Center for Management Science Research
ISSN: 3006-5291 3006-5283 Volume 1, Issue 2, 2023

THE ROLE OF MACHINE LEARNING IN ANALYZING TRAFFIC PATTERNS FOR IMPROVED ROAD DESIGN

Chaudhry Abu Bakar Imran[*1], **Malik Kamran Shakir**[2], **Muhammad Umer**[3], **Zaryab Imran**[4]

[1, 2, 3, 4] *NUST Institute of Civil Engineering, School of Civil and Environmental Engineering, National University of Sciences and Technology*

[*1]chabubakar2001@outlook.com

DOI: https://doi.org/10.5281/zenodo.16136657

Keywords
Intelligent transportation system, Machine learning, Traffic Pattern Analysis, Road Design, Intelligent Transportation Systems, Traffic Flow Prediction

Article History
Received: 23 November, 2023
Accepted: 23 December, 2023
Published: 30 December, 2023

Corresponding Author: *
Chaudhry Abu Bakar Imran

Abstract
Traffic blocking is becoming as a global concern. This study seeks to examine and evaluate the data mining and machine learning technologies utilized in studies and enterprises to address both direct and indirect traffic challenges affecting mankind and society. The data gathered is significant for traffic research groups, traffic software enterprises, and governmental traffic officials. It directly influences the formulation of explicit strategies for new traffic control proposals. The current study is among the greatest in terms of the work of examined literature pertaining to data mining and machine learning. This review will also highlight a novel method to traffic management.

INTRODUCTION

As urbanization advances and automotive usage increases, transportation issues are becoming increasingly problematic traffic blocking is prevalent, accidents occur often, and the traffic environment is degrading. The issue of enhancing the capacity of the road network has garnered interest from a growing number of researchers. In order to solve this issue, the primary option that comes to mind for many is to construct additional roadways, hence increasing the number of lanes available on the road (Boukerche and Wang. 2020). According to the study conducted about increasing road capacity will exacerbate traffic situations. One effective method to enhance the traffic environment is to implement a precise and efficient transportation system (Dechenaux *et al*., 2014). The corresponding author assists in optimizing transportation resources, alleviating traffic congestion prior to overload, and enhancing the availability of on-road entertainment. The Intelligent Transport System (ITS) is among the most renowned of these systems (Moura *et al*., 2019). ITS is a sophisticated system composed of several modern technologies, such as transportation communication systems. In the interim, ITS can enhance traffic efficiency, alleviate traffic congestion, expand road capacity, and lower environmental pollution and traffic accidents by using the advancements in 5G communication technology, a multitude of on-road sensors, etc. (Aljeri, and Boukerche, 2019).
Recently, the swift population growth and surge in vehicle numbers, together with traffic congestion and overcrowding, have emerged as substantial issues in

metropolitan areas. Traffic congestion adversely impacts daily living and directly increases transportation expenses. A research published on April 18, 2018, by The Boston Consulting Group (BCG) and commissioned by Uber indicates that India's major cities are incurring yearly losses of up to $22 billion owing to traffic congestion. Furthermore, the incidence of traffic deaths is escalating to a concerning degree (Jindal *et al.*, 2018). According to the WHO, it has reached up to 1.35 million individuals annually. Therefore, with completely automated equipment and traffic control personnel, effective, automated, and real-time traffic management is needed (Bansal, 2018).

The World Bank Group Report on The High Toll of Traffic Injuries indicates that a substantial reduction in road traffic injuries positively influences the nation's long-term economic growth. A systematic and clearly defined deterrence-based traffic police enforcement with fines will significantly reduce road traffic accidents. Although traffic regulations are relatively harsh, their enforcement is inadequate due to a shortage of traffic police on the roads (Gao, 2019). In 2018, 30% of 85,144 traffic policeman posts and 39% of 58,509 sanctioned traffic constable positions remained unfilled, according to data from the Bureau of Police Research and Development (Tanwar *et al.*, 2018). According to 2017 report, the overall number of traffic officers was just over 72,000, while the registered cars approached 200 million. Consequently, due to the limited presence of traffic police relative to the volume of cars, a methodical distribution and scheduling of traffic officers is essential to manage traffic chaos and ultimately diminish road accidents through rigorous enforcement of traffic citations (Nama *et al.*, 2021). Figure 1 showed a clear depiction of ITS and its applications associated to road safety and security.

