Drowsiness Detection

General Assembly DSI 31
Khoo Qi Xiang

TABLE OF CONTENTS

01

INTRODUCTION

Contents:

1. Problem Statement
2. Methodology

INTRODUCTION

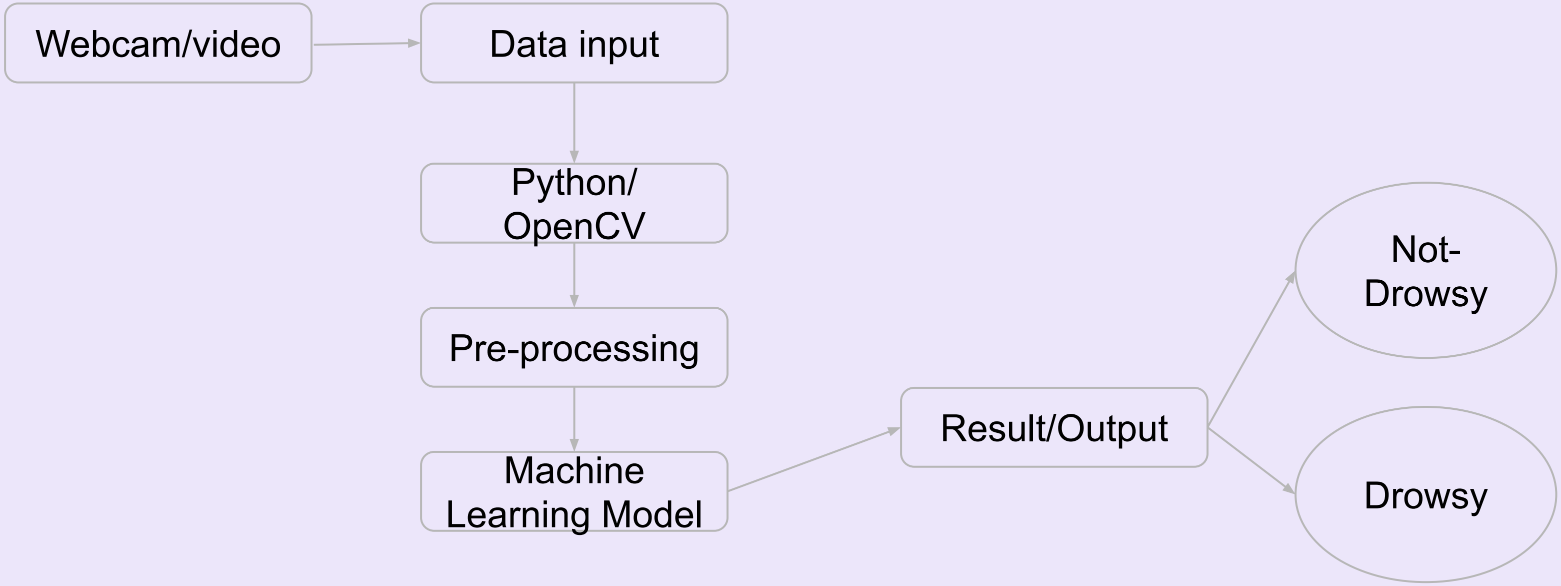
Singaporeans rank among the world's sleepiest people. (in 2022)
In fact, Singapore ranks first among the 43 cities evaluated in a recent survey for having the least amount of sleep.

Drowsy driving is a factor in 1 in 4 auto accidents, and 1 in 25 adult drivers say they have dozed off behind the wheel in the previous 30 days. (US Study)

Problem Statement

- To develop a prototype drowsiness detection system that can correctly and instantly track whether a driver is drowsy or not based on some of the facial features.
  As attempting to identify signs of drowsiness early on will help reduce the incidence of accidents if the driver is informed of them.

Methodology



02

Dataset

Contents:

1. Data Cleaning
2. Feature Engineering
3. EDA

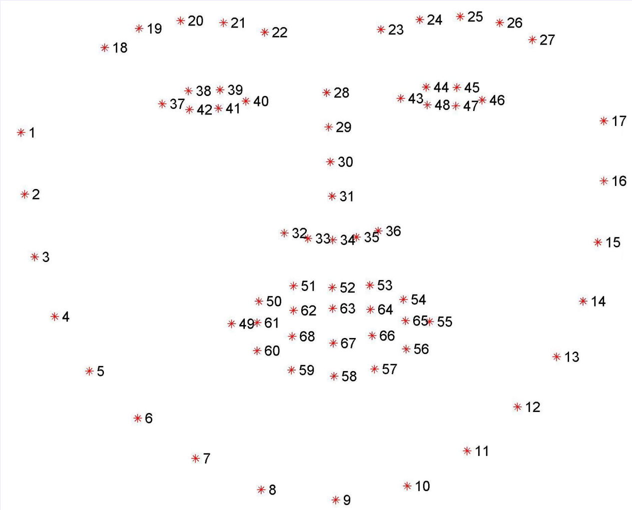
Dataset

The dataset consists of around 30 hours of videos of 61 unique participants including myself.

