Kathmandu University

Department of Computer Science and Engineering

Dhulikhel, Kavre

A Project Report
on

"Rennervate"

COMP 308

(For partial fulfillment of 3rd Year/ 2nd Semester in Computer Engineering)

Submitted by:

Aashish K.C. (19)

Urina Lama (25)

Neha Malla (27)

Shusant Sapkota (43)

Submitted to:

Dr. Gajendra Sharma

Department of Computer Science and Engineering

Submission Date: 28th September, 2020

CERTIFICATION

THIRD YEAR SECOND SEMESTER PROJECT REPORT

on

RENNERVATE

By:

Aashish K.C.

Urina Lama

Neha Malla

Shusant Sapkota

Approved by:

Project Coordinator	**Project Supervisor**
____________________	____________________
Dr. Gajendra Sharma	**Dr. Rajani Chulyadyo**
Associate Professor	**Lecturer**

ABSTRACT

Drowsy driving is one of the major causes of road accidents, leading to the death of people around the world. Since drowsy driving is prevalent all around the world, we decided to build a drowsiness detection model, which can be used to detect when a driver shows signs of drowsiness. The data was collected from videos of people showing drowsiness from NTHU-DDD, UTA-RLDD and YawDD datasets, using OpenCV and dlib's 68 landmarks face detector and predictor model. Then, various classification algorithms viz. Logistic Regression, Decision Tree, Random Forest, Naive Bayes, KNN and SVM, were applied on the collected data. Based on the accuracy, a KNN based classification model was obtained which can be used in real time systems for drowsiness detection.

Keywords: *Drowsy driving, Drowsiness detection, Machine learning, Classification*

ACKNOWLEDGEMENT

It is an immense pleasure to express our sincere gratitude to our project supervisor Dr. Rajani Chulyadyo and project coordinator Dr. Gajendra Sharma for providing their valuable guidance, support and useful suggestions throughout the course of this project. The moral support and continuous encouragement we were provided with, enabled us to complete our work successfully.

We are also thankful to Kathmandu University and the Department of Computer Science and Engineering (DoCSE) for providing us with this wonderful opportunity to work on this project.

SYMBOLS AND ABBREVIATIONS

Abbreviation	**Full Form**
NTHU-DDD	National Tsuing Hua University-Driver Drowsiness Detection
UTA-RLDD	University of Texas at Arlington Real-Life Drowsiness Dataset
YawDD	Yawning Detection Dataset
KNN	K-Nearest Neighbors
SVM	Support-Vector Machines
FDS	Fatigue Detection System
SEBR	Spontaneous Eye Blink Rate
EAR	Eye Aspect Ratio
MAR	Mouth Aspect Ratio
IoT	Internet of Things

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Background

Road accidents are a major cause of death of people around the world. Approximately 1.35 million people die each year as a result of road traffic crashes (World Health Organization, 2018). In Nepal alone, 22,461 lost their lives in road traffic accidents in the last 10 years, according to Nepal Police (Gartoula, 2019). According to Nepal Police data, 8,461 cases of accidents among 10,965 occurred due to driver's negligence in the last fiscal year (Paudel, 2019). We are aware that major causes of road traffic accidents are poor road conditions (in case of Nepal), over speeding, drunk driving, mechanical failure, fatigue and drowsiness, etc. While most people are aware of the dangers of driving while intoxicated, many do not know that drowsiness also impairs judgment, performance and reaction times just like alcohol and drugs.

In today's fast paced world, sleep and rest have been highly neglected. Sleep deprivation has become a major issue among people in most of the countries today. This is one of the reasons why drowsy driving has become a serious problem. According to a 2014 Centers for Disease Control and Prevention survey among nearly 150,000 adults in 19 states and the District of Columbia, about one in 25 adults reported that they had fallen asleep while driving at least once in the previous 30 days (Wheaton et al., 2013). According to National Sleep Foundation surveys, half of American adults consistently report that they have driven drowsy and approximately 20% admit that they have actually fallen asleep at the wheel in the previous year. The National Highway Traffic Safety Administration conservatively estimates that 100,000 police reported crashes each year are caused primarily by drowsy driving and that such crashes result in more than 1,550 deaths, 71,000 injuries and $12.5 billion in monetary losses (Drowsy Driving, nd.). And all of this is valid, all over the world.

All these statistics suggest that nowadays drivers are often distracted or driving under fatigue and with drowsiness leading to fatal accidents. So, it is essential to develop a system that detects driver drowsiness and consecutively decrease the number of road accidents.

1.2 Problem Statement

Drowsy driving is not just falling asleep at the wheel - it is a profound impairment resembling someone driving under influence. Drowsiness impairs judgment, processing and decision making and leads to slower reaction times. A driver shows common signs of drowsiness like sleepy eyes, yawning or blinking frequently. So, it is important to detect the signs of drowsiness in the driver early, so that measures can be applied to prevent any sort of accidents from happening.

1.3 Objectives

The objectives of the project are as follows:

- To develop a software which detects the signs of drowsy driving.
- To collect data / features from videos of people showing signs of drowsiness which can be used as representations of drowsy driving.
- To feed the features, and collected data into various classification algorithms to obtain a classification model.

