EEG Based Participant Independent Emotion Classification using Gradient Boosting Machines

Paper ID: 272

Sagar Aggarwal, Luv Aggarwal, Manshubh Singh Rihal and Swati Aggarwal
Department Of Computer Engineering, Netaji Subhas Institute of Technology
Sector-3, Dwarka, Delhi 110078, India

BENNETT UNIVERSITY
A TIMES GROUP INITIATIVE
IEEE

Introduction

Analysis of EEG (Electroencephalography) signals provides an alternative ingenious approach towards Emotion recognition. Nowadays, Gradient Boosting Machines (GBMs) have emerged as state-of-the-art supervised classification techniques used for robust modeling of various standard machine learning problems. In this paper, two GBM's (XGBoost and LightGBM) were used for emotion classification on DEAP Dataset. Furthermore, a participant independent model was fabricated by excluding participant number from features. The proposed approach performed well with high accuracies and faster training speed.

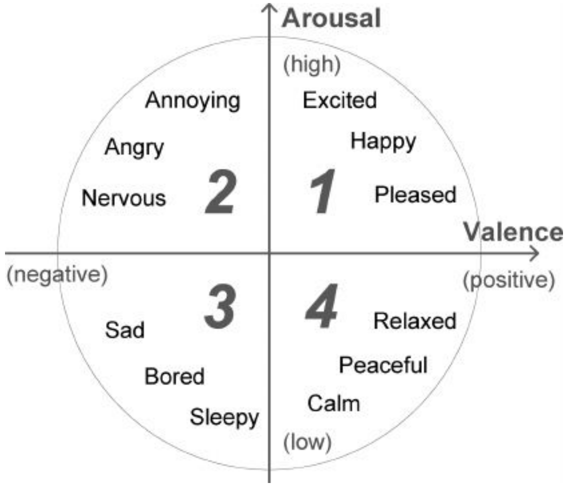
- What is EEG (*Electroencephalography*)?
- Emotion Classification
- Gradient Boosting Machines

Problem Statement

To perform Emotion Classification using algorithms that are computationally inexpensive and remove the participant dependency that has been part of all major researches in the past.

Literature work

- Ekman et al [21] pro- posed six basic emotions using the facial expression: Anger, Disgust, Fear, Happiness, and Sadness.
- Eugene et al [22] concluded a newly derived set of basic emotions with a 6.1% increase in distinctness using clustering of emotional tweets: Accepting, Ashamed, Contempt, Interested, Joyful, Pleased, Sleepy, Stressed.
- Two-dimensional model was put forward by Russell [23] : Arousal-Valence space (AVS)
- Yuan-Pin Lin et al [25] used Russell's 2D AVS to create EEG data for Emotion Recognition and Support Vector Ma- chines were used to classify four emotional states (joy, anger, sadness, and pleasure)



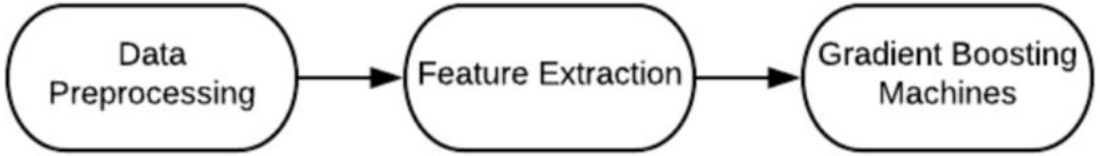
Experiment Design

- Dataset Used: DEAP (A Database for Emotion Analysis Using Physiological Signals)
- System Specifications: Intel(R) Core(TM) i7-7500 CPU @ 2.70 GHz x 4, 16 GiB RAM (Hardware) and Windows 10 with Python 3.6 (Software).
- XGBoost and LightGBM Parameters used: (max depth=1, min child weight=1, gamma = 0.1,colsample bytree = 0.55, scale and os weight = 0.9)

Dataset

- The DEAP (A Database for Emotion Analysis Using Phys- iological Signals) [26] dataset comprises of two parts:
  - 14-16 volunteers rated 120 one-minute extracts of music clips on valence, arousal, and dominance.
  - Out of these 120 videos, 40 videos were selected using a web-based subjective emotion assessment interface.
- Thirty-two participants watched these above selected 40 music clips and the physiological signals were recorded.
- Russell's AVS was used to quantitatively describe the emotions of 32 participants.

