AI for Safer Roads: A UK Study on Accident Severity Prediction and Driver Drowsiness Detection

Brikelda Sema
Leeds Beckett University, Leeds, UK

***Abstract*—Leveraging Artificial Intelligence (AI)'s recent advancements, this paper investigates how AI-powered driver assistance systems can be optimized to prevent car accidents, from a comprehensive analysis of UK collision data. Conducted using Google Colab, this research is divided into three parts: The first part focuses on a comparative analysis of Machine Learning Classification Models for predicting accident severity. Utilizing a UK collision dataset, we evaluate various models and identify XGBoost as the most effective for this task.The second part involves using ML Clustering Model DBSCAN to do spacial and spaciotemporal analysis in identifying accident hotspots in the city of Leeds. The last part delves deeper into human behavior, employing Deep Learning to develop a Driver Drowsiness Detection Mechanism. This mechanism analyzes eye position to determine drowsiness, which can be used afterwards to trigger an alert system. The study aims to contribute to the ongoing research efforts towards AI-powered car accident prevention systems and evaluate the best approaches that achieve the highest results.**

***Index Terms*—Machine Learning, Deep Learning, Artificial Intelligence, Accidents, Driver Drowsiness Detection, Classification, Clustering.**

I. Introduction

TRAFFIC safety has become an case of considerate importance during the last decades, this mostly related to and dependant on the driver's behaviour. Even though the UK has been near the lowest places in the world for road collisions (5th out of 38 countries) [7], the amount of fatalities is still considerable: 1,711 fatalities in total or 5 road fatalities per billion vehicle miles travelled for 2022 [7]. However, research initiatives are currently underway with the aim of using AI in lowering the number of car accidents and casualties in a smart and sustainable way.

A. Background of Context

Car accidents happen from different causes and they might have different levels of severity, depending on the number of casualties, level of damage or costs. The severity of an accident can also be related on weather conditions, whether it was a heavy rain making it difficult for the driver to follow the road; it can depend on light conditions, when low levels of light make it difficult for the driver to see road signs or other vehicles; it can be road structure, where a damaged road surface can lead into loss of control of the car, and so on.

Artificial Intelligence can help in determining prediction models for different target features, such as the severity of the accident, based on the factors that have a certain impact on it, as mentioned above. It can also go deeper and imitate the human brain by analysing the driver's behaviour to prevent any accidents.

For this case, different Machine Learning (ML) classification models were trained to predict in reasonable accuracy the severity of a car accident, clustering models to identify risky regions and Convolutional Neural Networks (CNN) were used to create a Driver Drowsiness Detection model that identifies if the driver is feeling drowsy based on eye positioning. This can alert the driver and help in decreasing the possibilities of an incident.

B. Aim and Research Objectives

The aim of this research is to train and do comparative analysis on different ML models that predict the severity of an accident and find hotspots of accidents based on a given dataset and its features, as well as applying Deep Learning in detecting driver's drowsiness using a given set of eye images. The research objectives are:

1) Perform literature review on UK car accidents, their causes and prevention mechanisms.
2) Select a dataset for UK accidents for accident severity predictions accidents areas location, and another dataset containing images of closed and opened eyes to train the CNN model for drowsiness detection.
3) Design experiments to test the models and evaluate their accuracy on prediction.
4) Conduct the experiments, capture the data and apply different strategies to improve the models.
5) Analyse the results obtained, create a discussion and give recommendations for future work.

C. Rationale

The rationale of this research is to identify the ML classification model with the highest accuracy in predicting the severity of an accident in a given test set of data, locate potential accident areas using clustering, as well as improving the CNN model for driver drowsiness detection by applying data augmentation (such as high and low exposure images), to enhance its practicality and better simulate real-world conditions.

D. Contribution of the Research

This research contributes on the effect of AI in human safety and accident prevention by using Deep Learning algorithms for detecting abnormal behaviours and adapt the training and test set as close as possible to real life situations.

