IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Simulated Driver Behaviour Analysis and Drowsiness Detection using Machine Learning

[1]Nageshwar Prasad Yadav, [2]Aniruddha Dewangan, [3]Krish Srivastava, [4]Paras Agarwal

[1]Student, [2]Student, [3]Student, [4]Student
[1]Computer Science and Engineering,
[1]SRM University AP, Amaravati, India

Abstract: Road safety is an extremely major issue around the world. There has always been an increase in the number of road accidents because of distracted driving, drunk driving, speeding, running red lights and stop signs, reckless driving and aggressive driving. India has the highest number of road accident fatalities in the world, with around 1.5 million deaths each year, which translate, on an average, into 1130 accidents and 422 deaths every day or 47 accidents and 18 deaths every hour. In 2015, there were about 500,000 road accidents in India, which killed about 150,000 people and injured about 500,000. [WHO2015] This research paper discusses an innovative approach to reduce the number of such cases significantly. The project aims to evaluate drivers according to their driving patterns and behaviours by synergizing the concepts of Machine Learning and Blockchain Technologies. The proposed application will be able to analyze and rate a driver based upon their driving behaviour and drowsiness patterns. The Machine Learning Algorithms are brought into action to process the data and give out the result in the desired format. The algorithms would be responsible for identification of rash driving patterns and keeping a count of it. Another Machine Learning model will be responsible for keeping a note of any drowsiness that a driver might exhibit during his/her drive. The Blockchain Technology is brought into this project to ensure the security and immutability of the results gathered after the analysis from the Machine Learning models. Furthermore, the motivation to work on this research project comes from the alarming rate of increase in road fatalities.

Index Terms **- Driver behaviour, Accidents, Machine Learning.**

I. INTRODUCTION

In recent years, the surge in road accidents globally, and particularly in India, has underscored the urgency to address road safety concerns. The World Health Organization (WHO) reports a staggering number of fatalities and injuries resulting from road accidents, with India ranking high in these statistics. Recognizing the critical need for innovative solutions to mitigate road accidents, this research project focuses on Simulated Driver Behaviour Analysis. This work presents an innovative approach to driver behaviour analysis using machine learning and blockchain technology to evaluate and rate drivers based on their driving patterns. The project aims to develop an application that utilizes data from a mobile sensor attached to a driver's car, capturing essential attributes. These attributes form the basis for training models to discern distinct driver behaviours. The final application also involves the inclusion of a Drowsiness Detection System to alert the drivers when they are detected to be feeling drowsy, so as to reduce the number of mis-happenings. Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), Long Short-Term Memory (LSTM), and conventional Artificial Neural Networks (ANN) are among the neural network architectures that are being studied in this research. The Random Forest Algorithm is eventually incorporated into the project in order to enhance the driver behaviour prediction model's resilience. The main objective is to establish a system that ensures accountability and transparency by securely and irrevocably recording sensor data on the blockchain.

The Drowsiness Detection System is based on computer vision methods, i.e., OpenCV's video processing capabilities. It actively monitors and analyses eye movement in real-time video broadcasts. The primary approach entails calculating the eye aspect ratio (EAR) for both eyes, which is a measure of ocular openness calculated from the distances between specified facial features. This EAR calculation assists in spotting drowsiness-related alterations.

The incorporation of blockchain technology ensures the immutability and transparency of the recorded driver behaviour ratings. By leveraging smart contracts, the sensor data is securely transmitted to an API housing the machine learning models. The results are then permanently stored on the blockchain, providing a tamper-resistant ledger of driver ratings. Due to the increasing integration of modern technology in vehicles, solutions capable of evaluating and tracking driver behaviour have been developed. In this context, there is a requirement to evaluate and rate drivers on the basis of a variety of standards, such as driving maneuvers and occurrences of drowsiness. Solidity, a programming language used largely for developing smart contracts on blockchain platforms such as Ethereum, offers a decentralized and secure environment in which to implement a ranking algorithm. Using the transparency and immutability of blockchain technology, a fair and secure system for evaluating and ranking drivers based on their driving behaviour and drowsiness occurrences can be built.

