and DT (100%), followed by KNN (95%), then LDA (78%), then SVM (75.5%). As for the second database, the highest accuracy was obtained using RF (86.47%), followed by SVM (65.09%), then KNN (73.37%), then LDA (65.71%), then DT (78.07%). Figure 2 represents a flow chart for using the machine learning techniques used in this study and comparing the percentage of accuracy between them, so that Figure 2(a) represents an explanation of the results of mental states data, and Figure 2(b) represents confusion situations.

Table 4. Machine learning models' accuracies are typically expressed as percentages

	SVM	DT	RF	KNN	LDA
DATA 1	75.5%	100%	100%	95%	78 %
DATA 2	65.09%	78.07%	86.47%	73.37%	65.71%

5.2. Confusion matrix

The confusion matrix, a summary of the machine learning model's performance on the test data used to predict category scores for input instances, serves as an additional tool to evaluate the effectiveness of our classification model. This matrix displays the proportions of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) generated by the model in the test data. It helps understand the unpredictability of classification model predictions, enabling us to identify the types of errors that have been made. The confusion matrix results for the top algorithms that achieved the best accuracy for both datasets are shown in Figure 3. Whereas the results for the mental states data are shown in Figure 3(a), Figure 3(b), and the result for the confusion states data are shown in Figure 3(c). When discussing the accuracy results, we find that the RF classifier is the best for both databases, as it achieved 100% accuracy for the first

database and 86.47% accuracy for the second database. RF is an ensemble learning method that combines the predictions of multiple decision trees, which often produces robust and accurate results, especially for complex data sets. For the first database, both the RF and DT classifiers achieved perfect accuracy (100%), indicating that these models were able to accurately classify all data points in the dataset. KNN achieved an accuracy of 95%, which indicates that it performed well but may have encountered some difficulties in some cases where the nearest neighbors were not representative of the class. LDA and SVM achieved lower accuracy of 78% and 75.5%, respectively. These results indicate that the data in the first database may not have clear linear separation, making it difficult for linear classifiers such as LDA and SVM to perform the same as tree-based methods. For the second database, RF still performs the best with an accuracy of 86.47%, indicating its robustness across different datasets. SVM, which is known for handling complex decision boundaries, achieved an accuracy of 65.09%. KNN and LDA achieved an accuracy of 73.37% and 65.71%, respectively. The decision tree was the least accurate at 78.07%.

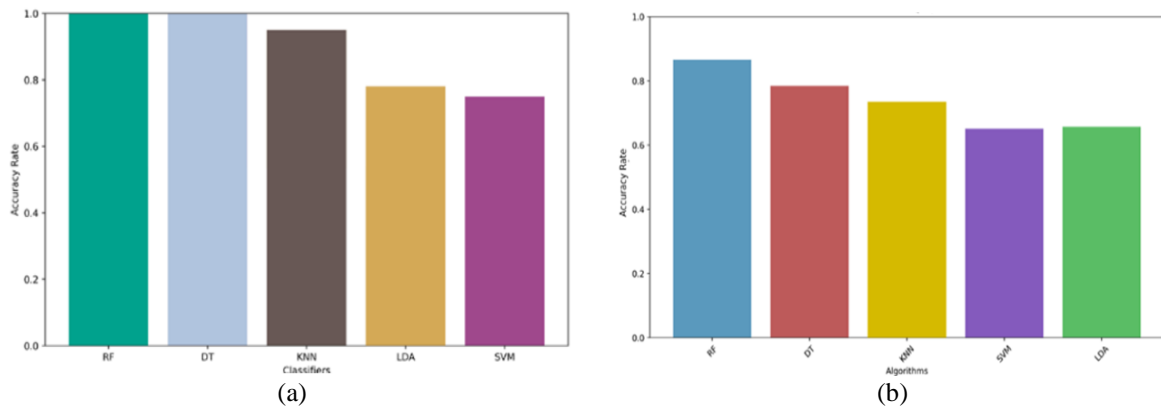


Figure 2. Flowchart of using ML techniques for (a) mental states and (b) confusion situations

