International Journal of Artificial Intelligence and Machine Learning (IJAIML)
Volume 3, Issue 2, July-Dec 2024, pp. 204-210. Article ID: IJAIML_03_02_016
Available online at https://iaeme.com/Home/issue/IJAIML?Volume=3&Issue=2
Impact Factor (2024): 3.56 (Based on Google Scholar Citation)
Journal ID: 9339-1263, DOI: https://doi.org/10.5281/zenodo.14010606

EVALUATING ENSEMBLING METHODS FOR EYE STATE CLASSIFICATION

Rahul Kavi
Independent Researcher, USA

Jeevan Anne
Independent Researcher, USA

ABSTRACT

Eye state tracking is one of the well-researched problems. EEG is a frequently used sensor modality to capture the state of human cognition. This is also a widely studied problem in thefield of Computer Vision. Tracking eye movements with the help of a camera may be obtrusive (camera placement may only sometimes be conducive in each environment). In such a scenario, approaches like EEG measurement are preferred to tracking eye movements using a camera. Since this is a non-intrusive approach (and painless), this is preferable. In this work, we explore a supervised classification approach to identifying eye states (open or closed). We compare using widely used approaches such as LightGBM, Random Forests (Decision Trees), and XGBoost. These three classification approaches use ensemble-based techniques to aggregate decisions (sequential tree building, bagging, and boosting). These ensemble methods are preferred over other classifiers as they aggregate decisions over several classifiers and improve generalization. There are over 14,000 samples in this dataset (EEG Eye State Dataset). This is relatively small. Hence, our approaches use simple algorithms (as Deep learning-based algorithms usually require large training datasets). We use k-fold cross-validation to evaluate our results over several folds. This way, we ensure that performance is generalized and not dependent on a specifically chosen train or test set.

Keywords: Classification, Decision Trees, Electroencephalography, Ensemble Classification,Random Forests

Cite this Article: Kavi, R., & Anne, J. (2024). Evaluating ensembling methods for eye state classification. International Journal of Artificial Intelligence & Machine Learning, 3(2), 204–210.
https://iaeme.com/MasterAdmin/Journal_uploads/IJAIML/VOLUME_3_ISSUE_2/IJAIML_03_02_016.pdf

I. INTRODUCTION

Some professions demand extended and odd working hours. This is sometimes a job requirement. Workers in these professions have a fatigue problem affecting their attentiveness and ability to focus. This may lead to accidents at work. There is a lot of research that indicates that EEG monitors are used to measure worker attentiveness to improve occupational safety (drivers, etc.) [1][2]. Monitoring workers in these environments is a challenging problem. Hence, Computer vision-based approaches and EEG sensors often help in such settings. Monitoring personnel/workers using cameras is a vast area of research.

However, this usually places strict requirements on the placement of cameras, which may only sometimes be feasible or acceptable for the worker. This also raises concerns about worker privacy, etc. In such settings, EEG sensors are often preferred as they are non- invasive, and workers may not have reservations about these sensors compared to cameras. Using EEG signals to measure fatigue has been extensively studied [3][4]. When the user's eyes are open or closed, there is usually a change in alpha wave activity [5]. Well-established research indicates that there is a correlation between attentiveness and EEG activity. This information can be leveraged to check whether the user is attentive or has closed their eyes. Incorrect classification may lead to decreased accuracy or degraded user attentiveness estimation. There is a lot of research on EEG-based eye state classification.

Most focus on building classification approaches using Decision Trees, Random Forests, Support Vector Machines, Multi-Layer-Perceptron or 1D Convolution Neural Networks (and other deep learning-based approaches). Modern approaches, such as deep learning-based approaches, can lead to overfitting. This scenario must be avoided to generalize the classifier to several users. However, this work mainly uses Ensemble-based approaches to showcase which techniques best classify a user’s eye state (open or closed). The methods chosen in this work rely on several weak classifiers (LightGBM, XGBoost, and Random Forest) to combine decisions from several classifiers to make a strong classification technique. Our approach to solving Eye state classification is based on the classification approach. Once the eye state is determined (open or closed), this information can be further used to map brain activity to the user's attentiveness (or if a user is sleeping or attentive). For the sake of simplicity, we only focus on building a classifier for determining eye state.