1.4 Motivation and Significance

In today's competitive world, there is very less time for people to settle down and take proper rest. There are lots of youth who work till late night having to drive to their office early in the morning. Similarly, there are lots of late night bus drivers, responsible to safely deliver passengers to their destination. These types of people are most vulnerable to meeting accidents while driving under fatigue and drowsiness. This problem is not going anywhere for some time. So we decided it would be great to have a drowsiness detection system which would eventually help people from meeting unfortunate accidents. Our classification model has real world application, and can be deployed / implemented on IoT devices on vehicles to prevent accidents and reduce deaths due to road traffic accidents all over the world.

CHAPTER 2: RELATED WORKS

2.1 Fatigue Detection System

The Fatigue Detection System (FDS) monitors the driver by using a mobile phone camera; the phone will be put on a stand in the car to make the driver feel comfortable(no wearables) and record video of the driver then process it for real time eye tracking. This system has fast and real time face and eye tracking, external illumination interference is limited, more robustness and accuracy allowance for fast head movement. It counts the number of eye blinks and compares them with a threshold value for a certain range of frames, then determining whether the driver is feeling drowsy if eye blink count exceeds the threshold calculated (Abulkhair et al., 2015). It gives highly accurate results but solely depends on the number of times a driver blinks, which can be not correct for people with high spontaneous eye blink rate (SEBR). Also there are other indicators like yawning, slow eyelid movement, frequent nodding, fixed gaze, sluggish facial expression, and drooping posture that should be taken into account.

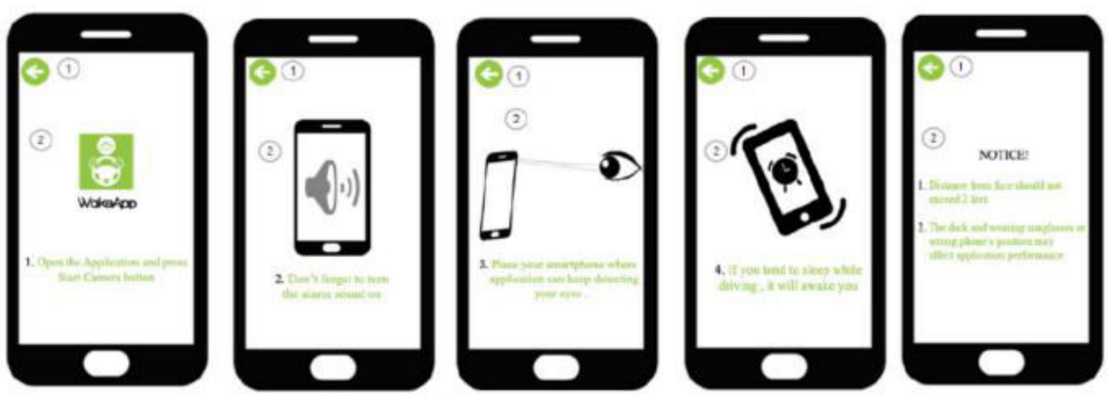


Figure 2.1.1: Fatigue detection system

2.2 SomnoAlert

The Caluna Bioengineering Institute and the Ficosa corporation developed SomnoAlert which is capable of detecting the early signs of drowsiness in the driver (Somnoalert, 2017). The creators have taken advantage of inertial sensors which incorporate the phone - the same that indicates the position of the screen - and the GPS data it provides to detect the characteristic movements of fatigue or drowsy drivers, such as lane deviation or sudden steering wheel corrections, which would occur when falling asleep, then the application emits an acoustic signal and a red display (GPS World Staff, 2016). The application is a prototype that is not highly efficient and currently works for iOS only.