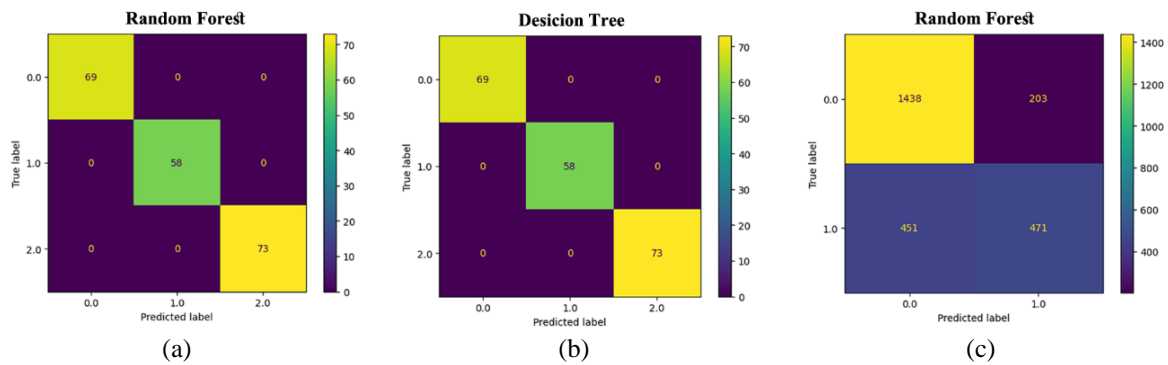


Figure 3. Confusion matrices of the best classification algorithms obtaining the (a) best accuracy, (b) mental states database, and (c) confusion situations database

The lower accuracy of all classifiers in the second database compared to the first database may indicate that the second dataset is more challenging, perhaps due to higher dimensionality, noise, or class imbalance, which may affect the performance of machine learning models. After RF achieved excellent results, it has proven that it is the ideal choice for classification problems of all types (multiple and binary) and that it is highly efficient when it comes to this type of data, which is represented by EEG signals and their various types. Therefore, it was chosen as an ideal classifier that can be relied upon in the classification stage through the BCI to determine states of attention (focused, unfocused, and sleepy) and to determine whether confusion exists or not, for the first database and the second database, respectively. This study investigated the effects of improving the performance of classifiers at the classification stage in EEG-based BCI systems. While previous studies have used machine learning classifiers, they have not exploited classifiers effectively, paid appropriate attention to selecting the optimal classifier, nor have they shown widespread interest in improving the performance of the models they have used. We found that improving

classification results is related to working on pre-normalizing the data and improving the hyperparameters of the machine learning algorithms. The method proposed in this study tends to have an unusually high percentage of “accuracy” similar to that achieved by the RF algorithm and DT. Our study indicates that high performance accuracy of BCI systems is not only associated with poor performance in the stages that precede classification. The proposed method may benefit from the characteristics of the original data without negatively affecting its quality or reducing its values, in contrast to some previous research. This study explored a comprehensive optimization method for brain-computer interface systems based on EEG with the use of efficient machine learning classification algorithms. Despite the positive findings of this paper, more in-depth studies may be needed to ensure that the research steps have positive outcomes for all, or at least most, work.

Our results update this data and make it recommended and of unlimited use. This also gives effectiveness to BCI systems and the possibility of rapid progress in research. Our outcomes encourage others to build on our findings, as the work steps in this study can be exploited on similar EEG data for the purpose of improving performance. Recent observations indicate that focusing on data preparation, such as normalizing it, prior imbalance, and optimizing the hyperparameters of machine learning algorithms, has a very significant impact in improving the results. In comparison with published works that relied on the same databases used in our study, as shown in the Table 5, our obtained results achieved better results, proving the effectiveness and superiority of our model, and this is what we sought in this study. Our final results provide conclusive evidence that this phenomenon is related to the change that occurs before and during the use of machine learning tools, and not due to improvement being limited to only one stage.

