NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF CIVIL ENGINEERING

MODELLING AND PREDICTING CAR FOLLOWING BEHAVIOR IN CONNECTED VEHICLES: A MACHINE LEARNING APPROACH

M.Sc.THESIS

Abdinasir Mohamed YUSUF

Nicosia

January, 2024

Abdinasir Mohamed YUSUF

MODELLING AND PREDICTING CAR FOLLOWING BEHAVIOR IN CONNECTED VEHICLES: A MACHINE LEARNING APPROACH

NEU
2024

NEAR EAST UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF CIVIL ENGINEERING

MODELLING AND PREDICTING CAR FOLLOWING BEHAVIOR IN CONNECTED VEHICLES: A MACHINE LEARNING APPROACH

M.Sc. THESIS

Abdinasir Mohamed YUSUF

Supervisor

Asst.Prof.Dr: Ikenna UWANUAKWA

Nicosia
January, 2024

Approval

We certify that we have read the thesis submitted by **Abdinasir Mohamed YUSUF** titled "**MODELLING AND PREDICTING CAR FOLLOWING BEHAVIOR IN CONNECTED VEHICLES: A MACHINE LEARNING APPROACH**" and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Educational Sciences.

Examining Committee	Name-Surname	Signature
Head of the Committee:	Assoc.Prof.Dr. Shaban Ismael Albrka	....
Committee Member:	Assoc.Prof.Dr. Mustafa Alas	....
Supervisor :	Assoc.Prof.Dr. Ikenna Uwanuakwa	....

Approved by the Head of the Department

.03/.03./2024

Prof. Dr. Kabir Sadeghi

Head of Civil Engineering Department.

Approved by the Institute of Graduate Studies

...../...../2024

Prof. Dr. Kemal Hüsnü Can Başer

Head of the Institute of Graduate Studies.

Declaration

I hereby declare that all information, documents, analysis and results in this thesis have been collected and presented according to the academic rules and ethical guidelines of the institute of graduate studies, Near East University. I also declare that as required by these rules and conduct, I have fully cited and referenced information and data that are not original to this study.

Abdinasir Mohamed YUSUF

11/03/2024

Acknowledgments

I would like to extend my sincere gratitude to my advisor ASST. PROF.DR: IKENNA UWANUAKWA for his kindness, motivation, and knowledgeable counseling throughout this thesis. It has been a privilege for me to work and learn under his helpful advice and without his support and advice, this research could not have been done.

I want to express my appreciation to all of the professors and instructors at Near East University for spreading knowledge and offering sincere and valuable support during the course, your expertise, understanding, and patience added considerably to my graduate experience. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

My sincere gratitude and appreciation to my parents for their encouragement and support in helping me finish my master degree both directly and indirectly.

Finally, I want to thank my brothers, sisters, and friends for helping me develop emotionally and physically throughout my life.

Abdinasir Mohamed YUSUF

Abstract

MODELING AND PREDICTING CAR FOLLOWING BEHAIOR IN CONNECTED VEHICLES: A MACHINE LEARNING APPROACH.
Abdinasir Mohamed YUSUF
MA, Department of Civil Engineering, Faculty of Civil and Environmental Engineering, Near East University, Nicosia.
January, 2024, 95 Pages

This research presents the results of a study on Modelling and Predicting Car Following Behavior in Connected Vehicles: A Machine Learning Approach with a particular focus on Gradient Boosting and Random Forest algorithms for modelling and predicting car-following behavior in the connected vehicles (CVs). It utilizes data-rich environment enriched by vehicle to vehicle.

The methodology encompasses a comprehensive data collection from car following experiment involving five platoon seniors, followed by the application of machine learning algorithm, the performance of the models was thoroughly evaluated using metrics such as R-Squared, RMSE, MAE, and NSE.

the Results shows that the models are performing well during the training phase with reasonably good accuracy and realistic R^2 values throughout the different vehicle platoon scenarios. Precisely, for Gradient Boosting, R^2 varies between 0.92 and 0.98, where it crosses 1.0 (near-perfect scores) for Random Forest with training. However, a noticeable decrease in the performance is observed in test phase, especially with less R^2 values in some scenarios like 0.87 for 2nd platoon using Gradient Boosting pointing out of potential overfitting issues. The below study assesses performance of machine learning modeling car following behavior but it also demonstrates a need of careful validation and adjustment of such models in practical applications. This entails that incorporating better machine learning in the CVs would go a long way in enhancing traffic safety and management to help advance intelligent transportation systems.

Keywords: Connected vehicle, Machine learning, Traffic simulation;

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CVs:	Connected Vehicles
V2V:	Vehicle-to-Vehicle
V2I:	Vehicle-to-Infrastructure
V2X:	Vehicle-to-Everything
RTI:	Real-Time Information
LWR:	Lighthill-Whitham-Richards
ACC:	Adaptive Cruise Control
CACC:	Cooperative Adaptive Cruise Control
IDM:	Intelligent Driver Model
RL:	Reinforcement Learning
ADRT:	Acceleration or Deceleration Reaction Time
CAVs:	Connected and Automated Vehicles
GB:	Gradient Boosting
GBRT:	Gradient Boosted Regression Trees
RF:	Random Forest
CART:	Classification and Regression Tree
RMSE:	Root Mean Squared Error
MAE:	Mean Absolute Error
NSE:	Nash-Sutcliffe Efficiency
SSR:	Sum of Squares due to Regression
SST:	Total Sum of Squares
FCN:	Fully Convolutional Network

CHAPTER I

Introduction

1.1 Background

Connected Vehicles (CVs) and recent advances of machine learning technologies open new opportunities for the analysis and prediction of the car following behavior; one of the key parameters in modern traffic dynamics. This study is aimed at modeling and predicting car following behavior in connected vehicles with the use of machine learning techniques.

Connected vehicles have unique advantages over the unconnected vehicles main of being able to share data and be interactive with other vehicles and also infrastructure. This ability, mainly facilitated through technologies such as Vehicle-to-Vehicle (V2V) and Vehicle-to-Everything (V2X), makes connected vehicles be in a position of accessing real-time information regarding traffic status, possible road dangers, and weather among others hence improving on the situational awareness and safety. Instead, the unconnected vehicles fully depend on the driver as well as the in-vehicle sensors since they cannot get any relevant information from a lot of external sources. This difference through connected vehicles being able to react to situations as they come up in a proactive manner compared to unconnected vehicles' reaction based only on immediate limited information. (Arvin et al., 2020).

Machine learning can bring a new way of modeling car-following behavior in CVs by handling large datasets and detecting intricate patterns. This makes this technique different from conventional statistical models with its live adaptive learning data (Nu et al., 2022). The implementation of machine learning in this context is thus not merely academic but also a response or solution to the increasing complexity and demands of modern-day traffic systems, especially in urban setups. This research is highly significant and relevant based on global road safety and efficiency improvement initiatives. That is, the World Health Organization has defined road traffic accidents as one of the major causes of mortality in young individuals, and with CVs getting integrated, still there lies a scope of appreciably decreasing such incidents by getting better apprehension and foresight of driver behavior principally in car-following circumstances (Session, 2020).

Another vantage of this study is to explore the data-rich environment provided by CVs along with the analytical prowess of machine learning to come up with more accurate and dynamically updated car-following models. These studies are expected to make it possible to improve the predictability of the connected vehicle system.

1.2 Problem Statement of the Study

There is significant gap in understanding and accurately predicting car following behavior particularly connected vehicle, presents a substantial problem, traditional models for this purpose are often inadequate primarily because they fail to fully leverage the advanced data and connectivity features of modern vehicles. These traditional approaches lack the sophistication needed to accurately model the complexities of modern vehicle behavior especially in scenarios where vehicles are equipped with advanced sensors and commination.

1.3 Aim of the Study

The aim of this study is to evaluate the performance of machine learning in modeling car following behavior

1.4 Objectives of the study

The objective of this research are as follows

- To apply Gradient Boosting and Random Forest Algorithms in Connected Vehicles
- To Analyze and Predict the Behavior of Following Vehicles
- To Assess the Performance of Various Vehicle Platoons

1.5 Significance of the Study

This study presents a great potential to advance connected vehicle technologies using advanced machine learning algorithms such as Gradient Boosting and Random Forest. This would deepen the understating of vehicle behaviors, in particular car-following modelling, enhancing traffic flow and road safety by evaluating vehicle platoon performance. The research could impact on transport policy and urban planning giving data-driven understandings for the strategy of managing traffic.

1.6 Scope of the Study

The scope of this study's focus is wide and mostly on the application of advanced machine learning algorithm such as that used by Gradient Boosting and Random Forest to connected vehicles. The study seeks to analyze and predict car-following behavior, with a view of understanding how vehicles modify their movement depending on the actions of the leading vehicle and general traffic conditions. A major element entails performance evaluation of platoon to ascertain its impact to traffic flow and safety. It is a research that aims to improve the safety of roads and efficiency of traffic to the goals of the overall intelligent transportation system through existing data from the connected vehicles such as speed, time headway, and acceleration.

1.7 Limitation of the Study

the study on the Modeling and Predicting Car Following Behavior in Connected Vehicles: A Machine Learning Approach include the constraints in data availability and quality which are paramount for the effectiveness of the machine learning models. This may potentially affect the performance of these models. Moreover, the complexity of algorithms such as Gradient Boosting and Random Forest that require boosting and bootstrapping may incur challenges in interpretation and implementation. It also fails to capture the randomness of human behavior as well as environmental factors thereby creating a potential gap between the simulated model result and real-life traffic conditions. The development in connected-vehicles technology and machine learning might therefore render the study irrelevant.

CHAPTER II

Literature Review

2.1 Foundations of Traffic Theory

The theories of traffic flow can be seen as the bedrock of Traffic Science. They aim to elucidate the phenomena associated with the motion of single vehicles on a road as they engage with adjacent vehicles. The outcomes of these engagements are what define the essential traits of roads, like their potential and their ability to handle different degrees of traffic flow movement.

Theories of traffic flow were some of the initial contributions to the field of Traffic Science, incorporating both broad-scale (macroscopic) and detailed (microscopic) approaches. From a macroscopic standpoint, traffic is perceived as a continuous flow similar to a liquid moving through a channel, which in this case is the road. From a microscopic standpoint, the focus is on the movements of individual vehicles and their interactions with one another

A traffic flow can be defined mathematically as;

$q=uk$ represents a fundamental relationship in traffic flow theory, where:

q is the flow rate of traffic on a highway lane, measured in vehicles per hour.

u is the average speed of the traffic, expressed in kilometers per hour.

k is the density of traffic, quantified in vehicles per kilometer.