II. Related Work

Many studies have been conducted to simulate driver behaviour for commercial interests, managerial tasks, or public awareness campaigns. Their major objective is to use their model to explain the relationship between driver behaviour and other circumstances. It is a complicated system with many factors, and it has been proved that the majority of accidents are caused by human mistakes such as conscious legal breaches, distraction, inattention, exhaustion, and so on. The advancement of data analysis methods throughout the years has enabled the evolution of this field of study. The development of these methodologies enhanced the quality of driver behaviour analysis while also opening up new sectors of application. However, no standard model has been suggested in the literature. Due to the lack of a uniform framework for evaluating driver behaviour, we propose this work, an attempt to collect the collection of factors regulating driving behaviours into a single model. We gathered a collection of quantitative and qualitative indicators that are critical in assessing driver behaviour. They are either connected to driving or to the driver. These elements are the outcome of a literature analysis that combines several models and generates the factors considered in relation to driver behaviour. These factors are then grouped based on their importance in the research in which they were referenced, as well as their frequency of presence. Table 2.1 organizes these elements into categories.

Table 2.1 – The set of Quantitative and Qualitative factors

Quantitative Factors	Qualitative Factors	
	Driving-related	**Driver-related**
Mileage	Law violation	Anger
Speed- Acceleration	Distraction	Narcissism
Braking	Maneuvers prediction	Impulsivity
Orientation	Attention	Sensation seeking
Road type		
Position		
Time range		

Driver behaviour modelling is a critical study area for the automotive and intelligent transportation industries, as well as vehicle insurance, and the government agencies in charge of infrastructure and public transportation. These disciplines are interested in understanding such behaviour in order to develop optimal solutions that increase mission performance. Several approaches are used in this regard, such as monitoring the driver's physical state (facial recognition, physical characteristics monitoring, etc.), collecting navigation data using onboard telematics, assessing driving styles, and so on. This document categorizes all of these applications based on their ultimate goal.

Applications for Vehicles: These approaches are largely concerned with the automobile as an end objective. They hope to improve the driving experience by developing intelligent technologies that help drivers on their journeys and solutions that interact with them in real time. Make the driving effort easier. These applications differ in terms of their level of autonomy, ranging from basic driving aid (parking assistance, lane keeping systems) to the highest level of autonomy, which is self-driving automobiles (Autopilot). Management-oriented apps: These are applications that try to optimize vehicle utilization, such as fleet management and traffic modelling.

Applications for Drivers: The driver is regarded as the entry point for all functionalities in this sort of program. The parameters gathered on the drivers include physical and mental status, vehicle operations, driving styles, and so on. The ultimate aim is not only to rate drivers but also to create a culture of safety and responsibility on the roads, contributing to a reduction in road accidents and ensuring a safer driving environment.

III. Dataset

3.1 Driver Behaviour Dataset

The dataset under consideration is aimed at classifying different driver behaviours and extracting driver patterns. The data was collected using the Raspberry Pi Model B. The collected data encompasses the driving behaviours of three distinct drivers on a predefined route with multiple turns. The primary objective of this dataset is to analyse driver behaviour and discern patterns that could have practical applications in areas such as autonomous vehicle development, driver monitoring systems, and road safety.

Data Source: The data was collected using the Raspberry Pi Model B equipment. Three different vehicles were used, driven by three different drivers. The vehicles were Ford Fiesta 1.4 Diesel Engine, Ford Fiesta 1.25 Gasoline Engine, and Hyundai i20 1.2 Gasoline Engine. The drivers were aged 27, 28, and 37 years. The readings were taken in Ankara and Isparta, two major cities in Turkey. The road was dry asphalt, and the weather was sunny. Vehicle speeds were set to 30-40 km/h for sudden lane changes and 70-80 km/h for sudden acceleration and deceleration.

Sensor Information: The dataset includes information from an MPU6050 Sensor, which has a 3-axis accelerometer and 3-axis gyroscope for recording data. The sensor was mounted onto the center of the dashboard with the black box's front face facing upwards.