Table 5. Comparison with research work using the same datasets

	Work	Classification method	Best accuracy
Data1	[22]	KNN, ANFIS, SVM	91.72
	[37]	KNN	97.5
	[38]	CNN	96.40%
	[39]	SVM, KNN, QDA	95.39%
	[40]	XGBoost, RF, KNN	98%
	[41]	RF, KNN, SVM	96%
	[42]	RF, SVM, XGBoost, Neural Networks	99.9%
	[43]	Optimizable Ensemble	97.8%
Our contribution	SVM, DT, RF, KNN, LDA	100%	
Data 2	[23]	Specific classifiers, independent classifier	67%
	[44]	LSTM	73.3%
	[45]	KNN	73.33%
	[46]	AlexNet, Custom CNN with Dropout	65%
	[47]	SVM	59.1%
	[48]	RF, XGBoost, LightGBM, Catboost	64.75%
	[49]	GTN, RNN, GCN	53.67%
	[50]	fMGTN, GRU, TTNN, GCN	56.10%
Our contribution	SVM, DT, RF, KNN, LDA	86.47%	

6. CONCLUSION

In this study, an important stage was worked on, which is classification, using two databases of different types, original, open source and available to everyone, with different classification (binary and multi-section). In our work, we relied on the latest and best machine learning algorithms most used in this field. In our work, we were keen to improve the performance of each algorithm by changing and modifying the input data for each of them several times until we reached the best. We then evaluated the effectiveness of the developed classifiers by measuring the accuracy percentage and then selecting the best one and displaying it in the confusion matrix. DT, RF, LDA, KNN, and SVM are five classifiers used in this work to classify our data. The RF classifier achieved the best results on both databases, with an accuracy of 100% on the first, and more than 86 % on the second. This makes it recommended as a suitable, effective, and ready-to-use classifier for researchers interested in working on the same databases used in this study. As an idea for subsequent work, we can rely on the idea of amplifying this same data to test the efficiency of deep learning techniques on it, and then modifying the inputs of these algorithms in order to improve them as well.

REFERENCES





- [1] S. Mantri, V. Dukare, S. Yeole, D. Patil, and V. M. Wadhai, “A survey : fundamental of EEG,” *International Journal of Advance Research in Computer Science and Management Studies*, vol. 1, no. 4, pp. 83–89, 2013.
- [2] G. Pfurtscheller *et al.*, “Current trends in Graz brain-computer interface (BCI) research,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 216–219, Jun. 2000, doi: 10.1109/86.847821.