Essentially, there are two primary methods for measuring traffic quantities. The first method involves observations at two or more stationary points along a road, where data on flow and speed are collected at the moments when traffic events happen at each point. Here, an event is defined as the arrival of a part of a vehicle, such as the front, at these observation points. The second method entails making observations at two or more different moments in time (Greenberg, 1959).

2.1.1 Background Information of Traffic Theory

The primary focus of the transportation sector is to facilitate the movement of people and goods in a manner that is efficient, safe, and environmentally friendly. Transportation engineers are involved in several key stages of a transportation facility's lifecycle, which include planning, design, construction, operation, and maintenance (Sherali, 2014).

The area of traffic flow theory is defined by attempts to define the correlations that exist amongst the three main variable of traffic current, velocity (v), density (ρ), and j. Only two of these variables are independent as they are related by $j=\rho v$. Possible units for these variables can be [v]=km/h, [ρ]=vehicles/km, therefore [j]=vehicles/h.

The study of the traffic flow dynamics can be thought of as studying individual vehicular movement but within some broad and frequently repeating details which span many vehicles. As it provides analysis of the interactions between individual vehicles on the road, car-following models are the basis of this area. These models are testing to find out how drivers adapt their speeds and positions towards the surrounding vehicles. It gives insight into factors influencing traffic flow which includes driver behavior and road and traffic. The development of macroscopic traffic flow models such as the Lighthill-Whitham-Richards (LWR) model that describes and predicts the characteristics associated with traffic has used car-following models. But the LWR model doesn't capture stochastic variabilities in traffic flow (Fan et al., 2023).

Additionally, there are complexities in traffic flow within interconnected systems like urban grid networks as demonstrated by (Daganzo, 2007) which highlights strategies for reducing urban gridlock. This complicates human factors that also need to be considered, with (Hoogendoorn & Bovy, 2001) hence emphasizing the critical position of drivers as the topmost element in the general traffic system. With the growth of technology today, autonomous vehicles are being developed with potential impacts on traffic dynamics within cities something that is predicted to (Shladover, 2018) affect or restructure our conventional understanding of traffic streams. Traffic congestion is a characteristic of large cities and major highway systems inflicting a large cost in terms of lost time, uncertainty, and aggravation to the movement of passengers and freight (de Palma & Lindsey, 2011).One of the primary challenges that are facing drivers and users of roads is traffic congestion, It has some detrimental effects on issues of security, physical and psychological health and as well as economics occasioned due to an rise in the quantity of available vehicles on roads, poor driving behavior, and inadequate infrastructures (Boluma Mangata et al., 2022).

Traffic congestion indicates the condition when there is more traffic demand than what the road system can accommodate (Aftabuzzaman, 2007), Roadway capacity represents a feature of road section that quantifies its aptitude in accommodating traffic flow within a certain span of time often expressed as the number of vehicles able to pass through the given section within an hour.

Such an important concept in transportation planning and design as it helps in establishing the performance and efficiency of roads. Determining road capacity is made through putting into consideration a number of factors such as geometric conditions of the road, environmental conditions, and traffic volume among others. The capacity value represents a maximum sustainable flow rate for given conditions and it's used for making road modifications and further improvement (Minkin & Whiting, 2018). The enormous traffic congestion on the roadway arises due to an increasing population and motorized vehicles.

Traffic congestion primarily stems from the growing population and increasing number of vehicles, as rural residents migrate to urban areas in search of improved prospects. The insufficiency of road infrastructure and the prevalence of narrow city streets are key factors leading to traffic jams. Additionally, poor public transportation systems, limited development in road infrastructure, and the rising reliance on private vehicles further exacerbate traffic congestion (Shah, 2020). Traffic congestion is categorized into two types: recurring and non-recurring. Recurring congestion typically results from capacity issues and behavioral factors, while non-recurring congestion arises from unexpected events like accidents, construction work, or emergencies. Although their causes differ, both types of congestion lead to similar outcomes, traffic congestion is defined as a condition in traffic where speeds are reduced, travel times are prolonged, and vehicles are more likely to form queues (Kumar et al., 2021). Congestion is a significant issue in many countries, impacting not only people's daily lives but also increasingly contributing to environmental pollution (Samal et al., 2020).

the impact of movement restriction policies like driving restrictions and setting up of low emission zones have varying impacts reduction in pollution emanating from traffic congestion.

(S. Jia et al., 2021) argue that such policies can effectively deal with traffic congestion and emission constriction in the short run, however, their long-term effectiveness is discussed.

Furthermore, the air quality effects of these policies are dependent on types and locations of pollution in space, further driving restrictions might unintentionally switch consumers' consumption decisions that might result in lower brick and mortar spending inside the regulated areas, also important is to take into consideration unintended results of such policies that include traffic diversion to other routes - driver adaptation through purchasing extra vehicles (Ren et al., 2018).

The COVID-19 pandemic outbreaks significantly improved air quality globally and led to a marked decrease in traffic, resulting in reduced traffic delays, energy use, and emissions (Du et al., 2020). Almost every country reported a decline in air pollution due to lockdown measures (Dantas et al., 2020), with particulate matter PM2.5 and PM10 concentrations dropping by 30% and 25%, respectively (Radoni & Davidovi, 2023). This period also saw a notable reduction in air pollutants and a substantial improvement in air quality during the lockdowns caused by COVID-19 (Addas, 2021).

On the other hand, (Bigazzi et al., 2015) further postulated with relation to traffic density that it is only through an increase in tail gas concentration expositional intensity and driver delay expositional time that traffic congestion increases total amount of traffic pollutants mainly worsen air pollution.

(Fox et al., 2018) showed that traffic congestion increases as the population of urban growth increases; however, this increase in private cars enhanced air pollution in the city of Yangon from Myanmar.

(Rajé et al., 2018) noted that the extensive use of motor vehicles resulted in more frequent and intense traffic jams, correlating with rapid urbanization in Nairobi, Kenya, and consequently leading to heightened air pollution. In their study, (Xie et al., 2018) utilized an improved STIRPAT (Stochastic Impacts by Regression on

Population, Affluence, and Technology) model to investigate the effect of traffic density on haze in urban areas of different sizes.

Both in the secondary and ground levels it can be observed that traffic density does not have a notable effect on haze in large- and medium-sized cities, but it may influence the air quality in small cities Traffic Congestion Urban accessibility affects mobility.

Traffic congestion affects performance to the organization and the travel time and fuel cost of the employees trying to ferry goods and services. Traffic congestion is the problem from which urban areas present such negative effects as delay in time, waste of energy, air pollution, and increase in transportation costs Besides, it influences everyday existence of people and causes an increase in tension wave and wasted time Traffic density is one of the most used congestion indicators, on the given section length of a roadway, reflecting the number of vehicles.

Several factors can be considered to assess traffic congestion, including the level of service, roadway congestion index, and the lane-kilometer duration index, which are measures of travel time reliability (Samal et al., 2020).

2.1.2 Traffic Flow Analysis

The traffic stream is then considered as a continuous fluid with the derived relationships between speed, density, and flow using fluid dynamic principles to produce a general model of traffic follow, the rapid advancement in vehicle design has led to traffic volumes that often exceed the capacities of national highways. The increasing traffic emphasizes the importance of comprehending traffic flow dynamics and formulating a mathematical representation of this phenomenon. This is especially vital during times of extremely high traffic when roads are at full capacity. While many traffic flow theories rely on statistical analysis, (Horňák & Přikryl, 2015) carried out experimental studies by recording actual traffic flows (vehicles per hour) and the speeds of the vehicles. They plotted speed against density (vehicles per kilometer), focusing specifically on single-lane traffic scenarios.

Traffic flow can be studied using different approaches, including fluid dynamics and other methods. For instance, according to (Hou, 2021) the fluid dynamics approach formulates continuity equations and differential analysis on traffic flow parameters with a view of modeling the motion of road traffic (Setiawan et al., 2016), also take a classic fluid dynamics approach and pay particular attention to the Lighthill-Whitham-Richards (LWR) model in an endeavor to simulate traffic flow in a single road, they describe traffic flow in terms of density and average speed of vehicles using the finite volume method for numerical approximation (Catalin et al., 2022), come up with a Bond Graph approach to model traffic flow considering vehicle density and velocity as independent variables quickly point out that traffic flow theory has evolved over time with many mathematical and empirical models available for traffic forecasting and designing traffic control frameworks, (Sean et al., 2017) developed a macroscopic heterogeneous purely car-following multi-class traffic flow model considering the interactions of the several classes of vehicles.

2.1.3 Traffic flow patterns

Determining traffic flow patterns is essential in the realm of traffic monitoring and management. This involves delving into the complex relationships between various elements such as the characteristics of vehicles, the design of roads, driver behavior, and strategies for managing traffic. This crucial aspect of highway traffic surveillance and control has attracted considerable interest over the last thirty years, emphasizing its role as a fundamental component in the effective management and control of traffic. While numerous studies on modeling traffic flow performance are available, only a select few go beyond theoretical approaches and are driven by a strong purpose to address real-world challenges (Celikoglu, 2013). These studies include methods that systematically observe and analyze traffic flow patterns, considering both spatial and temporal variables. There is the application of spatial-temporal information and traffic pattern similarity method for capturing on time-space dependences in traffic prediction models (L. Yang et al., 2021) Another way uses traffic data tensor-based data representation and tensor robust principal component analysis for detecting normal and abnormal flow in large urban networks, a matching method for the flow sequence pattern based on influence factor analysis can use the pattern to match the flow of series and judge the similarity of them, network traffic which is also can be reduced by aggregating packets for a flow and carrying out latency and throughput analysis for difference flows with a flow exporter and flow analyzer agents, flow level analysis of network traffic that is also can be reduced by aggregating packets for a flow and carrying out latency and throughput analysis for flows with flow exporter and flow analyzer agents, vision of computer executes out the flow analysis through and measures variables need engineers of traffic flow (Lykov & Asakura, 2018).