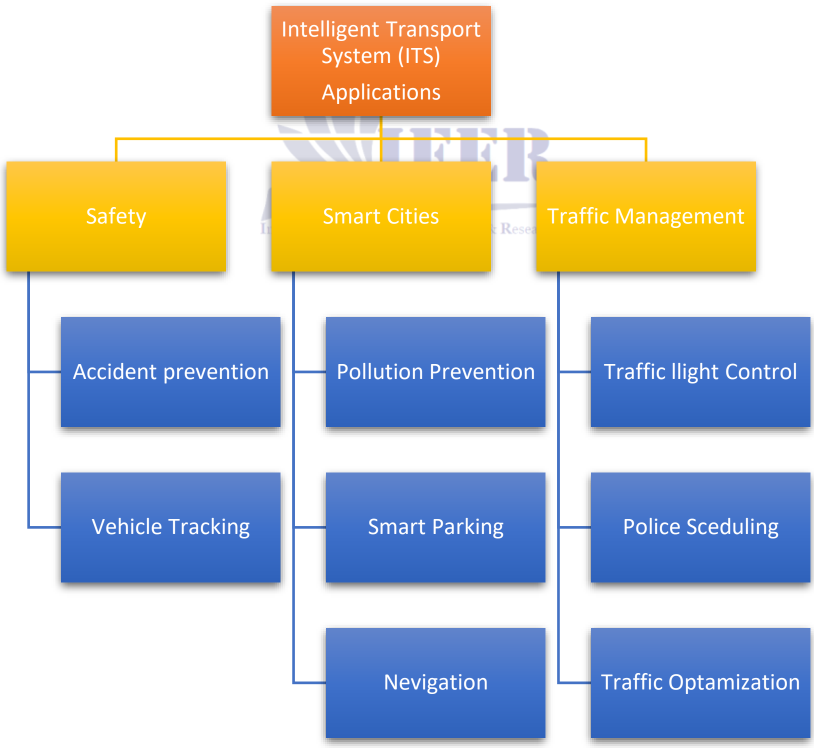

Figure 1: *Overview of ITS and Applications.*

These services include emergency vehicle alert systems, automatic road enforcement, adjustable speed limits, flexible traffic light sequence conflict resolution systems, pollution mitigation, parking advisories, notification systems, weather information, and bridge removal and relocation systems, all contributing to seamless vehicular flow. However, ITS is deficient in resource management and human resources management. The traffic police constitute the essential framework of the traffic management system in each city. Their scheduling and reallocation in response to fluctuating traffic density are imperative to address the exponentially escalating traffic congestion situation. In order to create a sustainable smart city, a review explained machine learning-based traffic control force scheduling, which guarantees the best possible use of non-renewable fuel, time, money, and human resources (Liu, 2019).

Cities, municipalities, governments, corporations, and scholars have offered several methods to address the issue of traffic congestion. Several of these ideas incorporate adaptive traffic signals, vehicle-to-infrastructure smart corridors, autonomous vehicle technology, real-time traffic feedback, pedestrian traffic monitoring, automobile sharing, and multi-modal approaches. There are many solutions focused on the principles of the Internet of Things (IoT), Wireless Sensor Networks (WSN), and Data Analytics (DA) methodologies. The following additional partial solutions were proposed:

1) Building new highways, bridges, tunnels, flyovers, and bypass routes.

2) Constructing rings and repairing existing roads.

Traffic congestion denotes an overabundance of cars on a segment of route at particular periods, leading to reduced speeds and extended travel durations, and therefore constitutes a significant difficulty in traffic management and transportation planning (Raj *et al.*, 2016). It cannot be entirely resolved, but it may be addressed to a certain degree. Road users can make better judgements when travelling and reduce the likelihood of traffic congestion by being informed in advance on the state of the roads. This information encompasses quantitative metrics of traffic congestion, expressed by estimated characteristics such as journey time and traffic density (Alsrehin *et al.*, 2019). Traffic produces substantial volumes of data collected from many devices, including intelligent cameras and sensors. The collection of data poses no difficulties; the challenge is in the storage, handling, processing, analysis, and management of the growing volumes of traffic data to render it valuable. The aforementioned methodologies primarily concentrate on the analysis of substantial traffic data to derive certain characteristics, but not limited to traffic speed, traffic volume, vehicle arrival rate, and average waiting time (Sundaram *et al.*, 2015). Traffic congestion for the traveler results in time loss, missed opportunities, and irritation. For the employer, it signifies diminished worker productivity, lost trade opportunities, delayed deliveries, and heightened costs. Mitigating traffic congestion will ensure safe transit for individuals, decrease accident rates, lower fuel consumption, assist in controlling air pollution, minimize waiting times, facilitate the smooth flow of vehicles on transportation routes, and supply essential data for future road planning and analysis. Traffic congestion arises from several factors, including inadequate capacity, excessive demand, prolonged red signal durations, and impediments such as accidents, unplanned vehicle halts, double parking, construction activities, and road narrowing (Alsrehin *et al.*, 2019).

