# EEG based Drowsiness Prediction Using Machine Learning Approach

**V. Vijay Priya**
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, India.

**M. Uma***
Department of Software Engineering, SRM Institute of Science and Technology, Kattankulathur, India.

*Received April 23, 2021; Accepted August 18, 2021*
*ISSN: 1735-188X*
*DOI: 10.14704/WEB/V18I2/WEB18351*

## Abstract

Drowsiness is the main cause of road accidents and it leads to severe physical injury, death, and significant economic losses. To monitor driver drowsiness various methods like Behaviour measures, Vehicle measures, Physiological measures and Hybrid measures have been used in previous research. This paper mainly focuses on physiological methods to predict the driver's drowsiness. Several physiological methods are used to predict drowsiness. Among those methods, Electroencephalography is one of the non-invasive physiological methods to measure the brain activity of the subject. EEG brain signal extracted from the human scalp is analysed with various features and used for various health application like predicting drowsiness, fatigue etc. The main objective of the proposed system is to early predict the driver drowsiness with high accuracy so that we have divided our work into two steps. The first step is to collect the publicly available dataset of EEG based Eye state as (Eye open and Eye closed) where the signal acquisition process was done from Emotiv EEG Neuroheadset (14 electrodes) and analysed various feature engineering techniques and statistical techniques. The second step was applied with the machine learning classification model as K-NN and performance-based predicting models are used. In the Existing System, they used various machine learning classification models like K-NN and SVM for EEG Eye state classification and produced results around 80% -97%. Compared to the Existing system our proposed method produced better classification models for predicting driver drowsiness using different Feature engineering process and classification models as K-NN produced 98% of accuracy.

## Keywords

Brain Waves, Electroencephalography (EEG), Emotiv EEG Neuroheadset, K-Nearest Neighbours.

## Introduction

According to the Ministry of Road Transport and Highways, India ranks first in road accidents death among 199 countries and the Road accident in India during 2019 were 449,002 accidents and 151,113 deaths and 451,361 Injuries and the state of Tamil Nadu reported the greatest number of road accidents (57,228) in 2019 the accident caused by Car, Taxi, and van leads to 18.6% of road accidents. Long-distance driving often leads a driver to feel drowsy at the wheel. The Biggest human fault is drowsy driving. Drowsy driving is dangerous for the driver and road users. The symptoms of Drowsiness are dropping eyelids, frequently yawning, Rubbing Eyes, Nodding Head. In Existing methods Research to monitor drowsiness is classified into four methods they are Vehicle based methods, Behaviour based method, Physiological based methods and Hybrid based methods (A. Sahayadhas. et. al, 2012). It examines physiological signals including Electroencephalogram (EEG), Electrooculogram (EOG), and Electrocardiogram (ECG) for detecting driver drowsiness in physiological-based techniques. Previous methods for detecting driver tiredness employed cameras, wearable sensors, and EEG signals. They used various computer vision methods to detect abnormal driving behaviour in camera-based drowsiness detection. They applied several signal processing and machine learning algorithms in wearable sensor-based sleepiness detection. The subject is processed using EEG signals and sleeping based analysis in EEG based drowsiness detection (Budak, U. et.al, 2019). EEG is related to record brain activities using one or more electrodes from the scalp. The non-invasive electrodes from the Brain-Computer Interface (BCI) are used in the position of the wearable Emotiv EEG Neuroheadset. In EEG, signals are received, amplified and digitized and forwarded for storage and data processing to computers or mobile devices. Emotiv EEG Neuroheadset is a low-cost multichannel wireless EEG designed for research. Emotiv EEG Neuroheadset contains 14 sensors from the brain wave electrodes. A physiological signal such as Electroencephalogram (EEG), electrooculography (EOG), Electrocardiography (ECG), and electromyography (EMG) among these signal Electroencephalogram (EEG) signal indicates reliable measures for detecting drowsiness. The electrical patterns observed from the EEG signals are called brain waves. EEG contains five brainwaves for identifying their different frequency ranges. The five brainwaves used in EEG are Delta, Theta, Alpha, Beta and Gamma. Delta waves range from 1-4Hz and are observed during deep sleep. Theta waves range from 4-8Hz from the frontal and temporal region and are observed during daydreaming. Alpha waves range from 8-12Hz from the occipital and parietal region and are observed during a relaxed wake state with the eye closed. Beta waves range from 13-30Hz from the pre-central and frontal region and are observed during stress state, less concentration and attention. Gamma waves range from

31Hz – 100Hz from the frontal region and are observed during high mental activities, problem-solving and creativity (Saichoo.et.al, 2019).