2.1.4 Congestion and bottlenecks

almost all the modern metropolis cities have suffered congestion as a serious problem towards increased usage of vehicular transportation, urbanization, and population increase

Traffic bottlenecks are identified as a major factor in congestion, accounting for approximately 40% of traffic jams (Yue et al., 2022). This underscores the importance of pinpointing traffic bottlenecks to understand and address congestion causes. Addressing these bottlenecks can provide effective and cost-efficient solutions for traffic improvements. Strategies go beyond just expanding road capacity at bottleneck points; Additionally, they incorporate sophisticated techniques such as controlling traffic lights and rerouting vehicles to reduce congestion in these key areas (Urban et al., 2018). Although a significant portion of existing research has been centered on pinpointing bottlenecks on freeways (Kerner, 2007), there is a growing consideration for bottleneck identification in the context of urban road networks. the problem has been quite challenging. Millisecond with urban network existence, the road topology is more intricate. Therefore, it results in difficulty for estimation not only of the vehicle travel pattern but also the pattern of congestion propagation. Urban roadways have much more traffic equally to that in expressway compared to the second one so more unexpected traffic condition exists in road networks. Third, the traffic signal and social activities have other factors that impact more significantly for the urban roadways than the freeways. Lately, bottleneck identification in the urban area received a lot of attention (Ma et al., 2016).

2.1.5 Recent identification methods to traffic flow patterns:

a. **License Plate Recognition (LPR):** LPR is based on optical character recognition which identifies and recognizes license plates in images or video feed. It is mainly used in monitoring and management of the traffic due to its ability to indicate movement of vehicles within a region by identifying their license numbers.
b. **Object Recognition through Machine Learning**: Recent developments in machine learning using deep learning techniques, in particular, have made feasible the development of high accuracy object recognition systems. It requires drawing of approaches to detect and recognize such vehicles, persons, or any other things from images, making real-time situations easier for the traffic understanding purpose.
c. **Connected Vehicle and IoT**: The entry of connected vehicles and Internet of Things makes it possible to extract real-time data out of vehicles. Such data stuff would comprise of location, speed, acceleration kinds of information that could enable completion of analysis related to traffic patterning and enhancement of traffic management.
d. **Mobile Apps and GPS Data**: Many people have navigation apps in their mobile phones where the application captures GPS data as he uses it. Aggregated and anonymized GPS data from this kind of application could answer various traffic pattern questions for various authorities and transport agencies when need be.
e. **Video Analytics**: More complex video analytic system based on computer vision pay attention to the video content from traffic cameras. This is a type of video analytic system designed for detection and tracking of cars, pedestrians, and other objects, which supply data to the departments responsible for traffic management.
f. **Blockchain for Traffic Management**: The technology can offer enhanced traffic data security and reliability by applying blockchain within transport-service firms. It could be used to build a decentralized as well as tamper-proof structure for logging every transaction associated with traffic, guaranteeing the integrity of mined data in traffic analysis.

These identification methods illuminate traffic flow patterns by providing real-time data of vehicle movement, congestion, and other relevant issues. By incorporating these technologies, improvement can be realized in traffic management techniques that further widens to intelligent transport systems development.

2.2 Traffic Simulation Frameworks

Macroscopic models solicit the traffic flow as an integral, while microscopic models that focus on the dynamics between individual vehicles and their interaction while mesoscopic models are situated in between these approaches. A microscopic model of traffic flow seeks to understand the movement of traffic by simulating interactions between driver and as well as between a driver and the road whereby the driver interacts with the other drivers as well as interacting with various features on the road. Traffic simulation is thus an accomplished method employed towards the study of traffic systems, therefore playing a central role in modelling the functions of dynamic traffic networks. Traffic simulation has thus been tremendously popular in the dynamic modelling of operations in a traffic system. With the use of traffic simulation, ideally, three main types of the traffic simulation models.

2.2.1 Microscopic traffic simulation modelling

Microscopic traffic simulation models are highly effective exactly in reproducing the real traffic conditions over the road network created virtually in the computers. These models add special significance while evaluating the impacts of Real-Time Information (RTI) systems. RTI systems render up to the minute information about traffic conditions, road works, accidents and all other relevant data. According to the above, these microscopic simulators have been custom-made to better emulate the field traffic and this makes them most suitable for studying the impacts of mixed traffic conditions with automation and treading driving vehicles. Undoubtedly high among these elements will be the importance of the precision with which the models can replicate traffic situations, since accurate impacts measurement like those of RTI systems can greatly affect vehicle dynamics and driver’s behaviors. The two most important key factors in determining success in simulations are:

Delineation to details of road geometric network including details of all traffic facilities that are associated with the road such as lights, detectors, Variable Message Signs (VMS) panels among others.

Comprehensive depiction of traffic characteristics faithful to the actual motion of individual vehicles.

This means that microscopic traffic simulation models are considered better evaluation tools for the effects of RTI systems on the behavior of traffic.

Microsimulation models are characterized by which can easily identify the vehicle types as well as imitating what the driver does. The models are very effective in reproducing traffic systems which allow analysis of flexible traffic situations. These models provide complete datasets in developing traffic control and management plans by capturing the dynamics of the traffic by focusing on the behavior of individual vehicles and how they interact with other vehicles and with road infrastructure.

These models work through use of particular rules of behavior that direct the way a vehicle slow down, speeds up and switch lanes. These rules are representative of which way, and at what time a vehicle would switch the lanes to reach its destination by first selecting the lane to be used. The interactions can be broadly categorized into the following types of interaction dynamics the vehicle portrays: Vehicle pursuit models, used for changing lanes while maneuvering and keeping the predestined travel path upon alteration intact, and models that select paths for traveling:

Vehicle Pursuit Model: This model uses how the car justifies its speed and the distance of the car equally from one that is in front of it via direct. different factors which include the road conditions and also that time's legal speed limit.

Model for Changing Lanes: This model explores the decision-making process of a driver to change lane regarding personal driving style for the situational features that in comparison to the speed of vehicles around him/her and there is enough space if he switches lanes.

Model on travel path selection: examines the method that drivers use in choosing the travel paths from the starting point to the destination, whether they adopt to either traffic jams or any other information about the route while on board.

Some of the simulation models with detailed and accurate frameworks include VISSIM (Khashayarfard & Nassiri, 2021), AIMSUN/2, MITSIMLab (Mutabazi, 1987) and CORSIM (Mutabazi, 1987) among others. The paper identifies accuracy and high level of detail (of the traffic behavior) as the hallmarks of the models.

2.2.2 Macroscopic simulation modelling

The macroscopic model is a mathematical structure that defines the interrelationships between primary traffic characteristics such as density, flow, and average velocity of a traffic stream (Francesco & Rosini, 2015).

Macroscopic Models:

LWR Model: The Light hill-Whitham-Richards (LWR) model represents a type of hyperbolic conservation law where the solution, typically reflecting the macroscopic density, demonstrates the average spatial distribution of vehicles (Levi-civita & Pura, 2022).

Microscopic Models:

The General Motors (GM) model, initially presented by Chandler and colleagues in 1958, stands as one of the most recognized stimulus-response models. Within this model, the stimulus is identified as the relative speed of vehicles, implying that each vehicle tends to maintain a velocity akin to the one ahead of it (Li & Sun, 2012).

Macroscopic models generally require less computational effort as they deal with aggregated traffic information. Microscopic models, on the other hand, provide more detailed information in such a way that it is even possible to analyze specific interaction or. More clearly or describe particular situations much more demonstratively. While macroscopic models emphasize mostly about the general behavior of traffic streams - by which they are largely just aggregates of individual vehicles - microscopic models focus on the interaction among the individual vehicles (Khan & Gulliver, 2018). Dynamic models have been developed to model the variation of the character of traffic flows in time. (Lighthill & Whitham, 1955) and (Logghe & Immers, 2008)presented LWR model as a macroscopic dynamic model. To date, the

LWR model is one of the most prevalently used in the traffic modeling can be differentiated by three primary variables:

1. Flow (q): The frequency at which vehicles cross a specific point.
2. Density (k): The number of vehicles per unit length of the road.

2.2.3 Mesoscopic Traffic Simulation Modelling

Mesoscopic traffic simulation models act as a missing link between the broad view macroscopic models and detailed microscopic models, offering an all-round perspective of traffic elements with more detail in sight. One of the approaches is the organization of the vehicles into groups or "packets" that move jointly along the road network behaving as a unit. Doing so, the speed of this unity on each road results from a specific function of speed-density and density of vehicles on that road. Density here is the ratios vehicles to road length. In case a road shows high density, meaning heavy congestion than the speed-density function decreases the vehicles' speed where the opposite happens for low density. On the other hand, mesoscopic models do not go to individual details of lane change or acceleration/deceleration, unlike microscopic models. Another mesoscopic paradigm arranging individual vehicles into "cells," controlling their behavior as they traverse the road network; allowing entry and exit but preventing overtaking. These models provide a more refined abstraction of traffic patterns, still at a high level of abstraction, regarding the movement of single vehicles (Adapted from "Traffic Simulation Models: System instance dictates vehicles' speed and not individually drivers' acceleration choices DYNAMIT (Atasoy & Akkinepally, 2019).

(Traoré et al., 2020) argued that the road is conceptualized into double separate mechanisms: the queuing section and the running section. Some considerable dynamic traffic theory supports the argument that each lane has got the possibility to be modeled independently but they are, however, often lumped together. Although the cars are adopted as individual entities represented over the screen following their performances but this simulation is not followed in detail. Vehicles traverse the operational segments of the roadways at a velocity determined by a macroscopic speed bump function integrated into the model. At the endpoint, a queue-server mechanism assists in directing vehicles onto connecting roads.

2.3 Traffic Flow Model

Traffic flow models, which have been developing since the early 20th century, are a part of the broader historical evolution of mathematical modeling of various systems. These models, used by scientists and engineers, serve as simplified representations of the real world. They have applications in diverse fields, from forecasting weather and understanding chemical reactions to studying material behavior, human behavior, fluid dynamics, and, notably, traffic flow ((Brackstone & Mcdonald, 2000).