General steps for the development of intelligent transportation and control systems:

The primary steps for designing intelligent transportation and control systems includes Collection, Pre-processing, Analysis. Storage, Communication, Maintenance and Archiving (Miller *et al.*, 2018).A brief information about these steps is mentioned here.

Collection: The Collection of data can be done by using various methods as mentioned in table 1.

Table 1: *Methods for the development of intelligent transportation and control systems.*

Methods	Uses in Controlling System
Image- or video-based methods.	Surveillance cameras are employed to visually monitor road traffic in a designated region and transmit or record the recorded images/videos to control rooms. It is extensively utilized in road traffic management because of its effectiveness and ease of maintenance. Nonetheless, video and picture material need

	substantial storage, network bandwidth, and computational complexity.
Hybrid-based methods	It amalgamates two or more of the aforementioned strategies.
Sensor-based methods	It include ultrasonic sensors, lasers, RFIDs, photoelectric sensors, radar, and vehicle probe data.
Vehicle to Vehicle (V2V) and Vehicle to infrastructure (V2I)	This method involve communications using Wi-Fi, GPRS, WiMAX and Bluetooth.

Pre-processing: The raw data that is collected using any of the aforementioned methods is susceptible to noise, absent values, and inconsistent data as a result of sensor failures, measurement errors, data link errors, or its vast size (Lopes *et al.*, 2010). Consequently, data manipulation is necessary, and some of the following methods are employed:

1. Data cleansing help in encompassing noise elimination, anomaly identification, and recovery of absent data.
2. Dimensionality reduction employing manifold learning, non-negative matrix factorization, or kernel dimension reduction techniques to decrease data dimensionality. This enhances the efficacy of learning-driven tasks inside the reduced dimensional space.
3. Sparsity Analysis involves eliminating redundant features from the original feature space by compressive sensing or heterogeneous learning.
4. Data fusion necessitates the integration of many data sources.

Analysis: Data analysis involves employing various analytical methods to yield valuable insights, such as estimating the total number of cars traversing a certain piece of roadway on any given day of the year. The identification of erroneous data pieces and assessing the effect of diverse data-driven processes may also be conducted to assure the quality of the analyzed data. Big traffic data might be analyzed using cloud computing and sophisticated data processing tools and methodologies to provide better real-time traffic choices. Furthermore, it employs various learning methods to instruct systems on managing traffic lights, lane signals, visual message systems (VMS), and traffic information. These methodologies are predominantly founded on machine learning, data mining, and artificial intelligence algorithms (Alsrehin *et al.*, 2019).

Storage: The fast increase in traffic data volume necessitates highly efficient storage systems. Cloud storage may be utilized to store and safeguard extensive traffic data, facilitating more efficient real-time traffic decision-making. The usage of data is more trusted and confident when it is safe and properly formatted (Miller *et al.*, 2018).

Communication: Data communication encompasses the use and dissemination of traffic data. Traffic data is utilized to analyze, strategize, build, construct, run, and oversee traffic systems. Traffic data communication assists researchers, policymakers, government officials, planners, and transportation agencies in comprehending traveler behavior and patterns, as well as in identifying methods to enhance system efficiency and cost-effectiveness. The use of this data is contingent upon the intended objective and the methods of its initial collection, processing, analysis, and storage. Disseminating traffic data sourced from diverse internal and external resources can assist agencies and researchers in acquiring a more thorough understanding, hence enhancing the clarity and quality of their decision-making. Nonetheless, disseminating and conveying public traffic data has other challenges, including openness, privacy, security, liability, and cooperation with various authorities and partners (Alsrehin *et al.*, 2019).

Maintenance and Archiving: Data maintenance is the process of continuous enhancement and systematic verification, encompassing ongoing correction and validation. Enhanced maintenance levels ensure the optimal operation of all required systems. Data archiving involves transferring and storing seldom accessed data from live systems and databases into specialized archive systems to enhance performance, implement a cost-effective strategy, and facilitate future retrieval (Alsrehin *et al.*, 2019).

Machine Learning and various Algorithm Theories:

Machine Learning (ML) is increasingly used for traffic flow prediction tasks. The rationale is attributed to limited previous knowledge regarding the interrelations among various traffic patterns for model construction, reduced constraints on predictive tasks, and enhanced non-linear characteristics (Arrieta *et al.*, 2020). This study will concentrate on the machine learning models employed in traffic prediction problems. It is essential to evaluate various model types based on their correctness, as well as their capacity to address specific issues, efficiency, and dependencies on technology and data. Machine learning is an extensive subject with several categorization techniques for models, categorized from various angles. A few machine learning algorithms are elucidated in this review.

