Tuijin Jishu /Journal of Propulsion Technology
ISSN:1001-4055
Vol. 44 No.04 (2023)

# A Comprehensive Review of the Literature on Machine Learning-Based Road Safety Prediction Techniques for Internet of Vehicles (IoV)-Enabled Vehicles

**[1]Amit Manglik, [2]Dr. JN Singh, [3]Prof. (Dr.) Munish Kumar Tiwari**

[1]*Research Scholar, Computer Science and Engineering Department, Galgotias University, Greater Noida, UP,*

[2]*Professor, Computer Science and Engineering Department, Galgotias University, Greater Noida, UP,*

[3]*Professor, Institute of Management, Commerce, and Economics, ShriRamswaroop Memorial University, Barabanki, UP*

***Abstract*:**

The goal of this comprehensive assessment of the literature is to ascertain the state-of-the-art in machine learning (ML)-based vehicle safety measure prediction, obstruction estimate, and road traffic analysis the vehicle-related internet (IoV). In particular, we concentrate on confirming the necessity and extent of federated studying in this area. A decentralized machine learning method called federated learning enables numerous edge devices to train a shared model together while storing the data locally. We looked through numerous scholarly databases and a few chosen peer-reviewed papers on the subject. The paper outlines the benefits and future prospects of federated learning in road traffic analysis and vehicle safety while highlighting the current drawbacks and restrictions of the conventional centralized ML systems. In addition, we examined the state of the art studies in federated learning for road traffic analysis at the time and noted any knowledge gaps as well as potential future study areas. The results of this paper show the extent to which federated learning is required in the subject of road traffic analysis and vehicle safety, as well as how it can be used to get around some of the drawbacks of centralized machine learning techniques.

***Keywords:*** *deep learning, risk assessment, driving behavior analysis, machine learning, collision prediction, and traffic accidents.*

## I. Introduction:

The potential for the "Internet of Vehicles (IoV)" to completely transform the transportation sector has drawn more attention to it recently. A promising technique that can be applied to several IoV applications, including road traffic analysis, obstacle estimation, and vehicle safety measure prediction, is machine learning (ML).However, the massive volumes of data produced by the Internet of Vehicles are too much for conventional centralized machine learning techniques to handle [1].In order to increase the accuracy and safety of road traffic analysis and vehicle safety measures for the Internet of vehicles, federated learning has emerged as a viable decentralized machine learning technique. It can also overcome the limits of centralized machine learning approaches. Federated learning keeps the data local while enabling numerous edge devices to work together to build a common model [2].

The goal of the planned systematic literature review is to confirm the necessity and extent of federated learning in this field while also examining the state-of-the-art in ML-based road traffic analysis, impediment assessment, and vehicle safety measure prediction for the Internet of Vehicles. A thorough search of scholarly databases will be carried out, and relevant peer-reviewed papers will be chosen. We'll examine the benefits of federated learning in road traffic analysis and vehicle safety, as well as the drawbacks and restrictions of current centralized machine learning methods. Furthermore, the field's research gaps and future orientations will be determined [1].The goal of this systematic literature review is to further the development of machine learning (ML)-based solutions that are more precise and effective for road traffic analysis, impediment estimate, and IoV vehicle safety measures. The results of this research should help federated learning become more widely used in the Internet of Vehicles (IoV) and enhance the precision and security of road traffic analysis and vehicle safety measures [3].

**II. Review of Literature:**

An important part of research is a review of the literature, which provides a comprehensive overview of the body of previously known facts as well as gaps in a particular field. A comprehensive assessment of the literature on "Machine Learning (ML)" based road traffic, vehicle hindrances estimation, and vehicle safety measure prediction for the Internet of Vehicles (IoV) is presented in this part. Three subsections make up the structure of the review: review approach, research questions, and review. The process for doing the literature search and choosing pertinent articles is described in the review strategy. The review's distinct focus and the reader's guide are furnished by the research questions. The review portion concludes with a thorough analysis of the chosen papers, emphasizing important discoveries and outlining areas in need of further research. The purpose of this paper is to present a thorough overview of the state-of-the-art in machine learning (ML)-based road traffic analysis, impediment assessment, and vehicle safety measure prediction for IOV.We also seek to identify areas in need of more investigation as well as prospective avenues for future investigation. We anticipate that this overview of the literature will be an invaluable resource for practitioners, scholars, and decision-makers interested in using machine learning in the transportation business.

**A. Analyze Strategy:**

A potential approach to doing a systematic literature review on "Automotive Obstacles and Road Traffic based on Machine Learning" Calculating and Forecasting Safety Measures for Vehicles for Internet of Vehicles," which confirms the extent and requirement of Federated learning might entail the following:

1. Specify what the study questions are. Describe the study inquiries depending on the gaps and the goals of the research recognized in scholarly works. For instance:

• What is the state-of-the-art right now for ML-based road traffic analysis, assessment of obstacles, and prediction of IoV-related vehicle safety measures?

• What are the benefits and potential gains of federated learning for vehicle safety and road traffic analysis in the Internet of Vehicles?

• What are the current drawbacks and restrictions of conventional centralized machine learning techniques for IoV vehicle safety and traffic analysis?

• In machine learning-based solutions for road traffic analysis, obstruction estimation, and vehicle safety measures for the Internet of Vehicles, what are the research gaps and future directions?

2. Specify the requirements for inclusion and exclusion: Based on the research questions, specify the inclusion and exclusion criteria for the literature search. As an illustration, the inclusion criteria can be:

• Research published in peer-reviewed journals or presented at conferences

• Research on ML-based road traffic analysis, assessment of obstacles, and forecasting of safety precautions for vehicles an IoV

• Research that employ federated learning as a method to managing the enormous volumes of data produced by the IoV

Among the possible exclusion criteria are:

• Research not released in English

• Research not centered on road traffic analysis using machine learning, assessment of obstacles and forecasting vehicle safety precautions in the IoV

• Research that don't employ federated learning as a methodology to manage the enormous volumes of data produced by an IoV

3. Perform the search for literature: Perform the literature look through pertinent scholarly databases like additionally, "Google Scholar" and "ACM Digital Library" as "Internet Explorer." Use a combination of keywords such as "machine learning," "road traffic analysis," "estimation of vehicle hindrances," "vehicle safety measures," "Internet of Vehicles," and "federated learning" that are relevant to the study objectives and the inclusion criteria.

4. Examine the research: Establish the inclusion and exclusion criteria prior to the studies' screening. Screen the studies first by looking at the title and abstract, and then look through the complete texts of the research that you have chosen.

5. Take the data out: Take out the pertinent information from the chosen studies, including the goals, methodology, datasets, machine learning approaches, and findings.

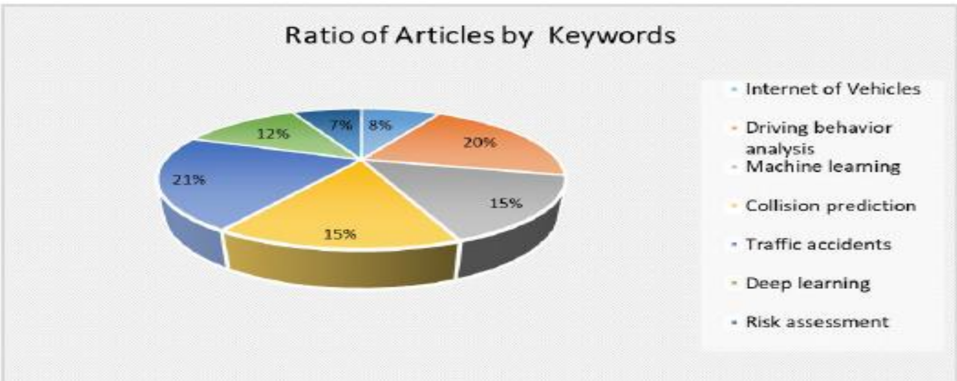
Analyze the information: Apply a methodical analysis technique, like a narrative synthesis or meta-analysis, to the acquired data. Determine the trends, patterns, and gaps in the literature by contrasting and comparing the findings of the chosen studies.

7. Draft the review of the literature: Write the review of the literature using the analysis of the chosen studies as a guide. Subheadings including "ML-based road traffic analysis," "Vehicle hindrances estimation," "Predicting vehicle safety measures," and "Federated learning" should be included in the literature study. List the main conclusions from each subsection and offer a critical assessment of the advantages and disadvantages of the body of current research.

8. Confirm the extent and necessity of federated learning: Using the literature review as a guide, confirm the extent and necessity of federated learning in ML-based road traffic analysis, obstacle estimate, and vehicle safety measure prediction for the Internet of vehicles. Determine the research gaps and future directions in this area and evaluate the benefits and possible advantages of federated learning to conventional centralized machine learning techniques.

9. Summarize the main conclusions and emphasize the study's contributions to wrap up the literature review section. Make suggestions for future study and discuss the practical ramifications for creating machine learning (ML)-based solutions that are more precise and effective for road traffic analysis, obstruction estimation, and IoV vehicle safety measures.

**Figure 1: The distribution of articles based on keywords**



**B. Questions for Research:**

1. What are the difficulties in estimating traffic and vehicle obstructions and forecasting vehicle safety measures for the Internet of Vehicles utilizing centralized machine learning approaches?

2. In the context of the Internet of Vehicles, how may federated learning be used to solve the shortcomings of centralized machine learning approaches?

