Mental Confusion Prediction in e-Learning Contexts with EEG and Machine Learning

No Author Given

No Institute Given

Abstract. In cognitive science, the term confusion is used to capture the decline in learners' cognitive ability, which affects their ability to think, solve a problem, learn and understand. Unlike classroom education, in online e-learning contexts (such as in Massive Open Online Courses (MOOC)), confusion hinders the smooth evolution of the learning process from the learners' side, as the educator can't immediately interact with the students in order to restore cognitive equilibrium. This paper presents a Machine Learning (ML) based approach by comparing several classifiers that were trained and tested exploiting Electroencephalogram (EEG) data (namely, band power, attention and mediation features) acquired by the MindSet device in order to efficiently distinguish "Confused" from "Not-Confused" subjects. In particular, the J48 was the dominant model reaching an optimal performance with accuracy, precision and recall equal to 99.9%, and an Area Under the Curve (AUC) of 100%.

Keywords: E-learning · MOOC · Confusion Detection · EEG· Machine-Learning.

1 Introduction

During the COVID-19 pandemic, classroom education was shifted to online education establishing the latter as an alternative learning mode. The rapid advances in information and communication technology have facilitated the migration of the majority of learning activities to the online mode of education but without deficiencies. Students' mental confusion while watching MOOC videos is among the drawbacks that should be properly handled [6].

Using EEG to quantify the confusion that occurs in the learning process as well as intervening has gained great interest from researchers [15], [4], [8]. Electroencephalogram is a physiological signal that records brain activity in different areas (called lobes) through the scalp. EEG is generally the most effective non-invasive method for assessing a subject's cognitive functions[16]. Wearable EEG-based devices, Artificial Intelligence and Machine Learning have come to support this trend facilitating the collection of enough data for the development of efficient prediction models. Apart from online education, its recognition can be beneficial in various domains such as healthcare, cognitive psychology, virtual online games, etc [8].

Since confusion is a dynamic process, an EEG-based recognition system can help educators quantify and monitor the students' cognitive state (which spans into attention, meditation, concentration, frustration and boredom, level of stress, anxiety, etc), early identify if students feel confused (due to difficulties in solving a problem or understanding the conveying knowledge), and accordingly adapt the teaching plan without the students' feedback or intervention.

Recently, various works [14], [9], [17] have experimented with EEG-based features that are fed as input to ML or Deep Learning (DL) models aiming to reason the mental state of participants (confused or not-confused) in online education platforms.

In the context of this study, we used an EEG-based publicly available dataset whose records are represented by a feature set related to the mean value of raw data, the mean power of five EEG frequency bands at the frontal channel Fp1 and two additional features that capture attention and meditation. The confusion detection was treated as a binary classification problem and solved by investigating three types of classifiers: distance-based, probabilistic and tree-based. Our aim is to find a model with high sensitivity and separation ability of the mental states (confused, not-confused). The main suggestion of this paper is from an ML perspective and, especially, a decision tree-based model for confusion detection. Using 10-fold cross-validation, J48 was the best-performing in all performance metrics achieving accuracy, precision, recall of 99.9% and AUC of 100%.

The remainder of the paper is organized as follows. In Section 2, a description of the dataset and adopted methodology is outlined. Furthermore, in Section 3, we discuss the ML models' performance outcomes. Finally, conclusions and future directions are presented in Section 4.

2 Methodology

In this section, the dataset and its characteristics are illustrated, the adopted methodology is noted, the ML models have been described, as well as the evaluation metrics with which the experimental evaluation was carried out.

2.1 Presentation of Data Collection Process

The dataset was derived from Kaggle [1]. For the data collection, ten students were arranged and assigned to watch ten MOOC video clips dedicating 2 minutes to each one. During this process, the students wore a single-channel wearable MindSet device that measured and recorded the brain's spontaneous electrical activity (over the frontal partial lobe - channel Fp1) for a specific time period. Raw data recorded from the Fp1 channel were used to quantify the mental state of 10 students and recognize the occurrence of confusion. The target class was confirmed and self-labelled by the students as confused or not. This process was repeated in all sessions of watching online videos. Initially, the subjects rated their confusion level on a scale of 1–7 from low to high and then quantized into binary, confused or not confused [13].

By using NeuroSky's API, the raw EEG signals were sampled at 512 Hz. Also, MindSet's proprietary "attention" and "meditation" signals measured (at 1 Hz) mental focus and calmness. Moreover, the MindSet device measures and outputs the average power at five frequency bands, i.e., delta (1–3 Hz), theta (4–7 Hz), alpha (8–11 Hz), beta (12–29 Hz), and gamma (30–100 Hz) [8].

2.2 Data Understanding

To characterize the attention and mediation level of learners, we advised the manual of the MindSet device[1]. More specifically, the eSense, NeuroSky's proprietary algorithm is used for characterizing mental states. First, the NeuroSky ThinkGear technology processes the raw brainwave signal in order to remove the ambient noise and muscle movement. Then, the eSense algorithm interprets eSense meter values, which describe ranges of activity. The meter value is reported on a relative eSense scale of 1 to 100, where the different categories are: i) 1-20: Strongly Reduced, ii) 20-40: Reduced, iii) 40-60: Neutral, iv) 60-80: Slightly Elevated and v) 80-100:Elevated.