NN Model:

NN model is a model constructed by emulating the transmission of information through neurons in the brain. The input data traverses several network architectures across distinct neural network models. The input data will be converted into an activation signal using the activation function inside the network. The ultimate forecast is determined by the activation signal. This study reviews fundamental neural network models, including the Feed Forward Neural Network (FFNN), Recurrent Neural Network (RNN), and Convolutional Neural Network (CNN). Simultaneously, fundamental neural network models may be consolidated into deep learning models with varying architectures, which will also be examined in this research (Boukerche and Wang, 2020).

Regression Model:

The regression model analyses the connection between the dependent and independent variables, employing a curve or line to match the dataset. Regression models have been employed for traffic prediction for many years. Recent years have seen a scarcity of traffic forecast models utilizing regression techniques. The regression model, while straightforward to apply, possesses a simplistic structure that inadequately captures the temporal characteristics of continuous traffic records. Nonetheless, the regression model remains a liable option for traffic prediction jobs in small and straightforward traffic networks due to its lower computational requirements (Boukerche and Wang, 2020). Simultaneously, regression functions like LASSO can ensure the model's capacity to identify non-linear characteristics within the traffic data.

Example-Based Model:

The example-based model addressed the prediction challenge by assessing the similarity between the input sequence and previous data samples, utilizing the identified samples to provide the final forecast. This research primarily examines the k-Nearest Neighbors (KNN) model whereas other models are illustrated in table 2.

Table 2: *Popular Example-Based Models with uses in Analyzing Traffic Patterns.*

MODEL	DESCRIPTION	TRAFFIC USE
K-NEAREST NEIGHBORS (K-NN)	Predicts by averaging outcomes of k most similar past examples	Real-time traffic speed prediction
LAZY LEARNING	Learns only when needed and avoids upfront training	Travel time estimation
MEMORY-BASED LEARNING	Stores all past data and predicts by similarity without abstraction	Congestion detection
CASE-BASED REASONING (CBR)	Solves new problems by adapting solutions of similar past cases	Incident response modeling

Kernel-Based Model:

In the kernel-based model, a kernel function is employed to transform the input data into a high-dimensional vector space, facilitating the resolution of prediction problems. This research primarily examines the Support Vector Machine (SVM) and Radial Basis Function (RBF) models.

Hybrid Model:

In the hybrid model, the ultimate forecast is derived by amalgamating the outcomes from two or more distinct predictive models.

Table 3: *Common Hybrid Approaches in Road Safety*

Hybrid Approaches	Components	Applications
Machine Learning + Statistical Methods	Random Forest + Poisson regression	Predict accident frequency based on traffic flow and road conditions
Supervised + Unsupervised Machine Learning	SVM + K-Means	Classify accident severity and cluster accident-prone locations
Machine Learning + Simulation	Machine Learning + Traffic simulation models (e.g., SUMO, VISSIM)	Test safety policies or road designs in simulated environments
Machine Learning + Fuzzy Logic	Neural Networks + Fuzzy Inference Systems	Model uncertainty in human driving behavior for crash risk prediction
Deep Learning + Feature Engineering	CNN + domain-specific features	Analyze video footage for near-miss detection or risky driving behavior

Weak Lane Discipline Modeling:

Driving behavior modelling in mixed traffic streams remains a difficulty. In developing-world cities, you may often see a wide variety of vehicles on the road, and drivers often ignore lane laws. Traditional microscopic models have a hard time recreating these features. Due to multiple-leader following, it is difficult to determine leader-follower couples in car following circumstances. Also, vehicles don't follow the actual lane markings, so it's hard to tell which way to go while changing lanes. One of the problems with existing models of driver behavior in mixed-traffic situations is that they don't take into account the whole spectrum of 12–10 Gipps, according to a review by Asaithambi *et al.* (2016). Munigety and Mathew (2016) showed that insufficient lane discipline leads drivers in mixed traffic streams to exhibit unusual behaviors, including shorter headways, swerving, and filtering. They have further suggested that the lane be split into narrow sections to accommodate virtual lane transitions. Li *et al.* (2015) suggested a car-following model that accounts for the influence of two-sided lateral gaps, demonstrating that their model exhibits a wider stable zone compared to a model that includes the lateral gap effects on only one side. Additionally, Parsuvanathan (2015) employed proxy lanes alongside the primary lanes. It is considered that tiny cars view free space as lanes.

The distribution and kinds of vehicles, however, may have an impact on the lanes' width. Mathew *et al.* (2015) focused their research on segments of traffic queues rather than conventional main lane queues. Kanagaraj *et al.* (2013) evaluated the performance of various car-following models in mixed traffic conditions. However, they have not considered that a vehicle may not align precisely with its leading vehicle because of inadequate lane discipline in mixed traffic conditions. Metkari *et al.* (2013) modified an existing car-following model to incorporate lateral movements and account for mixed traffic conditions. Choudhury and Islam (2016) developed a latent leader acceleration model. Papathanasopoulou and Antoniou (2017). Employed data-driven methodologies to model mixed traffic and introduced virtual lanes to address weak lane discipline conditions.