In the Existing System, the author used the Emotiv EEG Neuroheadset dataset and produced 94% accuracy using a k-star classifier to predict Eye state (Rosler et.al, 2013). The author (Wang Ting et.al, 2014), used the Emotiv EEG Neuroheadset dataset and produced 76% accuracy using Incremental Attribute Learning to predict eye state Identification. From the existing system, we proposed machine learning classifiers as K-Nearest Neighbour for predicting Driver Drowsiness Eye state using Emotive EEG Neuroheadset. To evaluate the performance of the proposed method Emotive EEG Neuroheadset is placed at the subject scalp and produces low cost with a smaller number of electrodes. We utilized performance-based forecasting models such as the Confusion Matrix, Receiver Operating Characteristic (ROC) curve, Sensitivity, Specificity, False Positive Error, True Positive Error, Precision, Recall, F1 score, and Accuracy to extract features. These measures were used to evaluate performance metrics with K-NN-based classification models that were used to extract features from the data. This paper focuses mostly on EEG Feature engineering techniques and compares their performance with K-NN classification models to determine the best ways for predicting driver tiredness. The following is the design of the paper. Section 2 depicts the Associated works based on existing approaches. Section 3 depicts the proposed method as well as the data gathering methods employed in the experiment. The Experiment Results and Discussions are displayed in Section 4. Section 5 discusses the conclusions and future efforts.

## Related Works

In (Ma, P. et.al, 2020), the authors used EEG data to predict the open or closed eye status based on the extraction of the feature and depth factorization machine model is used as a Factorizing Machine (FM) and Long Short-Term Memory (LSTM) to derive functionality from EEG data (LSTM). In (Sulaiman et.al,2011), the authors used EEG signals to extract stress features using two eye state as open eyes and closed eyes and used K-NN as a classifier and produced 88.89% of accuracy while extracting stress features. In (Mehmood et.al, 2017) the author proposed emotion recognition based on EEG signals to extract the features like happy, sad, calm and scared and produced good optimal feature selection as an ensemble method with 76.6% for detecting emotion recognition. In (Benitez, D.S. et.al, 2016), the author used Emotiv EPOC Neuroheadset for detecting simple and accurate eyewink identification using EEG signal and used classifier as an artificial neural network. In (Belakhdar, I. et.al, 2016) The authors developed a drowsiness detection system to evaluate the performance of two classifiers like Support Vector Machine and

Artificial Neural Network using EEG signals and finds the best classifier for detecting drowsiness. The authors of (Gang.Li.et.al, 2015) created a driver drowsiness detection system based on an SVM-based posterior probabilistic model and a smart watch-based wearable EEG device with a Bluetooth EEG band. In (Budak.et.al, 2019) the author employed EEG readings with Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) to detect tiredness with 94.31 % of accuracy. In (Rosler.et.al, 2013) the authors developed an Emotiv EPOC headset for predicting the brain waves using fourteen electrodes from EEG signals and tested the performance of the classifier using 42 different machine learning classifiers out of 42 classifiers K Star produced the best performance with 94% of accuracy. The authors (Saghafi.et.al, 2017) used three different classifiers to extract features from EEG signals with changes in eye state, using multivariate Empirical mode decomposition, Artificial Neural Network (ANN), Logistic Regression, and Linear Regression, measuring the performance of both Artificial Neural Network (ANN) and Logistic Regression produced 88.2 % of accuracy. In (Balam.et.al, 2021) the authors used single-channel EEG signals to detect driver drowsiness and used Convolutional Neural Network (CNN) for classifying the models as combined subject validation, cross subjects and subject wise validation procedures using various subjects with single-channel based EEG signals. In (Chaabene.et.al,2021) the author proposed two methods data acquisition and model analysis in data acquisition they used a single collection from Emotiv EPOC+ headset and signal annotation with different characteristics of brain waves like Gamma, Theta, Beta and Delta and used CNN for classification and produced 90.42% of accuracy. In (Wang Ting.et.al,2014) the author developed a time series classification using Incremental Attribute Learning (IAL) for detecting EEG Eye State Identification and used Emotiv EEG Neuroheadset as a benchmark dataset and formed a high classification error rate. In (Poorna, S. et.al, 2018) the author used EEG signals from 14 channel wireless neuroheadset for detecting driver drowsiness and drowsiness of pilots and used two classifiers K-Nearest Neighbours produced 80% of accuracy and Artificial Neural Network produced 85% of accuracy. In (Kim, Dajeong.et.al,2012) the authors developed EEG based power spectrum of alpha, beta, theta and delta signals that were analysed for detecting drowsiness from the power spectrum alpha signals produced significant changes when the eyes are opened during drowsiness times.

## Proposed Methodology

In the proposed method, the Physiological method is used for predicting drowsiness. In the physiological based method, Eye state (open or close) using EEG (brain wave) data are used to predict the driver drowsiness. In the proposed method, Emotiv EEG

Neuroheadset from the subject is collected from the brain wave signal and the different machine learning classifiers were applied for predicting the driver drowsiness based on the eye state (open or close) are depicted in Figure 1. When the eyes are closed it predicts 1 and when the eyes are open it predicts 0.