There are three categories for each 10-minute clip: Alert; Low vigilant; and Drowsy

Using openCV , was able to extract facial landmarks from 45 videos of 23 participants

Predefined Functions



Dlib:

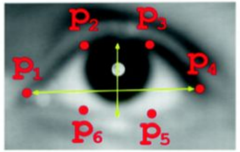
- It’s a landmark’s facial detector with pre-trained models.

- Dlib is used to estimate the location of 68 coordinates (x,y) that map the facial points on a person’s face.

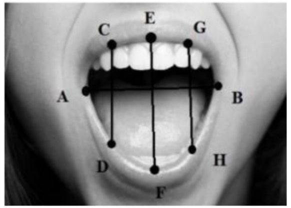
Predefined Functions

Eye Aspect Ratio (EAR):
- The proportion of eye length to eye width.
- Two separate vertical lines are averaged across the eyes to determine their length
-

Mouth Aspect Ratio (MAR):
- The proportion of mouth length to mouth width.
- Similar to EAR, two separate vertical lines are averaged across the mouth to determine their length



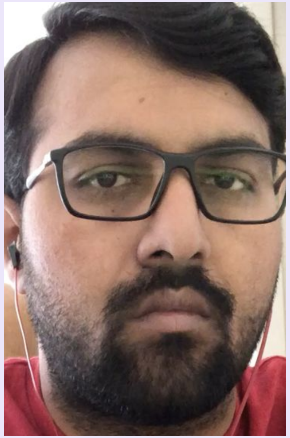
$$\text{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

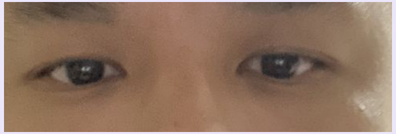


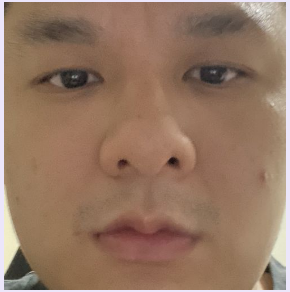
$$\text{MAR} = \frac{|\text{EF}|}{|\text{AB}|}$$

Data-Cleaning

Eyes open

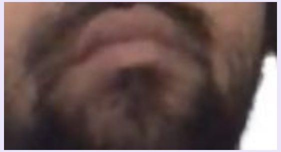


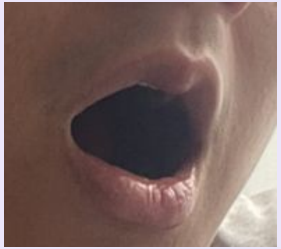




Data-Cleaning

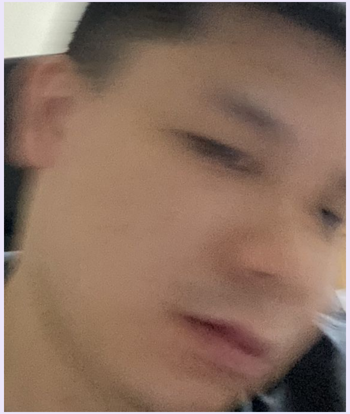
Eyes slightly close

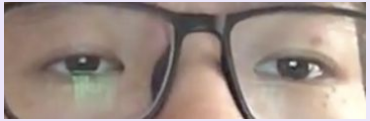




Data-Cleaning

Challenges




EDA

- Each individual has different core features in their default alert state.
- Different lighting conditions may disrupt the detection performance
- EAR will be small while MAR will be big

Feature Engineering

MAR / EAR (MOE):

- Simply the ratio of the MAR to the EAR.
- More responsive to changes than EAR and MAR

Pupil Circularity (PUC):

- A measure complementary to EAR but it places a greater emphasis on the pupil instead of the entire eye.
- Similar to EAR, it was hypothesized that tiredness would result in a decrease in a person's pupil circularity.

$$MOE = \frac{MAR}{EAR}$$

Mouth Over Eye Ratio (MOE)

$$Circularity = \frac{4 * \pi * Area}{perimeter^2} \qquad Area = \left(\frac{\text{Distance(p2, p5)}}{2}\right)^2 * \pi$$

$$\begin{aligned} Perimeter = {} & Distance(p1,p2) + Distance(p2,p3) + Distance(p3,p4) + \\ & Distance(p4,p5) + Distance(p5,p6) + Distance(p6,p1) \end{aligned}$$