Determination of Virtual Lanes:

Temporary virtual lanes are suggested for simulating mixed traffic circumstances. The diversity of vehicle types results in differing vehicle sizes and hence varying widths of virtual lanes. Figure 2 exemplifies a standard variation of virtual lane change. This figure depicts two cars. The initial car adheres to the virtual lane i. Although there are little lateral movements, it is deemed that there is no lane change. When the travel is restricted by the hatching vehicle at the breakpoint, it deemed to change lanes and subsequently adhere to virtual lane i + 1. The issue lies in the continuous lateral movement of vehicles. This may have approached in two separate manners. The initial task is to ascertain the threshold that signifies a lane change. The second method involves the utilization of change detecting methods. It is possible to utilize algorithms that can identify significant changes in the data sequence, such as the "structure change" paths (Papathanasopoulou *et al.*, 2019).

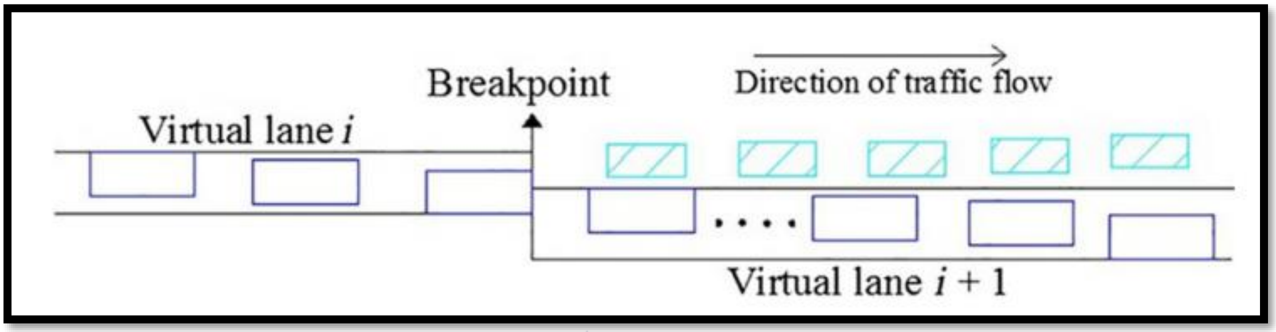

Figure 2: *Virtual lane change according to the Traffic Flow. (Papathanasopoulou et al., 2019)*

Simulators for Traffic network:

A real-world dataset is crucial for model training. Nonetheless, the available real-world data sets may not align with the requirements of the experiment. The scale of the traffic network in the specified real-world dataset is insufficient, or we may require additional driving behaviors that are absent from the dataset. A traffic simulation system is necessary to enhance the traffic dataset for issue resolution. The traffic simulation system replicates a realistic traffic environment by creating a computational model. Simulation models can be categorized as macroscopic and microscopic models. In macroscopic traffic simulation, traffic flow is seen as a continuous compressible fluid, consisting of several vehicles (Abdelgawad *et al.*, 2016). It emphasizes the volume, density, and velocity of traffic flow while disregarding the interactions among cars inside the flow. Macroscopic traffic simulators need lower computational resources and exhibit enhanced simulation speed.

Several well-known macroscopic traffic simulations include Trans CAD and EMME. The interactions within traffic flow are not represented in macroscopic traffic simulators. Conversely, tiny traffic simulators concentrate on the position and velocity of each vehicle inside the traffic flow. The technology emulates traffic flow by modelling the behavior of individual vehicles (Maciejewski *et al.*, 2010). In a microscopic traffic simulation system, the user may specify car-following behavior, lane-changing behavior, destination information, etc. These capabilities provide the user with the possibility to build a more intricate and realistic traffic scene. Below are notable tiny traffic simulators: SUMO, VISSIM, SimTraffic, CORSIM, Paramics, MITSimLab, TransModeler (Boukerche and Wang, 2020).

Identifying Traffic Patterns and Behaviors:

In addition to helping forecast short- and long-term traffic problems, recognizing vehicle movements and comprehending traffic patterns, behavior, and how traffic congestions arise and worsen across time and space can also help to lessen congestion. There have been numerous endeavors to analyses and identify traffic patterns and behaviors (Alsrehin *et al.*, 2019). Sekar and Shondelmyer concentrated on the detection and analysis of traffic violations based on driving behavior. The authors presented a method wherein an information processing system identifies a traffic violation committed by a driver operating a vehicle. The information handling system establishes an infraction detection zone including a collection of traffic control devices and transmits a series of setup settings to these devices.