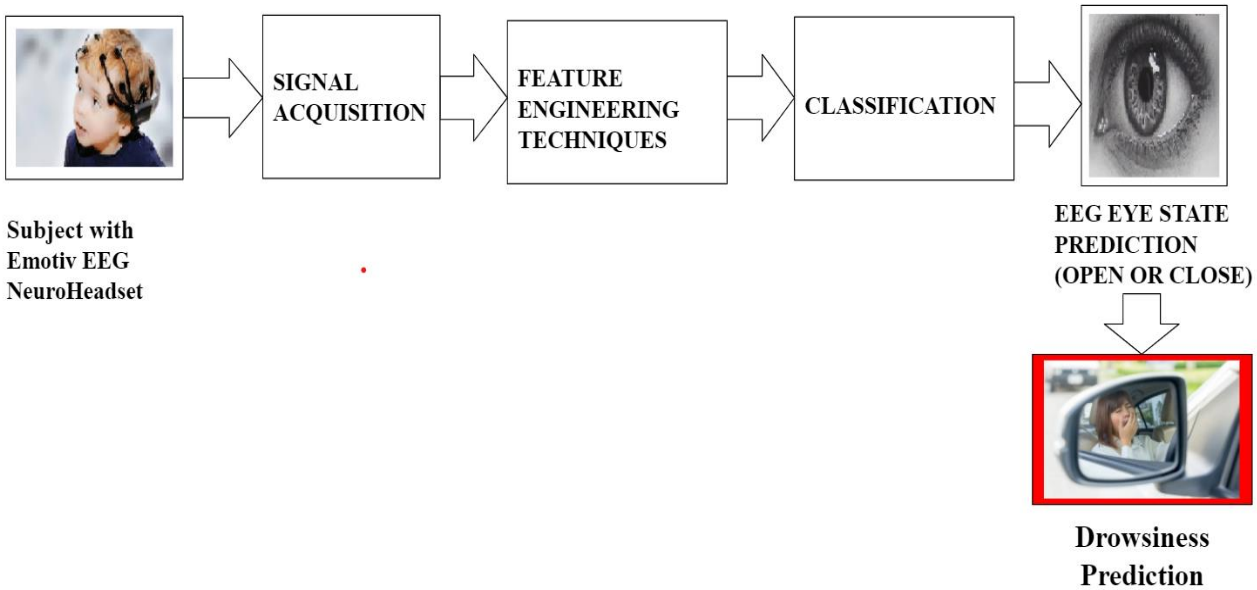


**Figure 1 The Architecture Diagram for Drowsiness prediction**

### 1) Emotiv EEG Neuroheadset

We propose Emotiv EEG Neuroheadset brain waves to predict the driver drowsiness either drowsy or non-drowsy by applying each value that indicates the driver Eye state (open or close). Emotiv EEG Neuroheadset records the brain waves from the electrodes fixed above the subject head. The electrodes are positioned at positions AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4 following the international 10-20 system, resulting in seven sets of symmetric channels (see Figure 2). This high-resolution neuro signal acquisition and wireless processing headset have 14 active electrodes located above the ears as well as two reference electrodes located above the ears, one for the left hemisphere of the head and another for the right hemisphere of the head. Letters denoting the surrounding brain regions are affixed to the electrodes: F denotes the Frontal, C denotes the Central, P denotes the Parietal, T denotes the Temporal, and O denotes the Occipital. The letters are followed by odd numbers for electrodes on the left side and even numbers for electrodes on the right side.


**Figure 2 Emotiv EEG Neuroheadset brain wave with 14 Electrodes position**

### 2) Data Collection

Experimenters used an Emotiv EEG Neuroheadset, which was placed on the subject's head to collect data on brain waves, to collect data during the experiment. Emotiv EEG Neuroheadset contains a single 16-bit ADC for converting raw EEG data into analog to digital format. It contains a sampling frequency of 128HZ. The programming language and tools used for experimenting are Python version 3.7.4. The datasets used for the experiment are open-source from the Machine Learning Repository, University of California, and Irvine (UCI). The number of observations taken from the subject is 14980 observations were made within 117 seconds. There was about 128 observation per second.

### 3) Signal Acquisition

The Emotiv EEG Neuroheadset device is used for signal acquisition, which is the process of acquiring signal data from the scalp to capture brain signals. During the EEG recording, the subject's eye state is captured on video during the procedure. Unwanted noises can interfere with the recording of EEG data, resulting in incorrect results from the signals. It removes the artefacts during the acquisition process. The subject did not know the exact start time of the EEG test to reduce the artefacts. Alternatively, the subject was ordered to sit down, gaze directly at the camera, and modify its eye condition at will. In the EEG Eye state, the subject data contains three occurrences with the number 899, 10387 and 11510 has some errors so the values are deleted from the subject data to avoid transmission errors.

### 4) Feature Engineering Techniques

In the proposed system the feature engineering techniques play a major role in detecting features from the EEG eye state data. Using the feature engineering in the proposed method selects and transforms the feature for creating a predictive model and improves the performance of the machine learning model. The process of Feature Engineering is presented in Figure 3.