03

Model

Contents:

1. Model
2. Live Demo
3. Evaluation

Models used

1. Logistic Regression
2. Naive Bayes
3. KNN (K-Nearest Neighbors)
4. MLP
5. Decision Tree
6. Random Forest
7. CNN (Convolutional Neural Network)
8. XG Boost

Results

Models / Metrics	Accuracy
Logistic Regression	58.8%
Naive Bayes	54.6%
KNN	60.2%
MLP	61.4%
Decision Tree	62.4%
Random Forest	60.2%
CNN	61.0%
XG Boost	62.2%

Results (After normalising)

Models / Metrics	Accuracy	F1 Score	ROC
Logistic Regression	0.888	0.904	0.933
Naive Bayes	0.865	0.886	0.930
KNN	0.877	0.892	0.936
MLP	0.881	0.895	0.941
Decision Tree	0.900	0.913	0.938
Random Forest	0.890	0.902	0.959
CNN	0.879	0.894	0.50
XG Boost	0.940	0.948	0.975

Further dive in

Models / Metrics	Accuracy	F1 Score	ROC
MLP	0.881	0.895	0.941
KNN	0.877	0.892	0.936
CNN	0.879	0.894	0.50

Final Model

MLP

83% Recall

94% ROC

Features	Random Forest	XGBoost
MOE_N	31.21%	37.16%
MOE	11.30%	14.24%
MAR_N	18.14%	13.55%
MAR	8.10%	8.72%
EAR_N	12.85%	7.52%
EAR	9.78%	8.21%
Circularity_N	3.70%	5.77%
Circularity	4.87%	4.79%

Live Demo

04

CONCLUSION

Contents:

Conclusion

Model Complexity

When performing tasks, less complicated models can be just as effective as more complex ones.But ultimately, it is better to employ the more sophisticated model with a lower false-negative rate than a simpler model.

Feature Importance

We understood that everyone's baseline for the eye and mouth aspect ratios is different which was why normalization was crucial to our performance.

Flexibility of the system

System permits different models to predict the outcome, which could be useful because we could use different models to predict certain attributes based on the system's accuracy.

Limitation

More different training data

For the models, more diverse images taken in a range of seating arrangements and lighting conditions would have been desirable.

Environment

There cannot be any reflective materials behind the driver, which is another limitation. The system becomes more reliable as the background becomes more uniform.

Future Enhancement

More Complex Models

By incorporating more sophisticated models, like transfer learning and LSTM, to see how I can add value to the system. After all, these models might have a lower false negative rate than our current one.

Behavior feature

Like sudden head movements, hand motions, head tilting (posture), or even monitoring eye movements, these are new, distinct signs of tiredness.

Fine-tune Model

How the false-negative rate for MLP and other straightforward models can be reduced.

THANK YOU!

Annex

Standardization is another scaling technique where the values are centered around the mean with a unit standard deviation. This means that the mean of the attribute becomes zero and the resultant distribution has a unit standard deviation.
Here’s the formula for our normalised feature/ standardization:

$$\textbf{Normalised Feature}_{n,m} = \frac{\textbf{Feature}_{n,m} - \mu_{n,m}}{\sigma_{n,m}}$$

where:
n is the feature
m is the person
$\mu_{n,m}$ *and* $\sigma_{n,m}$ *are taken from the first 3 frames of the “Alert” state*

Annex

Model/ Confusion Matrix	True Negative (TN)	False Positive (FP)	False Negative (FN)	True Positive (TP)
MLP	711	26	198	956
KNN	703	34	197	957
CNN	706	31	196	958
XG Boost	732	5	108	1046

MLP

Multilayer Perceptron (MLP)

A multilayer perceptron (MLP) is a feedforward artificial neural network that generates a set of outputs from a set of inputs. An MLP is characterized by several layers of input nodes connected as a directed graph between the input and output layers. MLP uses backpropogation for training the network. MLP is a deep learning method.

A multilayer perceptron is a neural network connecting multiple layers in a directed graph, which means that the signal path through the nodes only goes one way. Each node, apart from the input nodes, has a nonlinear activation function. An MLP uses backpropagation as a supervised learning technique. Since there are multiple layers of neurons, MLP is a deep learning technique.

MLP is widely used for solving problems that require supervised learning as well as research into computational neuroscience and parallel distributed processing. Applications include speech recognition, image recognition and machine translation.