E. Organisation of the Paper

The paper is organised in this way: The Literature Review section provides an overview of car accidents in the UK, statistics, as well as factors that impact these accidents. It incorporates the role of drowsiness as a contributor to the probabilities of collision. The Methodology section describes the macro and micro methodology used in this research, describing the experimental design and its implementation step-by-step. The Findings and Discussion section presents the results of the experiments and interprets these results. To finish with the Conclusion and Future Work section where the key findings of this research are summarized and recommendation for future improvements are given.

II. Literature Review

A. Road Accidents in the UK

Several studies and surveys have been conducted to explore the underlying reasons for car accidents in the UK. Some of these studies ([3], [10]) have identified key contributing factors:

- Excessive speed
- Misjudgement and perceptual errors in 'right of way'
- Alcohol or drug involvement
- Fatigue, tiredness or lack of sleep
- Loss of Control

These impact of these factors vary significantly, depending on age for example. Young drivers have the great majority of their accidents by losing control on bends or curves, showing high levels of speeding, alcohol involvement and recklessness, whereas older drivers had fewer accidents, but those fatalities they were involved in tended to associate misjudgement and perceptual errors in 'right of way' collisions [3].

Another important factor contributing in accidents is tiredness or fatigue. Police statistics show that fatigue contributes to about 4% of fatal road crashes and 2% of all collisions in Britain. However, it is likely that these figures are far higher since fatigue is hard to be detected and, unlike alcohol and drugs, police can't test for tiredness. Worldwide, it is estimated that between 10% and 20% of all road crashes are fatigue-related [2]. Based on [11] research, 29% of the drivers reported that they had felt close to falling asleep while driving in the last 12 months, and this probability increases with the proportion of time spent driving on motorways. Drivers who are prepared to drive for long periods without taking a break are also more likely to fall asleep at the wheel, especially during early hours of the morning or after work[11].

B. Artificial Intelligence in preventing car accidents

Driver inattention might be the result of a lack of alertness when driving due to driver drowsiness and distraction. Driver distraction occurs when an object or event draws a person's attention away from the driving task, whereas driver drowsiness is characterized by a progressive withdrawal of attention from the road and traffic demands [13]. To help prevent collisions or decrease their likelihood, AI comes at hand, with one of the most promising mechanisms like Driver Drowsiness Detection. Drowsiness Detection can be divided into 3 categories [13]:

- **Vehicle based:** Measures deviations from lane position, movement of the steering wheel, pressure on the acceleration pedal etc, and any change in these that crosses a specified threshold indicates a significantly increased probability that the driver is drowsy.
- **Behavioural based:** Measures yawning, eye closure, eye blinking, head pose, etc. which are monitored through a camera and the driver is alerted if any of these drowsiness symptoms are detected.
- **Physiological based:** Measures the correlation between physiological signals ECG (Electrocardiogram) and EOG (Electrooculogram). Drowsiness is detected through pulse rate, heart beat and brain information.

The most challenging, but at the same time most effective category is behavioural based, more specifically depending on eye positioning. Different architectures have been designed to capture eye movements in order to detect drowsiness. Some include attaching cameras in front of the driver's seat, implementing an optical sensor system that uses infrared or near-infrared LEDs to light the driver's pupils, passing the data to computer algorithms that analyse blink rate and duration to determine drowsiness [13]. Another way is using mobile phone's front camera, which can collect various visual parameters, including eye features, mouth features, and head movements [1]. Ultimately, all the data captured by these sensor are of no use if they are not trained and tested in AI models, like Convolutional Neural Networks (CNN) to make right predictions.

III. Methodology

A. Macro Methodology

The Macro Methothodology of this paper follows these steps:

1) Research problem identification and development of research questions.
2) Perform of literature review related to the research problem and questions.
3) Experiment design, including dataset selection, architecture infrastructure and resources to run the experiments.
4) Data Collection and Analysis, by running and modifying the experiments.
5) Interpretation of results, and comparing them to the research objectives.
6) Conclusions and Recommendations for future improvements.