This utilize vehicle identifying data within the configuration settings to ascertain the driver's behaviors in the infraction detection zone and then issues a ticket based on the detected behaviors (Sekar and Shondelmyer, 2018). Wang with fellows presented a unique concept termed "shadow traffic" for the unified modelling of traffic anomalies in simulations. They converted the characteristics of anomalies to those of shadow vehicles and subsequently elucidated the role of these shadow vehicles in traffic simulations. Furthermore, allowing for a cohesive representation of various traffic anomalies and a comprehensive description of the anomaly's evolution. They asserted that their model could be integrated into the majority of current traffic simulators with minimal computing burden. Furthermore, experimental findings indicate that the model may realistically and efficiently simulate a range of anomalous traffic behaviors. They proposed, as future work, the development of a real-time editing traffic anomaly simulation system that allows users to insert and modify anomalies at any time and location inside traffic simulations, as well as to investigate the modelling of intricate vehicle-crowd interactions (Wang *et al*., 2018).

Another research presented a system and methodology for traffic engineering in networks, specifically focusing on distributed traffic engineering in Software Defined Networks. According to one embodiment, a network component for dynamic zoning for traffic engineering (TE) in software-defined networking (SDN) comprises a processor and a computer-readable storage medium containing code for execution by the processor. The programming encompasses directives to acquire network information from at least one SDN controller among multiple SDN controllers within a network. This highlight the plurality of TE Zones for the network, designate a local Zone TE controller for each TE Zone, and select a master TE controller based on the network information and a Zoning scheme. The local Zone TE controller is chosen from one of the SDN controllers and the master TE controller is also selected from one of the SDN controllers. It helps to convey an indication of the Zone composition, the local Zone TE controllers, and the master controllers to a subset of the SDN controllers (Li *et al*., 2013).

Signal Control and Traffic Light:

The traffic light serves as the primary mechanism for regulating vehicular movement by designating intervals for stopping and proceeding; establishing fixed timings for traffic lights is an ineffective method of managing vehicle flow, resulting in system imbalance owing to the variable volume of cars on either side. A research presented an algorithm to regulate traffic lights depending on the amount of cars at each signal. The method utilizes picture data obtained from video shot by a field-installed camera and employs an artificial neural network alongside fuzzy logic to adjust the duration of each light. The algorithm is validated by a comparison of its outcomes with manual results. The produced outcomes will optimize traffic flow and decrease the time wasted on the roadways (Zaid *et al*., 2017). The video monitoring and surveillance systems are extensively utilized in traffic management and traffic signal control systems. Numerous efforts have been undertaken to create intelligent traffic signals. A separate research described a technique that utilizes pictures obtained from live video feeds from cameras positioned at traffic intersections to assess real-time traffic density. This approach adjust the lighting based on traffic density to mitigate congestion (Kanungo et. al., 2014). The probability of incidents increases when drivers linger in the traffic light queue for an extended period and the traffic light transitions from green to yellow. Consequently, the majority of drivers cross the road during the transition from yellow to red. This is known as the Red Signal Running (RLR) phenomena, which frequently arises due to the improper calibration of the traffic signal. The authors presented a method that utilizes data collected from a wireless sensor network to dynamically optimize waiting times in vehicle queues and mitigate the occurrence of the RLR phenomena at an isolated crossroads. This is accomplished by allocating an extended duration of green light to the route with the largest line (Collotta *et al*., 2014). Advanced Driver Assistance Systems (ADAS) include functionalities such as automated driving and parking, traffic sign recognition, and collision detection. Recent advancements seek to automate certain activities of drivers through the use of computer vision and technologies like machine learning and robotic navigation. ADAS significantly influences the industry and society, enhancing driving safety and assisting in vehicle maintenance (Velez and Otaegui, 2017). A review illustrates the present advancements and prospective trajectories in vision-based embedded Advanced Driver Assistance Systems (ADAS), effectively connecting theoretical concepts with practical applications. Furthermore, the authors examined several hardware and software alternatives employed in ADAS, along with issues about design, development, and testing. Additionally, some significant problems have been found. The authors provided a comprehensive classification and vocabulary for vision-based Advanced Driver Assistance Systems (ADAS). They suggested an abstract approach to formalize a top-down perspective on application development for scaling autonomous driving systems (Horgan *et al.*, 2015).