- [3] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, Jan. 2012, doi: 10.3390/s120201211.
- [4] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, "The development of Honda humanoid robot," in *Proceedings. 1998 IEEE International Conference on Robotics and Automation (Cat. No.98CH36146)*, IEEE, pp. 1321–1326. doi: 10.1109/ROBOT.1998.677288.
- [5] R. A. Brooks, C. Breazeal, M. Marjanović, B. Scassellati, and M. M. Williamson, "The cog project: building a humanoid robot," in *International Workshop on Computation for Metaphors, Analogy, and Agents*, 1999, pp. 52–87. doi: 10.1007/3-540-48834-0_5.
- [6] K. Blinowska and P. Durka, "Electroencephalography (EEG)," in *Wiley Encyclopedia of Biomedical Engineering*, Wiley, 2006. doi: 10.1002/9780471740360.ebs0418.
- [7] E. Başar, "Brain oscillations in neuropsychiatric disease," *Dialogues in Clinical Neuroscience*, vol. 15, no. 3, pp. 291–300, Sep. 2013, doi: 10.31887/DCNS.2013.15.3/ebasar.
- [8] D. Millett, "Hans Berger: from psychic energy to the EEG," *Perspectives in Biology and Medicine*, vol. 44, no. 4, pp. 522–542, Sep. 2001, doi: 10.1353/pbm.2001.0070.
- [9] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A new EEG Acquisition protocol for biometric identification using eye blinking signals," *International Journal of Intelligent Systems and Applications*, vol. 7, no. 6, pp. 48–54, May 2015, doi: 10.5815/ijisa.2015.06.05.
- [10] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
- [11] P. Rani, C. Liu, N. Sarkar, and E. Vanman, "An empirical study of machine learning techniques for affect recognition in human–robot interaction," *Pattern Analysis and Applications*, vol. 9, no. 1, pp. 58–69, May 2006, doi: 10.1007/s10044-006-0025-y.
- [12] F. Nasoz, K. Alvarez, C. L. Lisetti, and N. Finkelstein, "Emotion recognition from physiological signals using wireless sensors for presence technologies," *Cognition, Technology and Work*, vol. 6, no. 1, pp. 4–14, Feb. 2004, doi: 10.1007/s10111-003-0143-x.
- [13] C. Conati, "Probabilistic assessment of user's emotions in educational games," *Applied Artificial Intelligence*, vol. 16, no. 7–8, pp. 555–575, Aug. 2002, doi: 10.1080/08839510290030390.
- [14] G. F. Wilson and C. A. Russell, "Real-time assessment of mental workload using psychophysiological measures and artificial neural networks," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 45, no. 4, pp. 635–644, Dec. 2003, doi: 10.1518/hfes.45.4.635.27088.
- [15] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, Sep. 1936, doi: 10.1111/j.1469-1809.1936.tb02137.x.
- [16] Y.-H. Yu, P.-C. Lai, L.-W. Ko, C.-H. Chuang, B.-C. Kuo, and C.-T. Lin, "An EEG-based classification system of Passenger's motion sickness level by using feature extraction/selection technologies," in *The 2010 International Conference on Neural Networks (IJCNN)*, IEEE, Jul. 2010, pp. 1–6. doi: 10.1109/IJCNN.2010.5596739.
- [17] A. Bhattacharyya, R. K. Tripathy, L. Garg, and R. B. Pachori, "A novel multivariate-multiscale approach for computing eeg spectral and temporal complexity for human emotion recognition," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3579–3591, Feb. 2021, doi: 10.1109/JSEN.2020.3027181.
- [18] L. E. Brown, I. Tsamardinos, and C. F. Aliferis, "A novel algorithm for scalable and accurate Bayesian network learning," *Studies in Health Technology and Informatics*, vol. 107, pp. 711–715, 2004, doi: 10.3233/978-1-60750-949-3-711.