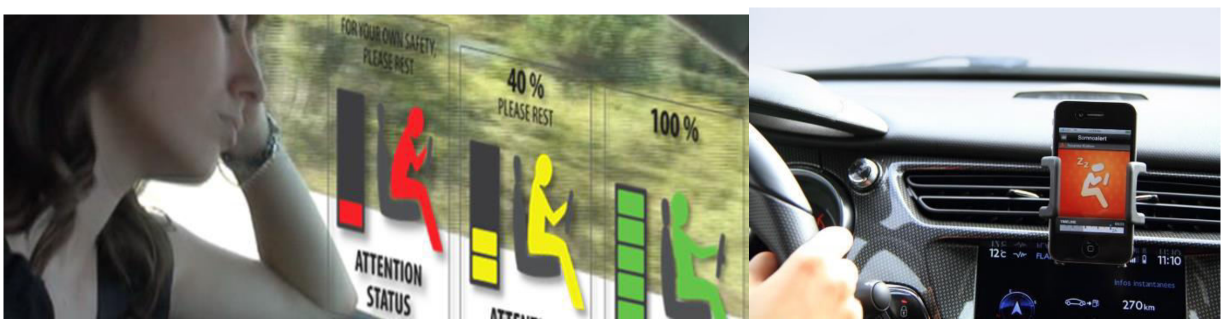


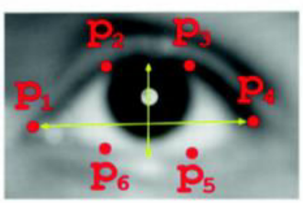
Figure 2.2.1: SomnoAlert

CHAPTER 3: DROWSINESS DETECTION

3.1 System Overview

3.1.1 Algorithm for data collection

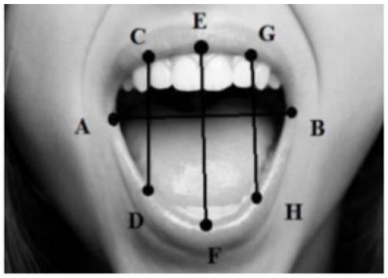
1. Load the saved video from the folder.
2. Start processing the frames.
3. Localize face.
4. If the face is detected then
   a. Apply facial landmark detection.
   b. Extract *(x, y)*-coordinates for both the left and right eye and mouth.
   c. Calculate eye aspect ratio (EAR) to determine if eyes are closed by computing the distance between the two sets of vertical eye landmarks and horizontal eye landmarks. The used formula is:



$$\mathrm{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Figure 3.1.1.1: EAR Formula

   d. Calculate mouth aspect ratio (MAR) to detect yawning by computing the distance between the two sets of vertical mouth landmarks and horizontal mouth landmarks. The used formula is:



$$MAR = \frac{|CD| + |EF| + |GH|}{3|AB|}$$

Figure 3.1.1.2: MAR Formula

   e. Save the value of EAR and MAR for each video in a list and then to a csv file.

5. Repeat the above steps until all the videos in the folder are read.

6. Stop

3.2 Experimental Study

3.2.1 Tools

These are the tools that were used while developing this drowsiness detection model, and these very tools must be available if one wishes to use or test the model prepared.

1. OpenCV

   OpenCV is a library of programming functions mainly aimed at real-time computer vision. Here, it is used to capture the video from the dataset, draw facial landmarks in the video and display the calculated EAR & MAR for each frame on the video.

2. Dlib

   Dlib is a modern C++ toolkit containing machine learning algorithms and tools for creating complex software to solve real world problems. Here, it is used to localise the face in the video and detect the key facial structures on the face using the dlib 68 landmarks predictor.

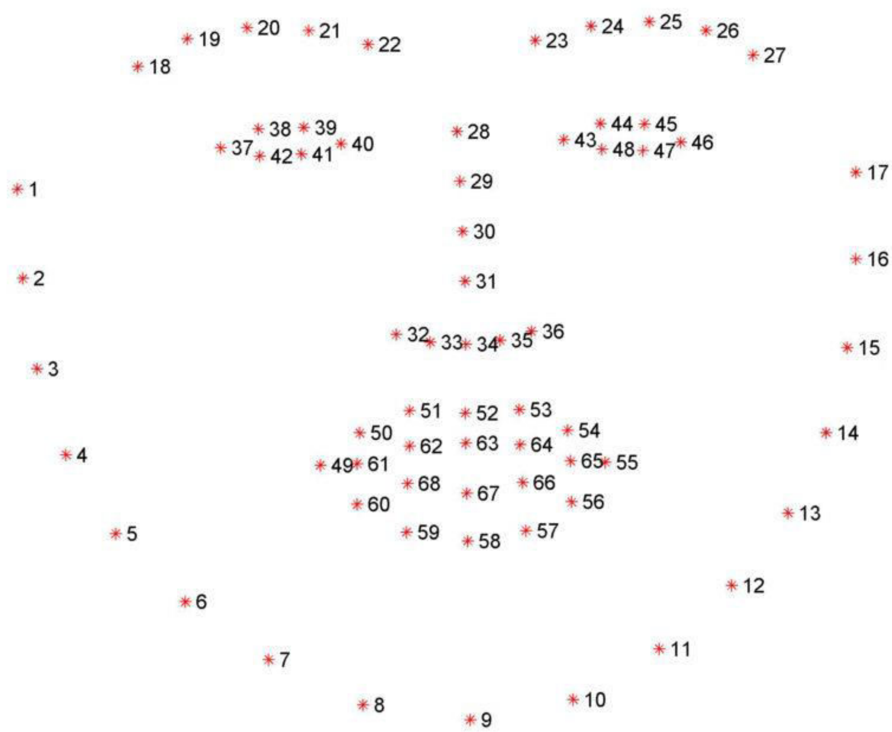


Figure 3.2.1.1: Dlib 68 facial landmarks

3. Numpy & Matplotlib

   NumPy is a library for Python to handle large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. Matplotlib is a plotting library for Python and its numerical mathematics extension NumPy. It is used here for visualizing the data.

4. Pandas

   Pandas is a library for Python to manipulate and analyse data. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is used here to work with data extracted from the videos.

5. Scikit-learn

   Scikit-learn (also known as sklearn) is a software machine learning library for Python. It features various classification, regression and clustering algorithms including support vector machines, random forests, k-means and many more. Sklearn integrates well with many other Python libraries, such as matplotlib for plotting, numpy for array vectorization and pandas dataframes.

3.2.2 Datasets

The model has been trained and tested on the selected videos from the following datasets that are specifically made for research and analysis purposes.

1. NTHU-DDD:

   The National Tsuing Hua University-Driver Drowsiness Detection (NTHU-DDD) video dataset (including training, evaluation, testing dataset) collected by NTHU Computer Vision Lab contains 36 subjects of different ethnicities recorded with and without glasses/sunglasses under a variety of simulated driving scenarios, including normal driving, yawning, slow blink rate, falling asleep, burst out laughing, etc., under day and night illumination conditions.