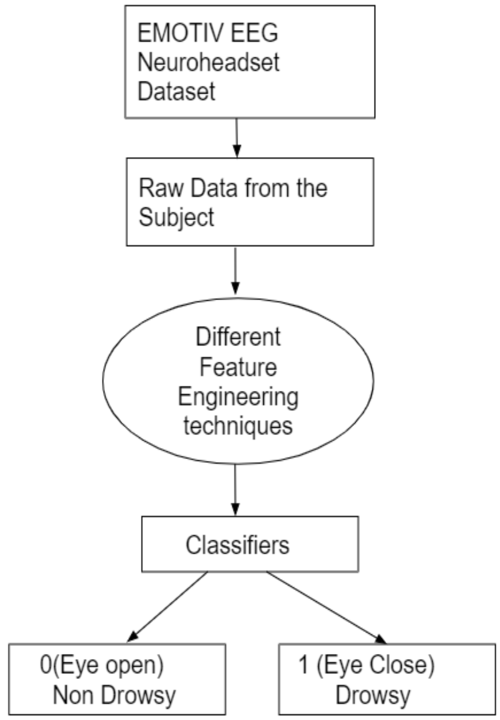


**Figure 3 EEG Feature Engineering**

The framework consists of 6 steps:

**Stage 1:** The system collects EEG brain wave data from the Emotiv EEG Neuroheadset.

**Stage 2:** In cleaning the dataset, we will be checking whether the dataset is balanced or not, label encoding is present or not and then take steps accordingly. It was found that the data did not have any missing values and unwanted parameters were removed.

**Stage 3:** Different feature engineering steps have been used to improve the performance models.

**Stage 4:** Training and Test Split: We split the data after the irrelevant variables have been removed. A split ratio of 70/30 is done.

**Stage 5:** The classification models are generated using machine learning classifiers as K-NN and the model is implemented on the data.

**Stage 6:** Finally, it predicts the driver is drowsy or non-drowsy based on the eye state using EEG brain waves.

The feature engineering techniques used in this paper are Data Exploration, Data Pre-processing and Scaling like Normalization and Standardization.

- **Data Exploration**

The dataset is collected from the UCI Machine Learning repository. The classification goal is to predict whether the driver is drowsy or non-drowsy using two values 0 and 1. The dataset contains 15 features from the Emotive EEG Neuroheadset. Data Exploration is used for checking the missing data.

- **Data Preprocessing**

Data pre-processing is used to convert raw data into a clean dataset. For data pre-processing, I have used Anaconda Navigator. The Dataset is loaded and the missing data are identified from the dataset using two categorical variables as Eye open and Eye closed state. The predicted categorical values as 0 and 1.

- **Scaling**

Scaling is the numerical features from the dataset to have a certain range and they differ from each other. Two ways of scaling used in this paper are:

- ***Normalization***

Normalization is a technique of scaling with all values ranging from 0 to 1. The formula for finding the normalization is

$$Y = \frac{X - X_{Min}}{X_{Max} - X_{Min}} \qquad (1)$$

The maximum and minimum values of the feature here are Xmax and Xmin. If X is the minimum column value, the numerator is 0, and so Y is 0. When X is the maximum column value, the numerator is equal to the denominator and Y equals 1. The value of X may be anywhere between the minimum and maximum values, in which case the value of Y will be somewhere between 0 and 1.

- ***Standardization***

In the context of scaling, standardization is a feature that exists between a given minimum and maximum value. Standardization scales the values to have mean values of the dataset as 0 and the standard deviation equal to 1. The statistics for finding data standardization is

$$Y = \frac{X - \text{Mean}}{\text{Standard Deviation}} \qquad (2)$$

When Y is 0 the observation is at the mean and when Y is 1 then the observation is the standard deviation.

The process of Feature Engineering is presented in Figure 3.

### 5) Classification Models

The following K-NN models are used to predict driver sleepiness with training and test data spread over the 70:30 ratio.

- **K-Nearest Neighbours**

K-NN is a supervised machine learning algorithm. K-NN uses all the data from the training data and classifies it to a new data instance. K-NN does not train before predictions can be simultaneously added to new data. The value of k and the distance function n are only two parameters. In K-Nearest Neighbours from the dataset, it contains 14 features and one label as eye state (open or close). K-Nearest Neighbours fits the model with the trainset named as fit () and forms prediction using a test set named as predict ().

## Experiment Results and Discussions

As an EEG sensor for the experimental model, we used the Emotiv EEG Neuroheadset to predict the state of the eye. The device acquires EEG brain waves as data by wearing like a normal headphone. The dataset contains EEG brain waves of 14977 occurrences from 15 features. In Emotiv EEG Neuroheadset, 14 features represent the sensor values from the electrodes and the other one is the eye state. During EEG measurement the eye state is identified from the camera and later included in the dataset file after the video frames have been analysed. The measurement time was 117 seconds. The 14 features from the electrodes are F7, F3, O2, P8, T8, FC6, F4, AF3, FC5, T7, P7, O1, F8, and AF4. In 14977 occurrences, 8255 occurrences represent Eye opened state and 6722 Occurrences represent Eye closed state is depicted in Figure 4.