3. In the context of the Internet of Vehicles, what are the drawbacks and possible advantages of federated learning for vehicle safety and traffic?

4. Which studies currently exist that have applied federated learning to car safety and traffic, and what are the conclusions and constraints of these studies?

5. In the context of vehicle safety and traffic for the Internet of Vehicles, what are the potential research avenues and uses for federated learning?

6. In the context of the Internet of Vehicles, what are the ethical and privacy issues surrounding federated learning, and how may they be resolved?

7. In the context of the Internet of Vehicles, how might federate learning be combined with other cutting-edge technologies like block-chain to improve road traffic and vehicle safety?

8. How can urban congestion reduction and traffic control be optimized with federated learning?

In order to integrate federated learning with machine learning-based road traffic and vehicle hindrances estimate and anticipate vehicle safety measures for the Internet of Vehicles, these research issues are crucial to demonstrating the need and scope of the literature study. These research questions will serve as a foundation for determining the shortcomings and restrictions of previous research in this field, as well as guide future research directions and opportunities for boosting road safety, streamlining traffic flow, and boosting vehicle performance within the context of the Internet of Vehicles. Every one of the 72 articles that were picked was chosen using pertinent keywords. An analysis was conducted on the frequency of the keywords and the number of relevant articles. Figure 1 shows the percentage of articles linked to each term.

A further filter applied to the articles based on the kind of publishing (journal, conference, book chapter, and other items). The distribution of the chosen articles according to their type is shown in Figure 2. The publisher then evaluated the remaining 72 items; no articles were dropped in this procedure. Additionally, a thorough analysis of the papers by publisher is shown in Figure 2.Two articles were eliminated in the last stage of the process when the articles were sorted according to the year of publication. Two further articles were eliminated after a qualitative synthesis factor assessment of the remaining 72 articles was conducted.The distribution of the papers by year of publication is shown in Figure 3.

**Figure 2 shows how articles are distributed by publishers.**

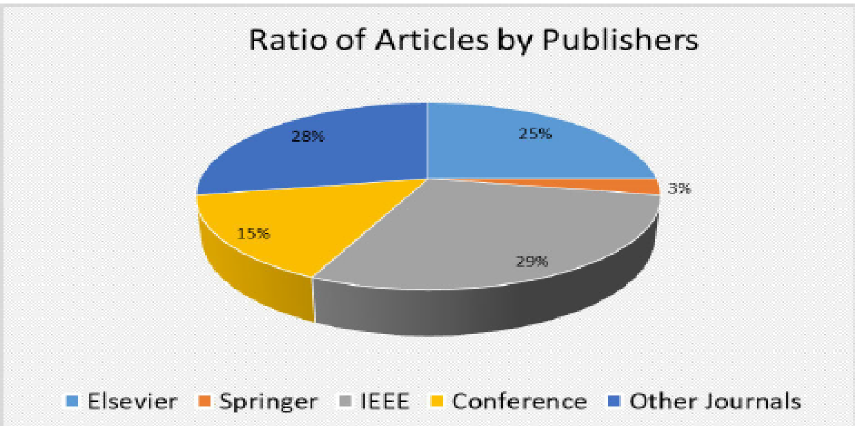
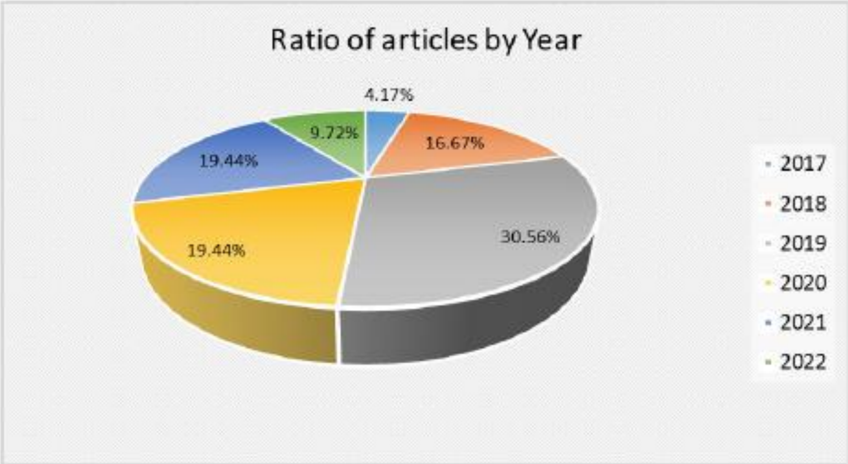


**Figure 3: Article fraction by year of publication**.




**C. Analyze**

**1) Driver Behaviour Assessment Based On Machine Learning**

The development of autonomous vehicles and driving safety both depend on the analysis of driver behavior. This section provides an overview and analysis of current papers that address different approaches to driver behavior analysis. Tabular 1.This survey of the literature demonstrates the many ways that machine learning (ML) is being applied to improve road safety.A novel approach called CRTI was suggested by Xiong et al. [4] to use machine learning and pattern recognition to predict traffic accidents.Jahangiri et al. [5] used basic safety message (BSM) data from connected vehicles to create an ML algorithm to identify driving behaviors that could be considered risky. Saleh and others [6] used smartphone sensor data to construct layered LSTM recurrent neural networks for the purpose of classifying driving behavior.

Gwak et al. [7] used ML algorithms in conjunction with driving performance variables to identify driver tiredness. An LSTM-NN model was presented by Huang et al. [8] to reproduce real-world traffic flow features and asymmetric driving behavior. Zhao et al. [9] used probing vehicle (PV) data to look at the travel habits of senior drivers. An adaptive Forward Collision Warning (FCW) system that reduces the frequency of false warnings and increases the acceptability of alerts was introduced by Iranmanesh et al. [10]. The review demonstrates how machine learning (ML) can improve road safety through various means such as accident prediction, hazardous driving action detection, driving behavior classification, driver drowsiness detection, replication of real-world traffic flow characteristics, investigation of senior driver travel behavior, and development of adaptive FCW systems.

A technique for employing smartphone sensors to identify driver distraction was presented by Xie et al. [11]. Using machine learning techniques, the data gathered from the accelerometer, magnetometer, and gyroscope sensors on the participants' smartphones was divided into categories for distracted and non-distracted driving.

With an F1 score of 0.91, the suggested process demonstrated high accuracy and was more exact than competing techniques. Nevertheless, the study's sample size was small, and external factors that can affect driver distraction were not taken into consideration. The CADSE system, a smartphone-based driving style evaluation system, was created by Bejani et al. [12]. Three subsystems make up the system, which classifies manoeuvres, calibrates smartphone sensors, and assesses driving patterns.Manoeuvres are classified by ensemble learning, and the distinction between safe and dangerous manoeuvres is made using a fuzzy inference system based on rules. The system's recall, F-measure, precision, and accuracy were all higher than 94%, 92%, 92%, and 93%, respectively. The suggested approach allows insurance firms to set driver risk-based prices and can be used by both the insurance and enforcement subsystems.

The Honda Research Institute Driving Dataset (HDD), introduced by Ramanishka et al. [13], offers a realistic and challenging chance to examine driver behavior and develop algorithms for driver assistance and self-driving capabilities. The collection seeks to understand how traffic situations and human driving interact. The authors offered a thorough comparison of the HDD dataset with other driving datasets as well as baseline techniques for identifying driver behavior. A system that combines dashcam and inertial sensors to detect unusual driving events was presented by Li et al. [14]. Anomalous circumstance identification capabilities of the system improve automated vehicle testing.The system uses an autoencoder-based technique to detect foreign viewpoints, and the unusual driving events aid in the retraining and improvement of the autonomous driving Model. It is possible that the suggested system will accelerate the development of robust systems for autonomous driving. Fugiglando et al. [15] proposed a procedure to look at and classifying the behavior of drivers using a certain subset of the CAN bus signaling. The utilization of unattended Learning was used to divide drivers into separate groups, and a validation strategy was used to evaluate the clustering results' durability across many testing situations. The current study offers a technique for categorizing the actions of drivers in unregulated situations in almost instantly.

According to Bouhoute et al. [16], numerical abstractionswithin the framework of cyber-physical system analysis,GauloisLinks are employed in the building of ongoingabstractions of system noises, and the writers offered both of theseNumerical intervals for driving dynamics and staticsinformation. The proposal described a process for using a continuous sound abstraction technique to analyze cyber-physical systems. Graph matching is used in similarity analysis to identify patterns and classify data. In Babol, Iran, Sheykhfard et al. [17] carried out a naturalistic driving study to look into interactions between drivers and pedestrians. In 216 vehicle-pedestrian encounters on divided highways and 485 such encounters on undivided roads, the study examined the conduct of 66 persons. The study discovered that the two main factors influencing vehicle-pedestrian interactions were the pedestrian's distance and the vehicle's speed.

Additionally, the study discovered that group crossings and pedestrians' awareness of traffic patterns before crossing the road improved driving performance. The application of machine learning (ML) to driving behavior analysis for auto insurance telematics was studied by Arumugam et al. [18]. The authors examined a variety of learning strategies used in the research of driver behavior, including deep, reinforcement, supervised, and unsupervised learning. According to the authors, pricing tactics, consumer involvement, and insurance risk evaluation could all benefit from the use of telematics. Papadimitriou et al.'s [19] study looked into the usage of cellphones while driving. To evaluate driving behavior and exposure, the study used a cohort of 100 drivers and a smartphone application. According to the study, using a phone while driving had a detrimental effect on one's driving style. The authors used machine learning algorithms to create predictive models that would detect when someone might use a mobile phone during driving.