B. Micro Methodology

This paper encompasses two types of experiments. Initially, it integrates various ML Classification models to predict accident severity using a provided dataset and a ML Clustering Algorithm to detect frequent accident locations in an area. Subsequently, it employs Deep Learning to detect drowsiness by analyzing labeled eye images as either 'closed' or 'open', serving as a mechanism for accident prevention. The experiments and analysis were carried using Google Colab environment.

1) Experiment Design: The first step in designing the experiments is the selection of the datasets. For the Machine Learning Classification and Clustering analysis, UK traffic accidents [5] was used, with a 9 year time-stamp (2005-2014, excluding 2008) and including 1.6 million rows of data. The columns of this dataset had information regarding the severity of the accident, latitute, lontitude, date, number of casualties, number of vehicles, weather conditions, road surface conditions, light conditions etc. Since the aim was to predict the severity of the accident based on these features, the classification models used to conduct this experiment were:

- **Decision Tree**, is a non-parametric supervised learning algorithm, which is utilized for both classification and regression tasks. It has a hierarchical, tree structure, which consists of a root node, branches, internal nodes and leaf nodes.It employs a divide and conquer strategy by conducting a greedy search to identify the optimal split points within a tree [9].
- **Random Forest**, is a machine learning algorithm, trademarked by Leo Breiman and Adele Cutler, that combines the output of multiple decision trees to reach a single result. It has three main hyperparameters: node size,number of trees, and number of features sampled [9].
- **ADABoost**, short for Adaptive Boosting, is a machine learning meta-algorithm used in conjunction with many other types of learning algorithms to improve performance. It is adaptive in the sense that subsequent weak learners are tweaked in favor of those instances misclassified by previous classifiers [4].
- **XGBoost**, or eXtreme Gradient Boosting, is a machine learning algorithm under ensemble learning. It is trendy for supervised learning tasks, such as regression and classification. XGBoost builds a predictive model by combining the predictions of multiple individual models, often decision trees, in an iterative manner [14].
- **KNN**, or the k-nearest neighbors is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. It is part of a family of “lazy learning” models, meaning that it only stores a training dataset versus undergoing a training stage [9].
- **Balance Bagging Classifier**, is a specialized ensemble learning algorithm that is specifically designed to address the problem of imbalanced classification. It combines the power of bagging with resampling techniques to achieve a better balance between the minority and majority classes [12].

The performance of each of the models was improved using hyperparameter tuning and evaluated using accuracy.

For the Clustering Analysis, the algorithm used is Density Based Clustering (DBSCAN). DBSCAN groups together points that are close to each other based on a distance measurement (usually Euclidean distance) and a minimum number of points and marks as outliers the points that are in low-density regions [6]. On the other hand, for the Deep Learning analysis, Drowsiness Detection Dataset [8] was used,

Fig. 1: Accidents across the UK

containing more than 5000 images of eyes, classified as "closed" or "opened". To detect the state of the eye in the image, Convolutional Neural Networks (CNN) were used. The model has 6 convolutional layers and data augmentation was applied, by adding images of high and low exposure in the training and testing set.

2) Experiment Implementation: For the implementation of the experiments, Google Colab and Jupiter Notebook was used. For the first part of the experiments (ML Classification analysis), the 7 classification models were run using the UK traffic accidents dataset. For each of the models, these steps where followed:

1) Import of the dataset
2) Cleaning the dataset from null values
3) Feature Engineering, dropping unnecessary columns, like: Accident_Index, Police_Force,1st_Road_Class, Junction_Detail, 2nd_Road_Number etc.
4) Feature Selection and encoding, all the remaining columns were used as features, but the ones containing object values (string) where encoded into integer numbers.
5) Target Selection, which in this case was Accident_Severity.
6) Splitting the model into training set (80%) and testing set (20%).
7) Running the model and computing the accuracy in predicting accident severity.
8) Hyperparameter Tuninig for each model, to improve the accuracy by changing model's parameters.