The process of traffic flow modeling is essentially inductive: it starts with a theory about individual driver and vehicle behavior or general traffic flow, which is based on observed traffic patterns. This theory is then used to construct a model, which is subsequently discretized for use in simulations (Taylor & Castillo, 2012). Gathering data is the foundational step in traffic flow modeling and involves various techniques like loop detectors, cameras, GPS units, and driving simulators. These collected data inform the construction of a theoretical structure, encompassing qualitative narratives and assumptions about driver behavior, such as the dependence of speed on the perceived distance to the car ahead and interactions between leading and following vehicles.

The evolution of these models culminates in the formulation of traffic flow equations and principles, typically depicted via the fundamental diagram or the car-following model. For computational simulation, converting continuous models into discrete forms is essential, often by segmenting time and occasionally space as well. Numerical techniques are then applied to simulate a discrete version of the traffic flow, integrating these traffic conditions into a simulation program. This program employs real-time traffic sensor data for estimating and predicting traffic conditions, and the models are validated by comparing their outputs to real-world observations.
Car traffic models have various applications, including:

- ❖ Estimating current conditions and providing short-term forecasts for travelers.
- ❖ Implementing short-term traffic control based on current assessments and future predictions.
- ❖ Providing guidance for (semi-)autonomous vehicles.

- ❖ Long-term evaluation of infrastructure projects, like changes to the transportation network.
- ❖ Assessing traffic's effect on safety and environmental emissions.
- ❖ Creating strategies for emergency evacuations.

2.3.1 Background Information on Car-Following Models

The car-following model is a simple theoretical framework of traffic research meant to decipher the behavior of a following vehicle along streets and highways. Particularly, it focuses on their through the reactions of a driver concerning the vehicle at the front within the same lane. It's premised upon the fact that if a vehicle finds itself constrained by some leading vehicle such that continuing to maintain its current speed would lead it to collide in one way or another, then such a model describes the behavior of a vehicle under this situation. This behavior often is guided by principles of safe driving and traffic regulations, stating this being a crucial necessity to maintain a sufficiently safe distance which would prevent such accidents. A vehicle is considered to have freedom of motion if it does not necessarily have another vehicle as a constraint and hence can make maneuvers at the desired velocity. Car-following has been considered quite important in traffic safety and constant research is being taken to improve or further refine the model (Y. Jia et al., 2021). The term car-following was first originated back in early 1950s, while the concept was first presented by (Pipes, 1953). It refers to how a vehicle maintains its lane while following the vehicles in front of it.

The longitudinal movement of vehicles within a lane is a fundamental element of traffic flow theory, where modeling car-following behavior is crucial (Han, Wang, et al., 2022). Over nearly 70 years, research into car-following behavior has evolved to encompass various models based on different theories and perspectives (Han, Shi, et al., 2022).

a. Driving Behavior

Driving behavior models are designed to accurately forecast the movement of individual vehicles, providing highly detailed insights that are crucial for various systems.

Driver behavior encompasses both the deliberate and unintentional actions and traits exhibited by a person while driving a vehicle. A multitude of factors can influence a driver's behavior, including age, driving experience, gender, personal attitudes, emotional state, fatigue, drowsiness, and the prevailing driving conditions. Both internal and external factors have the potential to vary the same driver's risk assessment and decision-making capabilities across different situations. Generally, driver behaviors are categorized along a spectrum that ranges from normal to risky and aggressive.

b. Acceleration behavior

Acceleration behavior is the manner in which a driver of a vehicle adjusts its speed in accordance with this of the leading vehicle

1. The acceleration is a strictly decreasing function of the speed. There is also the limiting behaviour that if it were not other vehicles or obstacles constraining it, the vehicle would accelerate towards a desired speed vo

$$\frac{\partial a_{\mathrm{mic}}(s,v,v_l)}{\partial v} < 0, \lim_{s\to\infty} a_{\mathrm{mic}}(s, v_0, v_l) = 0 \qquad \text{Eq 1.}$$

2. The In-traffic modeling, acceleration is represented as a non-decreasing function of the distances to neighboring vehicles. This representation implies equality when other vehicles or obstacles (including "virtual" obstacles like a stop line at a red traffic light) fall outside the interaction range and therefore do not influence driving behavior. This scenario is referred to as free-flow acceleration (Sun et al., 2023).

The mathematical expression for free-flow acceleration is given by:

$$a_{\mathrm{free}}\,(v) = \lim_{s\to\infty} a_{\mathrm{mic}}\,(s, v, v_l) =\geq a_{\mathrm{mic}}(s, v, v_l) \qquad \text{Eq 2.}$$

The response time in car-following situations is characterized by a lag time, which is the duration it takes for the following driver to respond to changes in the leading driver's behavior. This driver reaction time in car-following scenarios can be defined during acceleration and deceleration maneuvers. When engaging in acceleration/deceleration maneuvers, the car-following driver is either speeding up or slowing down to maintain a desired speed under the given kinematic conditions (Mehmood & Easa, 2017).

c. The desired speed

The "Desired speed" is defined as the maximum speed a following driver would opt for, given certain kinematic circumstances. Factors such as speed limits, weather conditions, and visibility can also influence this speed. In the context of car-following, a driver modulates their speed by either accelerating (pressing the gas pedal) or braking, particularly when the vehicle ahead is braking, indicated by its activated brake lights. The term "acceleration or deceleration reaction time" (ADRT) is used to describe the scenario where the driver's speed adjustment is achieved purely through acceleration, without the use of brakes (Mehmood & Easa, 2017).

d. Driver Heterogeneity

The car-following patterns of different drivers can vary even in identical conditions, a concept known as "external heterogeneity." Conversely, the same driver may exhibit varying car-following behaviors in the same circumstances at different periods, a phenomenon referred to as "internal heterogeneity."

c. External Heterogeneity

"External heterogeneity" refers to the variance in car-following behaviors among different drivers. This diversity in driving styles impacts not just the immediate behavior of vehicles (micro level) but also plays a key role in the nonlinear dynamics of overall traffic flow (macro level), as noted by (Kerner & Klenov, 2004). This variation in drivers' behavior manifests not only in reaction time delays but also across various characteristics. Analyzing the car-following patterns of each individual driver is complex, but certain patterns can be identified. To facilitate this analysis, some researchers, like (Constantinescu et al., 2010), categorized drivers into several groups, allowing for a more structured study of the external heterogeneity among these different driver types.

d. Internal Heterogeneity

The car-following behavior of an individual driver can vary under different circumstances or even in the same situation, influenced by psychological or physical factors. This variation is termed "internal heterogeneity," as described by (Hamdar, 2008). In their work, (Zhu & Dai, 2008) presented two delay factors and extended on the Newell model.

2.4 Model Components and Variability

Understanding car dynamics is hence major component of the car following models. This takes into account key factors such as the velocity of the vehicle, and ability of accelerating and decelerating. These mathematical models can be either simple and complex (i.e., linear and non-linear) equations to capture such intricacies in vehicular movement (Chu et al., 2003),This variability often emanates from the differences in driving styles. Drivers have their preferred following distances and speeds based on each one individual's driving habit. A broad spectrum of driving behaviors results from drivers' aggressive styles and others which place emphasis on safety.

2.4.1 Model Calibration

Calibration is needed for the car-following models, so as to obtain a set of parameters that minimize variations between truck-trajectories which are simulated and real. The choice of inputs among the model inputs specifically, the number of vehicles and the percentage mix of heavy vehicles, is needed to assure satisfactory traffic simulation. Such data on vehicle specifics and driver behavior are to a great extent difficult to be acquired from real observations. In this respect, the users of simulation models have to carefully calibrate their inputs regarding driving behavior and vehicle specifics in conformity with some absolute metrics. This entire process is termed "Calibration," as defined by (Dawson, 2019).Adaptive Cruise Control (ACC) is one of the technologies used to help in maintaining distance with leading vehicles and applying the speed limit through automatic regulation of the car's speed with reduced driver input (Rajamani & Zhu, 2002).In modeling the car-following behavior of vehicles equipped with ACC, calibration is crucial. This particular control law within the vehicle controller for ACC is often unknown hence the need to develop a model that mimics the traffic-level behavior of the ACC vehicle based on obtainable inputs and outputs. He did not mean to say that this approach aimed at reverse engineering the actual ACC real controller. He implied that the objective of model calibration is the identification of some key parameters which could best mimic the observed vehicle trajectories, as described by (Milanés & Shladover, 2014).

2.4.2 Traffic Parameters (e.g., lane discipline, speed limits)

Key elements such as lane discipline play a significant role in shaping the dynamics of road traffic systems. Lane discipline, in particular, implies the extent drivers do not deviate from the assigned lanes, that all is a vital issue both as a traffic safety as for an efficient flow. Several studies have featured the role that proper lane discipline takes in cutting down accidents and enhancing flow of traffic including by flow (Delpiano et al., 2020).Proper lane discipline comes highly because there are high levels of a Variable Speed Limit (VSL) system due to highway management. VSL systems require drivers adapt speed to evolving traffic situations like work zones, reduced visibility, wet roadways or queue formation. These systems usually display progressively decreasing speed limits in advance of a developing bottleneck and help manage the flow of traffic. They also serve the reminder job of alerting drivers to adjust their speed in view of unfavorable weather, wet roads or darkness (Lin et al., 2004).

2.5 Advanced Features in Car Following Models

Car-following models have advanced to incorporate dynamic characteristics beyond simple implementations of vehicle following algorithms. More advanced features, prevalent amongst contemporary vehicles and automated driver assistance systems (ADAS), are helping in boosting driving safety, as well as its efficiency as well as its comfort features (Shladover et al., 2012). These advanced features broadly fall into ADAS and automated features. For example, these ADAS features would include systems like Blind Spot Warning (BSW), Lane Departure Warning (LDW), Over Speed Warning (OSW), and Forward Collision Warning (FCW). On the other side, automatic features include technologies like Automated Cruise Control (ACC), Cooperative ACC (CACC), Lane Keep Assist (LKA) as well as Automated Emergency Braking (AEB). ACC and LKA, commonly known as active lane keeping, are common in Level 1 and Level 2 connected and automated vehicles. ACC keeps predetermined speed and distance from the preceding car, adjust speed if necessary, can stop completely if necessary as well. LKA partly controls steering helping to keep the automobile in its lane. The addition of advanced features, such as ACC and CACC, to vehicles greatly enhances traffic flow, boosts stability behavior, and influences road capacity (Arnaout & Arnaout, 2014). However, overall the impact of these features all depends on the percentage of the vehicles having such features on the composition of the aggregate traffic (Shladover et al., 2012).