A approach for identifying driving behaviors using sensor data from in-vehicle CANBUS was presented by Zhang et al. [20] by using deep learning (DL) techniques. This approach made automated acquisition of driving patterns and modeling temporal features without needing proficiency in feature modeling. The method was ability to record intricate temporal driving patterns that cannot be obtained using conventional techniques. Zhang and colleagues [21] presented a defensive alerting system to recognize and warn about potentially dangerous vehicles. The system made use of a cloud server and vehicle computation to assess the performance of the driver of all cars on the planet. The suggested system able to identify cars being driven carelessly and then alerted drivers in order to lessen the possibility of incidents.

The asymmetric behavior of car-following during traffic oscillation was studied by Wan et al. [22]. In Nanjing, China, an unmanned aerial vehicle (UAV) was used in the investigation to film the area around a roadway restriction. According to the research, drivers display a variety of unique yet consistent driving habits, and each driver's oscillation response depends on their pre-existing characteristics. A Long Short Term Memory Fully Convolutional Network (LTSM-FCN) was presented by Moukafih et al. [23] as a tool for detecting aggressive driving behavior. The suggested approach outperformed other approaches in terms of accuracy when assessed over a 5-minute period. The scientists noted that those with reduced inhibitory capacity have a greater tendency to break traffic laws and be involved in auto accidents.

In order to detect instances of tiredness and unusual driving behaviors, Chan et al. [24] performed a comparative examination of several sensing techniques and detection algorithms put forth by various authors. These methods make use of cellphones. The study highlights how important it is for active systems like ADAS and IDAS to warn drivers of possible dangers. The authors do, however, contend that there are still cutting-edge approaches in the literature that merit more investigation. Time limitations' impact on drivers' braking behavior and accident probability were studied by Pawar et al. [25]. The study discovered that braking behavior was significantly influenced by a number of characteristics, including gender, driving occupation, age, driving history, approach speed, and driving condition. The study provides insight into how time restrictions affect drivers' braking habits and accident risk, which might help with the creation of treatments meant to lessen the negative effects of time constraints on drivers' behavior and increase road safety.

Petraki et al. [26] used high-resolution smartphone sensor data to examine how drivers behaved at intersections and road segments. The study clarifies how traffic factors affect driver behavior and implies that using the findings to develop plans for improving road safety and reducing aggressive driving behavior may be helpful. Machine learning (ML) models were employed by AbouElassad et al. [27] to examine driver behavior (DB). The authors offer a theoretical framework that takes into account a number of aspects that are present in the "Driver-Vehicle-Environment (DVE)" system in order to conceptualize the phenomena of driver behavior (DB).The authors gave a summary of the ML literature and carried out a systematic literature review (SLR) on the idea of database investigation. This review highlights the higher performance of machine learning approaches over other models and their capacity to assess database performance. A review of naturalistic driving research on driver behavior and road safety was given by Singhet al. [28]. Key findings from previous naturalistic driving research have been provided by the authors, including the notion that driver conduct is the main factor causing the majority of traffic accidents. The writers have also made a substantial contribution to the field of naturalistic driving studies, which is useful for studying driver behavior and improving traffic safety.

Driver behavior analysis (DBA) underwent a thorough literature evaluation by Azadani et al. [29], who also looked at the field's obstacles and potential directions. The goal of the current proposal is to assist researchers in understanding the direction and difficulties of the DBA field, which improves road safety and reduces the likelihood of accidents. ModAL-IoCV was presented by Aboulola et al. [30] as a way to perform driver behavior analysis using DL approaches. The suggested method aims to provide itinerary planning and driver support in order to reduce accidents and improve traffic safety. In order to study driver behavior, the suggested technique uses a multimodal methodology that combines feature extraction, suggestion, route planning, and vehicle motion and lane change prediction.

Federated Machine Learning was used by Uprety et al. [31] to create an IoV misbehavior detection system that guarantees privacy preservation. The technology in question detects fabricated data in vehicular ad hoc networks (VANETs) while maintaining user privacy protection. The current idea uses the basic safety message (BSM) dataset of automobiles to develop a machine learning model that can detect occurrences of VANET position falsification attacks. The use of smartphone sensors to classify driver behavior in an efficient and cost-effective manner was investigated by Brahim et al. [32]. They discovered that smartphone sensors—like GPS, gyroscopes, and accelerometers—performed better at classifying activity than CAN-bus sensors. To find the best strategy, the study compared many machine learning methods for time series categorization. A number of sensors' interaxial data was combined to increase classification accuracy. The significance of driver behavior profiling for fleet management and insurance was emphasized by the writers.

In conclusion, the papers discussed in this section offer a variety of techniques for analyzing driver behavior. High accuracy, precision, and recall values are demonstrated by the suggested approaches, which may improve driver safety in real-world situations. Nevertheless, some restrictions were noted, including the small sample size, the particular driving tasks used, and the disregard for external factors. The datasets that the authors have made available are extremely helpful for researching driver behavior and creating algorithms for self-driving and driver assistance.

**Table 1: Synopsis of the reviewed literature on driving and driving behavior**

| Author | Methodology | Application | Key Findings |
|---|---|---|---|
| Xiong et al. [4] | ML and pattern recognition (CRTI framework) | Predicting road traffic accidents | ML has the potential to predict road accidents |
| Jahangiri et al. [5] | ML algorithm analysing basic safety message (BSM) data from connected vehicles | Detecting hazardous driving actions | ML can detect hazardous driving actions |
| Saleh et al. [6] | Stacked LSTM recurrent neural networks using smartphone sensor data | Classifying driving behavior | ML can classify driving behavior through smartphone sensor data |
| Gwak et al. [7] | ML algorithms using driving performance metrics, behavioral characteristics, and physiological responses | Detecting driver drowsiness | ML can detect driver drowsiness using a combination of metrics |
| Huang et al. [8] | LSTM-NN model | Replicating authentic traffic flow characteristics and asymmetric driving behavior | ML can replicate authentic traffic flow characteristics |
| Zhao et al. [9] | Probe vehicle (PV) data analysis | Investigating travel behaviors of elderly drivers | ML can investigate travel behaviors of elderly drivers |
| Iranmanesh et al. [10] | Adaptive Forward Collision Warning (FCW) system | Enhancing FCW system performance | Adaptive FCW systems can minimize false warnings and enhance warning acceptability |
| Xie et al. [11] | Smartphone sensors classification using ML algorithms | Detecting driver distraction | ML can accurately detect driver distraction using smartphone sensors |
| Bejani et al. [12] | Ensemble learning and rule-based fuzzy inference system using smartphone sensors | Evaluating driving styles | Smartphone-based evaluation systems can be used to evaluate driving styles for insurance and enforcement purposes |
| Ramanishka et al. [13] | Honda Research Institute Driving Dataset (HDD) and baseline algorithms | Investigating driver behavior for driver assistance and self-driving capabilities | The HDD dataset provides a practical and demanding opportunity for investigating driver behavior |
| Li et al. [14] | Dashcam and inertial sensor data analysis | Identifying atypical driving occurrences for autonomous driving testing | ML can identify atypical driving occurrences for the advancement of autonomous driving systems |
| Fugiglando et al. [15] | Unsupervised learning using CAN bus signals | Classifying driver behavior in uncontrolled environments | Unsupervised learning can be used to classify driver behavior in uncontrolled environments |
| Bouhoute et al. [16] | Continuous sound abstraction technique using Galois connections and graph matching | Analyzing cyber-physical systems in driving data | Continuous sound abstraction techniques can be used to analyze cyber-physical systems in driving data |
| Sheykhfard et al. [17] | Naturalistic driving study | Investigating driver-pedestrian interactions | Speed and distance primarily influence vehicle-pedestrian interactions |
| Arumugam et al. [18] | Review of ML techniques for motor insurance telematics | Analyzing driver behavior for motor insurance telematics | Telematics can enhance insurance risk evaluation, pricing strategies, and customer engagement |
| Papadimitriou et al. [19] | Mobile application and ML algorithms | Identifying instances of mobile phone usage during driving | Mobile phone usage while driving negatively impacts driving behavior |
| Zhang et al. [20] | DL methodology using in-vehicle CAN-BUS sensor data | Recognizing complex driving behaviors | DL can capture complex temporal driving behaviors |
| Zhang et al. [21] | Defensive alerting mechanism using cloud server and vehicle computation | Notifying drivers of hazardous automobiles | Defensive alerting mechanisms can mitigate the likelihood of accidents |
| Wan et al. [22] | Employed an unmanned aerial vehicle (UAV) to capture footage | Investigated the asymmetric behavior of car-following during traffic oscillation | Motorists exhibit a range of distinct yet uniform driving behaviors, and the oscillation response of each driver is contingent upon their pre-existing traits. |
| Moukafih et al. [23] | Introduced a Long Short Term Memory Fully Convolutional Network | Identifying instances of aggressive driving behavior | Individuals with lower inhibitory capacity exhibit a higher propensity for violating traffic laws and experiencing vehicular accidents. |
| Chan et al. [24] | Conducted a comparative analysis of various sensing schemes | Identifying instances of drowsiness and atypical driving actions | Emphasizes the significance of active systems such as ADAS or IDAS in alerting drivers of potential hazards. |
| Pawar et al. [25] | Investigated the effects of time constraints | Drivers' braking behavior and accident risk | Various factors, including gender, driving profession, age, driving history, approach speed, as well as driving |