Using the same dataset for the ML Clustering, these steps were followed:

1) Import of dataset.
2) Feature Engineering, using only the columns Latitude, Longitude, Number_of_Vehicles, Time,

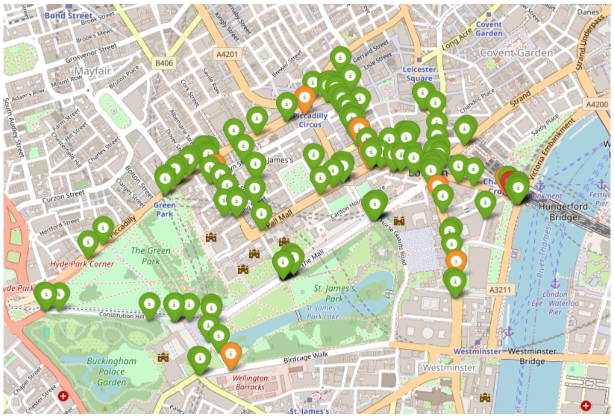


Fig. 2: Map of accident severity

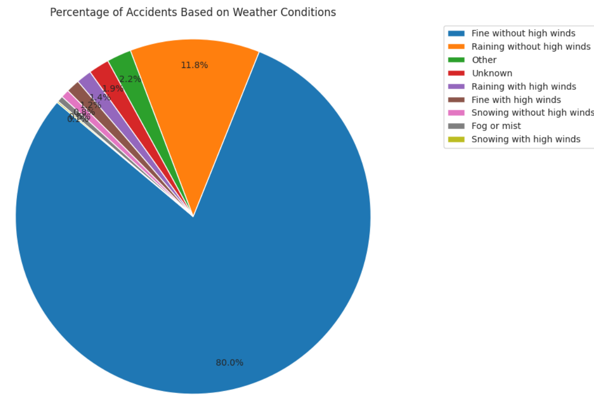


Fig. 3: Weather in relation to accidents

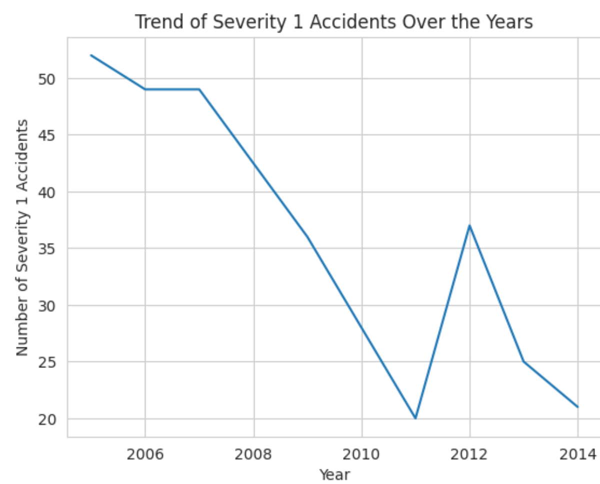


Fig. 4: Fatal Accidents Trend

Local_Authority_(Highway) and Year.

3) Defining Year=2014 and Local_Authority_(Highway) = "E08000035", which is the city of Leeds.
4) Applying spacial clustering.
5) Applying spaciotemporal clustering.

For the third part, which involves the CNN model in detecting whether the eyes if the people in the images in the dataset were opened (indicating non-drowsy) or closed (indicating drowsy), these steps were followed:

1) Import of dataset.
2) Image pre-processing (no data split was necessary, since the training set and test set were already split in the dataset).
3) Define CNN model architecture, with 6 convolutional layers and output layer with 2 neurons for binary classification (opened or closed).
4) Model training and evaluation.
5) First data augmentation, by copying 2000 images from the dataset, applying high exposure on 1000 and low exposure on the rest, and adding them back to the training set and evaluate the new model using the original test set.
6) Second data augmentation, by copying 800 images from the dataset, applying high exposure on 400 of them and low exposure on the rest, and adding them to the test set and evaluate the new model using the original training set but new test set.
7) Third model, using the new (augmented) training set and the new (augmented) test set for evaluation.