2.5.1 Robustness

The effectiveness of these models lies in their ability to sustain desired performance in diverse conditions and uncertainties. These models are utilized to simulate how vehicles adjust their speed and position in relation to the car ahead (Wu et al., 2020).

(Aslani et al., 2018) initially explored the resilience of different Reinforcement Learning (RL) algorithms against system disruptions in multi-intersection environments. It was noted that value-based reinforcement learning algorithms exhibit a lack of robustness in the presence of disturbances. Nevertheless, during an extended period of training, policy-based reinforcement learning algorithms demonstrated superior proficiency in handling these disruptions. It is noteworthy that in these works, tile coding was employed as approximation technique of the value-function thus raising concerns on accuracy and its implications on stability characteristics of Reinforcement Learning (RL)-based control algorithms. (Rodrigues & Azevedo, 2019) came up with the RL-based controller for the management of one traffic intersection. This was tested for effectiveness under various conditions, using different levels of traffic demand and with sensor malfunctioning. Therefore, their study established that mostly improved performances of the RL algorithm came with changing the sequence of phases of traffic lights. Also, they emphasized to include as part of offline training process a more complete set of data in terms of traffic state. This is to ensure the development of a more reliable and effective RL-based algorithm for traffic signal control.

2.5.2 Adaptive Cruise Control ACC

This system is engineered to aid vehicles in keeping a safe distance from the car ahead and adhering to speed limits. It automatically regulates a vehicle's speed, reducing the need for driver intervention (Rajamani & Zhu, 2002).

ACC operates using sensory technologies like cameras, lasers, and radar installed in vehicles. These technologies gauge the proximity of one car to another or to different objects on the road, laying the groundwork for future advancements in car intelligence. These sensory tools enable the vehicle to detect potential forward collisions and alert the driver. In such instances, red lights flash and a "brake now!" alert may appear on the dashboard, accompanied by an audible warning.

Currently, various car manufacturers are developing adaptive cruise control (ACC) systems. These systems will augment traditional cruise control by adding the functionality to maintain a specific distance from the car ahead detected in the same lane (Martin, 1993). These ACC systems are designed to be "autonomous," relying solely on on-board sensors, such as radar for distance and speed measurements, to maintain the desired spacing.

2.5.3 Advantages of adaptive cruise control

(Rajamani & Zhu, 2002) highlights several benefits of adaptive cruise control, such as enhanced road safety. Cars equipped with this technology maintain appropriate distances from other vehicles, which helps in avoiding accidents caused by obstructed views or tailgating. Additionally, ACC contributes to smoother traffic flow due to its ability to be spatially aware. For drivers, this means less concern about speed control, allowing more attention to be paid to the surrounding environment.

2.5.4 Sensitivity analysis in traffic Simulation

Sensitivity analysis is crucial in scientific modeling, yet it's sparsely represented in traffic modeling literature. This analysis typically follows two distinct methodologies:

- ✓ **One-at-a-Time (OAT)**: This approach evaluates the impact of individual model inputs (parameters) on model outputs by varying one parameter at a time while keeping others constant. However, OAT has two main limitations. First, it fails to capture the interactive effects between parameters, making it only reliable for purely additive models, which is rarely the case with traffic models. Second, OAT is a local method, focusing only on the vicinity of a specific point and not providing insights into the broader input space. The VISSIM model's application in studies by(Lownes & Machemehl, 2006) (Mathew & Radhakrishnan, 2010). utilized this approach.

- ✓ **Analysis of Variance (ANOVA) based on Design of Experiment (DoE)**: This method seeks to determine the contribution of each parameter and their combinations to the variance in the model output. Models are tested across various parameter combinations derived from a DoE. This technique was employed in traffic simulation models as noted in studies by (Bartin et al., 2005) (Punzo, 2016). However, ANOVA's major drawback is its inefficient exploration of the model inputs' space, which can lead to misleading results.

Elementary impact methods proved to be a brilliant alternative to the One-at-a-Time (OAT) method. Elementary impact and all its modification methods provide the genuine estimation of total order sensitivities without the necessity for further model evaluation and the necessity to rearrange the experimental setup, as it was questioned by (Campolongo et al., 2011). In the traffic simulation field, this method has been notoriously used in a particular case by (Ge et al., 2014). this method is even more interesting for models needing high computer power were other methods can appear impossible to achieve.

2.5.5 Traffic Hysteresis

Traffic hysteresis on freeways basically refers to the loops in the density-speed-volume relationship of vehicles, particularly under disturbances or post-incident conditions. First explicitly recognized back in 1974, it later was derived from two Greek words, hústeros and numeín meaning 'lagging behind' (D. Chen et al., 2012). This usually involves a delay in the recovery of speed, but years of study still had not identified what these causes were exactly for traffic hysteresis.(Newell, 2002) proposed that this hysteresis phenomenon arises from asymmetry in behavior between acceleration and deceleration, resulting in two distinctive branches in congested traffic flow. (Zhang, 1999) defined his model mathematically where he divided the flow into three phases:

acceleration, deceleration, and strong equilibrium. And he showed that transitions between these phases could give the rise for formation hysteresis loops some of his theoretical predictions correlated with real data. (Zhang & Kim, 2005) proposed a car-following model, which relates the speed of the driver to the current phase of traffic and elongation in time required for reaching the lead vehicle. This model seemingly can be justifiable, through gap-time functions, the hysteresis in traffic, though is yet empirically not justified. Further elaborating on this concept, (Yeo & Skabardonis, 2009) indicated that traffic conditions will be grouped into five states according to accelerations and decelerations, and obsessed with asymmetry as being the most important reason for hysteresis. They, however, said that the theory cannot account for hysteresis loops in real traffic conditions that are counter clockwise. On a whole, these models give the cause of traffic hysteresis to be as a result of asymmetries present amongst various traffic phases but the base foundation on the reason for this asymmetry is still open for continued research.

2.6 Behavioral Aspects and Interactions

The interactions and dynamics of behavior in car following models certainly require careful attention because play critical elements in an attempt to enhance both the road safety and traffic flow efficiency. These complexes include the intricacies in individual driver behaviors as well as minute patterns of their interactions with the other drivers and the outside environment. Hence, the optimization of car-following models is somewhat in the background of this complex web of behavior and vehicle interactions.

2.6.1 Interactions and Behavior

Interaction refers to a situation in which the activities of at least two road users are believed to be influenced by the possibility of occupying the same space at the same time in the near future. It's important to note that an interaction, as per this definition, requires mutual adjustment in behavior. For instance, consider a case where:

1. A pedestrian wait for a car to pass before crossing the road.
2. The car driver, however, continues without altering speed and without visibly acknowledging the pedestrian, indicating no change in behavior.

In this scenario, the pedestrian displays interactive behavior, but the car driver does not. Under the provided definition, this would not constitute an interaction.

Nonetheless, this interpretation can vary. For example, if it's assumed that the car driver did notice the pedestrian initially but chose not to make further eye contact as a strategy to signal non-yielding, then, in this perspective, it might still be considered an interaction (Clark, 2020).

2.6.2 Platoon behavior

A vehicle platoon is a formation of vehicles traveling closely together at a uniform speed with minimal gaps, facilitated by wireless vehicle-to-vehicle communication and automation. Such platoons are effective in enhancing traffic throughput and reducing jams, as noted by (Hall & Chin, 2005). Additionally, platooning can lower fuel consumption and emissions (Liang et al., 2016) and has the potential to increase driving comfort and road safety (Xu et al., 2014). Due to these transportation benefits, platooning is garnering significant interest in academic and industrial sectors. However, implementing platoons in actual traffic conditions remains a significant challenge.

Diverse behaviors of human drivers can greatly impact traffic safety and efficiency, necessitating thorough investigation. Car-following models, which focus on longitudinal vehicle interactions, are central to both microscopic simulations and traffic flow theory. Research shows that driver interactions are influenced by various factors such as vehicle speed, time-to-collision, traffic density, gaps between vehicles, road characteristics, weather, lighting, presence and behavior of other road users, as well as drivers' demographics, experience, knowledge, cognitive state, and feelings of safety or insecurity (Aramrattana et al., 2021) (Endsley, 1995) emphasized the importance of a shared understanding among road users to avoid conflicts and breakdowns in interactions.

Platoons occur mostly in cases of busy highways, perhaps more often on two-lane highways where restricted heavy vehicles are unable to overtake (M. Wang et al., 2019), It goes through various behavioral adaptations that in turn compromise the safety and efficiency of traffic such as poor gap perception increased risk of rear-end accidents, uncertainty in travel time. Additionally, the platoons of drivers take more risks during the changing of lanes since they are not in a position to adjust their speeds at the target lane. This sort of impulsive lane changing can create shock waves in the left lane increasing the risk of collision, overburdening the lane, reducing overall highway speed and negatively affects the road capacity and traffic flow. Besides, this would lead to attempts of right-side overtaking and create unpredictable conditions as well as loss in traffic safety. Concerning highway on-ramp blending, interaction exists between the merging vehicles and with the mainline drive-through vehicles (Y. Chen et al., 2022). Studies have showed that the mainline drivers usually exhibit cooperative behavior, such as lane changing or stopping when signaling for merges (Aramrattana et al., 2021). Lane-target drivers vary their speeds according to the lane that is targeted. During busy traffic, however, this causes inefficient and unsafe behavior of early merging at lower speeds or abruptly stopping at the end of acceleration lanes that disrupt the traffic.

2.6.3 Platoon of Vehicles

platoon consists of a group of vehicles traveling in close formation, one behind the other, at highway speeds. This formation is led by a lead vehicle, with subsequent vehicles closely matching the speed and movements of this leader.

Vehicle platooning is a strategy aimed at coordinated vehicle movement, proposed as a solution to various contemporary transportation challenges such as traffic congestion, road safety, energy consumption, and pollution (Maiti et al., 2020). This concept involves vehicles following each other closely, yet without any physical connection, while maintaining a safe distance. The primary goals of platooning include reducing fuel consumption (Steven & Thompson, 2015) and enhancing road capacity and throughput. Additionally, it focuses on road safety by minimizing collision risks through coordinated vehicle behavior (Maiti et al., 2020).