The dataset contains the following features:

Driver IDs: The "DriverID" column identifies the different drivers, with values [1, 2, 3] corresponding to the three distinct drivers involved in the data collection.

Driving Behaviours: The dataset classifies a total of 4 risky driving behaviours: Sudden Acceleration (Class: 1), Sudden Deceleration (Class: 2), Sudden Left Turn (Class: 3), Sudden Right Turn (Class: 4).

Timestamp: The data was collected over a period of 2 weeks, and the dataset includes the Hour, Minute, and Seconds as attributes to provide a precise time reference for each recorded event.

3.2 Drowsiness Detection Dataset

The "*shape_predictor_68_face_landmarks.dat*" file, which is an essential component of the dlib library, is the dataset extensively utilized in this study to enable precise facial landmark identification. This pre-trained model is crucial for accurately identifying 68 different facial features, including key regions such as the mouth, nose, eyes, and jawline. Its integration into the research methodology forms the foundation for comprehensive facial analysis, significantly aiding in the tracking of facial structure, the identification of facial emotions, and the computation of eye aspect ratios, among other tasks.

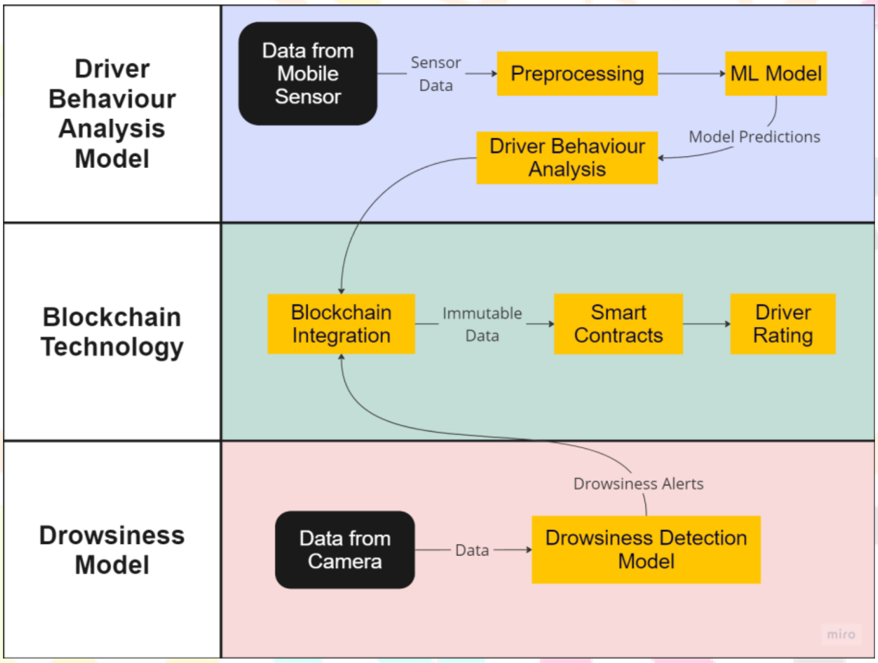
The research employs several datasets containing a broad range of facial images and videos during the data collection phase. The "*.dat*" file is an essential data collection tool that assists with the precise localization and annotation of facial landmarks within the captured images and video frames. Ensuring the availability of accurately labeled data for subsequent research and model training depends heavily on the effective execution of this approach.

IV. Research Methodology

The methodology section outlines the plan and method that how the study is conducted. This includes Collection and Sources of Data, Universe of the study, study's variables and analytical framework. The details are as follows.

Driver Behavior Analysis, Drowsiness Detection, and Blockchain Integration are the three main components of the proposed system. The flowchart in Figure 4.1 clearly depicts the system's sequential evolution, emphasizing the interdependence of data processing, machine learning analysis, drowsiness detection, and blockchain integration. This comprehensive method provides a robust solution for assessing and rating drivers based on their behavior, thus improving road safety and accountability.

Figure 4.1 - Flowchart of Complete Analysis



4.1 Data Collection for the Driver Behaviour Analysis Model

4.1.1. Using a Mobile Sensor

The procedure begins with the acquisition of raw data from a mobile sensor mounted on the driver's vehicle. This sensor captures critical features including AccX, AccY, AccZ, GyroX, GyroY, and GyroZ, offering a comprehensive view of the driving environment.