Based on the aforementioned scaling, we obtained Figures 1. Interpreting these outcomes, it is observed that attention and mediation levels don't considerably differ among confused and not-confused subjects verifying the complex nature of the specific mental state and the need for diverse features not only from EEG but also from eye tracking [11].

It should be noted that, from this dataset, we excluded subjects with zero attention and meditation. As a result, the final dataset consisted of 5463 Not-confused and 5925 Confused subjects. The age of 10 subjects varied between 24 and 31, with a mean age of 25.54 and a standard deviation of 2.27 years. Further statistical details of the EEG features in the final dataset are captured in Table 1.

2.3 Machine Learning Models and Evaluation Metrics

The evaluation of our ML models was carried out with a widely known free software, namely WEKA [2], which contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. As for the ML methodology, we selected Logistic Regression [5] and Naive Bayes [7], which are probabilistic models. Also, a distance-based classifier and especially k-NN [18] was evaluated. Moreover, three tree-based models were assessed, namely, Random Tree, Random Forest (an ensemble of decision trees where the final prediction is based on voting) [3] and Decision Tree (J48) [12].

In order to evaluate the ML models, we relied on metrics [10] commonly used in the ML field, namely accuracy, precision, recall, and AUC. Note that

[1] https://developer.neurosky.com/docs/doku.php?id=mindset_instruction_manual

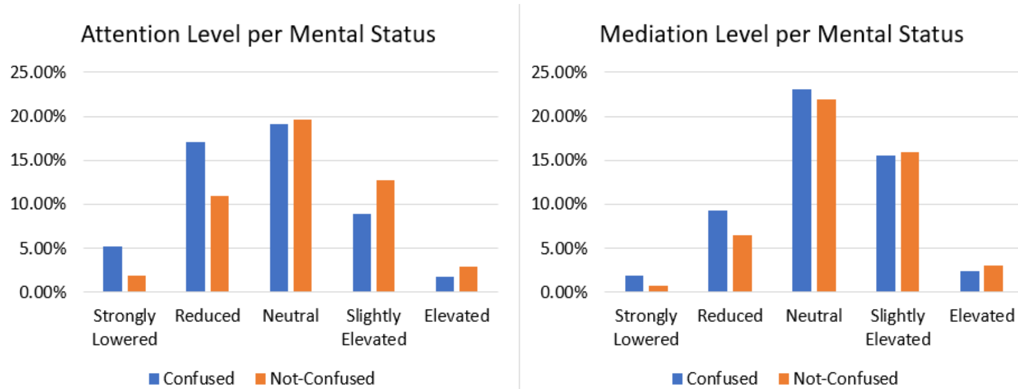

Fig. 1. Attention and Mediation Level as measured by MindSet device.

Table 1. Statistical description of the Dataset.

Features	**Mean±std**	**Min**	**Max**
Attention	46.47±19.06	1	100
Meditation	53.08±16.26	1	100
Raw	34.18±131.55	-2048	1440
delta	588937.42±634988.04	440	3960000
theta	159174.09±236645.72	17	2570000
low-alpha	38520.89±67646.02	2	1370000
high-alpha	28216.66±48541.36	2	1020000
low-beta	20330.76±29239.49	3	841000
high-beta	24283.97±36256.25	2	1080000
low-gamma	16961.54±25819.75	1	658000
mid-gamma	8291.37±11585.23	2	284000

the final score in each metric is derived by averaging the scores from all folds. The definition of these metrics is based on the confusion matrix consisting of the elements true positive(Tp), true negative (Tn), false positive (Fp) and false-negative (Fn). Hence, the aforementioned metrics are defined as follows:

– $\text{Acc} = \frac{\text{Tn+Tp}}{\text{Tn+Fn+Tp+Fp}}$, $\text{Prec} = \frac{\text{Tp}}{\text{Tp+Fp}} + \frac{\text{Tn}}{\text{Tn+Fn}}$, $\text{Recall} = \frac{\text{Tp}}{\text{Tp+Fn}} + \frac{\text{Tn}}{\text{Tn+Fp}}$

To evaluate the distinguishability of a model, the AUC is exploited. It is a metric that varies in [0, 1].

3 Results and Discussion

In this section, our aim is to present and analyze the performance behaviour of the selected ML models. As a first approach, the models were trained and tested with all available features (apart from demographics) in the dataset.

Table 2. Performance Results of ML Models.