2.6.4 Advantages associated of Vehicle Platoons:

- Road capacity is increased as vehicles can drive in extremely
- Improved traffic speed and fuel economy could lead to lower freight costs.
- The accident rate may decrease as vehicles in a platoon are in sync with each other's actions.

2.6.5 Disadvantages of Vehicle Platoons:

- ✓ The closely following vehicle may get a reduced cooling airflow over their radiators thus increase the likelihood of overheating.
- ✓ At traffic lights, the hazard of not all cars in the platoon clearing the junction in sequence, or there not being enough room available on the far side of the junction for all vehicles.
- ✓ Platooning is not practical for stopping at stop signs.
- ✓ Potential rear-end crashes may occur when the following vehicles have different braking and acceleration characteristics

The perception of the drivers to their traffic environment is very important in the analysis of driving behaviors, decision processes, and the possible risks for collisions. Understanding how the drivers are perceiving and interpreting their environment ensures measures to be successfully implemented regarding road safety, driver training, and proper establishment of infrastructures.

2.6.7 Reaction time of the car following vehicle

Moreover, reaction time typically refers to the duration it takes for a driver to respond to a situation, such as applying brakes during an emergency. Longer reaction times are generally linked with higher risks, especially if the lead vehicle performs sudden maneuvers. Conversely, shorter reaction times are crucial for enhancing the safety of road users. In existing simulation models, solutions are derived for the current time step only, meaning that reaction times cannot surpass the simulation's time step size. As a result, the chosen increment for simulation time must be carefully considered for its impact on reaction times. This leads to a compromise between the time step size and feasible reaction times in simulations.

To address these limitations, it's necessary to develop solutions that make reaction times independent of simulation time steps. The car-following algorithm developed for this purpose provides continuous, independent solutions that don't rely solely on simulation time steps. This is made by the use of linear acceleration model and continuous solution of car-following logic pass solving only at present time step.

Continuous solutions enable generating a set of successive chain reaction times. At each simulation time step it is ascertained whether occurrence of the scenario of following a car arises and if a reaction required to be scheduled for the car to follow in case exists. Such sequentially scheduled reactions may cause chain reactions. In practical applications, reaction time decreases in consecutive time steps as once alerted, the driver would most probably have experienced reduced reaction times following the initial response.

2.7 Simulation, Validation, and Analysis

Simulation, validation and analysis make the crux in the intricate process on how to develop the car-following models and evaluate them which are pivotal on simulating studies of vehicle interaction with close proximity on roadways. This comprehensive literature review explores the meaning of these three interrelated phases that may help one understand traffic behavior, may help progress the field of transportation engineering, and most importantly improve road safety.

2.7.1 Traffic Simulation Scenarios

Traffic Simulation widely uses transport analysis through transportation engineering which is regarded as the essential tool for guiding traffic-related decision-making activity and policy development. According to (Lopez et al., 2018), most models of traffic simulators often use heuristic models of car-following, which take general characteristics of flow and density into account but do not provide the detailed insight at the level of separate streets. Due to this limitation, it results in a huge gap on their domain and thus it is not suitable to the offline evaluations of autonomous systems. Increasingly accurate modeling and simulating of transportation systems present challenging problems by the fact that diverse and variable traffic behaviors as well as temporal-spatial variations are highly complex. To solve the above problem, several traffic simulation models have been developed among them SUMO, as pointed by (Lopez et al., 2018).

2.7.2 Simulation and Validation

The enhancement of the simulation tool poses a lot of challenges. Some of the major goals include enhancing the driving behavior models since such models have very low accuracy, especially under congested and heavily trafficked states. There is an increasing interest in employing simulation models for predicting traffic safety and environmental impacts. Such models should also incorporate other modes in the transport system that include freight, public transit, cycling as well as walking, all with consideration to their interactions. Other areas that further research and development will touch on are multi-resolution modeling and real time traffic flow simulation. Since trajectory data collection is progressively getting more feasible, it will be important to develop methods of its collection as well as for the calibration and validation of these data. The current research efforts try to simulate and analyze operation ties with the deployment of transportation systems and services, with a significant focus of research on Connected and Automated Vehicles (CAVs). In particular, challenges will involve simulating actions involving the CAVs together with the human-operated vehicles and non-motorized traffic, dealing with sensor accuracies under different infrastructures, and observing impacts of the detection errors on CAVs. Simulation modeling incorporated in decision support systems for real-time improvements in transportation systems brings challenges such as ensuring that simulations exhibit minimal latency for timely predictions of performance and impacts. With validation, the objective is to guarantee that a model performs its purpose. However, when a model is perfectly calibrated with an ideal data fit, it may not validate accurate, if it does not seem to jibe accurately with other observed patterns during validation. On the other hand, calibration identifies the optimal parameter set while validation defines a specific target for which the model is being developed and validates the range of applications. The range therefore covers the circumstances in which the model operates at some set quality level, and may entail different days, different locations, countries, or lane configurations depending on user specification. It is imperative to appreciate that a model is only validated under the specific conditions it has been tested. For example, a model validated for one road layout of some day may not be valid in another road layout. Purpose of validation - The core intent is to confirm the fact that the model attains necessary performance level, without which no such assertions on quality about model can be made (May, 1997).

2.7.3 Microscopic Simulation

Microscopic traffic simulation systems have emerged as key instruments for analyzing and managing transportation systems. Their primary function is to predict or assess the performance of existing or planned traffic scenarios or measures. In an effort to enhance the effectiveness of these simulation systems, substantial work has been done to develop detailed microscopic models that help understand the laws governing driving behavior, including car-following rules for longitudinal interactions (Q. I. Yang & Koutsopoulos, 1996).

In recent years, there has been a growing focus on microscopic road traffic simulation. Traditional analytical tools often fall short in addressing complex issues like congestion, incident management, signal control optimization, and public transport priority, especially within the intricate urban road transport systems. Microscopic simulators enable transportation planners to model these complex systems in their entirety and evaluate different traffic management options. This helps in identifying the most effective solutions for various traffic scenarios. To this end, several traffic simulation tools have been developed and utilized in the field (Hidas, 2005).

2.7.3 Validation with real data

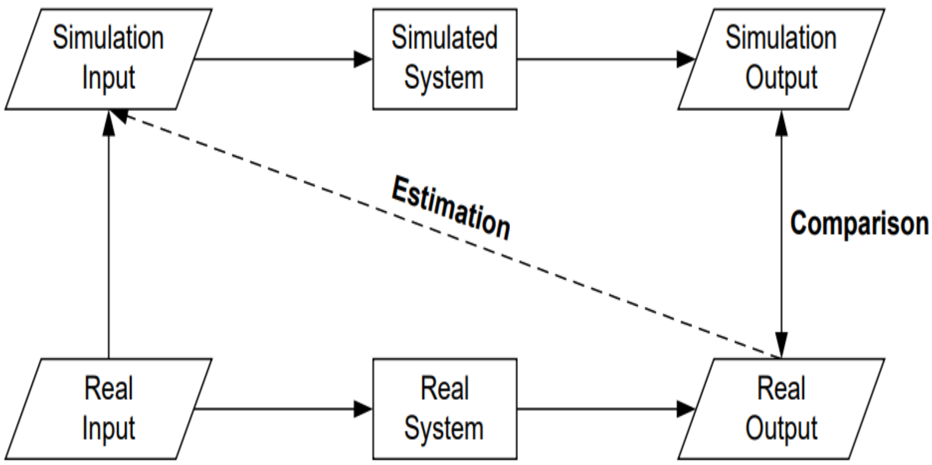
The aim of aggregate validation is to determine how accurately a simulation model replicates a real-world system, using data sources such as loop detectors. The ideal process involves comparing the outputs of both real and simulated systems under the same input conditions. Therefore, it's important to record not only the outputs of the real system but also its input variables for replication in the simulation. This method reduces the disparity between observed and simulated outputs, thereby improving the effectiveness of the comparison. An important input in traffic simulations is travel demand, typically represented by dynamic origin-destination (O-D) matrices. However, these matrices are often not directly observed and need to be estimated, making aggregate validation a complicated process, as shown in the figure by (Hazelton, 2001). Validation frequently focuses on O-D reconstruction the process of determining the demand scenario most likely responsible for the observed traffic counts.

When faced with multiple sets of traffic data (for example, from different days), a decision is required regarding the use of all available data. In congested networks, even minor changes in demand can have a significant impact on simulation results. As such, separate O-D matrices should be estimated for each data set and applied individually in the simulation model, with the resulting outputs compared against the actual observed data. Relying on average traffic counts to estimate average O-D flows may lead to a biased assessment of model performance since both the O-D estimator and the simulation model function nonlinearly. The degree of this bias depends on the level of nonlinearity and the variability in daily demand.

Model validation is based on how closely the performance measures (MOPs) from the real and simulated systems align. These MOPs should not be derived from data used for calibration or input estimation. Since O-D flows are typically estimated by reducing the difference between observed and simulated traffic counts, validating the simulation model solely on traffic counts could lead to an overestimation of the model's accuracy. For example, (Toledo & Koutsopoulos, 2004)observed that while the simulated flows in the INTEGRATION model closely matched real observations, the simulated speeds did not.

Figure 1

Validation process



2.7.4 Stability Analysis

Numerous research efforts have focused on the stability of both simple car-following models and acceleration models (Chang et al., 2020). As outlined by (Ward, 2009)a general car-following model can be mathematically represented by the equation:

$$\begin{cases} \dot{x}_n = v_n \\ a_n = f(\Delta x_n, \Delta v_n, v_n) \end{cases} \qquad \text{Eq 3.}$$

In this equation, x_n denotes the displacement of vehicle n, v_n represents the velocity of vehicle n, and a_n is the acceleration of vehicle n.