The total time of the entire dataset is about 9 and a half hours. The training dataset contains 18 subjects with 5 different scenarios (BareFace, Glasses, Night BareFace, Night Glasses, Sunglasses). The sequences for each subject including yawning and slow blink rate with nodding are each recorded for about 1 minute long. The sequences corresponding to two most important scenarios, combination of drowsiness-related symptoms (yawning, nodding, slow blink rate) and combination of non-drowsiness related actions (talking, laughing, looking at both sides) are each recorded for about 1.5 minutes long. The evaluation and testing datasets contain 90 driving videos (from the other 18 subjects) with drowsy and non-drowsy status mixed under different scenarios.

The subjects are recorded while sitting on a chair and playing a plain driving game with simulated driving wheels and pedals; meanwhile, being instructed to perform a series of facial expressions. NTHU has used an active infrared (IR) illumination to acquire IR videos in the dataset collection. The video resolution is 640x480 in AVI format. The videos are captured at 15/30 frames per second (Weng et al., 2016).

2. UTA-RLDD:

The University of Texas at Arlington Real-Life Drowsiness Dataset (UTA-RLDD) was created for the task of multi-stage drowsiness detection, targeting not only extreme and easily visible cases, but also subtle cases when subtle micro-expressions are the discriminative factors. UTA-RLDD dataset is the largest to date realistic drowsiness dataset.

The RLDD dataset consists of around 30 hours of a total of 180 RGB videos of 60 participants. For each participant, there is one video of around 10 minutes long, and is labeled as belonging to one of three classes: alert (labeled as 0), low vigilant (labeled as 5) and drowsy (labeled as 10). All participants are over 18 years old. There are 51 men and 9 women, from different ethnicities (10 Caucasian, 5 non-white Hispanic, 30 IndoAryan and Dravidian, 8 Middle Eastern, and 7 East Asian) and ages (from 20 to 59 years old with a mean of 25). The subjects are wearing glasses in 21 of the 180 videos, and have considerable facial hair in 72 out of the 180 videos.

Videos are taken from roughly different angles in different real-life environments and backgrounds. Each video is self-recorded by the participant, using their cell phone or web camera. The frame rate is always less than 30 fps. The 60 subjects were randomly divided into five folds of 12 participants, for the purpose of cross validation. The dataset has a total size of 111.3 Gigabytes (Ghoddoosian, et al., 2019).

3. YawDD:

Yawning Detection Dataset (YawDD) of videos is recorded by an in-car camera, of drivers in an actual car with various facial characteristics (male and female, with and without glasses/sunglasses, different ethnicities) talking, singing, being silent, and yawning. It can be used primarily to develop and test algorithms and models for yawning detection, but also recognition and tracking of face and mouth. The videos are taken in natural and varying illumination conditions.

The videos come in two sets, in the first set, the camera is installed under the front mirror of the car. This set provides 322 videos, each for a different situation: normal driving (no talking), talking or singing while driving, and yawning while driving. Each subject has 3 or 4 videos. In the second set, the camera is installed on the driver's dash. This set provides 29 videos, one for each subject, and each video contains all of driving silently, driving while talking, and driving while yawning.

The video resolution is 640x480 of 24-bit true color (RGB) in AVI without audio. The frame rate is 30 frames per second. The dataset has a total size of 5.1 Gigabytes (Abtahi et al., 2014).

3.2.3 Dataset Selection

The selection process was manual, only clearly drowsy-state videos of people with/without glasses in different lighting conditions were selected for the drowsy part and similarly for the non drowsy videos. From the above datasets, for training and testing, we have selected videos which are listed below with their characteristics:

Dataset	Duration per video	Number of People	Gender	Drowsy/ Not Drowsy	Total videos	Glasses
RLDD	10mins	2	M	D + ND	4	Y
RLDD	10mins	2	M	D + ND	4	N
NTHU	1-2 mins	9	M	D + ND	18	N
NTHU	1-2 mins	2	F	D + ND	4	N
YawDD	1-2mins	5	M	D	5	N
YawDD	1-2mins	5	F	D	5	N

Table 3.2.3.1: Dataset for Rennervate

3.2.4 Data Filtering

The videos are labelled as drowsy or not drowsy in the datasets. These videos are fed into dilib library through OpenCV. EAR and MAR are extracted every second from the videos and stored as .csv file. Since the frames were taken at an interval of one second, the particular frame may be not drowsy though it may be from the drowsy labelled video and vice versa. This creates inconsistencies in the data set, and must be filtered. An interval was chosen according to video data structures. For Drowsy State, minimum value of EAR and maximum value of MAR in that interval is taken into account for cleaned dataset. In case of Not Drowsy State, maximum value of EAR and minimum value of MAR was considered.