IV. Findings and Discussion

Before starting with ML and DL experiments, some data visualisation is necessary to explore the dataset. The first dataset (UK Accidents) can give us insights on the density of car accidents across the UK, as shown in Figure 1. Each white dot represents an accident. Based on the map, it can be seen that Southern (especially London) and Central England (Liverpool, Manchester and Birmingham) are the cities or locations with the high frequency of accidents throughout the years (2005-2014).

According to the dataset, accidents are classified based on their severity. Figure 2 shows the accidents happened in 2014 in a part of London based on their severity: Green means Slight, Orange means Serious and Red means Fatal. As it can be seen, the majority of the accidents during that year for that particular area were of a slight severity. Figure 4 shows the trend of fatal accidents (Accident_Severity = 1) through the years. The graph shows a downward trend in the frequency of these accidents over time. However, there is a slight uptick in the year 2012, interrupting the overall decline. This decrease can likely be attributed to advancements in road safety measures, such as improved infrastructure and stricter police enforcement.

Weather can be a factor that can impact the possibility of an accident. Figure 3 shows the percentage of accidents in accordance to weather conditions. It's evident from the chart that the majority of the accidents (80%) have happened in normal weather conditions, followed by rainy weather with 11.8%.

The UK accidents dataset provides information regarding the time of the accident as well. To better visualize the relation between the number of accidents and their time of occurrence, we create 4 time categories:

- **Early Morning:** From 00:00 to 07:00
- **Morning to Noon:** From 07:00 to 12:00
- **Afternoon:** From 12:00 to 19:00

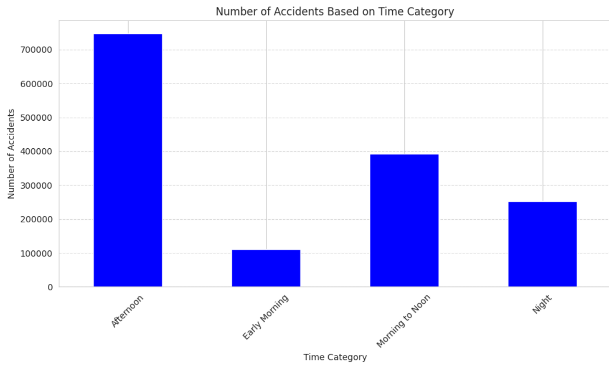


Fig. 5: Number of accidents in different time stamps

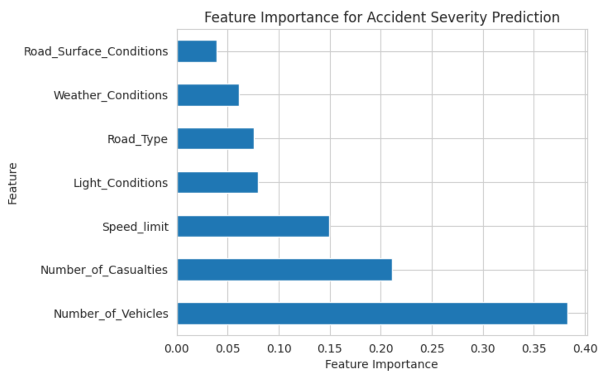


Fig. 6: Feature Importance for Accident Severity

- **Night:** 19:00 to 00:00

The results of this classification are shown in Figure 5. Judging by the chart, the time category with the most accidents is afternoon, so between 12:00 and 19:00, followed by morning to noon category. The category experiencing the fewest accidents is early morning, which may be attributed to the reduced frequency of cars during that time.

To understand the weight of the features that determine the severity of an accident, feature importance analysis is used. The features taken into account are Number of Vehicles, Road Surface Conditions, Weather Conditions, Light Conditions, Number of Casualties, Speed limit and Road Type. According to Figure 6, the feature that has the most impact to the target variable, which is Accident Severity, is Number of Vehicles, with a weight 0.383077. This suggests that the more vehicles involved in a collision, the greater the likelihood of a severe accident.