II. RELATED WORK

EEG signal-based activity monitoring has been extensively studied. Most techniques include statistical approaches, such as SVM-based, Neural Network-based, and Deep Learning-based classification techniques. Electroencephalogram (**EEG**) signals provide neuron activities in the form of electricity. Even during sleep, these signals from brain cells are active. EEG- based analysis has been explored with applications in the following domains: sleep state analysis, sleep disorder detection, sleep stage analysis, fatigue detection, and blink rate calculation. EEG signals have been covered thoroughly in the literature survey. Research work presented in [6] lists EEG-based BCI (Brain Computer Interface) in their study (including a list of features, classification, system types, etc.). Subha. Et al. [7] lists detailed information about EEG signals (including their non-linear, non-stationary nature) and how they can be used to observe human mental states and diseases. Jetoi et al. [8] discussed EEG source localization and how that provides physiological and functional abnormalities. Gu et al. [9] present a survey of BCI interfaces, their use for continuous monitoring, and the field's future direction. This makes EEG a beneficial modality to measure sleep disorders. This is presented in Derex's work [10], where they study sleep disorders using a statistical approach.Sharma et al. [11] present an approach to score sleep stages in over 80 subjects. Behzad et al.

[12] presented an approach to detect sleep disorders using a clinical EEG band and how signals from an EEG headband can be processed using statistical approaches. Aboalayon et al. [13] presented a multi-class SVM-based approach using EEG signals to classify different sleep stages. Similarly, in [14], modern EEG sleep signal analysis approaches to analyze sleep stages have been explored. This includes several preprocessing, classification, and analysis of data collected with EEG devices for sleep stage analysis. Diykh. et al. [15] explored approaches such as time domain features and K-means clustering. These approachesare also compared to SVM-like classification algorithms. EEG signals can also be used to measure blinks in a human. Chang. et al. [16] explored eye blink artifact detection with a single-channel EEG. Soomro. et al. [17] proposed an ICA-based approach to detect and remove eye blink artifacts from EEG signals. Lenskiy. et al. [18] present an approach to analyzing blink rate variability while reading & resting. EEG signals have been used in driverdrowsiness detection as driver fatigue detection is a widely studied problem. A thorough review of EEG signal features with applications in Drowsiness detection in work presented by Stancin. et al [19] and Hussein. et al. [20]. Sheykhivand. et al. [21] explored Deep Neural Network based approaches. VGGNet, AlexNet & LSTM-based approaches have been employed by Budak. et al. [22]. Rundo. et al. [23] also presented an approach based on the Discrete Cosine Transform of EEG signals combined with deep learning approaches to detect fatigue. Our previous work has explored driver fatigue detection in coal mine safety [2].

III. OUR APPROACH

Oliver Roesler [24] collected this open-source dataset using an Emotive EEG Neuro Headset. A state of '1' indicates eye-closed and '0' is eye-open. The eye state was detected using a camera. The data was collected with frontal, central, and temporal lobes sensors, etc. The datais verified to have no missing values. We have used algorithms such as Random Forests, Gradient Boosted trees (XGBoost), and LightGBM. In this section, we explore the data, algorithms, their history (of usage in the EEG domain), their application in our domain, and results. This dataset contains 14 channels of EEG data. The distribution of the entire dataset data is shown in Figure 1. The correlation matrix is added in Figure 2. There is some positive correlation and a negative correlation between certain variables. We have explored the robust scaling approach available in sklearn library [25] to normalize the data. Figure 1 showcases the distribution of the dataset.

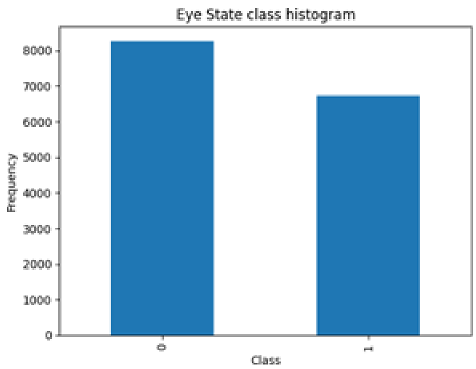

Figure 1: Distribution of Data

Random Forest classifier is widely used in supervised classification, pattern recognition tasks, etc. It combines several decision trees and aggregates the decision. Random Forests area natural fit for classification problems in the EEG domain. A Random Forest is a bagging ensemble classifier approach combining decisions from several training data subsets. The predictions from individual classifiers are then combined to create a more robust classifier.

Our approach includes using 100 estimator decision trees with gini impurity (to measure thedataset's quality). Every constructed tree is expanded until all leaves are pure. **XGBoost** (ExtremeGradient Boosting) classifier is a popular off-the-shelf classifier for several classification problems. XGBoost builds an ensemble of decision trees sequentially. Newly added trees focus on the residual error from the results of the previously added trees (boosting). **LightGBM** (Light Gradient Boosting Machines) uses tree-based algorithms. It combines several decision trees using a boosting approach. Gradient descent is used to optimize the loss function.