This framework has been used to analyze the string stability of mixed traffic flow, considering various rates of technology penetration and different maximum platoon sizes (Qin et al., 2018)

2.7.6 Gap Acceptance Analysis

typically comes into play at junctions where a minor road intersects a major road. Here, only the first vehicle in line (the leader) can attempt to make a try or irritated, contingent on the movement type. If the leading vehicle has just reached the junction, it might turn while still in motion. Else, it initiates the turn from a stationary position. In situations involving straightforward crossing or merging maneuvers, drivers on the minor street face a sequence of gaps between vehicles in the conflicting traffic flow on the major street. These gaps are formed due to varying arrival times of vehicles on the major street, resulting in gaps of different lengths. When encountering these gaps, drivers on the minor street make decisions based on certain behavioral factors. Each driver has an associated 'critical gap' – a minimum time gap they require to make their move safely. This critical gap varies among drivers and can typically be represented by an empirical distribution. When a gap larger than this critical gap presents itself, the driver decides to take it and proceed with their maneuver (Kaysi & Alam, 2000)

2.8 Existing Traffic Simulation Models

Table 1

Categorizes some well-known car fallowing model

modelling	Car-fallowing model	parameters	Advantage	disadvantage	reference
microscopic	Intelligent Driver Model (IDM)	Requested temporal gap, desired sensitivity to minimal spacing, acceleration exponent	1. The accessibility and simplicity of use of its parameters. 2.Accurate modeling, 3. Seamless traffic tracking 4. Record different traffic conditions	1. Responsive to parameter configuration, 2. Absence of assertive behaviormodeling, 3. Surpassing the actual vehicle deceleration.	(Derbel et al., 2013)
microscopic	Gipps' Model	Peak acceleration, Peak deceleration, Optimal separation between vehicles, The desired	1. An optimal balance between precision and the quantity of variables for calibration.	1. Not well-suited for intricate situations	(Cattin et al., 2018), (Peng, n.d.)

		velocity of unrestricted movement, the duration it takes to respond.	2. Ideal for integrating scenarios		
microscopic	Pipes' Model	Response time, velocity difference	Small number of parameters Ease of understanding and implementation Rooted in behavior	Inadequate Depth Constant Reaction Time Linear Feedback	(Pipes, 1953)
macroscopic	Lighthill-Whitham-Richards(LWR)	V_0 the desired speed. The density of traffic The flow of traffic	Coverage is extensive.	Inadequate granularity	(Lighthill & Whitham, 1955)
macroscopic	Greenshields Model	Traffic velocity The density of traffic	Foundational	Overly simplified	(Bureau & Highway, n.d.)

mesoscopic	Gas-kinetic Model	Groups of vehicles	It strikes a balance between detail and scalability.	Hybrid techniques are required.	(Nelson, 1998)
mesoscopic	Cellular Automata Model	Distinct vehicle locations	Simplicity	The unique character	(Nagel et al., 1992)

a. Intelligent Driver Model (IDM)

The equation

$$\dot{v}_n = a_n \cdot \left(1 - \left(\frac{v_n}{v_n^0}\right)^4 - \left(\frac{s^*(v_n, \Delta v_n)}{s_n}\right)^2\right), \qquad \text{Eq 4.}$$

describes a dynamic in traffic flow modeling, where:

where a_n is the maximum acceleration of the vehicle n measured in meters per second squared ($\mathrm{m.s}^{-2}$),

v_n^0 is the desired velocity of vehicle ($\mathrm{m.s}^{-1}$), s_n the distance gap(in meters)

$$s_n = \Delta x_n - l_{n+1} \qquad \text{Eq 5.}$$

The desired minimum gap of the vehicle n, s_n , is given by

$$s^*(v_n, \Delta v_n) = s_n^0 + T_n v_n - \frac{v_n \Delta v_n}{2\sqrt{a_n b_n}} \qquad \text{Eq 6.}$$

where s_n^0 is the jam distance for vehicle n(m), T_n the safety time gap for vehicle n(second) and $b_n$is the desired deceleration of the vehicle n ($\mathrm{m.s}^{-2}$).

This model accounts for various factors such as the vehicle's current speed, desired speed, distance to the vehicle ahead, and the driver's safety considerations, represented by the acceleration, deceleration, and time gap parameters.

b. Gipps Model

$$v_n(t+\tau_n) = min\begin{Bmatrix} v_n^a(t+\tau_n) \\ v_n^d(t+\tau_n) \end{Bmatrix}$$

$$\min\begin{cases} v_n(t) + 2.5a_n\tau_n\left(1 - \frac{v_n(t)}{V_n^{\text{des}}}\right)\sqrt{0.025 + \frac{v_n(t)}{V_n^{\text{des}}}}; \\ b_n\left(\frac{\tau_n}{2} + \theta\right) + \sqrt{b_n^2\left(\frac{\tau_n}{2} + \theta\right)^2 - b_n\left[2(x_{n-1}(t) - x_n(t) - S_{n-1}) - v_n(t)\tau_n - \frac{v_{n-1}(t)^2}{\hat{b}_{n-1}}\right]} \end{cases}$$

Eq 7.

$n-1$ refers to its leading vehicle

a_n is the maximum acceleration rate of the follower;

$S_{n-1} = \mathrm{l}_{n+1} +$ safety margin represents the length of the leader vehicle including a minimum safe distance

τ_n is the reaction time of the follower;

v_n^{des} is its desired speed

b_n is its maximum braking rate;

$\hat{b}_{n-1}$ is the assumed braking rate of the leader;

c. Lighthill-Whitham-Richards (LWR)

$$\frac{\partial k(t,x)}{\partial t} + \frac{\partial Q(t,x)}{\partial x} = 0 \qquad \text{Eq 8.}$$

$k(t,x)$ traffic density at time t and position x. The variable k typically denotes the number of vehicles per unit length at a given location and time

$Q(t,x)$ traffic flow rate at time t and position x. This is the product of traffic density and velocity.

d. Greenshields Model

$$S = S_F - \frac{S_F}{D_J} D \qquad \text{Eq 9.}$$

$S =$ Average Speed (km/hour)

S_F =Free Flow Speed(km/hour)

$D =$ Average Density(unit/km)

$D_J =$ Density jam(unit/km)

e. Gas-kinetic Model

$$\frac{\partial \mathrm{f}}{\partial \mathrm{t}} + v \frac{\partial \mathrm{f}}{\partial \mathrm{x}} + \frac{1}{\tau}\left(f - f_{eq}\right) = Q(f, f) \qquad \text{Eq 10.}$$

- $f(v, x, t)$ is the probability distribution function of vehicle speeds;
- τ is a relaxation time towards equilibrium;
- f_{eq} is the equilibrium distribution function;
- $Q(f, f)$is the binary interaction term (akin to the collision term in the Boltzmann equation), representing the interactions between vehicles;

f. Cellular Automata Models

The cellular automaton model is a microscopic approach to traffic modeling. In this method, a roadway is depicted as a sequence of cells, similar to points on a grid or squares on a checkerboard, with time being divided into discrete intervals. In this framework, vehicles transition from one cell to the next. (Nagel & Schreckenberg, 1992) were the pioneers in applying the Cellular Automaton model to traffic simulation. Their work involved simulating traffic flow on a single-lane highway using a stochastic Cellular Automaton (CA) model. A key principle of this model is that each vehicle advances 'v' cells at every time interval. The vehicle's velocity 'v' increments by 1 if there are no vehicles within 'v' spaces ahead, but it reduces to 'i-1' if another vehicle is located 'i' spaces ahead. Additionally, there's a random chance, represented by probability 'p', that the vehicle's velocity might decrease. Following this initial development, a variety of CA models have been created and implemented in traffic simulation studies (Ding, 2011).

g. Other Traffic Simulation Models

Simulation modeling is winning acceptability as a very useful and powerful tool to address complex transportation problems, which otherwise are too complex to be analytically modeled by the conventional methods of analysis. But a broad agreement in the venue of transportation simulation recognizes that microscopic simulation, affording such detailed computational analysis at the level of the individual traveler, is not just in the realm of possible approaches but all too often the only practical method toward a variety of intricate problems. This has been a catalyst to high completion microscopic simulation development models manufacture due to advancements in the computer technology, and great examples include AIMSUN, MITSIM, PARAMICS, and VISSIM.A typical microscopic traffic simulation model consists of a few physical elements that include a roadway network, traffic control systems, and driver-vehicle units. Also, these models encapsulate the representative behavioral characteristics such as driving behavior as well as route choice models. They necessitate intricate data inputs and involve numerous parameters. Though most simulators guide on input data and give default values of parameters, these models require calibration specific adapted to the network being analyzed and the purpose they intend to use the application as pointed by (Chu et al., 2003).

2.9 Special Topics

2.9.1 Speed of the car following vehicle

Influence on the speed of the vehicle and the need for the safe distance. Changes in the leading vehicle speed prompts adjustment in entering vehicle. Speed-flow and flow-density relationships show speed ranges between zero to a maximum free flow speed whilst density within zero to jam density. At maximum density, speed is zero and as traffic density decreases, the reliance of the speed on density lessens, though not necessarily the point the derivative of the prior was zero at zero density (Letters, 2013).

2.9.2 Acceleration of the car following vehicle

Acceleration is influenced by the behavior of the vehicle ahead, and fast adjustment in acceleration is very important to follow safe following distance as well as to respond efficiently in imitating speed changes that are happening with the leading vehicle. This is important in the prediction of the traffic flow and likely congestion. In an effort of realistic simulation of traffic flow in Intelligent Transportation Systems (ITS), a linear acceleration-based car-following model has been designed especially for real-time applications and systems as Autonomous Intelligent Cruise-Control Systems (AICCS). The linear acceleration model has some benefits: it guarantees that the vehicle acceleration profiles are continuous and hence that each time step's acceleration adds to that of the previous one; permits for the solutions that are continuous over time; and actually, the model represents driver behavior. For this reason, it is very suitable for continuity in acceleration which leads to car-following smooth behaviors devoid of sudden vehicle states changes, and hence the algorithm is very suitable for real-time applications.

2.9.3 Desired space of the car following vehicle

Desired speed is the speed at which the driver of a vehicle would wish to have travelled in the absence of any influences that would influence his choice of speed, with regard to obstacles or other traffic on the road. This preference includes personal preferences of driving, capabilities of the vehicle and conditions of the road. Drivers might change lanes or make other maneuvers if their actual speed falls below this desired speed. According to (Wilhelm & Lian, 2019), research notes that drivers adjust the speed with that of the leading car in thus maintaining a near-constant time headway referred to as "desired space." According to observation, this desired space is obtained by multiplying the speed by the ideal time headway (t), represents the interval a driver maintains following in steady-state car-following, hence an ability to decelerate at a similar rate applied to the leading vehicle and peak reaction time. The study findings revealed that the time headways between one vehicle and another, when following a lead vehicle at constant speeds of 50, 60 or 80 km/h, among different drivers were similar, irrespective of speed of the lead vehicle. Though time headways for each of the latter driving conditions appeared similar for individual drivers, there was some variability in time headway between drivers. This regularity in time headways implies that within the speed range studied (40 to 70 km/h), each driver's time headway is steady.