Here below is the visual representation of the data before and after filtering:

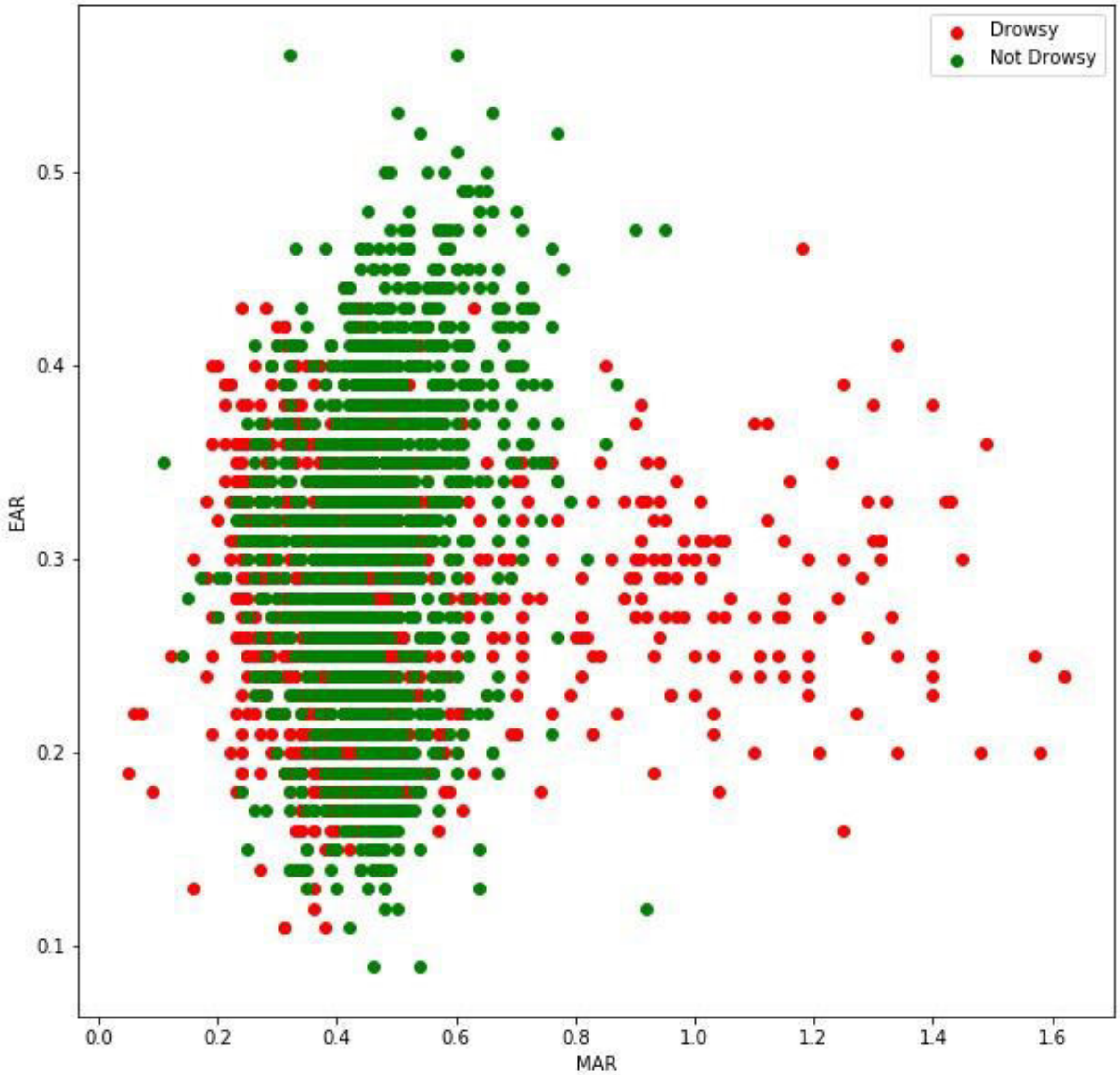


Figure 3.2.4.1: Unfiltered Data

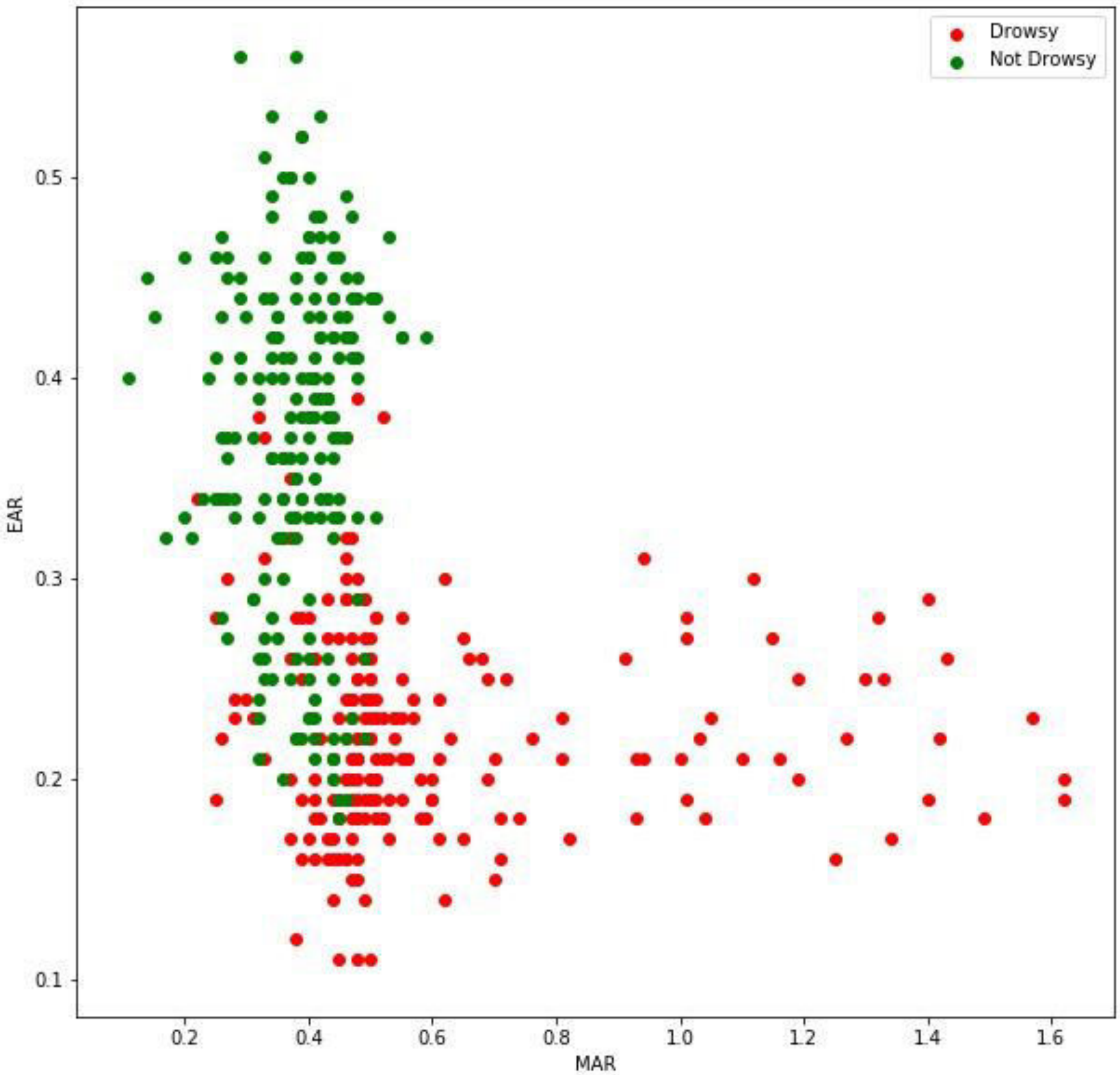


Figure 3.2.4.2: Filtered Data

3.2.5 Classification Algorithms

The obtained final data is divided into training data (70%) and testing data (30%) and a model is prepared. This process is done by using the train_test_split() function from Sklearn. The six different algorithms are used for preparing the model using the training data. These algorithms are:

a) Logistic Regression

b) Decision Tree

c) Random Forest

d) Naive Bayes

e) K-Nearest Neighbor (KNN)

f) Support Vector Machine (SVM)

The testing dataset is used to test the accuracy of different models prepared using different algorithms.

CHAPTER 4: DISCUSSION AND ACHIEVEMENTS

4.1 Discussion

KNN model tops all other models with an accuracy of 89.32%. The performances of all models from the used algorithms is shown below in bar graph:

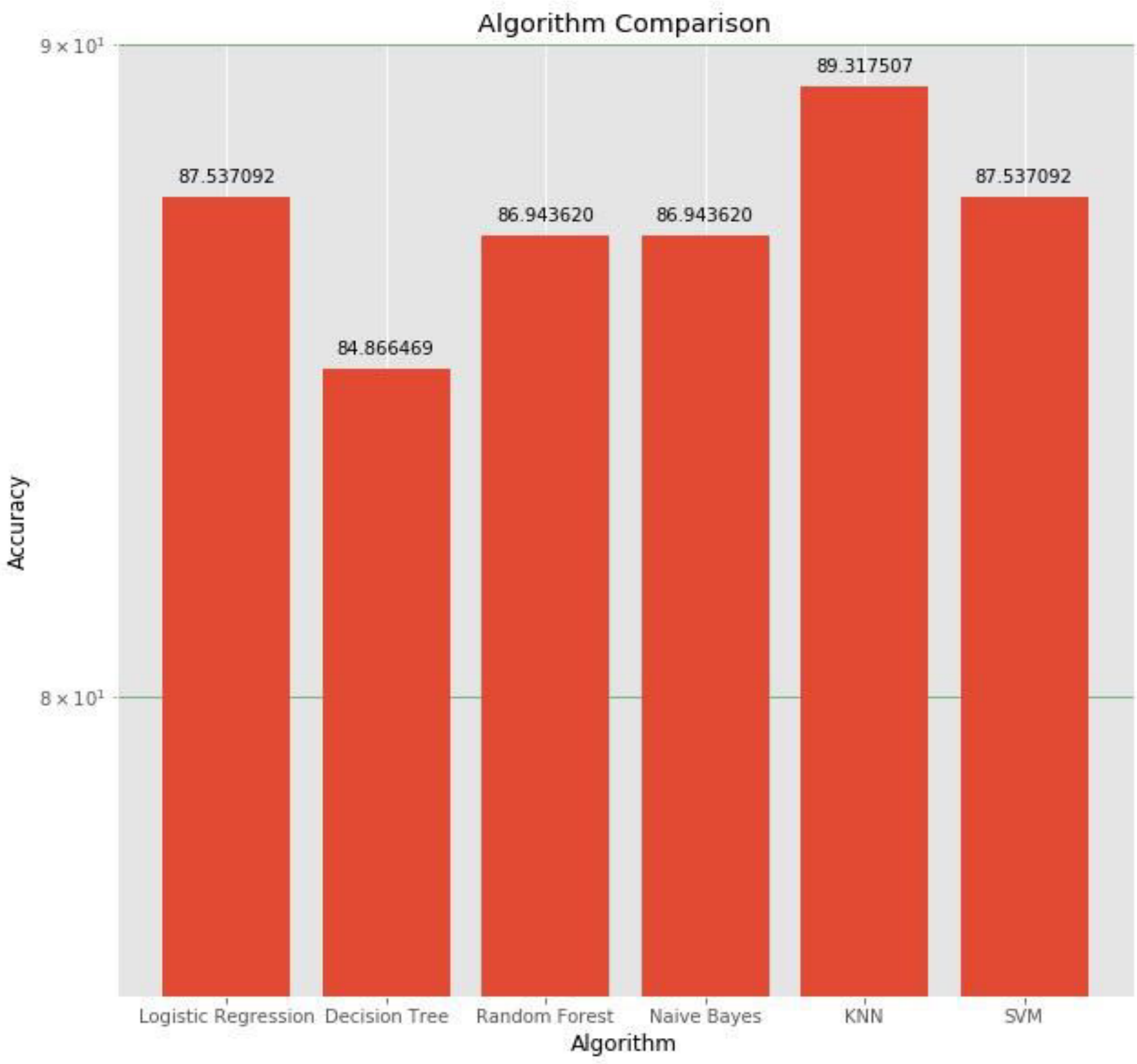


Figure 4.1.1: Model Performance in Different Algorithms

4.2 Achievements

The data which are not used while preparing the model was used to test the efficiency of the trained model. These independent data features were used to predict the target using the KNN model since it was most accurate. The testing was done for four different person's drowsy and not drowsy videos and the model performance is as shown below:

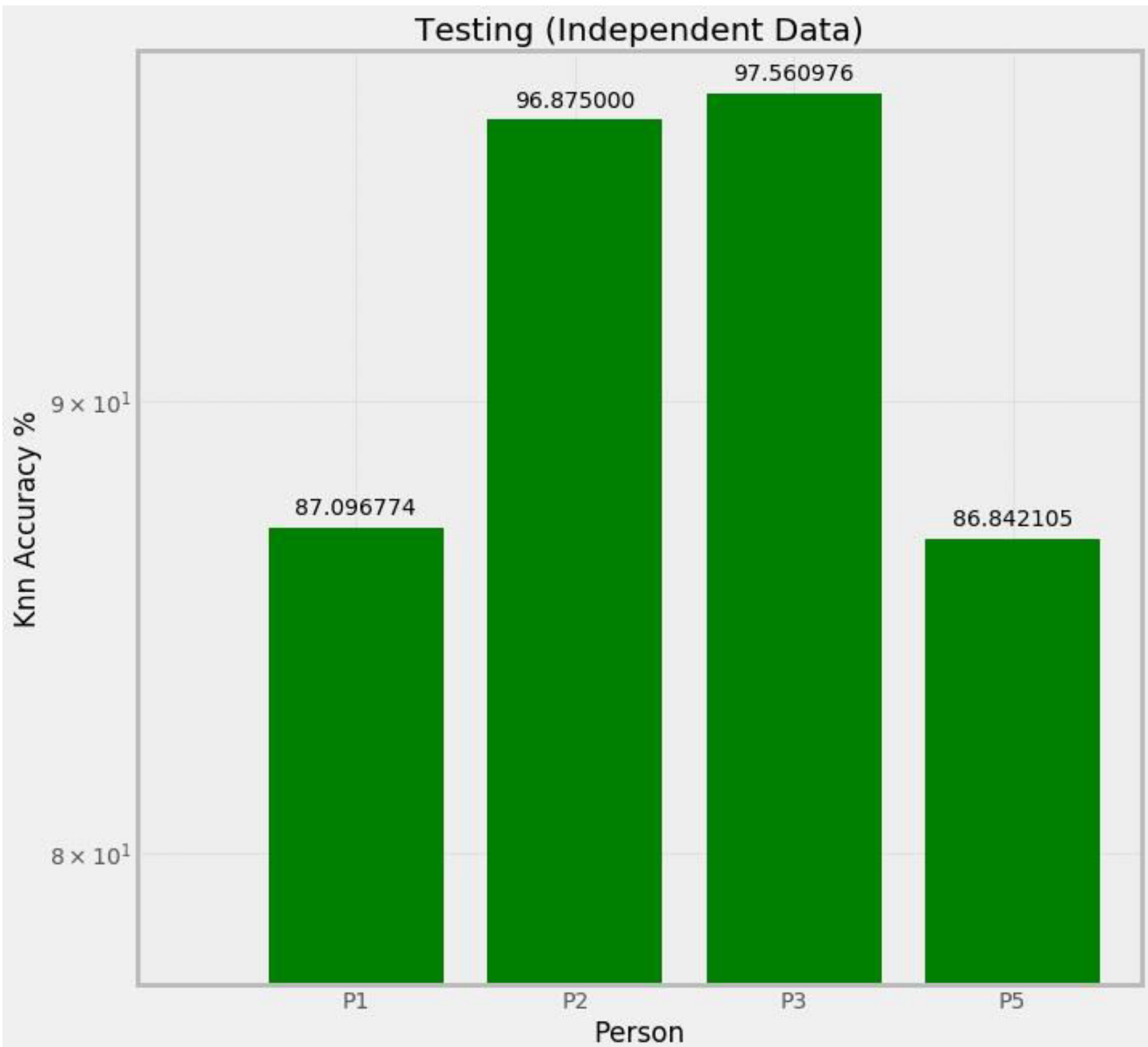


Figure 4.2.1: Independent Data Testing

CHAPTER 5: CONCLUSION AND RECOMMENDATION

5.1 Conclusion

As we all know, drowsy driving can lead to road accidents and later on death as well. All around the world, lots of people lose their lives due to drowsy driving. So, in case of drowsiness it is much better to detect it early and alert the driver. For that purpose, we have built a drowsiness detection model. The data of EAR and MAR collected from videos of people in different lighting conditions were extracted as features. Those features were fed into various classification algorithms like random forest, logistic regression, Naive Bayes, KNN, decision tree. Their accuracies were evaluated and based on that we have concluded that a KNN based classification model is the best for drowsiness detection.

The model we prepared has real world applications. It can be used in mobile phone applications or implemented in IoT devices to build systems for real time drowsiness detection. It can be used to detect drowsiness states of drivers and alert them beforehand, to prevent any fatal accidents from ever happening.

5.2 Limitation

Rennervate has following limitations:

1. In the data collection phase, only two features - EAR and MAR were extracted and used.
2. Face detection and landmark detection only work when the person is directly facing the camera and do not work otherwise.
3. Does not work for saved videos which are rotated, i.e. not upright.

5.3 Future Enhancement

1. In the future, we will be considering more input features than EAR and MAR for training the model. We have thought of including features like angle of head titling, blink frequency, etc.
2. We will be implementing the model ourselves into a proper drowsiness detection system.

REFERENCES

Abtahi, S., Omidyeganeh, M., Shirmohammadi, S., Hariri, B.. (2014). YawDD: A yawning detection dataset. Proceedings of the 5th ACM Multimedia Systems Conference, MMSys 2014. 24-28. 10.1145/2557642.2563678.

Abulkhair, M., Alsahli, A. H., Taleb, K. M., Bahran, A. M., Alzahrani, F. M., Alzahrani, H. A., & Ibrahim, L. F. (2015). Mobile Platform Detect and Alerts System for Driver Fatigue. Procedia Computer Science, 62, 555–564. https://doi.org/10.1016/j.procs.2015.08.531

Drowsy Driving. (nd.). National Safety Council. https://www.nsc.org/road-safety/safety topics/fatigued-driving

Gartoula, G. (2019, January 18). Nepal's deadly roads take their toll. Nepali Times. https://www.nepalitimes.com/banner/nepals-deadly-roads-take-their-toll/. GPS World Staff. (2016, April 25). Wake up! Smartphone App Aims to Alert Drowsy Drivers. GPS World. https://www.gpsworld.com/wake-up-smartphone-app-aims-to-alert-drowsy drivers/

Ghoddoosian, R., Galib, M. and Athitsos, V., (2019). A Realistic Dataset And Baseline Temporal Model For Early Drowsiness Detection. [online] arXiv.org. Available at: <https://arxiv.org/abs/1904.07312> [Accessed 28 September 2020].

GPS World Staff. (2016, April 25). *Wake up! Smartphone App Aims to Alert Drowsy Drivers*. GPS World. https://www.gpsworld.com/wake-up-smartphone-app-aims-to-alert-drowsy-drivers/

Paudel, N. (2019, February 10). Police to launch app aimed at making night buses safer. The Kathmandu Post. https://kathmandupost.com/valley/2019/02/10/police-to-launch-app aimed-at-making-night-buses-safer.

Somnoalert. (2017, May 31). Ficosa. https://www.ficosa.com/products/adas/somnoalert/

Weng, C.H., Lai, Y.H. & Lai, S.H. (Nov. 2016). “Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network”, In Asian Conference on Computer Vision Workshop on Driver Drowsiness Detection from Video, Taipei, Taiwan.

Wheaton, A. G., Chapman, D. P., Presley-Cantrell, L. R., & Croft, J. B. (2013). Drowsy Driving—19 States and the District of Columbia, 2009-2010. Jama, 309(8), 760. https://doi.org/10.1001/jama.2013.105

World Health Organization. (2018). Global status report on road safety 2018 (ISBN 978-92-4-156568-4).

APPENDIX

Gantt Chart

We divided our whole project into seven basic tasks and planned it to complete within sixteen weeks. The timeframe for this project is divided as:

ID	Title	Week										
		1	2	3	4	5	6	7	8	9	10	11
1	Literature review											
2	Datasets collection											
3	Coding											
4	Testing & debugging											
5	Documentation											

Figure A: Gantt Chart

Project Started: July 12, 2020

Project Deadline: September 27, 2020