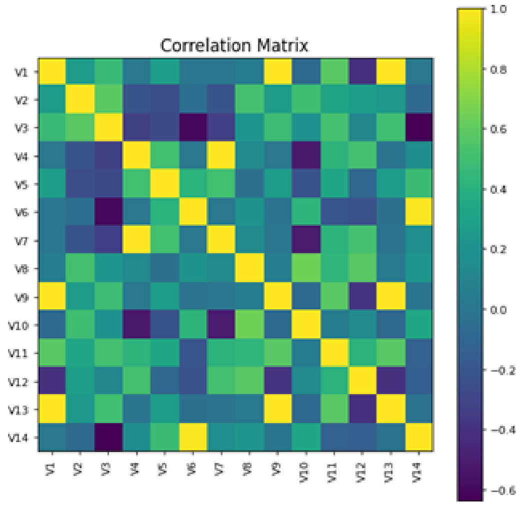

Figure 2: Correlation Matrix

In Figure 2, the correlation matrix summarizing the eye-state dataset is shown. We can see that there are 14 channels and some correlation between channels. Figure 3 shows the feature importance per channel based on F-Score. This shows that some channels are more important than others.

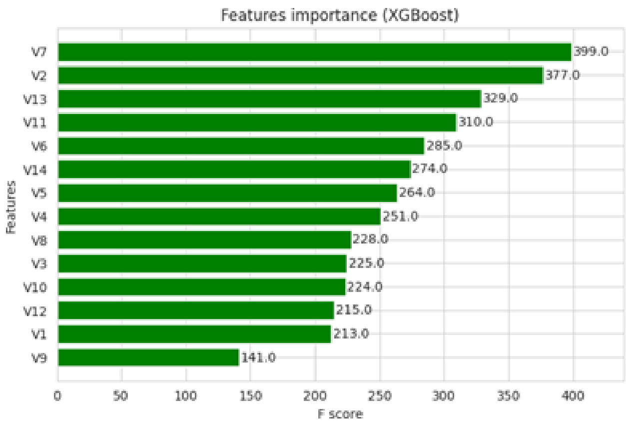

Figure 3: Feature Importance

IV. RESULTS AND CONCLUSION

In this section, we discuss the results and conclusion. Since the dataset is considerably small, we used K-Fold cross-validation (K=5). This helps us generalize the results in a small dataset. The reported Precision, Recall and F1-score are weighted. We trained the classifiers with K=5 (using a train-test split of 80:20). Table 1 presents the results from the Random-Forest classifier. We can see that on the train set, it performs perfectly over all folds (in train sets).

However, the test set sees a drop in performance to 0.92 (for precision, recall and f1). In Table 2, we see the results from the LightGBM classifier. The performance is slightly loweron the train set (compared to Random Forest). However, the test set is even lower (at 0.90).

When we compare these results with XGBoost, we see a much more balanced performance (but close) compared to Random Forest and LightGBM models.

Split	Precision	Recall	F1-Score
Train	1.0000	1.0000	1.0000
Test	0.9229	0.9224	0.9221

Table 1: Random-Forest classifier performance

Split	Precision	Recall	F1-Score
Train	0.9610	0.9608	0.9608
Test	0.9042	0.9040	0.9038

Table 2: LightGBM classifier performance

Split	Precision	Recall	F1-Score
Train	0.9960	0.9960	0.9960
Test	0.9246	0.9245	0.9244

Table 3: XGBoost classifier performance

REFERENCES

[1] Ben Coxworth. SmartCap monitors workers' fatigue levels by reading their brain waves.New Atlas.

[2] Rohit, F., Kulathumani, V., Kavi, R., Elwarfalli, I., Kecojevic, V., & Nimbarte, A. (2017).Real-time drowsiness detection using wearable, lightweight brain sensing headbands. IET Intelligent Transport Systems, 11(5), 255-263.

[3] Bongers, D. (2009). Operator fatigue detection and countermeasure development.

[4] Rogado, E., Garcia, J. L., Barea, R., Bergasa, L. M., & López, E. (2009, February). Driverfatigue detection system. In 2008 IEEE International Conference on Robotics and Biomimetics (pp. 1105-1110). IEEE.

[5] Barry, R. J., Clarke, A. R., Johnstone, S. J., Magee, C. A., & Rushby, J. A. (2007). EEG differences between eyes-closed and eyes-open resting conditions. Clinical neurophysiology, 118(12), 2765-2773.

[6] Hwang, H. J., Kim, S., Choi, S., & Im, C. H. (2013). EEG-based brain-computerinterfaces: a thorough literature survey. International Journal of Human-Computer Interaction, 29(12), 814-826.

[7] Subha, D. P., Joseph, P. K., Acharya U, R., & Lim, C. M. (2010). EEG signal analysis: asurvey. Journal of medical systems, 34, 195-212.

[8] Jatoi, M. A., Kamel, N., Malik, A. S., Faye, I., & Begum, T. (2014). A survey of methodsused for source localization using EEG signals. Biomedical Signal Processing and Control, 11, 42-52.