A. Machine Learning - Classification Models

Proceeding to the next step, which is Accident Severity Prediction using ML Classification Models, the 6 before-mentioned models where run and tuned to produce the highest accuracy. Table I shows the accuracy level for each model before and after tuning. For most of the models, the hyper-parameter tuning has improved their performance, except for Random Forest, whose accuracy decreased at a very low level. The Classification model with the mighest accuracy before tuning is Random forest and after tuning XGBoost. Overvall, XGBoost has the highest accuracy at predicting the severity class of an accident, with accuracy 0.86367. The values of XGBoost parameters that give the highest performance are: subsample = 1, reg_lambda = 10, reg_alpha = 1, n_estimators = 150, min_child_weight = 1, max_depth = 6, learning_rate = 0.5, gamma = 0.3 and colsample_bytree = 0.5. The model that is the least performing is BBC, with accuracy after tuning at 0.56350. As it can be seen from the table, the accuracy of the models is around 0.86 and does not go higher even with tuning. An explanation to this might be the structure of the dataset. Similarly to Figure 2, most of the accidents are of severity 3, or slight severity. Being overpopulated by severity type 3, and having much less cases with severity 1 (fatal) and 2 (serious), the model finds difficulties in making the difference. The confusion matrix II below of XGBoost shows this in a better way:

Model Name	Accuracy before tuning	Accuracy after tuning
Random Forest	0.86172	0.86127
Decision Tree	0.76402	0.86314
ADABoost	0.85631	0.86344
KNN	0.83932	0.86303
XGBoost	0.85390	0.86367
BBC	0.53058	0.56350

TABLE I: ML Classification Models Comparison

3	3	69
1	26	853
1	45	5652

TABLE II: Confusion Matrix

According to this matrix, the main diagonal represents the right predictions of the model for each severity category. Put simply, the model accurately predicted 3 accidents with severity 1, 26 accidents with severity 2, and 5652 accidents with severity 3. It can see seen that the times that the model predicted the accidents as type 3 of severity is very high. This is why the model has difficulties in identifying severity 1 or 2 of accidents, consequently giving a lower, but still good accuracy.

B. Machine Learning - Clustering

In this paper, clustering is used to identify accident hotspots in the city of Leeds. To achieve this, density based clustering (DBSCAN) is used. As a result, by leveraging this algorithm, areas with a frequent concentration of accidents are identified, grouped into clusters and plotted using folium.

As mentioned before, distance between points is measured. To do this, latitude and longitude are taken to determine a point (x,y) and the great_circle function from geopy ensures accurate distance calculations by considering the Earth's curvature. We determined the maximum distance between points to be considered as neighbours to be 100 meters (so the accidents should happen at most at 100 meters of distance to be considered as part of the same cluster) and the minimum elements/samples

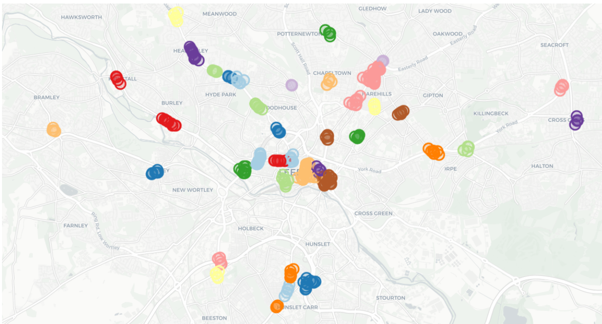


Fig. 7: Spacial clusters for the city of Leeds

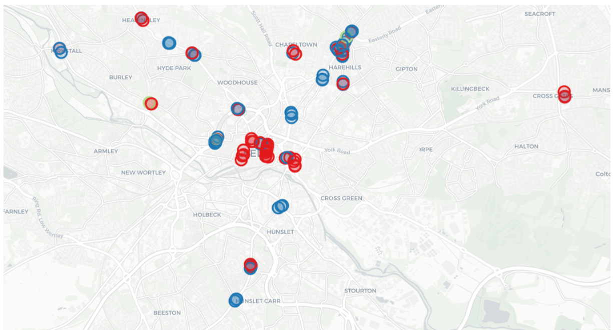
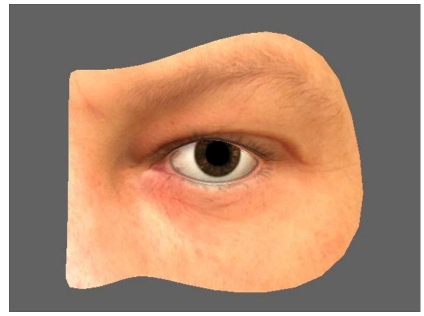