| | | | |
|---|---|---|---|
| | | | condition, had a significant impact on braking behavior. |
| Petraki et al. [26] | Analyzed the driving behavior of drivers | Driving behavior of drivers at road segments and junctions | The impact of traffic characteristics on driver behavior and suggests that the results may be utilized to formulate strategies aimed at enhancing road safety and mitigating aggressive driving behavior. |
| Abou Elassad et al. [27] | Utilized machine learning (ML) models | Analyze driver behavior (DB) | Machine learning techniques have the ability to evaluate database performance and exhibit superior performance compared to alternative models. |
| Singhet et al. [28] | Provided a summary of naturalistic driving studies | Driver behavior and road safety | Driver behavior is the primary cause of most road accidents. |
| Azadani et al. [29] | Conducted a comprehensive literature review | Driver Behavior Analysis (DBA) examining its challenges and future trends | Facilitate researchers in comprehending the research direction and challenges pertaining to the field of DBA, which enhances road safety and mitigates occurrences of accidents. |
| Aboulola et al. [30] | Proposed MODAL-IoCV as a means of conducting driver behavior analysis | Driver support and itinerary preparation with the objective of mitigating accidents and enhancing traffic safety | The proposed approach employs a multimodal methodology that integrates vehicle motion and lane change prediction, feature extraction, recommendation, and route planning to analyze drivers' behavior. |
| Uprety et al. [31] | Federated Machine Learning | Misbehavior detection in Vehicular Ad hoc Networks (VANET) | The proposed system detects falsified data in VANET while preserving user privacy, using basic safety message (BSM) dataset. |
| Brahim et al. [32] | Smartphone sensors and ML algorithms | Driver behavior profiling in insurance and fleet management | Smartphone sensors were found more effective than CAN-bus sensors in categorizing driver behavior, and integrating inter-axial information from multiple sensors improved classification accuracy. |

**2) Prediction Of Machine Learning Based Collision**

**Scope**

Research on traffic safety is now receiving more attention as a result of developments in the Internet of Vehicles (IoV). The current research on IoV-based traffic safety is examined in this section Table 2. According to Hyder et al. [33], governments must assume responsibility for maintaining road safety and providing funding for efficient interventions to reduce accidents. Road traffic injuries and fatalities are a major global public health concern. Zong et al. [34] created a severity causation network that forecasts severity indices using information entropy and Bayesian networks. This helps managers assess the severity of traffic accidents to improve safety protocols and reduce casualties and property damage. For the purpose of predicting short-term traffic flow, Yin et al. [35] presented a hybrid model and algorithm that combines GPSOWNN with ARIMA, demonstrating improved performance in predicted accuracy. A real-time object identification and dynamic prediction methodology for forecasting vehicle positions in Internet of Vehicles (IoV) frameworks was presented by Chang et al. [36].

The TASP-CNN method was presented by Zheng et al. [37] as a unique way to forecast the severity of traffic accidents. Combination relationships are used to improve predictive accuracy. Compared to other modern approaches; the Traffic Predict model, created by Ma et al. [38], shows better prediction accuracy. It forecasts traffic-agent trajectories in a variety of traffic situations by utilizing long short-term memory (LSTM). In order to detect vehicle-pedestrian confrontations, Lv et al. [39] suggested using high-resolution traffic trajectories from roadside LiDAR sensors in conjunction with SDP geographical distribution. In order to enhance classification outcomes, Peng et al. [40] explored the assessment of the importance of condition qualities inside equivalency classes and produced decision rules based on equivalency classes. In order to estimate collision risk in autonomous vehicles, Katrakazas et al. [41] developed an integrated approach that forecasts collisions on road segments using machine learning approaches, traffic simulation, and collision risk assessment. Ayuso et al. [43] presented a count data regression model for frequency, incorporating telematics data to accommodate driver behavior and other variables that could potentially impact insurance expenses. Gao et al. [42] developed a predictive model for insurance claims based on driver behavior. The application of telematics technology holds promise for improving the way traffic regulatory agencies determine insurance rates.

Numerous research projects have been carried out to improve traffic safety through the prediction and mitigation of traffic accidents. A traffic accident prediction technique based on convolutional neural networks (CNN) was presented by Zhao et al. [44]. To automatically extract multi-dimensional features from Vehicle Ad-hoc Networks (VANET) data, they used deep learning techniques. When compared to other machine learning techniques, the CNN-based model demonstrated better prediction accuracy and reduced loss during training and testing using a dataset of simulated car accidents. It is anticipated that the theoretical underpinnings of the model will aid in the advancement of intelligent vehicle path planning optimization, anti-collision guidance, and vehicle safety assisted driving.A collision warning model for Advanced Driver Assistance Systems (ADAS) in a Vehicle-to-Vehicle (V2V) communication context was proposed by Lyu et al. [45]. Both driving simulator and in-person vehicle testing were used to assess the model's lane-change behavior and trajectory prediction models.

The suggested model offered new modeling ideas and theoretical support for the optimization of ADAS cut-in functionality, outperforming traditional models in terms of warning confusion matrix and warning time. Using deep learning techniques, Zhao et al. [46] presented an algorithm for forecasting the likelihood of traffic accidents in vehicular edge networks. The approach extracts multi-dimensional features from a significant amount of traffic data coming from the edge network of automobiles by using a CNN inside an edge server. The vehicle collision prediction model alerts drivers to slow down and be cautious. Four popular machine learning techniques and two statistical methods were compared to predict the severity of collision injuries. The simulations showed that the algorithm under consideration shows reduced loss and improved prediction accuracy compared to other machine learning algorithms.

For the assessment of emergency driving safety, Peng et al. [47] recommended using real driving data. They used a variable precision rough set (VPRS)-based classification technique to take a condensed core subset—which includes the most important attributes for assessing driving safety—out of the driving dataset. The proposed method demonstrates a high degree of accuracy and consistency, making it appropriate for determining timely emergency braking actions, increasing the effectiveness of collision avoidance systems (CASs). In order to estimate the benefits of improving vehicle design, Bhalla et al. [48] evaluated road safety improvements in the Latin America and Caribbean (LAC) region.

**Table 2: Summary of the reviewed articles pertaining to Collision Scope Prediction using Machine Learning**

| Author | Methodology | Application | Key Findings |
|---|---|---|---|
| Hyder et al. [33] | Review | Road safety | Governments need to ensure road safety and provide resources for effective interventions to mitigate accidents. |
| Zong, Fang, et al. [34] | Information entropy, Bayesian networks | Traffic accident severity forecasting | Severity causation network aids managers in scrutinising the severity of traffic accidents to enhance safety measures and minimise casualties and property damages. |
| YIN, Lisheng et al. [35] | ARIMA, GPSOWNN | Traffic flow prediction | Hybrid model and algorithm exhibit superior performance in predictive accuracy. |
| Chang, Che-Cheng et al. [36] | Real-time object detection, dynamic prediction methodology | Vehicular position prediction | Real-time object detection and dynamic prediction methodology for predicting vehicle positions in IoV frameworks. |
| Zheng, Ming, et al. [37] | TASP-CNN method | Traffic accident severity prediction | TASP-CNN method enhances predictive accuracy through the incorporation of combination relationships. |
| Ma, Yuexin et al. [38] | LSTM | Traffic-agent trajectory forecasting | TrafficPredict model exhibits superior predictive accuracy compared to other contemporary techniques. |
| Lv, Bin et al. [39] | SDP spatial distribution | Vehicle-pedestrian conflict identification | High-resolution traffic trajectories obtained from roadside LiDAR sensors used to identify vehicle-pedestrian conflicts through the use of SDP spatial distribution. |
| Peng, Liqun, et al. [40] | Variable precision rough set (VPRS) | Emergency driving safety evaluation | Classification technique based on VPRS extracts a condensed core subset from the driving dataset, encompassing the most significant attributes for evaluating driving safety. |
| Katrakazas, Christos et al. [41] | Traffic simulation, collision risk assessment, machine learning | Collision risk assessment in autonomous vehicles | Integrated method for assessing collision risk in autonomous vehicles using traffic simulation, collision risk assessment, and machine learning techniques to forecast collisions on road segments. |
| Gao, Guangyuan , et al. [42] | Predictive model for insurance claims | Driver behaviour and insurance expenses | Predictive model for insurance claims based on driver behaviour. |
| Ayuso, Mercedes et al. [43] | Count data regression model | Driver behaviour and insurance expenses | Count data regression model incorporating telematics data to accommodate driver behaviour and other variables that could potentially impact insurance expenses. |
| Zhao et al. [44] | CNN | Traffic accident prediction | CNN-based model exhibits higher prediction accuracy and lower loss compared to other machine learning algorithms. |
| Lyu et al. [45] | Collision warning model | Advanced Driver Assistance Systems (ADAS) | Collision warning model for ADAS in a Vehicle-to-Vehicle (V2V) communication setting outperforms conventional models in terms of warning confusion matrix and warning time. |
| Zhao et al. [46] | Deep learning techniques | Risk of traffic accidents prediction | Algorithm utilises a CNN within an edge server to extract multi-dimensional features from a substantial amount of traffic data originating from the edge network of vehicles. |
| Peng et al. [47] | Variable precision rough set (VPRS) | Emergency driving safety evaluation | Suggested approach exhibits a notable level of precision and consistency, rendering it suitable for deducing prompt emergency braking maneuvers. |
| Bhalla et al. [48] | Counterfactual analysis | Vehicle safety interventions | Evaluation of road safety interventions estimating the advantages of enhancing vehicle design. |
| Walugembe et al. [49] | Investigation | Road safety in Tanzania | Investigation suggests enhancing the road transportation infrastructure to guarantee the |
| Wang et al. | DL techniques, CNN model | Rear-end collision prediction | RCPM exhibits superior performance compared to other algorithms. |
| Pawar et al. | Driving simulator | Braking behaviour and likelihood of accidents | Various factors influence the braking behaviour of drivers. |
| Lin et al. | Crowdsourced data | Traffic flow prediction | Enhancing traffic condition data leads to improved precision. |
| Gutierrez-Osorio et al. | Review of contemporary ML techniques | Accident analysis | Bayesian networks are identified as the most precise models for accident analysis. |
| Yu et al. | DSTGCN model | Traffic accident forecasting | DSTGCN exhibits superior performance compared to traditional and contemporary techniques. |
| Fang et al. | SCAFNet model | Driver attention forecasting | SCAFNet demonstrates exceptional performance on three distinct datasets. |
| Lee et al. | Driving behaviour model | Collision avoidance for two-wheeled vehicles | Theoretical framework offers a predictive model for avoiding collisions by riders who engage in risky behaviour. |
| Li et al. | Risk assessment algorithm | Autonomous vehicle decision-making | The proposed approach addresses the needs of both drivers and passengers by implementing a collision avoidance strategy that takes into account driving style preferences. |
| Zhao et al. | Integration of online and offline data | Driver risk classification for UBI | Integrating both types of data enhances the accuracy of risk assessments, with offline consumer behavior variables being particularly significant. |
| Lin et al. | Predictive model using Bayes' theorem | Intersection accident risk assessment | Road dimensions, designated speed limits, and markings along the roadside are significant determinants of the likelihood of traffic |