- [19] M. Macas, M. Vavrecka, V. Gerla, and L. Lhotska, "Classification of the emotional states based on the EEG signal processing," in *2009 9th International Conference on Information Technology and Applications in Biomedicine*, IEEE, Nov. 2009, pp. 1–4. doi: 10.1109/ITAB.2009.5394429.
- [20] G. Chen and R. Hou, "A new machine double-layer learning method and its application in non-linear time series forecasting," in *2007 International Conference on Mechatronics and Automation*, IEEE, Aug. 2007, pp. 795–799. doi: 10.1109/ICMA.2007.4303646.
- [21] W. Y. Huang, X. Q. Shen, and Q. Wu, "Classify the number of EEG current sources using support vector machines," *Proceedings of 2002 International Conference on Machine Learning and Cybernetics*, vol. 4, 2002, doi: 10.1109/icmlc.2002.1175348.
- [22] Ç. İ. Aci, M. Kaya, and Y. Mishchenko, "Distinguishing mental attention states of humans via an EEG-based passive BCI using machine learning methods," *Expert Systems with Applications*, vol. 134, pp. 153–166, Nov. 2019, doi: 10.1016/j.eswa.2019.05.057.
- [23] H. Wang, Y. Li, X. Hu, Y. Yang, Z. Meng, and K. M. Chang, "Using EEG to improve massive open online courses feedback interaction," *CEUR Workshop Proceedings*, vol. 1009, pp. 59–66, 2013.
- [24] T. C. Nokeri, "Dimension reduction and multivariate analysis using linear discriminant analysis," in *Data Science Revealed*, Berkeley, CA: Apress, 2021, pp. 117–128. doi: 10.1007/978-1-4842-6870-4_6.
- [25] H. Zhou, "Linear discriminant analysis," in *Learn Data Mining Through Excel*, Berkeley, CA: Apress, 2023, pp. 53–70. doi: 10.1007/978-1-4842-9771-1_4.
- [26] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 141–144, Jun. 2003, doi: 10.1109/TNSRE.2003.814441.
- [27] M. R. N. Kousarizi, A. A. Ghanbari, M. Teshnehlab, M. A. Shorehdeli, and A. Gharaviri, "Feature extraction and classification of EEG signals using wavelet transform, SVM and artificial neural networks for brain computer interfaces," in *2009 International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing*, IEEE, 2009, pp. 352–355. doi: 10.1109/IJCBS.2009.100.
- [28] S. Ray, "A quick review of machine learning algorithms," in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, IEEE, Feb. 2019, pp. 35–39. doi: 10.1109/COMITCon.2019.8862451.
- [29] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," *Informatica (Ljubljana)*, vol. 31, no. 3, pp. 249–268, 2007.
- [30] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*, vol. 173, no. 3. in Springer Series in Statistics, vol. 173. New York, NY: Springer New York, 2009, doi: 10.1007/978-0-387-84858-7.
- [31] M. I. Prasetyowati, N. U. Maulidevi, and K. Surendro, "Determining threshold value on information gain feature selection to increase speed and prediction accuracy of random forest," *Journal of Big Data*, vol. 8, no. 1, p. 84, Dec. 2021, doi: 10.1186/s40537-021-00472-4.
- [32] M. G. HR, A. MV, G. P. S, and V. S, "Development of anti-phishing browser based on random forest and rule of extraction framework," *Cybersecurity*, vol. 3, no. 1, Dec. 2020, doi: 10.1186/s42400-020-00059-1.
- [33] T. A. Rather, S. Kumar, and J. A. Khan, "Multi-scale habitat selection and impacts of climate change on the distribution of four sympatric meso-carnivores using random forest algorithm," *Ecological Processes*, vol. 9, no. 1, Dec. 2020, doi: 10.1186/s13717-020-00265-2.
- [34] G. Abdelkader, A. Alipacha, and N. H. Said, "Evolutionary approach to secure mobile telecommunication networks," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 1, Oct. 2021, doi: 10.11591/ijeecs.v24.i1.pp357-366.