4.1.2 Preprocessing

Following data gathering, the raw data is preprocessed. This phase includes timestamp synchronization, feature extraction, and the critical process of labeling the data with driver IDs and behavior classes. The latter entails categorizing driving habits such as Sudden Acceleration, Sudden Right Turn, Sudden Left Turn, and Sudden Brake.

4.2 Machine Learning Models and Integration of Blockchain

4.2.1 Model for Driver Behavior Analysis

The preprocessed data is fed into machine learning models that utilize neural network architectures such as DNN, CNN, GRU, LSTM, and ANN. These models are trained to assess driving patterns and predict the behavior classes. The end result is a system that can analyze and score driver conduct based on recorded data.

4.2.2 Detection of Drowsiness

In parallel, a second module actively monitors and analyzes eye movement in real-time video broadcasts using computer vision algorithms, specifically OpenCV. The eye aspect ratio (EAR) for both eyes is calculated using this Drowsiness Detection System, which measures ocular openness based on the distances between facial features. EAR changes caused by drowsiness are then recognized, adding an extra layer of protection by warning drivers when sleepiness is detected.

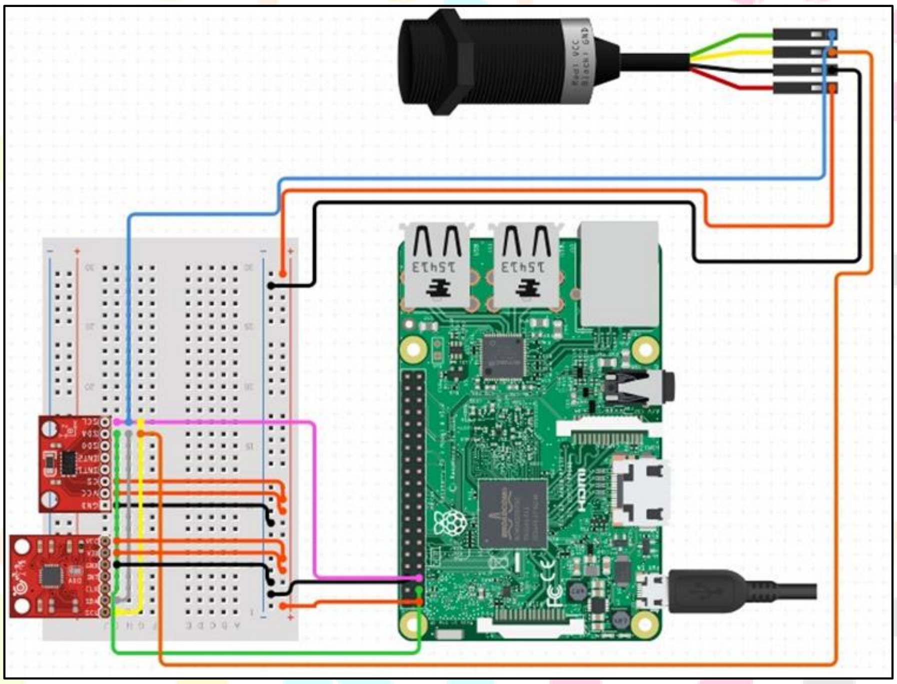
4.2.3 Integration of Blockchain

Both the Driver Behavior Analysis and Drowsiness Detection data are securely sent to the blockchain using smart contracts. This integration ensures the immutability and transparency of recorded data, such as driver ratings and drowsiness occurrences. The blockchain acts as a tamper-proof ledger, securely preserving the results of the study and enhancing accountability.