Fig. 8: Spaciotemporal clusters for the city of Leeds

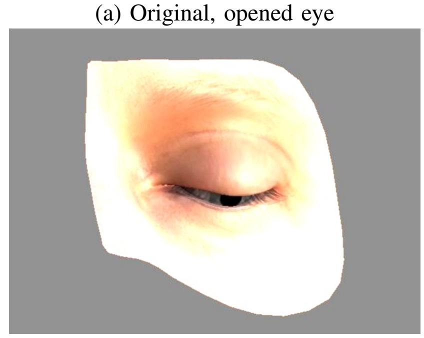
a cluster can have as 5. After running the model, there were found 40 clusters and plotted in a map represented by Figure 7. To uncover high-risk regions and times, spatiotemporal clustering analysis were applied. The maximal spatial distance was set to 500 meters and the maximal temporal distance to 5 hours. This time, the minimum samples per clusters was set to 3, since the range shrunk, which resulted in 31 clusters. Based on the feature importance analysis, the number of vehicles involved in an accident emerged as the most significant factor. To visualize this relationship, clusters were colored according to the number of vehicles: red for accidents with 1 vehicle, blue for 2 vehicles, green for 3 vehicles, and orange for 4 vehicles, as shown in Figure 8. According to the map, accidents involving more than 2 vehicles are most likely to happen in big and busy junctions, like Inner Ring Road, whereas one-vehicle accidents occur more frequently in areas with high pedestrian activity, like the city center. This clustering model identifies the most dangerous areas in the city, pinpointing locations with a high probability of car accidents. This valuable insight allows for targeted interventions, enabling authorities to reinforce safety measures in these high-risk zones.

C. *Deep Learning*

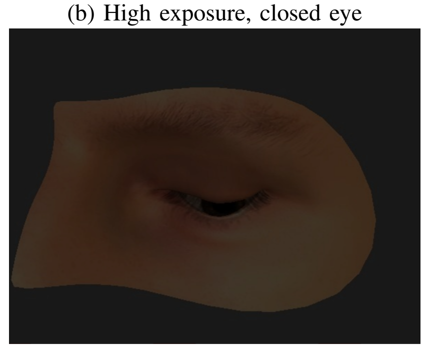
Lastly, the analysis progress to Deep Learning Models in detecting driver's drowsiness, utilizing the eye images dataset with the TensorFlow library. The initial CNN model was trained and tested using the original dataset. Achieving an accuracy of 0.96739, this model performed exceptionally well. To better simulate real-life scenarios, both high and low exposure settings were applied to a subset of images in the dataset. High exposure represents conditions like direct sunlight affecting the driver's visibility, while low exposure reflects situations where lighting is limited, such as driving in dimly lit areas or during nighttime. Figure 9 shows a comparison between an original image of the dataset and two images of the dataset with high and low exposure.



(a) Original, opened eye



(b) High exposure, closed eye



(c) Low exposure, closed eye

Fig. 9: Data augmentation using different levels of exposure in DDD dataset

For the second model, data augmentation was applied only for the training set, keeping the original test set. After running the model, it gave an accuracy of 0.95534, which is lower than the previous model, but still at a very good level. The reason why is because we are training the new model on a test set that does not know high and low exposure images. This is why we conduct the third model. The third CNN model involves applying data augmentation on the test set as well. A total of 400 images were copied, edited with high and low exposure, and added to the test set. After running the model with augmented training and test set, the accuracy was 0.95852, which is slightly higher than the previous one. The value has increased due to the model's improved ability to detect whether the eyes are open or closed under varying lighting conditions. Now the Driver Drowsiness Detection Model is more performing when it comes to real-life situations.

