Brain Computer Interface for Eye State Classification using Tree Based Algorithm

Ayushma Joshi | Prof. Abdul Gaffar H| SCOPE

Introduction

BCI has low recognition accuracy and calibration time, especially for motor imagery tasks and motion control tasks. We have presented an end-to-end supervised learning framework for EEG classification and EEGEMG fusion analysis, which efficiently extracts distinctive feature embedding from unlabeled EEG samples and improves the performance of hybrid BCI.

Motivation

One of the primary sources used to build Brain-Computer Interface (BCI) technology is EEG signal. The BCI is a non-muscle communication between the brain and an external device that is often meant to allow people with neurological conditions to communicate with others using their brain signals. This inspires us.

SCOPE of the Project

- To evaluate the performance of a combination of classification algorithms to predict the state of the human eye based on brain signals.
- To develop a classification method which is more robust and higher performance.
- To provide accurate eye condition differentiation.
- To determine if there has been a change in the status of the eyes.

Methodology

- Data Preprocessing: To start, the code goes through a number of data preprocessing operations on the EEG data. It comprises signal processing methods including bandpass filtering and Hilbert transformation as well as interpolation, independent component analysis (ICA), and filtering outliers using z-scores.

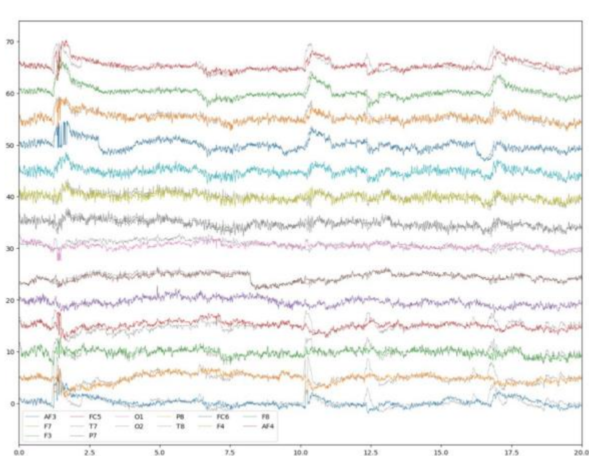

Fig: Independent Component Analysis

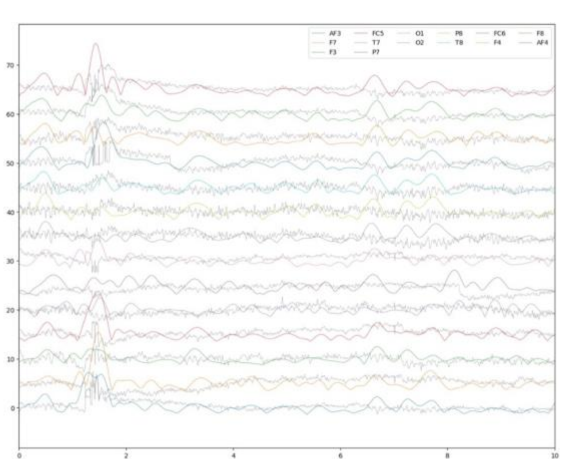

Fig: Bandpass Filter and Hilbert Transform

- Exploratory Data Analysis: To understand the connections between variables and find correlations, the algorithm then visualizes the data using line plots and correlation heatmaps.
- Feature Selection: Recursive Feature Elimination with Cross-Validation (RFECV) is a technique the model uses to choose the most crucial characteristics for categorization.

```
Selected Features:
Index(['AF3', 'F7', 'F3', 'FC5', 'T7', 'P7', 'O1', 'O2', 'T8', 'FC6', 'F4',
       'F8', 'AF4'],
      dtype='object')
```

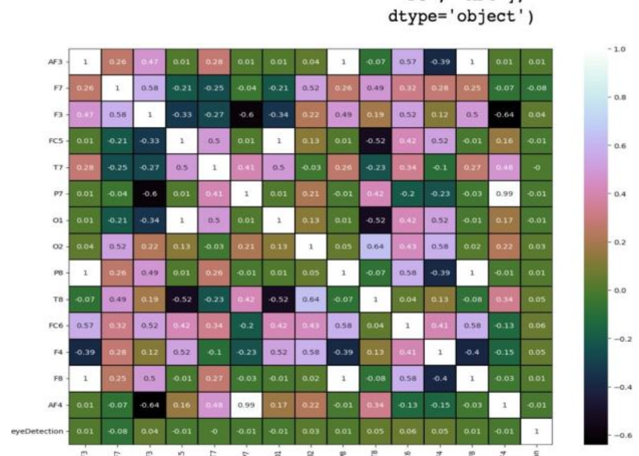

Fig: Correlation Matrix(Heatmap)