	Accuracy%	**Recall%**	**Precision%**	**AUC%**
Naive Bayes	54.76	54.8	59.5	65.0
Logistic Regression	61.47	61.5	61.4	65.4
Random Tree	83.39	83.4	83.4	83.3
Random Forest	93.79	93.8	94	98.6
1-NN	95.64	95.6	95.6	95.6
J48	99.90	99.9	99.9	100

In Table 2, we demonstrate the average values of performance metrics, which were acquired assuming 10-fold cross-validation. Note that for the k-NN classifier, we experimented with the parameter k and verified that for k=1 the highest performance was achieved. Comparing the assessed models, we see that the 1-NN classifier outperformed the Random Tree and Random Forest in terms of accuracy, recall and precision, while Random Forest indicated higher AUC. The higher AUC of Random Forest reveals that for the specific model, there is a 98.6% chance that the model will be able to distinguish between confused class and not-confused class. Also, Logistic Regression and Naive Bayes were the least efficient models in identifying confused subjects. For the specific dataset, J48 was the most efficient model for confused/non-confused students' prediction as this specific classifier managed to keep the Fp, Fn at the lowest level (just a few subjects were misclassified with Fp = 3 and Fn = 7).

4 Conclusions

In the content of this study, we relied on an EEG-based publicly available dataset which helped us to identify a robust and powerful ML model for confused subjects detection. A limitation of the specific dataset is that we don't have access to the whole time series data from the Fp1 channel in order to extract several other EEG features in the time, frequency or time-frequency domain. However, it is adequate to train efficient models for the prediction of the human mental state of subjects who attended lectures in an online education environment. From our analysis, J48 was the prevailing model with accuracy, precision and recall equal to 99.9% and AUC of 100%.

In future work, we anticipate applying various feature selection techniques in order to understand the significance and correlation of the features to the specific task and reevaluate the ML models' performance. Also, our study will focus on the design of personalized confusion detection models which will be compared with the current global models. Finally, our research on confusion detection in online education platforms will be directed to eye-tracking-based datasets [11].

References

1. Confused student eeg brainwave data. `https://www.kaggle.com/datasets/wanghaohan/confused-eeg?select=EEG_data.csv` ((accessed on 12 May 2023))

2. Weka. https://www.weka.io/ ((accessed on 12 May 2023))
3. Alexiou, S., Dritsas, E., Kocsis, O., Moustakas, K., Fakotakis, N.: An approach for personalized continuous glucose prediction with regression trees. In: 2021 6th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM). pp. 1–6. IEEE (2021)
4. Balamurugan, B., Mullai, M., Soundararajan, S., Selvakanmani, S., Arun, D.: Brain–computer interface for assessment of mental efforts in e-learning using the nonmarkovian queueing model. Computer Applications in Engineering Education **29**(2), 394–410 (2021)
5. Boateng, E.Y., Abaye, D.A.: A review of the logistic regression model with emphasis on medical research. Journal of data analysis and information processing **7**(4), 190–207 (2019)
6. Chakraborty, P., Mittal, P., Gupta, M.S., Yadav, S., Arora, A.: Opinion of students on online education during the covid-19 pandemic. Human Behavior and Emerging Technologies **3**(3), 357–365 (2021)
7. Chen, S., Webb, G.I., Liu, L., Ma, X.: A novel selective naïve bayes algorithm. Knowledge-Based Systems **192**, 105361 (2020)
8. Daghriri, T., Rustam, F., Aljedaani, W., Bashiri, A.H., Ashraf, I.: Electroencephalogram signals for detecting confused students in online education platforms with probability-based features. Electronics **11**(18), 2855 (2022)
9. He, S., Xu, Y., Zhong, L.: Eeg-based confusion recognition using different machine learning methods. In: 2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE). pp. 826–831. IEEE (2021)
10. Hossin, M., Sulaiman, M.N.: A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process **5**(2), 1 (2015)
11. Khan, A.R., Khosravi, S., Hussain, S., Ghannam, R., Zoha, A., Imran, M.A.: Execute: Exploring eye tracking to support e-learning. In: 2022 IEEE Global Engineering Education Conference (EDUCON). pp. 670–676. IEEE (2022)
12. Maulana, M.F., Defriani, M.: Logistic model tree and decision tree j48 algorithms for predicting the length of study period. PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic **8**(1), 39–48 (2020)
13. Ni, Z., Yuksel, A.C., Ni, X., Mandel, M.I., Xie, L.: Confused or not confused? disentangling brain activity from eeg data using bidirectional lstm recurrent neural networks. In: Proceedings of the 8th acm international conference on bioinformatics, computational biology, and health informatics. pp. 241–246 (2017)
14. Reñosa, C.R.M., Bandala, A.A., Vicerra, R.R.P.: Classification of confusion level using eeg data and artificial neural networks. In: 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). pp. 1–6. IEEE (2019)
15. Tahmassebi, A., Gandomi, A.H., Meyer-Baese, A.: An evolutionary online framework for mooc performance using eeg data. In: 2018 IEEE Congress on Evolutionary Computation (CEC). pp. 1–8. IEEE (2018)
16. Trigka, M., Dritsas, E., Fidas, C.: A survey on signal processing methods for eeg-based brain computer interface systems. In: Proceedings of the 26th Pan-Hellenic Conference on Informatics. pp. 213–218 (2022)
17. Xu, T., Wang, J., Zhang, G., Zhang, L., Zhou, Y.: Confused or not: decoding brain activity and recognizing confusion in reasoning learning using eeg. Journal of Neural Engineering **20**(2), 026018 (2023)
18. Zhang, Z.: Introduction to machine learning: k-nearest neighbors. Annals of translational medicine **4**(11) (2016)