V. Conclusions and Future Work

In this paper, the contribution of AI in preventing car accidents was studied. Machine Learning Classification Models help in predicting the severity of an accident depending on some factors, like weather conditions, road surface conditions, speed limit, light conditions and so on. The most performing model in predicting the severity of the accident was XGBoost after hyperparameter tuning, with accuracy 86.37%. The DBSCAN clustering model helped in identifying danger zones by grouping the accidents based on their occurrence. Precautions can be taken to decrease the probability of accidents, judging by these factors. For example, recognizing rain as a significant contributor to accidents, implementing rainwater canals or drainage systems along roads with high levels of accidents, can mitigate wet and slippery conditions, enhancing road safety.

Based on the literature review, a prominent factor contributing to accidents — distinct from environmental factors — is the driver's personal state, particularly sleep deprivation, fatigue, or overall tiredness. Lack of sleep contributes in driver's loss of control and increases the chances of accidents or collisions. To prevent this, a Driver Drowsiness Detection Mechanism was developed in this study with accuracy 95.85%, using CNN. This model is trained to detected whether the eyes of the driver are closed or opened in different lighting conditions. A camera can be attached in the car, facing the driver and when the model detects that the eyes are closed, then an alert system can be incorporated to warn the driver to wake up.

Some recommendations that might improve the performance of the Driver Drowsiness Detect Mechanism could involve integrating a more complex dataset. This expanded dataset would encompass not only eye movements but also factors like mouth movements to detect yawning, variations in heart rates — which tend to decrease when individuals are tired — and other relevant physiological indicators. Using a CNN model with a different number of layers can also be an approach to increase accuracy. In addition, enhance driver safety, the system can automatically adjust car settings, like lowering speed, if drowsiness is detected.

Appendix A
Datasets and Source Code

The datasets used for this research, as well as the source code can be found the the following GitHub link : DAV

Acknowledgment

We would like to thank Erasmus Mundus Joint Masters Degree in Green Networking and Cloud Computing (GENIAL) and colleagues for the support, which made this research possible.

References

[1] Y. Albadawi, M. Takruri, and M. Awad. A review of recent developments in driver drowsiness detection systems. *Sensors*, 22(5), 2022.

[2] t. r. s. c. Brake. Driver fatigue, 2023. Accessed on 2024-04-13.

[3] D. D. Clarke, P. Ward, C. Bartle, and W. Truman. Killer crashes: fatal road traffic accidents in the uk. *Accident Analysis & Prevention*, 42(2):764–770, 2010.

[4] I. T. Community. Classification with adaboost. Accessed on 2024-04-17.

[5] K. DAVE FISHER-HICKEY. 1.6 million uk traffic accidents, 2017. Accessed on 2024-04-07.

[6] K. S. do Prado. How dbscan works and why should we use it? Accessed on 2024-05-02.

[7] G. B. D. for Transport. Reported road casualties great britain, annual report: 2022, 2022. Accessed on 2024-04-13.

[8] K. HAZEM FAHMY. Drowsiness detection dataset, 2023. Accessed on 2024-04-12.

[9] IBM. Ai and machine learning. Accessed on 2024-04-17.

[10] D. Léger, V. Bayon, M. M. Ohayon, P. Philip, P. Ement, A. Metlaine, M. Chennaoui, and B. Faraut. Insomnia and accidents: cross-sectional study (equinox) on sleep-related home, work and car accidents in 5293 subjects with insomnia from 10 countries. *Journal of Sleep Research*, 23(2):143–152, 2014.

[11] G. Maycock. Sleepiness and driving: the experience of uk car drivers. *Journal of sleep research*, 5(4):229–231, 1996.

[12] Medium. Balanced bagging classifier (bagging for imbalanced classification). Accessed on 2024-04-17.

[13] V. Saini and R. Saini. Driver drowsiness detection system and techniques: a review. *International Journal of Computer Science and Information Technologies*, 5(3):4245–4249, 2014.

[14] A. Vidhia. Introduction to xgboost algorithm in machine learning. Accessed on 2024-04-17.