4.3 Experimental Setting

The "Simulated Driver Behaviour Analysis" implementation takes a comprehensive approach, combining sensor technologies, machine learning, and blockchain. The SparkFun ITG-3200 Gyro Breakout and SparkFun ADXL345 Accelerometer Breakout are employed to measure rotational velocity and acceleration along the three axes, respectively, to record critical driving characteristics. These sensors are linked to a microcontroller or Single Board Computer (SBC) such as the Raspberry Pi, laying the groundwork for data collection. Figure 4.2 shows a schema of the sensor setup that can be mounted on the automobile. A camera module, such as the Raspberry Pi Camera, will be integrated in parallel to gather additional contextual data. This trio of sensors serves as the foundation for a comprehensive dataset that includes GyroX, GyroY, GyroZ, AccX, AccY, AccZ, and potentially visual data from the camera. The acquired data is synchronized using timestamps to ensure consistency across the various sensors.

Figure 4.2 - Gyroscope, Accelerometer and Camera on a Raspberry Pi Board



4.3.1. Random Forest Algorithm for the Driver Behaviour Model

The decision to implement the Random Forest Algorithm over other models stems from its unique strengths. Random Forest is an ensemble learning method that combines multiple decision trees, each trained on different subsets of the dataset. This approach results in a robust and accurate model, well-suited for complex and multidimensional data, such as that derived from driving sensors. Random Forest excels in handling noisy datasets and offers resilience to overfitting, a crucial consideration given the intricacies of driver behaviour analysis. Its ability to provide feature importance scores aids in understanding which attributes contribute significantly to the model's decisions. In the context of the project, this feature is invaluable for pinpointing specific sensor readings that correlate with distinct driving behaviours.

The ensemble nature of Random Forest introduces an additional layer of reliability. By aggregating the outputs of multiple trees, the model becomes less susceptible to outliers and biases present in individual trees. This robustness is especially vital in real-world scenarios where driving behaviour can be influenced by diverse factors.

4.3.2. Drowsiness Detection Model

The system detects drowsiness based on the eye aspect ratio (EAR), which is calculated from the distances between specific facial landmarks. When the EAR goes below a predetermined threshold value (set at 0.25 in this case) for several consecutive frames, the system identifies it as a potential indicator of sleepiness. If these criteria are met, the system detects and records timestamps indicating the start and end times of these possible drowsiness events. These timestamps are stored in a text file for further analysis and documentation. The recorded data provides valuable insights into the frequency, duration, and temporal patterns of detected drowsiness, allowing for future investigation or intervention measures.

Overall, by continually monitoring eye activity using facial landmark identification and EAR analysis, the system offers a viable technique for real-time drowsiness detection. This technology has potential applications in various fields, including driving safety systems, human-machine interaction, and occupational safety, by providing a proactive mechanism to detect and prevent fatigue-induced lapses in attention.

4.3.3. Rating Algorithm and Blockchain

In an innovative twist, the project integrates blockchain technology into the core execution, with the goal of ensuring immutability and transparency in the assessment and storage of driver ratings. The process involves sending sensor data directly to a smart contract, which then forwards the data in a specified format to a Machine Learning model running inside an API. The results generated by the ML model are stored in an immutable ledger on the blockchain.

The adoption of blockchain technology introduces several advantages. Firstly, it enhances data integrity and security by creating an immutable record of driver ratings, ensuring that the ratings are tamper-proof and can be trusted by all stakeholders involved. Secondly, the transparency offered by the blockchain allows for a clear audit trail of how the ratings were derived, promoting accountability in the assessment process.

IV. RESULTS AND DISCUSSION

4.1 Results of Driver Behaviour Analysis Model

The model’s accuracy, which represents the percentage of accurately predicted cases in the test set relative to the total number of instances, was approximately 98.65%. This high accuracy indicates strong predictive performance, suggesting that the model is highly effective in classifying different driver behaviours.

The result also provides the counts of accurately predicted occurrences for each behaviour type during testing for each driver. The behaviour types are categorized as: Class 1 for Sudden Acceleration, Class 2 for Sudden Right Turn, Class 3 for Sudden Left Turn, and Class 4 for Sudden Brake. The result of a sample case is presented in Table 4.1, detailing the number of correctly classified instances for each behaviour type, further demonstrating the model's effectiveness in accurately identifying specific driving behaviours.