- [35] G. Abdelkader, A. Adda, and H. S. Naima, "Chao Authentication and ciphering approach to secure mobile networks," *International Journal of Networks and Communications*, vol. 10, no. 1, pp. 20–32, 2020, doi: 10.5923/j.ijnc.20201001.03.
- [36] G. Abdelkader, H. S. Naima, and A. P. Adda, "Secure authentication approach based new mobility management schemes for mobile communication," *Journal of Information Processing Systems*, vol. 13, no. 1, pp. 152–173, 2017, doi: 10.3745/JIPS.03.0064.
- [37] R. S. Kumar, K. K. Srinivas, A. Peddi, and P. A. H. Vardhini, "Artificial intelligence based human attention detection through brain computer interface for health care monitoring," in *2021 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON)*, IEEE, Dec. 2021, pp. 42–45. doi: 10.1109/BECITHCON54710.2021.9893646.
- [38] D. Zhang, D. Cao, and H. Chen, "Deep learning decoding of mental state in non-invasive brain computer interface," in *Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing*, New York, NY, USA: ACM, Dec. 2019, pp. 1–5. doi: 10.1145/3371425.3371441.
- [39] E. Akman Aydın, "EEG sinyalleri kullanılarak zihinsel iş yükü seviyelerinin sınıflandırılması," *Politeknik Dergisi*, vol. 24, no. 2, pp. 681–689, Jun. 2021, doi: 10.2339/politeknik.794655.
- [40] K. Suwida, S. C. Hidayati, and R. Sarno, "Application of machine learning algorithm for mental state attention classification based on electroencephalogram signals," in *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, IEEE, Feb. 2023, pp. 354–358. doi: 10.1109/ICCoSITE57641.2023.10127825.
- [41] A. Al-Nafjan and M. Aldayel, "Predict students' attention in online learning using EEG data," *Sustainability*, vol. 14, no. 11, May 2022, doi: 10.3390/su14116553.
- [42] Y. Wang, R. Nahon, E. Tartaglione, P. Mozharovskiy, and V.-T. Nguyen, "Optimized preprocessing and Tiny ML for Attention State Classification," in *2023 IEEE Statistical Signal Processing Workshop (SSP)*, IEEE, Jul. 2023, pp. 695–699. doi: 10.1109/SSP53291.2023.10207930.
- [43] S. K. Khare, V. Bajaj, N. B. Gaikwad, and G. R. Sinha, "Ensemble wavelet decomposition-based detection of mental states using electroencephalography signals," *Sensors*, vol. 23, no. 18, p. 7860, Sep. 2023, doi: 10.3390/s23187860.
- [44] Z. Ni, A. C. Yuksel, X. Ni, M. I. Mandel, and L. Xie, "Confused or not confused?," in *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, New York, NY, USA: ACM, Aug. 2017, pp. 241–246. doi: 10.1145/3107411.3107513.
- [45] H. Jamielatuththooah, "Klasifikasi video pembelajaran daring yang membingungkan siswa dengan algoritma k-star nearest neighbor," *MATHunesa: Jurnal Ilmiah Matematika*, vol. 9, no. 1, pp. 22–26, Jan. 2021, doi: 10.26740/mathunesa.v9n1.p22-26.
- [46] I. P. A. E. D. Udayana, M. Sudarma, and N. W. S. Ariyani, "Detecting excessive daytime sleepiness with CNN and commercial grade EEG," *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, vol. 12, no. 3, Nov. 2021, doi: 10.24843/LKJITI.2021.v12.i03.p06.
- [47] D. Ganepola, M. Maduranga, and I. Karunaratne, "Comparison of machine learning optimization techniques for EEG-based confusion emotion recognition," in *2023 IEEE 17th International Conference on Industrial and Information Systems (ICIIS)*, IEEE, Aug. 2023, pp. 341–346. doi: 10.1109/ICIIS58898.2023.10253515.
- [48] M. J. Cheon, D. H. Lee, J. W. Park, H. J. Choi, J. S. Lee, and O. Lee, "CTGAN VS TGAN? Which one is more suitable for generating synthetic EEG data," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 10, 2021.
- [49] Y. L. Xu, K. Konstantinidis, and D. P. Mandic, "Graph tensor networks: An intuitive framework for designing large-scale neural learning systems on multiple domains," *ArXiv (Preprint)*, Mar. 2023.
- [50] Y. L. Xu, K. Konstantinidis, and D. P. Mandic, "Tensor networks for multi-modal non-euclidean data," *ArXiv (Preprint)*, Mar. 2021.

BIOGRAPHIES OF AUTHORS



Ferdi Ahmed Yassine     he is a Ph.D. student in computer science, specializing in artificial intelligence and its applications, at Tahri Mohammed University (UTMB), Algeria. He obtained a Master's Diploma in Engineering in Computer Science, specializing in artificial intelligence and decision making, from Tahri Mohammed University in Algeria in 2019. His research interests focus on the field of artificial intelligence, such as machine learning, deep learning, brain computer interface and data analysis. He can be contacted at email: ferdi.ahmed@univ-bechar.dz.



Ghazli Abdelkader     he is a lecturer at Tahri Mohammed University in Bechar, Algeria. He obtained an engineering diploma in computer science from University of Science and Technology in Oran (USTO), Algeria in 2005. He obtained a teaching diploma in computer science from University of Science and Technology in Oran (USTO), Algeria in 2009. His research interests are artificial intelligence, mobile networks, cyber security, brain computer interface, and embedded systems. He can be contacted at email: ghazek@gmail.com.