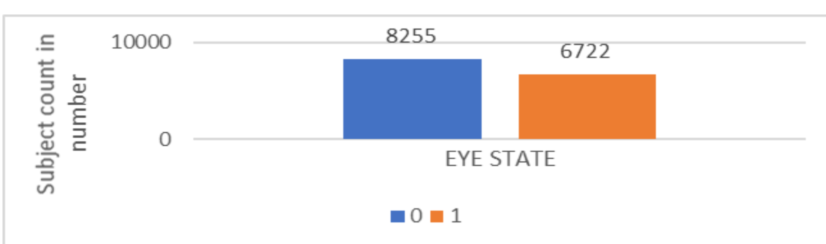


**Figure 4 Eye state (0, 1) samples**

During the analysis of its range of values for slightly different states, the amplitude of the sensors was significantly different. It shows the minimum and maximum values of the eye closed state and eye opened state with means and standard deviations individually. Based on data in Table 1, the mean of two eye states with similar characteristics is equal in both, and the standard deviation with two eye states is significantly different in the eye-opened state. The attributes F7, F3, O2, P8, T8, FC6 and F4 are the minimum eye-opening value. Tables 1 AF3, FC5, T7, P7, O1, F8 and AF4 show an eye closure with max values.

**Table 1 Statistical Data Analysis for the Eye state**

| **Electrode** | **AF3** | **F7** | **F3** | **FC5** | **T7** | **P7** | **O1** | **O2** |
|---|---|---|---|---|---|---|---|---|
| Count | 7923 | 7923 | 7923 | 7923 | 7923 | 7923 | 7923 | 7923 |
| Mean | 4344.58 | 4007.31 | 4264.63 | 4121.98 | 4341.07 | 4661.98 | 4075.46 | 4615.39 |
| Std | 3426.03 | 32.87 | 50.66 | 25.55 | 28.56 | 4021.41 | 41.06 | 19.94 |
| Min | 4198.97 | 3797.95 | 1040.00 | 3733.85 | 4304.62 | 4002.05 | 2086.15 | 4567.18 |
| 25% | 4284.62 | 3990.26 | 4251.28 | 4108.21 | 4330.77 | 4607.18 | 4057.44 | 4603.59 |
| 50% | 4298.46 | 4005.13 | 4263.08 | 4120.51 | 4337.44 | 4615.90 | 4073.33 | 4613.33 |
| 75% | 4316.92 | 4020.00 | 4272.82 | 4132.31 | 4346.67 | 4623.59 | 4090.26 | 4624.62 |
| Max | 309231 | 5500.51 | 6880.51 | 5416.41 | 6040.51 | 362564 | 6350.26 | 5361.54 |

| **Electrode** | **P8** | **T8** | **FC6** | **F4** | **F8** | **AF4** | **EYE STATE** |
|---|---|---|---|---|---|---|---|
| Count | 7923 | 7923 | 7923 | 7923 | 7923 | 7923 | 7923 |
| Mean | 4234.46 | 4232.34 | 4203.49 | 4281.04 | 4628.03 | 4456.63 | 0.848542 |
| Std | 2937.37 | 29.27 | 39.43 | 32.17 | 1660.38 | 7993.80 | 0.358517 |
| Min | 1357.95 | 3914.87 | 3273.33 | 2257.95 | 276.41 | 4246.15 | 0 |
| 25% | 4190.26 | 4221.03 | 4190.26 | 4269.74 | 4591.28 | 4346.15 | 1 |
| 50% | 4200 | 4229.74 | 4200.51 | 4278.97 | 4605.64 | 4358.97 | 1 |
| 75% | 4210.77 | 4241.03 | 4212.31 | 4289.23 | 4621.54 | 4378.46 | 1 |
| Max | 265641 | 6215.38 | 6823.08 | 4368.72 | 152308 | 715897 | 1 |

## 1) Performance of Evaluation Metrics

We have used various evaluation metrics in the proposed method to predict the various models, including the Confusion Matrix, Receiver Operating Characteristic (ROC) curve, Sensitivity, Specificity, False Positive Rate, False Negative Rate, Precision, Recall, F1 score, and Accuracy definition, which are all listed below. The accuracy definition is given below.

$$\text{Sensitivity} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)+ False Negative (FN)}} \tag{3}$$

$$\text{Specificity} = \frac{\text{True Negative (TN)}}{\text{True Negative (TN)+ False Positive (FP)}} \tag{4}$$

$$\text{False Positive Rate} = \frac{\text{False Positive (FP)}}{\text{False Positive (FP) + True Negative (TN)}} \tag{5}$$

$$\text{False Negative Rate} = \frac{\text{False Negative (FN)}}{\text{False Negative (FN)+True Positive (TP)}} \quad (6)$$

$$\text{Precision} = \frac{\text{True positive (TP)}}{\text{True Positive (TP)+ False Positive (FP)}} \quad (7)$$

$$\text{Recall} = \frac{\text{True positive (TP)}}{\text{True Positive (TP)+ False Negative (FN)}} \quad (8)$$

$$\text{F1 Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision+ Recall}} \quad (9)$$

$$\text{Accuracy} = \frac{\text{True Negative (TN) + True Positive (TP)}}{\text{TN+FP+TP+FN}} \quad (10)$$

The Confusion Matrix forms the accuracy of the machine learning models to classify the data with appropriate labels. Confusion Matrix is defined as an error matrix to visualize the performance of a classifier. In the proposed system confusion Matrix is found for K-NN based classifier models. The classifier correctly predicted a closed eye state with 193 cases and it wrongly predicted 28 closed eye state instance as opened eye state. It incorrectly predicted 15 instances as opened eye state and 1349 instance had been correctly predicted as opened eye state is depicted in Table 2.