| | | | |
|---|---|---|---|
| | | | accidents at intersections. |
| Chang et al. | Machine learning architecture | Vehicular position prediction for collision avoidance | Their algorithm outperforms a prior study's linear algorithm, demonstrating the potential of machine learning techniques and vehicle dynamics principles to enhance future vehicular position prediction. |
| Lin et al. [60] | Predictive model using Bayes' theorem and traffic accident data | Risk assessment for intersection accidents | Road dimensions, designated speed limits, and roadside markings significantly impact likelihood of intersection accidents |
| Chang et al. [61] | Vehicle dynamics and machine learning architecture | Predicting vehicular positions for collision avoidance in IoV | Machine learning techniques and vehicle dynamics principles can enhance future vehicular position prediction |
| Yang et al. [62] | Transfer learning based on Internet of Vehicles data | Forecasting vehicular collisions | High degree of precision in forecasting vehicular collisions |
| Gutierrez-Osorio et al. [63] | Ensemble deep learning approach with "Gated Recurrent Units" and "CNN" | Predicting traffic collisions using social media and open data | Deep learning ensemble model exhibits superior performance compared to other methods |
| Morimoto et al. [64] | Comparative analysis of global traffic safety objectives and strategies | Enhancing road traffic safety | Comprehensive strategy needed encompassing infrastructure, technology, culture, and behaviour |
| Malawade et al. [65] | Spatio-temporal scene graph embedding using visual perception, GNN, and LSTM | Forecasting collisions involving autonomous vehicles | SG2VEC model demonstrates superior predictive accuracy and earlier detection capabilities compared to current leading approach |
| Suat-Rojas et al. [66] | Named entity recognition and geocoding | Extracting data related to traffic accidents from Twitter | Twitter has the potential to furnish information regarding traffic accidents, with commercial and industrial areas of the city being impacted by Twitter |
| Pan et al. [68] | Communication network architecture using LoRa technology and LSTM/ANN | Early warning system for pedestrian-vehicle collision in uncertain situations | Probability of confidence is a determining factor for issuing warnings for pedestrian-vehicle collisions |

The study evaluated the potential reduction in fatalities and disability-adjusted life years lost in LAC nations if eight well-established automobile safety technologies were more widely available through a counterfactual analysis.The study calculated how many lives would be spared and years of disability-adjusted life that would be avoided if these technologies were installed in every car. The incidence of road traffic accident (RTA)-related deaths in prehistoric Ilala and two other municipalities in Tanzania's Dar es Salaam Region were examined by Walugembe et al. [49]. According to the research, the infrastructure for road transportation should be improved to ensure the safety of users by enforcing existing laws, enforcing them more strictly, and introducing swift and severe punishments.

In order to improve the prediction of rear-end crashes, which are a major cause of traffic accidents, Wang et al. [50] developed the Rear-end Collision Prediction Mechanism (RCPM), which makes use of DL approaches. In order to train the CNN model, training and testing sets from the preprocessed dataset are utilized. The study shows that when it comes to rear-end collision prediction, RCPM performs better than other algorithms. Using a driving simulator, Pawar et al. [25] looked at the relationship between driver braking techniques and the chance of accidents in reaction to elevated time constraint. Their study revealed that a number of variables, such as age, driving experience, approach speed, driving occupation, gender, and driving circumstances have a big impact on the braking actions of motorists. Lin and colleagues [51] employed crowd sourcing data to predict patterns of traffic flow following an accident and discovered that improving data on traffic conditions results in an improved with accuracy.

In this study, Gutierrez-Osorio et al. [52] A thorough analysis of modern machine learning methods for analyzing and forecasting traffic incidents and concluded that it has been determined that Bayesian networks are the most accurate accident analysis models. According to Yu et al. [53], a "Deep Spatio-Temporal Graph" Utilizing Convolutional Network (DSTGCN) to predict traffic incidents. Their model performs better when compared to both modern and conventional methods when applied to datasets from the real world. The "Semantic Context-Induced Attentive Fusion Network (SCAFNet)" was developed by Fang et al. [54] to forecast driver attention in the setting of "driving accident scenarios (DADA)". Their method performed remarkably well on three different datasets. A model of driving behavior was developed by Lee et al. [55] with the goal of predicting how riders of two-wheeled vehicles would avoid hazardous collisions. Their theoretical framework provides a predictive model for motorcyclists who participate in risky behavior to avoid collisions.

A method for making decisions in autonomous vehicles based on risk assessment was presented by Li et al. [56]. The installation of a collision avoidance technique that considers driving preferences is their suggested method for meeting the needs of both drivers and passengers. Lastly, driving behavior in collisions and near-misses between cars and bikes was examined by Zhao et al. [57]. It was discovered that the "Bus Rapid Transit (BRT)" parameter is crucial in averting crashes. After analyzing traffic incidents on Shenzhen's roads, Li et al. [58] recommended that China use "advanced driver assistance systems (ADASs)" and traffic management countermeasures.The effect of combining data from online and offline channels on driver risk categorization for usage-based insurance (UBI) products was examined by Zhao et al. [59]. LR, NN, RF, and SVM were utilized to categorize driver risk using driving behavior data from On-Board Diagnostics (OBD) loggers and offline consumer behavior data from 4S dealerships. The accuracy of risk evaluations was found to be improved by merging the two forms of data, with offline customer behavior characteristics being especially important.

The insurance industry's approach to UBI pricing and cost control may be affected by these findings. Lin et al. [60] used traffic accident data and the Bayes theorem to create a predictive model for the risk of intersection accidents. The size of the road, posted speed limits, and signs beside the road were all determined to be important factors in predicting the chance of traffic accidents at crossings. The predictive model is able to evaluate possible risks in order to reduce the frequency of car crashes and offer guidance for intersection design and environmental improvement. A novel method for forecasting vehicle positions for collision avoidance in the Internet of Vehicles was proposed by Chang et al. [61].

Building on the YOLOv4 real-time object recognition project, the authors optimized future vehicle location prediction through the application of machine learning architecture and vehicle dynamics. The study discovered that their system performed better than the linear approach used in a previous study, highlighting the potential for machine learning techniques and concepts of vehicle dynamics to improve the prediction of future vehicular positions. A vehicle accident prediction model based on transfer learning and utilizing data from the Internet of Vehicles was presented by Yang et al. [62]. By analyzing and recreating the vehicle's operating mode, the model used historical performance data to anticipate future collision occurrences and their corresponding timings. The algorithm demonstrated a high level of accuracy in predicting car crashes, which improves road safety and upcoming driving technologies.

In order to forecast traffic collisions, Gutierrez-Osorio et al. [63] suggested employing social media and open data together with an ensemble deep learning technique made up of "Gated Recurrent Units" and "CNN." According to the study, their Deep Learning ensemble model performed better than baseline algorithms and other deep learning techniques. A theoretical framework for improving road traffic safety through a comparative analysis of global traffic safety objectives and tactics was developed by Morimoto et al. [64]. The research emphasized the need for a thorough approach that addresses infrastructure, technology, culture, and behavior in order to guarantee road safety. A spatio-temporal scene graph embedding model named SG2VEC was developed by Malawade et al. [65] to predict collisions involving autonomous vehicles (AVs). To predict collisions, the model combines LSTM and GNN layers with visual scene perception. In comparison to the state-of-the-art method, the SG2VEC model showed better predicted accuracy and early detection capabilities.