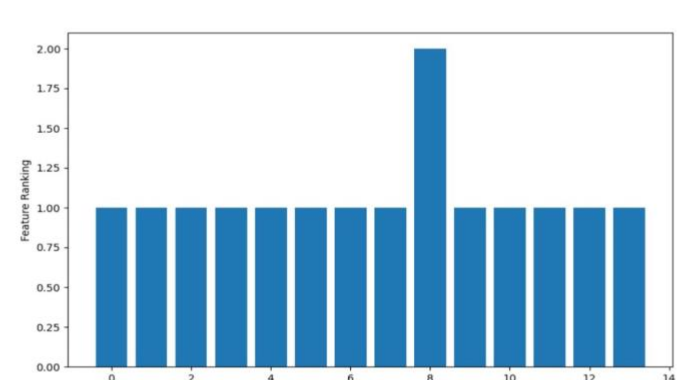

Fig: Feature ranking and index

- Model Training and Evaluation: Using the chosen features, the code divides the data into training and testing sets, trains a Random Forest Classifier, and assesses the model's performance using the classification report, confusion matrix, and accuracy metrics.

	precision	recall	f1-score	support
0	0.91	0.96	0.93	2642
1	0.95	0.89	0.92	2302
accuracy			0.93	4944
macro avg	0.93	0.92	0.93	4944
weighted avg	0.93	0.93	0.93	4944