Another research investigation examines the potential for enhancing commuter journey times through Vehicle-to-Infrastructure (V2I) technology and proposes a Belief-Desire-Intention architecture that utilizes local knowledge and data gathered from surrounding infrastructures, including vehicles and traffic light controllers (TLCs), to model vehicular behavior on the road. This design facilitates the exchange of beacons among cars to ascertain their ideal speed and location inside a road segment, aiming to traverse intersections with minimal delays and to avoid stoppages wherever feasible. The preliminary modelling outcomes indicate a substantial decrease in average trip duration. Nevertheless, the suggested architecture fails to effectively address cars exhibiting selfish behavior or the implementation of reactive and proactive remedies (Soufiene *et al.*, 2015). An additional effort was made to facilitate the transition of the traffic signal from red to green contingent upon traffic congestion. The authors integrated current technology with artificial intelligence to create and deploy a cost-effective, real-time, sensor-based dynamic traffic signal management system aimed at minimizing the Average Trip Waiting Time (ATWT) (Jagadeesh *et al.*, 2015). Another suggested system employed dynamic control, infrared sensors, low-power embedded controllers, comparators, and storage devices. Vehicle-to-vehicle and vehicle-to-infrastructure communications establish a framework known as VENET, which is crucial for smart cities by relaying information on traffic conditions and assisting drivers in making informed decisions to avoid congestion (Khekare and Sakahra, 2013).

Conclusion:

This review sought to enhance the comprehension of the current advancements in traffic management technology, particularly through the application of data mining and machine learning. It categorized the existing studies into approaches based on real-time traffic parameter measurement, approaches for detecting moving objects, approaches focused on routing identification, approaches aimed at analyzing driver and pedestrian patterns and behaviors, and finally, approaches centered on traffic light signals.

References:

Asaithambi, G., Kanagaraj, V., & Toledo, T. (2016). Driving behaviors: Models and challenges for non-lane based mixed traffic. Transportation in Developing Economies, 2, 1-16.

Munigety, C. R., & Mathew, T. V. (2016). Towards behavioral modeling of drivers in mixed traffic conditions. Transportation in Developing Economies, 2, 1-20.

S. Djahel, N. Jabeur, R. Barrett, and J. Murphy, "Toward V2I communication technology-based solution for reducing road traffic congestion in smart cities," in Proc. Int. Symp. Netw. Comput. Commun. (ISNCC), Hammamet, Tunisia, May 2015, pp. 1–6.

Y. M. Jagadeesh, G. M. Suba, S. Karthik, and K. Yokesh, "Smart autonomous traffic light switching by traffic density measurement through sensors," in Proc. Int. Conf. Comput., Commun., Syst. (ICCCS), Kanyakumari, India, Nov. 2015, pp. 123–126.

G. S. Khekare and A. V. Sakhare, "A smart city framework for intelligent traffic system using VANET," in Proc. Int. Mutli-Conf. Automat., Comput., Commun., Control

Compressed Sens. (iMac4s), Kottayam, India, Mar. 2013, pp. 302–305

M. Collotta, G. Pau, G. Scatà, and T. Campisi, "A dynamic traffic light management system based on wireless sensor networks for the reduction of the red-light running phenomenon," Transp. Telecommun., vol. 15, no. 1, pp. 1–11, 2014

Li, Y., Zhang, L., Peeta, S. *et al.* Non-lane-discipline-based car-following model considering the effects of two-sided lateral gaps. Nonlinear Dyn 80, 227–238 (2015)

Parsuvanathan, C. (2015). Proxy-lane algorithm for lane-based models to simulate mixed traffic flow conditions. International Journal of Traffic and Transportation Engineering, 4(5), 131-136.

A. Kanungo, A. Sharma, and C. Singla, "Smart traffic lights switching and traffic density calculation using video processing," in Proc. Recent Adv. Eng. Comput. Sci. (RAECS), Chandigarh, India, Mar. 2014, pp. 1–6

Maciejewski, M. (2010). A comparison of microscopic traffic flow simulation systems for an urban area. *Transport Problems*, *5*, 27-38.

Mathew, T. V., Munigety, C. R., & Bajpai, A. (2015). Strip-based approach for the simulation of mixed traffic conditions. *Journal of Computing in Civil Engineering*, *29*(5), 04014069.

Kanagaraj, V., Asaithambi, G., Kumar, C. N., Srinivasan, K. K., & Sivanandan, R. (2013). Evaluation of different vehicle following models under mixed traffic conditions. Procedia-Social and Behavioral Sciences, 104, 390-401.

Metkari, M., Budhkar, A., Maurya, A.K., 2013. Development of simulation model for heterogeneous traffic with no lane discipline. Procedia Soc. Behav. Sci. 104, 360–369.

Papathanasopoulou, V., & Antoniou, C. (2017). Flexible car-following models on mixed traffic trajectory data (No. 17-06671).

Choudhury, C. F., & Islam, M. M. (2016). Modelling acceleration decisions in traffic streams with weak lane discipline: A latent leader approach. Transportation research part C: emerging technologies, 67, 214-226.