Table 4.1: Descriptive Statics for Output of Driver Behaviour Model

Driver-ID	Class 1	Class 2	Class 3	Class 4
Driver-1	12	14	25	7
Driver-2	14	12	20	11
Driver-3	10	20	10	21

4.2 Results of Drowsiness Detection Model

The output of the model indicates a detected episode of drowsiness, flagged with the label “*Feeling drowsy*”. It includes a “Drowsiness count” which shows the number of identified sleepy episodes, with this being the first instance. Additionally, it records the “*Duration*” of the drowsy episode, which lasted 6 seconds. Each occurrence of observed sleepiness is systematically labeled, followed by the count of sleepy episodes and their respective durations. Table 4.2.

In the next stage of our research, we aim to develop a small in-vehicle rating device that will smoothly interact with various vehicles and provide real-time data on driver movements and behaviours. The purpose of this device is to enable a dynamic, real-time driver rating system, with collected data accessible to passengers via an accessible interface. Our goal is to work with cab-hailing companies to integrate the rating system into the booking process, allowing customers to make educated decisions based on the driver’s performance. Continuous machine learning model development, stringent privacy safeguards, and pilot testing will be critical in guaranteeing the system’s accuracy, security, and user acceptance. While the dataset used in the training of the Drowsiness Detection model, i.e., “*shape_predictor_68_face_landmarks.dat*” file greatly helped the research, it is critical to recognize its limits. The model’s performance may vary depending on occlusions, lighting conditions, or facial position. Exploring and overcoming these constraints may pave the way for future research approaches targeted at improving the model’s resilience and flexibility in a variety of real-world circumstances.

V. ACKNOWLEDGMENT

We really appreciate the contributions made by individuals who assisted in the successful conclusion of this project. This has been made achievable by their commitment, diligence, compassion, and vision. We would especially want to thank our mentor who has supported and advised us throughout the course of this project. Their expertise and insights have been instrumental in shaping the direction and scope of this study. We also sincerely thank everyone else who has contributed to our initiative, including our friends, for their support. It would have been a lot more difficult to accomplish the final aim without them. Since it was first envisioned, the project has advanced substantially, mostly due to the efforts of every individual involved. Finally, we would like to express my appreciation to my family and friends for their unwavering support and encouragement throughout this journey. Their love and encouragement have been a constant source of motivation and inspiration. This research paper would not have been possible without the contributions of all those mentioned above, and We are deeply grateful for their support and assistance.

References

[1] Asim Sinan Yuksel, Serafettin Atmaca, " Driving Behavior Dataset," Version 3, 2021, DOI: 10.17632/jj3tw8kj6h.3.

[2] Government of India, Ministry of Road Transport and Highways, Transport Research Wing, " Road Accidents in India 2021," 2021, Available: https://morth.nic.in/sites/default/files/RA 2021 Compressed.pdf.

[3] Kawtar Zinebi, Nissrine Souissi, Kawtar Tikito, " Driver Behavior Analysis Methods: Applications oriented study," 2018, Available: https://www.researchgate.net/publicatio n/324153684 Driver Behavior Analysis Methods Applications oriented study.

[4] Akshay Bahadur, " Drowsiness Detection," 2022, Available: https://github.com/akshaybahadur21/Drowsiness Detection.

[5] H. Ueno, M. Kaneda, M. Tsukino, " Development of drowsiness detection system," Proceedings of VNIS' 94 - 1994 Vehicle Navigation and Information Systems Conference, Yokohama, Japan, 1994, pp. 15-20, DOI: 10.1109/VNIS.1994.396873.

[6] S. Kaplan, M. A. Guvensan, A. G. Yavuz, Y. Karalurt, " Driver Behavior Analysis for Safe Driving: A Survey," IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 6, pp. 3017-3032, Dec. 2015, DOI: 10.1109/TITS.2015.2462084.

[7] Q. Zheng, Y. Li, P. Chen, X. Dong, " An Innovative IPFS-Based Storage Model for Blockchain," 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Santiago, Chile, 2018, pp. 704-708, DOI: 10.1109/WI.2018.000-8.

[8] Monir Shaker, et al., " Online Rating System Development Using Blockchain-Based Distributed Ledger Technology – Wireless Networks," SpringerLink, Springer US, 8 Jan. 2021.