**Table 2 Confusion matrix for predicted Eye state**

| **Confusion Matrix** | | **Predicted Eye state** | | **Total** |
|---|---|---|---|---|
| | | 1(close) | 0(open) | |
| Actual Eye state | 1(close) | 193 | 28 | 221 |
| | 0(open) | 15 | 1349 | 1364 |
| | Total | 208 | 1377 | 1585 |

It is possible to predict the probability of a binary classifier using the Receiver Operating Characteristic Curve. Figure 5 depicts the data with a False Positive rate on the X-axis and a True Positive rate on the Y-axis. A binary classifier's sensitivity and specificity are used to calculate the ROC. The fraction of Positives properly predicted by the classifier is called sensitivity. The fraction of Negatives properly predicted by the classifier is known as specificity.

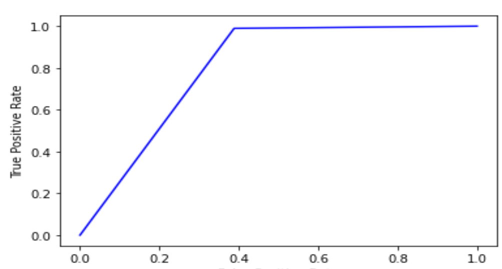


**Figure 5 ROC curve Results with False positive rate and True Positive rate**

The False-Positive error rate is defined as the percentage of Negatives predicted incorrectly by the classifier. When the classifier correctly predicts a Positive, the False-Negative error rate is equal to the fraction of Positives that were correctly predicted by the classifier. When the classifier correctly predicts a Positive, the False-Negative error rate is equal to the fraction of Positives that were correctly predicted by the classifier. True Positive represents the correctly predicted closed eye state. False-Positive represents the incorrectly predicted closed eye state. True Negative represents the correctly predicted open eye state. False Negative represents the incorrectly predicted open eye state. Precision is used to correctly identify both True Positive and False Positive. It is the percentage of positives that were correctly predicted by a classifier that is known as recall of the classifier. The F1 Score is the average of the accuracy and recall predictions made by the classifier as projected by the classifier. As stated by the Emotive EEG Neuroheadset dataset, accuracy can be defined as the ratio between the number of occurrences that were successfully classified as correctly identified by a classifier divided by the total number of instances in the dataset. Accuracy is a proportion of positives the classifier correctly predicts. In this paper to evaluate the performance metrics, we used different models such as Precision, Recall, F1 measure and accuracy. It is decided whether or not to classify an eye state as open or closed based on the EEG data using the Evaluation methods for binary prediction. In the Emotiv EEG Neuroheadset dataset for predicting Eye state, as shown in Table 3 and Figure 6, these four performance measures are used to compare the different classification models. These four performance measures are listed in Table 3 for reference.

**Table 3 Results obtained from Proposed Prediction methods**

| **Predicted Method** | **Predicted Result** |
|---|---|
| Sensitivity | 98.90 |
| Specificity | 87.33 |
| False Positive error rate | 12.67 |
| False Negative error rate | 1.09 |

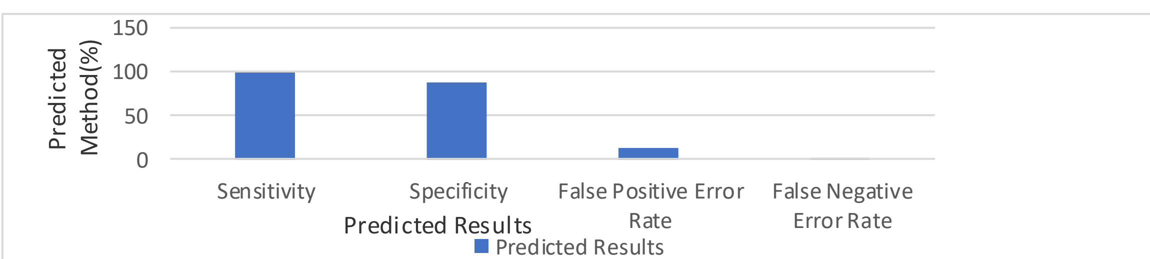


**Figure 6 Represents Prediction Results based on the prediction method**

### 2) Comparison of Classified Prediction Result

For our classification model, we implemented the following model as K-Nearest Neighbours. Table 4 and Table 5 compares the Existing method (Ma, P. et.al, 2020) and (Poorna, S. et.al, 2018) classification model as K-NN, Artificial Neural network and Factorization Machine (FM) and Long Short-Term Memory (LSTM) using the Emotive EEG Neuroheadset dataset are compared with the proposed method based on their performance and classification prediction results K-NN produced 98% of accuracy and good performance evaluation metrics.