A mechanism for extracting information about road accidents in Spanish from Twitter was proposed by Suat-Rojas et al. [66]. According to the study, Twitter can provide information on traffic accidents, and it has an impact on the city's commercial and industrial areas. The suggested approach makes use of geocoding and named entity recognition to address problems with informal language and misspellings. Last but not least, Pan et al. [67] proposed an early warning system and communication network architecture for Vehicle to Pedestrian (V2P) using LoRa technology. When pedestrian trajectories are unclear, the system uses LSTM and ANN to forecast the area of risk for a collision between a pedestrian and a vehicle. One of the factors that determine whether to issue warnings for pedestrian-vehicle incidents is the probability of confidence.This section's research offer insightful information on a range of topics related to road safety, accident prevention, machine learning analysis and prediction of incidents, and the creation of models for hazardous collision avoidance prediction. These studies can help with the creation of practical plans and tools that will increase traffic safety and lower accident rates.

**3) Predicting traffic congestion using machine learning**

A number of studies conducted recently have concentrated on creating machine learning-based traffic collision and congestion prediction algorithms. In order to increase the accuracy of driver risk assessment in usage-based insurance products, Zhao et al. [59] combined online and offline data. A hybrid CNN and BLSTME model was presented by Kothai et al. [68] to predict traffic congestion levels. This model successfully handled the overfitting problem and achieved a high degree of accuracy. In their analysis of the causes and effects of traffic congestion in urban areas, Onyeneke et al. [69] put out a number of policy suggestions for alleviation. In order to anticipate short-term traffic flow, Du et al. [70] developed a hybrid multimodal deep learning technique that outperformed other baseline approaches.

An AutoEncoder Long Short-Term Memory method for traffic flow prediction was proposed by Wei et al. [71] and showed better performance than previous methods. In Montreal, Canada, Hébert et al. [72] developed high-resolution accident predicting tools using big data analytics and found multiple important predictors of automobile collisions. A methodology for vehicle traffic prediction using machine learning techniques was developed by Moses et al. [73]. In comparison to previous DNNs, Ranjan et al.'s [74] unique neural network architecture demonstrated better computing efficiency and prediction performance when used to forecast traffic congestion levels over the course of an entire urban road network. Although these research provide intriguing solutions to traffic collision and congestion issues, each model has some limitations that should be considered before using them in practical situations. Additional investigation is required to resolve these constraints and maximize the effectiveness, Considering these suggested approaches' viability in intelligent systems of transportation Table 3.

## III. Observations

Centralized methods for machine learning in the automotive road safety and traffic in relation to the Internet of Vehicles (IoV) encounter major difficulties because of centralized gathering and storing of info. These methods might not be scalable for big datasets, requiring a lot of computer resources, and might not be strong enough to withstand cyber-attacks. One networked machine learning technique is federated learning, may overcome these obstacles by making it possible for data to be processed and instructed locally on personal gadgets, so addressing security and privacy issues. Moreover, Federated Learning lowers communication expenses, increases scalability, and is stronger in the face of hostile onslaught. Research studies that have already been conducted have shown how beneficial federated learning is for road traffic and vehicle safety in the context of the Internet of Vehicles (IoV). However, there are several drawbacks, such as the requirement for substantial processing resources, data heterogeneity, and restricted data sharing among devices. In the context of road traffic and vehicle safety, future research directions for federated learning include examining real-time decision-making, streamlining communication protocols and models, and resolving privacy and ethical issues.

In the context of the Internet of Vehicles, federated learning can be used with other cutting-edge technologies, like blockchain, to improve road traffic and vehicle safety. Blockchain technology can guarantee safe and private data sharing by providing a decentralized platform for exchanging and storing data across various entities, such as cars and traffic management systems. The creation of a blockchain-based platform for data exchange among cars to enhance traffic flow and lessen congestion is one possible use. Using the shared data, federated learning can create predictive models that are then utilized to optimize traffic flow in real time. Federated learning can help urban areas manage traffic better and experience less congestion.Federated learning can produce predictive models that offer other routes, optimize traffic signal timings, and reroute vehicles. Creating a federated learning-based platform to optimize public transportation in cities is one possible use case. This platform would integrate data from various sources to create predictive models for public transportation demand and recommend the best bus and train routes and times.

### 1. Scope of Research

Some potential research goals for a study on federated learning in the context of road traffic and vehicle safety for the Internet of Vehicles are as follows, based on the research questions:

1. To recognize and devise solutions for the ethical and privacy issues related to federated learning in the Internet of Vehicles.

2. To investigate how, in the context of the Internet of Vehicles, federated learning may be integrated with other cutting-edge technologies, including blockchain, to improve traffic and vehicle safety.

3. To look into the viability and efficiency of federated learning for improving traffic control and easing urban congestion.

4. To perform an extensive analysis of the literature on previous studies that have employed federated learning for traffic and vehicle safety, and to assess the advantages and disadvantages of those studies.

5. To determine potential areas for future study and applications of federated learning for the Internet of Vehicles in the areas of road traffic and vehicle safety.

**TABLE 3: A summary of the publications examined under the heading "Traffic Congestion Prediction based on Machine Learning."**

| Author | Methodology | Application | Key Findings |
|---|---|---|---|
| Kothai et al. [68] | Hybrid BLSTME and CNN model for predicting traffic congestion levels | Traffic congestion prediction | The proposed model effectively addressed the issue of overfitting and attained a high level of accuracy. |
| Onyeneke et al. [69] | Examination of factors contributing to and consequences of congestion in road traffic in metropolitan regions | Policy measures for congestion mitigation | The study proposed several policy measures for mitigating congestion. |
| Du et al. [70] | Hybrid multimodal deep learning method for predicting short-term traffic flow | Short-term traffic flow prediction | The proposed method surpassed various baseline methods. |
| Wei et al. [71] | AutoEncoder Long Short-Term Memory approach for traffic flow prediction | Traffic flow prediction | The proposed approach demonstrated superior performance compared to prior techniques. |
| Hébert et al. [72] | High-resolution accident forecasting using big data analytics | Accident forecasting | Several significant indicators of vehicle collisions were identified. |
| Moses et al. [73] | Methodology for vehicular traffic prediction using machine learning techniques | Vehicular traffic prediction | The proposed methodology utilised machine learning techniques. |
| Ranjan et al. [74] | Neural network architecture for predicting traffic congestion levels across an entire urban road network | Urban road network congestion prediction | The proposed neural network architecture exhibited superior computational efficacy and prediction performance compared to other DNNs. |

**B. purpose of The Research**

In order to prevent consequences like accidents and death from road transportation accidents, it is advised that more research be done on developing machine learning strategies to estimate the obstacles caused by traffic, vehicle health, driver cognitive levels, and driving behavior. To make this happen, the following major issues must be resolved: The data that collects from cars in an unstructured format is first managed and accessed via Federated Learning, which is based on blockchain technology. The data that IoV continuously gathers is then utilized to do feature engineering under federated learning, which is used to forecast driving behavior, driver cognitive levels, traffic and vehicle obstacles, and road transportation situations. These forecasts will be used to estimate the likelihood of accidents, automobile accidents, and safe. Lastly, the difficulties with intelligent traffic signals, road congestion management, and safe route evaluation will make an effort to use Federated Learning to address moves in. Enhancing road safety is the ultimate aim of this research to lower the number of road-related accidents and fatalities incidents. By tackling these difficulties by means of the goal of machine learning is to identify problems before they arise.

**IV. Conclusion:**

The application of machine learning techniques to traffic and safety of vehicles in relation to the Internet of Vehicles (IoV) has received a great deal of attention lately. But still, Centralized data gathering and archiving offer important obstacles, such as communication and privacy issues prices. Federated learning has come to light as a possible remedy for dealing with these restrictions. In the context of the Internet of Vehicles, this study of the literature offers an overview of current developments in the application of federated learning for road traffic and vehicle safety. A number of advantages of federated learning were noted in the assessment, such as enhanced security and privacy, scalability, and resilience to adversarial attacks. Nevertheless, there are drawbacks as well, such as the requirement for substantial computer power and the possibility of overfitting or model divergence. The paper outlines a number of previous experiments that have applied federated learning to car safety and traffic, but more investigation is required to confirm the technology's efficacy and scalability in practical settings. Overall, the work highlights how federated learning can overcome the shortcomings of centralized machine learning techniques and provides suggestions for additional research and advancements in this area.

**References:**

[1] M. N. Sadiku, M. Tembely, and S. M. Musa, ‘‘Internet of Vehicles:An introduction,’’ Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 8, no. 1,p. 11, 2018.

[2] O. Kaiwartya, A. H. Abdullah, Y. Cao, A. Altameem, M. Prasad, C.-T. Lin,and X. Liu, ‘‘Internet of Vehicles: Motivation, layered architecture,network model, challenges, and future aspects,’’ IEEE Access, vol. 4,pp. 5356–5373, 2016.

[3] S. Abbasi, A. M. Rahmani, A. Balador, and A. Sahafi, ‘‘Internet ofVehicles: Architecture, services, and applications,’’ Int. J. Commun. Syst.,vol. 34, no. 10, p. e4793, Jul. 2021.

[4] X. Xiong, L. Chen, and J. Liang, ‘‘A new framework of vehicle collisionprediction by combining SVM and HMM,’’ IEEE Trans. Intell. Transp.Syst., vol. 19, no. 3, pp. 699–710, Mar. 2018.