Accuracy: 92.64%

Fig:Random Forest classification report

```
print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0	0.76	0.81	0.79	2642
1	0.77	0.70	0.73	2302
accuracy			0.76	4944
macro avg	0.76	0.76	0.76	4944
weighted avg	0.76	0.76	0.76	4944

Fig: Adaboost classification report

- Hyperparameter Tuning: To determine the best set of hyperparameters for the Random Forest Classifier, the code uses GridSearchCV and BayesSearchCV.

- Final Model Evaluation: Using the accuracy and confusion matrices, the code assesses how well the top model produced via hyperparameter tuning performs on the test set.

Results

In this section, we provide the findings of the eye state detection using an EEG dataset and discuss them. Using the provided EEG data, the evaluation was carried out, and several analysis approaches were used to identify the eye state. The quantitative analysis of the EEG data for eye state detection yielded significant findings:

- Algorithms Computed: We computed two algorithms Random Forest and AdaBoost which gave us the following confusion matrix:

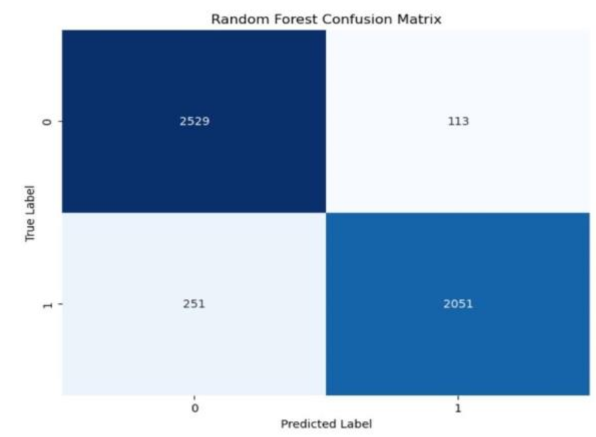

Fig: Random Forest Confusion Matrix

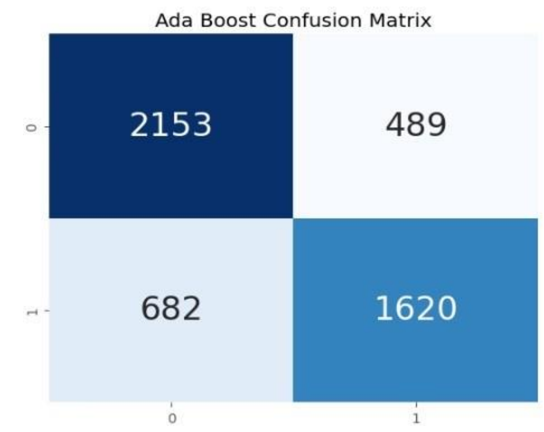

Fig: AdaBoost Confusion Matrix

Finally, the trained Random Forest classifier has a higher accuracy than the Adaptive Boosting classifier, achieving a 92.64% accuracy, when the model outputs the features chosen by the RFECV method. Random forest uses a large number of decision trees to combine their predictions, which helps to lessen overfitting and improve model accuracy . In contrast, a single decision tree(in this case AdaBoost), especially a deep one, may be prone to overfitting the training set. The test data's accuracy may suffer as a result.

- Hyperparameter Optimization: The number of trees in the forest (n_estimators), the depth of each tree to its maximum (max_depth), the minimum number of samples needed to split an internal node (min_samples_split), and the minimum number of samples needed to be at a leaf node (min_samples_leaf) are the hyperparameters that are being optimized. The Scikit-Learn library's GridSearchCV and BayesSearchCV functions are utilized in the given code for hyperparameter optimization

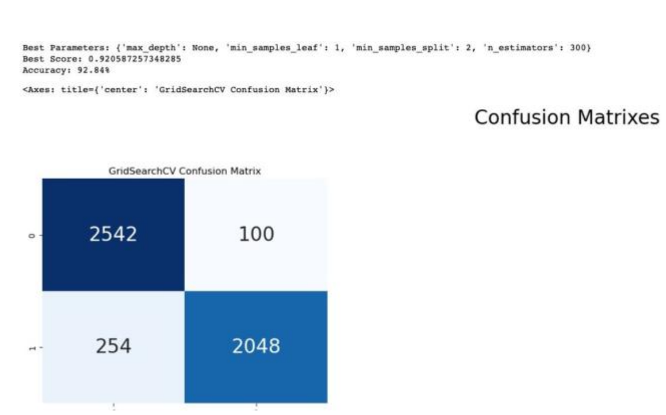

Fig: GrivSearchCV Confusion Matrix

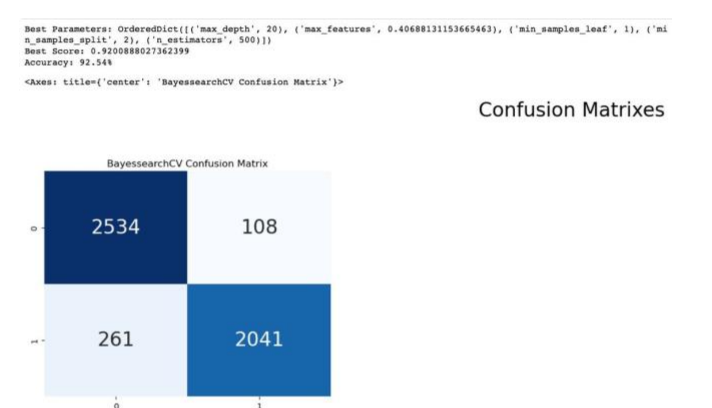

Fig: BayesSearchCV Confusion Matrix

The code uses GridSearchCV and BayesSearchCV to regulate the Random Forest classifier's hyperparameters. It is helpful in determining the ideal set of hyperparameters for the model to employ in order to improve accuracy. A dictionary of hyperparameters is available for searching in the param_grid variable of this function. To find the combination that produces the best accuracy, GridSearchCV and BayesSearchCV thoroughly tests all possible combinations of these hyperparameters. As a result, the Random Forest classifier for the provided dataset is improved.

Conclusion

Thus, Significant findings were produced by the quantitative analysis of the EEG data for the detection of eye states. The system's accuracy of 92.84% demonstrated its capacity to correctly identify the eye state (open or closed) based on the EEG data which was obtained by the optimization using GridSearchCV. Also calculated were precision and recall scores. These metrics offer information on the system's propensity to correctly identify positive eye state detections as well as its capacity to cover actual eye state occurrences.

References

- Automatic Sleep Stage Classification With Single Channel EEG Signal Based on Two-Layer Stacked Ensemble Model-Jinjin Zhou , Guangsheng Wang , Junbiao Liu , Duanpo Wu , Weifeng Xu , Zimeng Wang , Jing Ye , Ming Xia , Ying Hu and Yuanyuan Tian
- Improving Outcome Prediction for Traumatic Brain Injury From Imbalanced Datasets Using RUSBoosted Trees on Electroencephalography Spectral Power- Nor Safira Elaina Mohd Noor , Haidi Ibrahim , Muhammad Hanif Che Lah and Jafri Malin Abdullah
- A Hybrid Asynchronous Brain-Computer Interface Combining SSVEP and EOG- Yajun Zhou , Shenghong He , Qiyun Huang and Yuanqing Li