**Table 4 Classification prediction results for K-NN using Emotive EEG Neuroheadset Dataset**

| Reference No | Model | Accuracy | Precision (%) | | Recall (%) | | F1 score (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 0 | 1 | 0 | 1 |
| Proposed Method | K-Nearest Neighbour | 0.98 | 0.92 | 0.99 | 0.92 | 0.99 | 0.92 | 0.99 |
| [7] | Factorization Machine (FM) and Long Short-Term Memory (LSTM) | 0.93 | 0.94 | | 0.90 | | 0.92 | |

**Table 5 Compared Classification prediction results for K-NN with Existing Method**

| Reference No | Model | Accuracy | Sensitivity | Specificity | Precision |
|---|---|---|---|---|---|
| Proposed Method | K-NN | 0.98 | 98.90 | 87.33 | 0.99 |
| [19] | K-NN | 0.80 | 33.35 | 81.95 | 84.36 |
| [19] | Artificial Neural Network | 0.85 | 58.21 | 83.24 | 65.73 |

The Prediction results are compared with the different classification models based on the existing methods and depicted in Table 6. Compare to the different classifier (Ma, P. et.al, 2020), (Wang, Ting et.al, 2014), (Benitez, D.S., et.al, 2016), (Chaabene, Siwar, et.al, 2021), (Poorna, S. S., et.al, 2018), (Sotelo et.al,2018) used the same Emotiv EEG Neuroheadset as their dataset to predict the eye state and produced less accuracy are depicted in Figure 7. In the proposed method we used feature engineering to extract the features and produced better accuracy. In the proposed method compared to the existing classifiers, using different feature engineering techniques K-NN produced better accuracy.

**Table 6 Comparison of Different classifier based on the Accuracy result**

| Reference No | Classifier | Accuracy |
|---|---|---|
| [18] | Incremental Attribute Learning | 76% |
| [17] | Convolutional Neural Network | 90% |
| [10] | Artificial Neural Network | 87% |
| [7] | Factorization Machine (FM) and Long Short-Term Memory (LSTM) | 93% |
| [14] | K- Star | 94% |
| [19] | K-Nearest Neighbour, Artificial Neural Network | 80%-85% |
| [21] | Support Vector Machine | 80%-91% |
| Proposed Method | K-Nearest Neighbour | 98% |

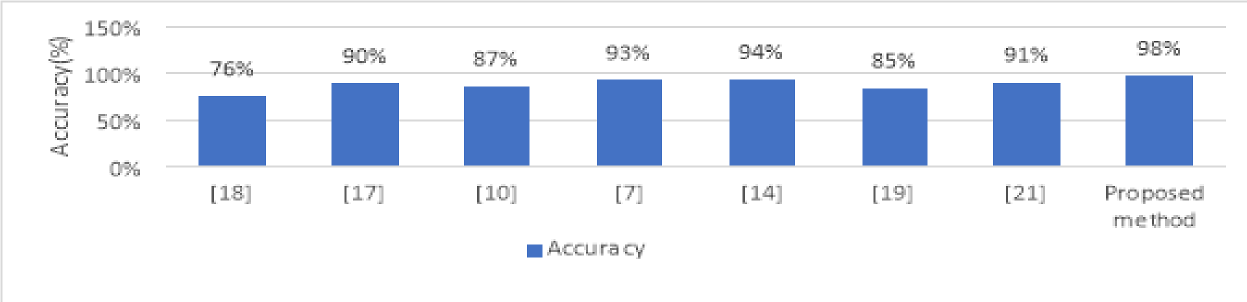


**Figure 7 Different Existing classifiers compared with the Proposed classifier methods**

## Conclusion

Road accidents in India are increasing day by day. The number of road accident under the category of car, Jeep, Taxi, and van produced a greater number of accidents compared to other categories. In this paper, the physiological method is proposed for EEG based drowsiness prediction. By analysing EEG data, it is possible to distinguish between drowsy and non-drowsy drivers based on their eye movements (either open or closed eyes) and so forecast drowsiness. In this paper, we used Emotiv EEG Neuroheadset data from the public dataset and used machine learning classifiers as K-NN. The comparison of different classifier shows that the K-NN classifier outperformed better when compared to the existing classifiers. Based on the results K-NN produced 98% of accuracy for predicting EEG Eye state using Emotiv EEG Neuroheadset device. Compared to the other methods physiological methods predict accurate drowsiness other than image-based behaviour methods. It is not possible to conduct an ideal brain wave experiment in a real vehicle because high-frequency noise interferes with the acquisition of the EEG signal. To simulate driving conditions at this stage, we can use a simulation test. In future work, a combination of Single Electrode EEG devices and Behavioural-based methods will be used to predict drowsiness.