[5] A.Jahangiri, V. J. Berardi, and S. G. Machiani, ‘‘Application of real fieldconnected vehicle data for aggressive driving identification on horizontalcurves,’’ IEEE Trans. Intell. Transp. Syst., vol. 19, no. 7, pp. 2316–2324,Jul. 2018.

[6] K. Saleh, M. Hossny, and S. Nahavandi, ‘‘Driving behavior classificationbased on sensor data fusion using LSTM recurrent neural networks,’’ inProc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC), Oct. 2017, pp. 1–6.

[7] J. Gwak, M. Shino, and A. Hirao, ‘‘Early detection of driver drowsinessutilizing machine learning based on physiological signals, behavioral measures,and driving performance,’’ in Proc. 21st Int. Conf. Intell. Transp.Syst. (ITSC), Nov. 2018, pp. 1794–1800.

[8] X. Huang, J. Sun, and J. Sun, ‘‘A car-following model consideringasymmetric driving behavior based on long short-term memory neural networks,’’Transp. Res. C, Emerg. Technol., vol. 95, pp. 346–362, Oct. 2018.

[9] Y. Zhao, T. Yamamoto, and T. Morikawa, ‘‘An analysis on older driver’sdriving behavior by GPS tracking data: Road selection, left/right turn, anddriving speed,’’ J. Traffic Transp. Eng., vol. 5, no. 1, pp. 56–65, Feb. 2018.

[10] S. M. Iranmanesh, H. N. Mahjoub, H. Kazemi, and Y. P. Fallah,‘‘An adaptive forward collision warning framework design based ondriver distraction,’’ IEEE Trans. Intell. Transp. Syst., vol. 19, no. 12,pp. 3925–3934, Dec. 2018.

[11] J. Xie, A. R. Hilal, and D. Kulic, ‘‘Driver distraction recognition basedon smartphone sensor data,’’ in Proc. IEEE Int. Conf. Syst., Man, Cybern.(SMC), Oct. 2018, pp. 801–806.

[12] M. M. Bejani and M. Ghatee, ‘‘A context aware system for driving styleevaluation by an ensemble learning on smartphone sensors data,’’ Transp.Res. C, Emerg. Technol., vol. 89, pp. 303–320, Apr. 2018.

[13] V. Ramanishka, Y.-T. Chen, T. Misu, and K. Saenko, ‘‘toward drivingscene understanding: A dataset for learning driver behavior and causalreasoning,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.,Jun. 2018, pp. 7699–7707.

[14] H. Li, H. Wang, L. Liu, and M. Gruteser, ‘‘Automatic unusual drivingevent identification for dependable self-driving,’’ in Proc. 16th ACM Conf.Embedded Networked Sensor Syst., Nov. 2018, pp. 15–27.

[15] U. Fugiglando, E. Massaro, P. Santi, S. Milardo, K. Abida, R. Stahlmann,F. Netter, and C. Ratti, ‘‘Driving behavior analysis through CAN bus datain an uncontrolled environment,’’ IEEE Trans. Intell. Transp. Syst., vol. 20,no. 2, pp. 737–748, Feb. 2019.

[16] A. Bouhoute, R. Oucheikh, K. Boubouh, and I. Berrada, ‘‘Advanced drivingbehavior analytics for an improved safety assessment and driver fingerprinting,’’IEEE Trans. Intell. Transp. Syst., vol. 20, no. 6, pp. 2171–2184,Jun. 2019.
[17] Sheykhfard and F. Haghighi, ‘‘Performance analysis of urban driversencountering pedestrian,’’ Transp. Res. F, Traffic Psychol. Behaviour,vol. 62, pp. 160–174, Apr. 2019.
[18] S. Arumugam and R. Bhargavi, ‘‘A survey on driving behavior analysis inusage based insurance using big data,’’ J. Big Data, vol. 6, no. 1, pp. 1–21,Dec. 2019.
[19] E. Papadimitriou, A. Argyropoulou, D. I. Tselentis, and G.Yannis, ‘‘Analysisof driver behaviour through smartphone data: The case of mobile phoneuse while driving,’’ Saf. Sci., vol. 119, pp. 91–97, Nov. 2019.
[20] J. Zhang, Z. Wu, F. Li, C. Xie, T. Ren, J. Chen, and L. Liu, ‘‘A deeplearning framework for driving behavior identification on in-vehicle CANBUSsensor data,’’ Sensors, vol. 19, no. 6, p. 1356, Mar. 2019.
[21] L. Zhang, L. Yan, Y. Fang, X. Fang, and X. Huang, ‘‘A machine learningbaseddefensive alerting system against reckless driving in vehicularnetworks,’’ IEEE Trans. Veh. Technol., vol. 68, no. 12, pp. 12227–12238,Dec. 2019.
[22] Q. Wan, G. Peng, Z. Li, F. Inomata, Y. Zheng, and Q. Liu, ‘‘Using asymmetrictheory to identify heterogeneous drivers’ behavior characteristicsthrough traffic oscillation,’’ IEEE Access, vol. 7, pp. 106284–106294,2019.
[23] Y. Moukafih, H. Hafidi, and M. Ghogho, ‘‘Aggressive driving detectionusing deep learning-based time series classification,’’ in Proc. IEEE Int.Symp.Innov.Intell. Syst. Appl. (INISTA), Jul. 2019, pp. 1–5.
[24] T. K. Chan, C. S. Chin, H. Chen, and X. Zhong, ‘‘A comprehensive reviewof driver behavior analysis utilizing smartphones,’’ IEEE Trans. Intell.Transp. Syst., vol. 21, no. 10, pp. 4444–4475, Oct. 2020.
[25] N. M. Pawar, R. K. Khanuja, P. Choudhary, and N. R. Velaga, ‘‘Modellingbraking behaviour and accident probability of drivers under increasingtime pressure conditions,’’ Accident Anal.Prevention, vol. 136, Mar. 2020,Art.no. 105401.
[26] V. Petraki, A. Ziakopoulos, and G. Yannis, ‘‘Combined impact of road andtraffic characteristic on driver behavior using smartphone sensor data,’’Accident Anal. Prevention, vol. 144, Sep. 2020, Art.no. 105657.
[27] Z. E. A. Elassad, H. Mousannif, H. Al Moatassime, and A. Karkouch,‘‘The application of machine learning techniques for driving behavioranalysis: A conceptual framework and a systematic literature review,’’ Eng.Appl. Artif. Intell., vol. 87, Jan. 2020, Art. no. 103312.
[28] H. Singh and A. Kathuria, ‘‘Analyzing driver behavior under naturalisticdriving conditions: A review,’’ Accident Anal. Prevention, vol. 150,Feb. 2021, Art.no. 105908.
[29] M. N. Azadani and A. Boukerche, ‘‘Driving behavior analysis guidelinesfor intelligent transportation systems,’’ IEEE Trans. Intell. Transp. Syst.,vol. 23, no. 7, pp. 6027–6045, Jul. 2022.
[30] O. Aboulola, M. Khayyat, B. Al-Harbi, M. S. A. Muthanna, A. Muthanna,H. Fasihuddin, and M. H. Alsulami, ‘‘Multimodal feature-assisted continuousdriver behavior analysis and solving for edge-enabled Internetof Connected Vehicles using deep learning,’’ Appl. Sci., vol. 11, no. 21,p. 10462, Nov. 2021.
[31] A. Uprety, D. B. Rawat, and J. Li, ‘‘Privacy preserving misbehaviordetection in IoV using federated machine learning,’’ in Proc. IEEE 18thAnnu.Consum.Commun.Netw. Conf. (CCNC), Jan. 2021, pp. 1–6.
[32] S. B. Brahim, H. Ghazzai, H. Besbes, and Y. Massoud, ‘‘A machinelearning smartphone-based sensing for driver behavior classification,’’ inProc. IEEE Int. Symp. Circuits Syst. (ISCAS), May 2022, pp. 610–614.
[33] A. A. Hyder, ‘‘Measurement is not enough for global road safety: Implementationis key,’’ Lancet Public Health, vol. 4, no. 1, pp. 12–13, Jan. 2019.
[34] F. Zong, X. Chen, J. Tang, P. Yu, and T. Wu, ‘‘Analyzing traffic crashseverity with combination of information entropy and Bayesian network,’’IEEE Access, vol. 7, pp. 63288–63302, 2019.
[35] L. Yin, S. Tang, S. Li, and Y. He, ‘‘Traffic flow prediction based onhybrid model of auto-regressive integrated moving average and geneticparticle swarm optimization wavelet neural network,’’ vol. 41, no. 9, pp. 2273–2279, 2019.
[36] C.-C. Chang,W.-M. Lin, and C.-A. Lai, ‘‘Using real-time dynamic predictionto implement IoV-based collision avoidance,’’ Appl. Sci., vol. 9, no. 24,p. 5370, Dec. 2019.