Methodology



- Brain signals can be severely affected due to the unwanted potentials or artefacts generated by non-physiological sources or physiological sources. Hence, ***Data Pre-processing*** is needed to remove such artefacts from the raw EEG data to improve the performance of classifiers.
- In DEAP dataset, for each trial, there is a hyperplane of 322560 features. Hence, ***feature extraction*** was performed on 8064 readings of 40 channels to reduce the feature size down to 4040.
- For any supervised machine learning problem, the bias- variance trade-off plays an important role in reducing the difference between the actual values and the predicted values of a model. The simultaneous minimization of these sources of error helps the model to generalize beyond the training data. To reduce these errors efficiently, a gradient boosting machine algorithm is used.

Results

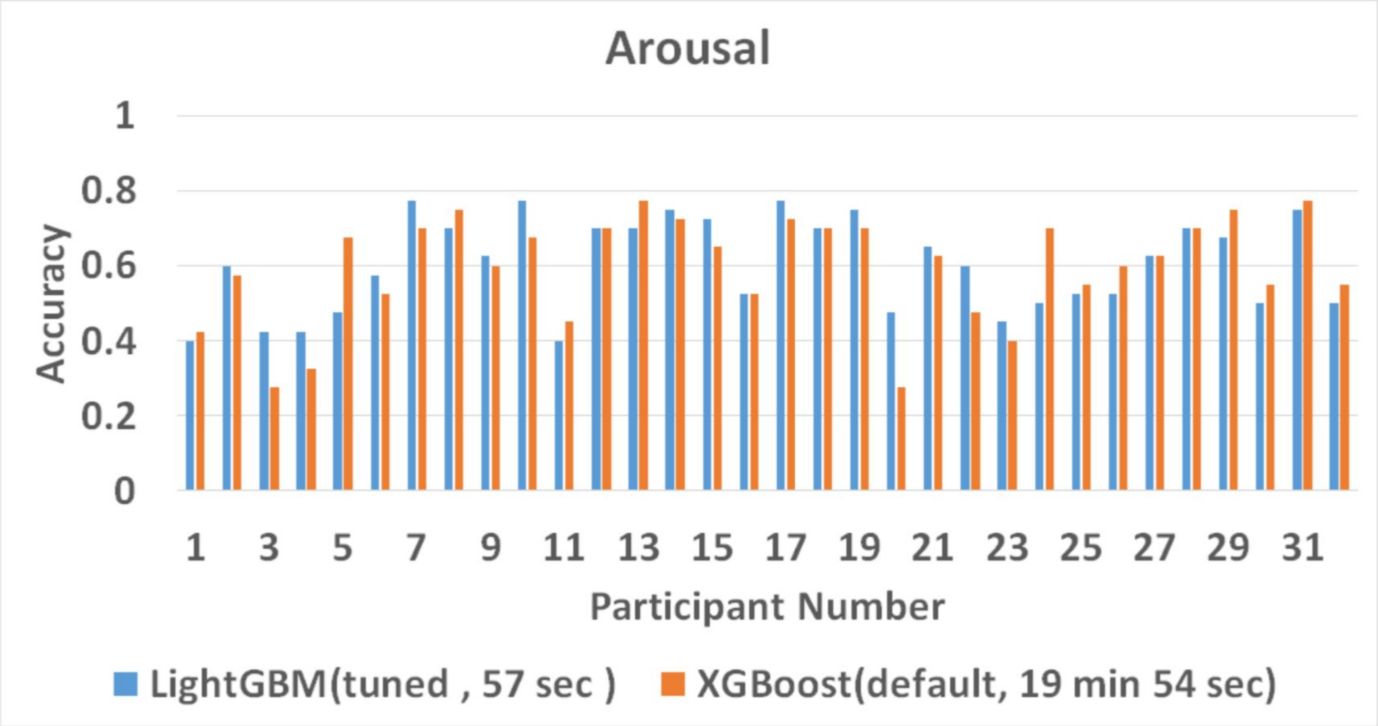


Fig. 3. Accuracy and Execution Time obtained for Arousal by XGBoost and LightGBM. LightGBM Parameters used: (max_depth=1, min_child_weight=1, gamma = 0.1,colsample_bytree = 0.55, scale_ and os_weight = 0.9)

Analysis

Classification Model 2 classes	Valence	Arousal
Chung et al	66.6%	66.4%
Rozgic et al	76.9%	68.4%
Samarth Tripathi (DNN)	75.78%	73.125%
XGBoost	76.25 % (0.088)	59.53% (0.139)
LightGBM	77.11% (0.081)	60.25% (0.122)

Average accuracy and standard deviation for valence and arousal.