Papathanasopoulou, V., Antoniou, C., & Koutsopoulos, H. N. (2019). Data-driven traffic simulation models: Mobility patterns using machine learning techniques. In Mobility Patterns, Big Data and Transport Analytics (pp. 263-295). Elsevier.

E. Dechenaux, S.D. Mago, L. Razzolini, Traffic congestion: an experimental study of the Downs-Thomson paradox, Exp. Econ. 17 (3) (2014) 461–487.

Boukerche, A., & Wang, J. (2020). Machine learning-based traffic prediction models for intelligent transportation systems. Computer Networks, 181, 107530.

D.L.L. Moura, A.L.L. Aquino, A.A.F. Loureiro, Towards data VSN offloading in VANETs integrated into the cellular network, in: ACM MSWIM, in: MSWIM '19, 2019, pp. 235–239.

Aljeri, N., & Boukerche, A. (2019). Movement prediction models for vehicular networks: an empirical analysis. *Wireless Networks*, *25*, 1505-1518.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion, 58, 82-115.

K. Abdelgawad, S. Henning, P. Biemelt, S. Gausemeier, A. Trächtler, Advanced traffic simulation framework for networked driving simulators, IFAC PapersOnLine 49 (2016) 101–108.

Nama M, Nath A, Bechra N, *et al.* Machine learning-based traffic scheduling techniques for intelligent transportation system: Opportunities and challenges. Int J Commun Syst. 2021; 34:e4814.

Jindal A, Aujla GS, Kumar N, Chaudhary R, Obaidat MS, You I. Sedative: SDN-enabled deep learning architecture for network trafficcontrol in vehicular cyber-physical systems. IEEE Netw. 2018;32(6):66-73.

Bansal K. Peak traffic hours in 4 Indian cities costing 22%bn a year: Uber. https://bit.ly/2Imtcy0. Accessed: 10-12-2020; 2018.

Honghao Gao XY. Applying probabilistic model checking to path planning in an intelligent transportation system using mobility trajectories and their statistical data. Intell Autom Soft Comput. 2019;25(3):547-559.

Tanwar S, Tyagi S, Kumar S. The role of internet of things and smart grid for the development of a smart city. In: Intelligent Communica-tion and Computational Technologies. Singapore: Springer; 2018:23-33.

K. Miller, M. Miller, M. Moran, and B. Dai, "Data management life cycle," Texas A&M Transp. Inst., College Station, TX, USA, Tech. Rep. 1, Mar. 2018.

A. A. Zaid, Y. Suhweil, and M. A. Yaman, "Smart controlling for traffic light time," in Proc. IEEE Jordan Conf. Appl. Elect. Eng. Comput. Technol. (AEECT), Aqaba, Jordan, Oct. 2017, pp. 1–5.

Jiping Liu CDFZ. Simulation of real-time path planning for large-scale transportation network using parallel computation. Intell AutomSoft Comput. 2019;25(1):65-7.

Alsrehin, N. O., Klaib, A. F., & Magableh, A. (2019). Intelligent transportation and control systems using data mining and machine learning techniques: A comprehensive study. IEEe Access, 7, 49830-49857.

Sundaram, S., Kumar, S. S., & Shree, M. D. (2015). Hierarchical clustering technique for traffic signal decision support. International Journal of Innovative Science, 2(6), 72-82.

Lopes, J., Bento, J., Huang, E., Antoniou, C., & Ben-Akiva, M. (2010, September). Traffic and mobility data collection for real-time applications. In 13th international IEEE conference on intelligent transportation systems (pp. 216-223). IEEE.

Sekar, D., & Shondelmyer, W. J. (2018). U.S. Patent No. 10,037,691. Washington, DC: U.S. Patent and Trademark Office.

H. Wang, M. Xu, F. Zhu, Z. Deng, Y. Li, and B. Zhou, "Shadow traffic: A unified model for abnormal traffic behavior simulation," Comput. Graph., vol. 70, pp. 235–241, Feb. 2018.

X. Li, P. Djukic, and H. Zhang, "Traffic behavior driven dynamic zoning for distributed traffic engineering in SDN," U.S. Patent 9 432 257 B2, Aug. 30, 2013.

G. Velez and O. Otaegui, "Embedding vision-based advanced driver assistance systems: A survey," IET Intell. Transp. Syst., vol. 11, no. 3, pp. 103–112, Apr. 2017.

J. Horgan, C. Hughes, J. McDonald, and S. Yogamani, "Vision-based driver assistance systems: Survey, taxonomy and advances," in Proc. 18th IEEE Int. Conf. Intell. Transp. Syst. (ITSC), Las Palmas, Spain, Sep. 2015, pp. 2032–2039.