## References


Ministry of road Transport and Highways. https://morth.nic.in/road-accident-in-india

Article “Drowsy driving-safety challenge” from motor India online. https://www.motorindiaonline.in/driver-welfare/drowsy-driving-a-safety-challenge

Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2012). Detecting driver drowsiness based on sensors: a review. *Sensors, 12*(12), 16937-16953. https://doi.org/10.3390/s121216937

Budak, U., Bajaj, V., Akbulut, Y., Atila, O., & Sengur, A. (2019). An effective hybrid model for EEG-based drowsiness detection. *IEEE Sensors Journal*, *19*(17), 7624-7631.

Article “Emotiv Headset”. https://www.emotiv.com/glossary/eeg-headset/

Saichoo, T., & Boonbrahm, P. (2019). Brain computer interface for real-time driver drowsiness detection. *Thai Journal of Physics*, *36*(1), 1-8.

Ma, P., & Gao, Q. (2020). EEG signal and feature interaction modeling-based eye behavior prediction research. *Computational and Mathematical Methods in Medicine*, *2020*.

Sulaiman, N., Taib, M.N., Lias, S., Murat, Z.H., Aris, S.A., & Hamid, N.H.A. (2011). Novel methods for stress features identification using EEG signals. *International Journal of Simulation: Systems, Science and Technology*, *12*(1), 27-33.

Mehmood, R.M., Du, R., & Lee, H.J. (2017). Optimal feature selection and deep learning ensembles method for emotion recognition from human brain EEG sensors. *IEEE Access*, *5*, 14797-14806.

Benitez, D.S., Toscano, S., & Silva, A. (2016). On the use of the Emotiv EPOC neuroheadset as a low-cost alternative for EEG signal acquisition. *In IEEE Colombian Conference on Communications and Computing (COLCOM)*, 1-6.

Belakhdar, I., Kaaniche, W., Djmel, R., & Ouni, B. (2016). A comparison between ANN and SVM classifier for drowsiness detection based on single EEG channel. *In 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, 443-446.

Li, G., Lee, B.L., & Chung, W.Y. (2015). Smartwatch-based wearable EEG system for driver drowsiness detection. *IEEE Sensors Journal*, *15*(12), 7169-7180.

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

Rösler, O., & Suendermann, D. (2013). A first step towards eye state prediction using EEG. *Proc. of the AIHLS*.

Saghafi, A., Tsokos, C.P., Goudarzi, M., & Farhidzadeh, H. (2017). Random eye state change detection in real-time using EEG signals. *Expert Systems with Applications*, *72*, 42-48.

Balam, V.P., Sameer, V.U., & Chinara, S. (2021). Automated classification system for drowsiness detection using convolutional neural network and electroencephalogram. *IET Intelligent Transport Systems*, *15*(4), 514-524.

Chaabene, S., Bouaziz, B., Boudaya, A., Hökelmann, A., Ammar, A., & Chaari, L. (2021). Convolutional Neural Network for Drowsiness Detection Using EEG Signals. *Sensors*, *21*(5), 1734.

Wang, T., Guan, S.U., Man, K.L., & Ting, T.O. (2014). EEG eye state identification using incremental attribute learning with time-series classification. *Mathematical Problems in Engineering*, *2014*.

Poorna, S.S., Arsha, V.V., Aparna, P.T.A., Gopal, P., & Nair, G.J. (2018). Drowsiness detection for safe driving using PCA EEG signals. In *Progress in computing, analytics and networking*, 419-428.

Kim, D., Han, H., Cho, S., & Chong, U. (2012). Detection of drowsiness with eyes open using EEG-based power spectrum analysis. *In 7th International forum on strategic technology (IFOST)*, 1-4.

Sotelo, D.I., Benitez, J.A.P., & Hernández, J.H.E. (2018). Identification and classification of eyes movement using EEG signals. *In International Conference on Electronics, Communications and Computers (CONIELECOMP)*, 25-30.

Thamer, K.A. (2020). Method of artificial neural networks teaching. *Webology*, *17*(1), 43-64.

## Biographies of Authors

| | |
|---|---|
|  | V. Vijay Priya, Research Scholar in Department of computer science and engineering in SRM institute of science and technology, Kattankulathur. Received M.E degree in Computer Science and engineering from Valliammai Engineering college, Chennai, India. she has 2 years' experience in teaching as teaching Associate in SRM institute of science and technology, Kattankulathur. Her research interests include Internet of things, Image processing, Machine learning, Computer vision |
|  | Dr. Uma Mohan, received an M.Tech in Computer Science from SRM University, Chennai, India, an MCA from Bharathidasan University, Trichy and is completed her Ph.D in the area of Brain Computer Interface from Bharathiyar University, Coimbatore, India. She has 18+ years of teaching and research experience. Currently, she is an Assistant Professor (Selection Grade) in Department of Software Engineering, SRM Institute of Science and Technology, kattankulathur, Tamil Nadu, India. She is the author of 23 journal publications, 17 national and international conferences and 4 patents published. Her research interests include brain computer interface, EEG, P300, Data Mining, Personalization, Recommendation System, Machine learning, Web services and Data Analytics. |