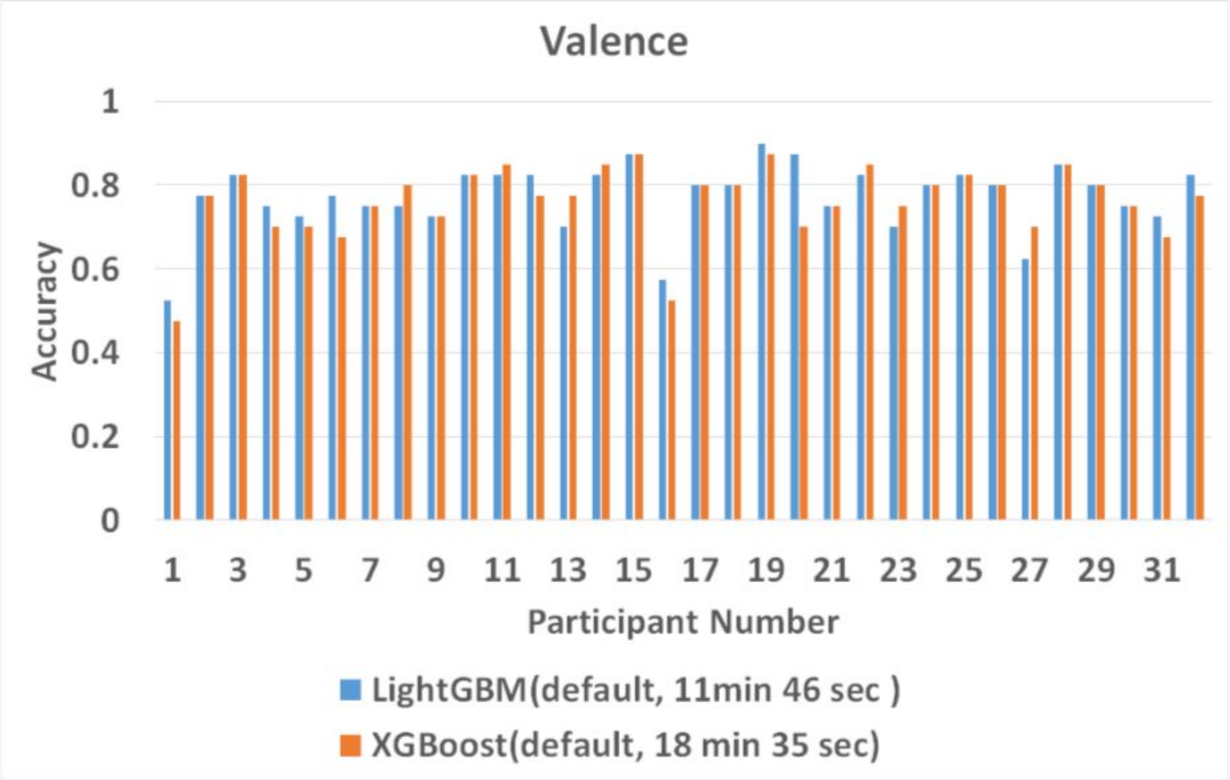


Fig. 2. Accuracy and Execution Time obtained for Valence by XGBoost and LightGBM.

Conclusion

- The size of the feature set was reduced by 21.55% per trial.
- Participant number was dropped as a feature and accuracies higher than previous works were achieved.
- GBMs were observed to be computationally inexpensive as compared to other models.

Future work

- Other prominent approaches for feature extraction such as Power Spectral Density (PSD), Fractal Dimensions, Differential, and Rational Asymmetry can be used to extract more powerful features