[9] Gu, X., Cao, Z., Jolfaei, A., Xu, P., Wu, D., Jung, T., & Lin, C. (2020). EEG-based Brain-Computer Interfaces (BCIs): A Survey of Recent Studies on Signal Sensing Technologies and Computational Intelligence Approaches and their Applications. ArXiv. https://arxiv.org/abs/2001.11337

[10] Peter-Derex, L., Berthomier, C., Taillard, J., Berthomier, P., Bouet, R., Mattout, J., ... &Bastuji, H. (2021). Automatic analysis of single-channel sleep EEG in a large spectrum of sleep disorders. Journal of Clinical Sleep Medicine, 17(3), 393-402.

[11] Sharma M, Tiwari J, Acharya UR. Automatic Sleep-Stage Scoring in Healthy and SleepDisorder Patients Using Optimal Wavelet Filter Bank Technique with EEG Signals. International Journal of Environmental Research and Public Health. 2021; 18(6):3087. https://doi.org/10.3390/ijerph18063087

[12] Behzad, R., & Behzad, A. (2021). The role of EEG in the diagnosis and management ofpatients with sleep disorders. Journal of Behavioral and Brain Science, 11(10), 257-266.

[13] Aboalayon, K. A., & Faezipour, M. (2014, October). Multi-class SVM based on sleepstage identification using EEG signal. In 2014 IEEE healthcare innovation conference (HIC) (pp. 181-184). IEEE.

[14] Aboalayon, K. A., Faezipour, M., Almuhammadi, W. S., & Moslehpour, S. (2016). SleepStage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation. Entropy, 18(9), 272. https://doi.org/10.3390/e18090272

[15] Diykh, M., Li, Y., & Wen, P. (2016). EEG sleep stages classification based on time domain features and structural graph similarity. IEEE Transactions on Neural Systems andRehabilitation Engineering, 24(11), 1159-1168.

[16] Chang WD, Cha HS, Kim K, Im CH. Detection of eye blink artifacts from single prefrontal channel electroencephalogram. Comput Methods Programs Biomed. 2016;124:19- 30. doi:10.1016/j.cmpb.2015.10.011

[17] Soomro, M. H., Badruddin, N., Yusoff, M. Z., & Malik, A. S. (2013, March). A method for automatic removal of eye blink artifacts from EEG based on EMD-ICA. In 2013 IEEE 9th international colloquium on signal processing and its applications (pp. 129-134). IEEE.

[18] Lenskiy, A., & Paprocki, R. (2016, July). Blink rate variability during resting and reading sessions. In 2016 IEEE Conference on Norbert Wiener in the 21st Century (21CW)(pp. 1-6). IEEE.

[19] Stancin, I., Cifrek, M., & Jovic, A. (2021). A review of EEG signal features and theirapplication in driver drowsiness detection systems. Sensors, 21(11), 3786.

[20] Hussein, M. K., Salman, T. M., Miry, A. H., & Subhi, M. A. (2021, April). Driver drowsiness detection techniques: A survey. In 2021 1st Babylon International Conference onInformation Technology and Science *(BICITS)* (pp. 45-51). IEEE.

[21] Sheykhivand S, Rezaii TY, Meshgini S, Makoui S, Farzamnia A. Developing a DeepNeural Network for Driver Fatigue Detection Using EEG Signals Based on Compressed Sensing. Sustainability. 2022; 14(5):2941. https://doi.org/10.3390/su14052941

[22] Budak, U., Bajaj, V., Akbulut, Y., Atila, O., & Sengur, A. (2019). An effective hybridmodel for EEG-based drowsiness detection. IEEE sensors journal, 19(17), 7624-7631.

[23] Rundo F, Rinella S, Massimino S, Coco M, Fallica G, Parenti R, Conoci S, Perciavalle V. An Innovative Deep Learning Algorithm for Drowsiness Detection from EEG Signal. Computation. 2019; 7(1):13. https://doi.org/10.3390/computation7010013

[24] EEG Eye State. https://archive.ics.uci.edu/dataset/264/eeg+eye+state

[25] Kramer, O., & Kramer, O. (2016). Scikit-learn. Machine learning for evolutionstrategies, 45-53.

Citation: Kavi, R., & Anne, J. (2024). Evaluating ensembling methods for eye state classification. International Journal of Artificial Intelligence & Machine Learning, 3(2), 204–210

Article Link:
https://iaeme.com/MasterAdmin/Journal_uploads/IJAIML/VOLUME_3_ISSUE_2/IJAIML_03_02_016.pdf

Abstract:
https://iaeme.com/Home/article_id/IJAIML_03_02_016

✉ **editor@iaeme.com**