[37] M. Zheng, T. Li, R. Zhu, J. Chen, Z. Ma, M. Tang, Z. Cui, and Z. Wang,''Traffic accident's severity prediction: A deep-learning approach-basedCNN network,'' IEEE Access, vol. 7, pp. 39897–39910, 2019.
[38] Y. Ma, X. Zhu, S. Zhang, R. Yang,W.Wang, and D. Manocha, ''TrafficPredict:Trajectory prediction for heterogeneous traffic-agents,'' in Proc. AAAIConf. Artif. Intell., Jul. 2019, vol. 33, no. 1, pp. 6120–6127.
[39] Lv, R. Sun, H. Zhang, H. Xu, and R.Yue, ''Automatic vehicle-pedestrianconflict identification with trajectories of road users extracted from roadsideLiDAR sensors using a rule-based method,'' IEEE Access, vol. 7,pp. 161594–161606, 2019.
[40] L. Peng, M. A. Sotelo,Y. He,Y. Ai, and Z. Li, ''Rough set based method forvehicle collision risk assessment through inferring driver's braking actionsin near-crash situations,'' IEEE Intell. Transp. Syst. Mag., vol. 11, no. 2,pp. 54–69, Summer. 2019.
[41] Katrakazas, M. Quddus, and W.-H.Chen, ''A new integrated collisionrisk assessment methodology for autonomous vehicles,'' Accident Anal.Prevention, vol. 127, pp. 61–79, Jun. 2019.
[42] G. Gao, M. V. Wüthrich, and H. Yang, ''Evaluation of driving riskat different speeds,'' Insurance, Math. Econ., vol. 88, pp. 108–119,Sep. 2019.
[43] M. Ayuso, M. Guillen, and J. P. Nielsen, ''Improving automobile insuranceratemaking using telematics: Incorporating mileage and driver behaviourdata,'' Transportation, vol. 46, no. 3, pp. 735–752, Jun. 2019.
[44] H. Zhao, H. Cheng, T. Mao, and C. He, ''Research on traffic accidentprediction model based on convolutional neural networks in VANET,''in Proc. 2nd Int. Conf. Artif. Intell. Big Data (ICAIBD), May 2019,pp. 79–84.
[45] N. Lyu, J. Wen, Z. Duan, and C. Wu, ''Vehicle trajectory prediction andcut-in collision warning model in a connected vehicle environment,'' IEEETrans. Intell. Transp. Syst., vol. 23, no. 2, pp. 966–981, Feb. 2022.
[46] H. Zhao, H. Cheng, Y. Ding, H. Zhang, and H. Zhu, ''Research on trafficaccident risk prediction algorithm of edge Internet of Vehicles based ondeep learning,'' vol. 42, no. 1, pp. 50–57, 2020.
[47] L. Peng, M. Angel Sotelo, Y. He, Y. Ai, and Z. Li, ''A method for vehiclecollision risk assessment through inferring driver's braking actions in nearcrashsituations,'' 2020, arXiv:2004.13761.
[48] K. Bhalla and K. Gleason, ''Effects of vehicle safety design on road trafficdeaths, injuries, and public health burden in the Latin American region:A modelling study,'' Lancet Global Health, vol. 8, no. 6, pp. 819–828,Jun. 2020.
[49] F. Walugembe, F. Levira, B. Ganesh, and D. W. Lwetoijera, ''A retrospectivestudy on the epidemiology and trends of road traffic accidents,fatalities and injuries in three municipalities of dares Salaam Region, Tanzaniabetween 2014–2018,'' Pan Afr. Med. J., vol. 36, pp. 1–10, May 2020.
[50] X.Wang, J. Liu, T. Qiu, C. Mu, C. Chen, and P. Zhou, ''Areal-time collisionprediction mechanism with deep learning for intelligent transportation system,''IEEE Trans. Veh. Technol., vol. 69, no. 9, pp. 9497–9508, Sep. 2020.
[51] Y. Lin and R. Li, ''Real-time traffic accidents post-impact prediction:Based on crowdsourcing data,'' Accident Anal. Prevention, vol. 145,Sep. 2020, Art.no. 105696.
[52] Gutierrez-Osorio and C. Pedraza, ''Modern data sources and techniquesfor analysis and forecast of road accidents: A review,'' J. Traffic Transp.Eng., vol. 7, no. 4, pp. 432–446, Aug. 2020.
[53] L. Yu, B. Du, X. Hu, L. Sun, L. Han, and W. Lv, ''Deep spatio-temporalgraph convolutional network for traffic accident prediction,'' Neurocomputing,vol. 423, pp. 135–147, Jan. 2021.
[54] J. Fang, D. Yan, J. Qiao, J. Xue, and H. Yu, ''DADA: Driver attentionprediction in driving accident scenarios,'' IEEE Trans. Intell. Transp. Syst.,vol. 23, no. 6, pp. 4959–4971, Jun. 2022.
[55] J. Lee, S. Kishino, and K. Suzuki, ''Prediction of collision avoidance abilityof two-wheeled vehicle riders using driving behaviors and emotionalstates,'' Int. J. Automot. Eng., vol. 12, no. 2, pp. 32–40, 2021.
[56] G. Li, Y. Yang, T. Zhang, X. Qu, D. Cao, B. Cheng, and K. Li, ''Riskassessment based collision avoidance decision-making for autonomousvehicles in multi-scenarios,'' Transp. Res. C, Emerg. Technol., vol. 122,Jan. 2021, Art.no. 102820.

[57] Y. Zhao, T. Miyahara, K. Mizuno, D. Ito, and Y. Han, ''Analysis of cardriver responses to avoid car-to-cyclist perpendicular collisions based ondrive recorder data and driving simulator experiments,'' Accident Anal.Prevention, vol. 150, Feb. 2021, Art.no. 105862.

[58] G. Li, Y. Liao, Q. Guo, C. Shen, and W. Lai, ''Traffic crash characteristicsin Shenzhen, China from 2014 to 2016,'' Int. J. Environ. Res. PublicHealth, vol. 18, no. 3, p. 1176, Jan. 2021.

[59] X. Zhao, T. Lu, and Y. Dai, ''Individual driver crash risk classificationbased on IoV data and offline consumer behavior data,'' Mobile Inf. Syst.,vol. 2021, pp. 1–10, Jun. 2021.

[60] D.-J. Lin, M.-Y.Chen, H.-S. Chiang, and P. K. Sharma, ''Intelligent trafficaccident prediction model for Internet of Vehicles with deep learningapproach,'' IEEE Trans. Intell. Transp. Syst., vol. 23, no. 3, pp. 2340–2349,Mar. 2022.

[61] C.-C. Chang, Y.-M.Ooi, and B.-H. Sieh, ''IoV-based collision avoidancearchitecture using machine learning prediction,'' IEEE Access, vol. 9,pp. 115497–115505, 2021.

[62] L. Yang, Z. Wang, L. Ma, and W. Dai, ''Transfer learning-based vehiclecollision prediction,'' Wireless Commun. Mobile Comput., vol. 2022,pp. 1–9, Aug. 2022.

[63] C. Gutierrez-Osorio, F. A. González, and C. A. Pedraza, ''Deep learningensemble model for the prediction of traffic accidents using social mediadata,'' Computers, vol. 11, no. 9, p. 126, Aug. 2022.

[64] A. Morimoto, A.Wang, and N. Kitano, ''A conceptual framework for roadtraffic safety considering differences in traffic culture through internationalcomparison,'' IATSS Res., vol. 46, no. 1, pp. 3–13, Apr. 2022.

[65] V. Malawade, S.-Y. Yu, B. Hsu, D. Muthirayan, P. P. Khargonekar,and M. A. A. Faruque, ''Spatiotemporal scene-graph embedding forautonomous vehicle collision prediction,'' IEEE Internet Things J., vol. 9,no. 12, pp. 9379–9388, Jun. 2022.

[66] N. Suat-Rojas, C. Gutierrez-Osorio, and C. Pedraza, ''Extraction andanalysis of social networks data to detect traffic accidents,'' Information,vol. 13, no. 1, p. 26, Jan. 2022.

[67] R. Pan, L. Jie, X. Zhang, S. Pang, H. Wang, and Z. Wei, ''A V2P collisionrisk warning method based on LSTM in IOV,'' Secur. Commun.Netw.,vol. 2022, Jul. 2022, Art. no. 7507573.

[68] G. Kothai, E. Poovammal, G. Dhiman, K. Ramana, A. Sharma,M. A. AlZain, G. S. Gaba, and M. Masud, ''A new hybrid deep learningalgorithm for prediction of wide traffic congestion in smart cities,'' WirelessCommun. Mobile Comput., vol. 2021, pp. 1–13, May 2021.

[69] Onyeneke, C. Eguzouwa, and C. Mutabazi, ''Modeling the effects oftraffic congestion on economic activities-accidents, fatalities and casualties,''Biomed. Statist.Informat., vol. 3, no. 2, pp. 7–14, 2018.

[70] S. Du, T. Li, X. Gong, and S.-J. Horng, ''A hybrid method for traffic flowforecasting using multimodal deep learning,'' 2018, arXiv:1803.02099.

[71] W. Wei, H. Wu, and H. Ma, ''An AutoEncoder and LSTM-based trafficflow prediction method,'' Sensors, vol. 19, no. 13, p. 2946, Jul. 2019.

[72] A. Hébert, T. Guédon, T. Glatard, and B. Jaumard, ''High-resolution roadvehicle collision prediction for the city of Montreal,'' in Proc. IEEE Int.Conf. Big Data, Dec. 2019, pp. 1804–1813.

[73] Moses and R. Parvathi, ''Vehicular traffic analysis and prediction usingmachine learning algorithms,'' in Proc. Int. Conf. Emerg. Trends Inf.Technol. Eng. (IC-ETITE), Feb. 2020, pp. 1–4.

[74] N. Ranjan, S. Bhandari, H. P. Zhao, H. Kim, and P. Khan, ''City-wide trafficcongestion prediction based on CNN, LSTM and transpose CNN,'' IEEEAccess, vol. 8, pp. 81606–81620, 2020.