2.9.4 Human Factors in Traffic Modeling

Integrating human features into traffic modeling presents significant challenges. Yet, some studies have managed to include various human aspects in the context of operational and tactical driving maneuvers. These aspects are 1) risk-taking behavior, 2) willingness to cooperate, 3) capacity to learn, 4) levels of impatience, 5) aggressiveness, 6) susceptibility to distraction, 7) driving experience, and 8) the element of uncertainty (Hamdar et al., 2015). It is important to emphasize that the majority of current traffic flow models still fail to consider collisions, which highlights the critical role that human factors play in traffic accidents.

2.10 Connected vehicles

A Connected Vehicle or CV, is a purposely designed vehicle for transmitting information about the nearby traffic to the drivers through some procedures of communication which include vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) and vehicle-to-everything (V2X). V2X communication comprises of vehicles, infrastructure, pedestrians, as well as aftermarket devices that may cause an influence on the behavior of CVs. The main benefits of CVs are the increasing derivatives of traffic as well as reducing occurrences of both traffic congestion and accidents besides mitigating traffic emissions (Abbas et al., 2015). With time, improvements have been embraced, notably those involving connected automated vehicle. Such improvements aim at enhancing comfort and safety in driving through the introduction of vehicle automation technologies. However, initial research on CAV development has primarily focused on its automation and to some less extent attended to the influence of traffic flow dynamics through which the vehicular impacts were found to be beneficial and adverse as for dissipation and generations of vehicular flow. Recent studies on CAVs were more rigid in terms of the focus of vehicle control, however, some more recent ones broadened their focus away from individual vehicle control to more complex areas such as vehicle platooning and cooperative autonomous driving like coordinating during merging. This trend has largely been driven by developments in new connectivity and automation technologies that have enabled the investigations and development of feasible holistic strategies for vehicle control.

CHAPTER III

Methodology

3.1 Overview

In this chapter, we present the improvement made to car following models based on a machine learning technique. The data used in this study was obtained from of (Ciuffo et al., 2021), study towards comprehensive analysis of ACC adaptive cruise control for commercial vehicles. This database constitutes a precious dataset collection that was specially designed for car following simulations. The current study using in-depth analysis and applying the state-of-the-art machine learning techniques was expected to assist us to ameliorate car-following algorithms by amplifying the accuracy and efficiency.

3.2 Data Collection and Processing

This section of the research work deals with conducting the car-following experiment and organizing, collecting, analysing data from relevant resources. The car-following experiment was executed at the AstaZero test track in July 2019 with Oxford Technical Solutions for two days through high-alpha vehicles operative with their own RT-Range S system. Researchers analysed trajectory data from five different vehicles. This experiment was done from a 5.7 km rural road section, whereby through the use of the lead car that sustained a constant speed via Adaptive Cruise Control (ACC), different subsequent vehicles showed the car platoon sort of behaviours since they either sustained constant speeds or had variations in speed. Firstly, the data was collected at more than 100 Hz but reduced to 10 Hz for practical analysis. The data contained time, speed, latitude, longitude, altitude, ENU coordinates, relative distance between two vehicles and status of the driver (ACC or manual). The files were very aptly named and had metadata included within them giving information regarding the vehicle behaviour under different settings of control highlighting specially the contributions made by ACC systems on the dynamics of a vehicle in a rural road scenario. On the other hand, various key metrics such as time headway, acceleration, and inter-vehicle spacing were annotated into CSV files to ensure proper arrangement for storage as well as an easy accessibility.

3.3 Machine Learning Model

a machine learning model is built through an orderly process that starts with the collection of relevant data, then cleaning, and transformation for analysis. Then one can go ahead to select a proper algorithm based on the task at hand i.e., classification versus regression or clustering. A model is trained with some fraction of data so that it comes to learn about the patterns and relationships. Then, it is tested and validated over another subset of data to verify how efficiently the model is working. Now, a well-scoring model can be put into real-world practice where it applies its learned patterns to new data which never saw before and makes predictions or decisions about it. Almost always for the accuracy of model over time continuous monitoring as well as updating is needed since with constant flow of data in most real-world applications many things change out there.

3.3.2 Gradient boosting (GB)

GB is an ensemble model, which uses boosting techniques within the ensemble framework. Often individual decision trees are not accurate and robust. Typically, in ensemble methods, multiple decision trees get combined for enriching learning each case. In boosting, the sample weights of training set samples are adjusted with every iteration. The adjustment is such that more attention is directed to observations which together constitute a sample that is difficult to predict and less emphasis is put on those already predicted accurately. The model is a set of predictors (x) and attempts to estimate for the response variable through a function f(x), expressed as a cumulative series of functions (Friedman, 2001).

$$f(\boldsymbol{x}) = \sum_{k=1}^{K} f_k(\boldsymbol{x}) = \sum_{k=1}^{K} \boldsymbol{\beta}_k h(\boldsymbol{x}; \boldsymbol{a}_k) \qquad \text{Eq 11.}$$

The term "a_k" refers to the average value of the terminal nodes in the kth individual regression tree, whereas "β_k" indicates the weights assigned to the terminal nodes of the kth tree. The function h(.) represents the process of combining basis functions in an additive manner.

The parameters a_k and β_k undergo a gradual process of improvement and enhancement. Each iteration involves the addition of a new tree to minimize the objective function L(.) of the entire model. After t iterations, the model is described by the following equation, which allows for the calculation of the optimal β_t, as depicted in Equation (13).

$$f_t(\boldsymbol{x}) = f_{t-1}(\boldsymbol{x}) + \beta_t h(\boldsymbol{x}; \boldsymbol{a}_t) \qquad \text{Eq 12.}$$

$$\beta_t = \operatorname{argmin} \sum_{n=1}^{N} L\big(y_n, f_t(\boldsymbol{x}_n)\big) = \operatorname{argmin} \sum_{n=1}^{N} \big(y_n - f_t(\boldsymbol{x}_n)\big)^2 \qquad \text{Eq 13.}$$

where L(.) stands for the objective function used where in this case it is squared loss. N stands for total number of training observation with x_n and y_n representing predictor variables and their corresponding outcome variable at nth observation, respectively. To avert overfitting, the algorithm incorporates a learning rate, ζ, that controls the impact of each added tree. A smaller value of ζ leads the model to be more stable and it also reduces the risk of overfitting, but at a cost of more iterations for reaching a desired level of accuracy, therefore computation time is increased. The method known as shrinkage is introduced in Gradient Boosted Regression Trees (GBRT). This approach does not seek full optimization at each step but it allows incrementally improving the outputs over iterations as what was explained by (T. Wang et al., 2021).

$$f_t(\boldsymbol{x}) = f_{t-1}(\boldsymbol{x}) + \zeta \beta_t h(\boldsymbol{x}; \boldsymbol{a}_t), \zeta \in (0,1) \qquad \text{Eq 14.}$$

With increasing complexity in algorithms, there's often an improvement in accuracy at the expense of interpretability. However, tree-based models strike a balance as they can determine the importance of predictors and compute their partial dependence, as a result, the combination of accuracy and interpretability is provided.

3.3.3 Random Forest

The Random Forest regression technique involves a large number of decision trees, each acting independently as a model for regression. The final outcome in such setups is akin to taking the average of all these individual regression trees' outputs. This method is an extension of the decision tree model, specifically the Classification and Regression Tree (CART), initially proposed by (Breiman, 1984). RF Regression, as an advanced version of CART, offers improved predictive capabilities. During the training phase of RF, a collection of decision trees is generated, each functioning independently from the others. The term "Random Forest" refers to the method's approach of building each tree with a subset chosen at random of forecasters.

3.4 Evaluation of model performance

Performance of the models is on four major metrics. Root Mean Squared Error (RMSE) is an accuracy metric expressed by taking the square root of average squared differences between predicted values and actual values, a low value of RMSE means more accuracy lying in a model. The (MAE) Mean Absolute Error measures the error between the values predicted by a model and the actual values, hence delivering an understandable measure of how far off the predictions are from real observations. R-Squared R^2 is a statistical metrical aiming to culminate the quantification of the proportion of the variance in the dependent variable which can be predicted from independent variables, taking value between 0 (no explanatory power) and 1 (perfect fit), hence, depict the quality through which the fit quality is through the model. Lastly, the Nash-Sutcliffe Efficiency (NSE) measures the model's predictive ability by summarizing the utility of the instances predicted against those observed, with a value of 1 suggesting perfect prediction and values less than zero implying that the model is not as accurate as a simple average over this collection. These metrics collectively provide a comprehensive picture of the model's predictive performance and accuracy as stated by (Chai & Draxler, 2014).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(Y_{\text{Pred},i} - Y_{\text{Obs},i}\right)^2}{N}} \qquad \text{Eq 15.}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| Y_{\text{Obs},i} - Y_{\text{Pred},i} \right| \qquad \text{Eq 16.}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n} \left(Y_{\text{Obs},i} - Y_{\text{Pred},i}\right)^2}{\sum_{i=1}^{n} \left(Y_{\text{Obs},i} - \bar{Y}_{\text{Obs}}\right)^2} \text{y} \qquad \text{Eq 17.}$$

$$R^2 = \frac{\text{SSR}}{SST} \qquad \text{Eq 18.}$$

Where SSR=$\sum\left(Y_{\text{Pred},i} - \bar{Y}\right)^2$, $SST = \sum\left(Y_{\text{Obs},i} - \bar{Y}\right)^2$, $\bar{Y}$ is the mean of y value; N = number of observed value, Y_{Predi} predicted value Y_{Obi} =observed value.

CHAPTER V

Results and Findings

4.1 Introduction

This chapter presents the results of predicting the behavior of following vehicles within a platoon. The data have been divided into four categories: one-platoon, two-platoon, three-platoon, and four-platoon scenarios. This classification is essential in predicting vehicle acceleration behavior and speed motion in various platoon scenarios.

4.2 predication of following vehicle behavior with platoon vehicle

a. Vehicle Acceleration behavior

Figure 2

Comparison of Actual vs. Predicted Acceleration Over Time for one platoon Test Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

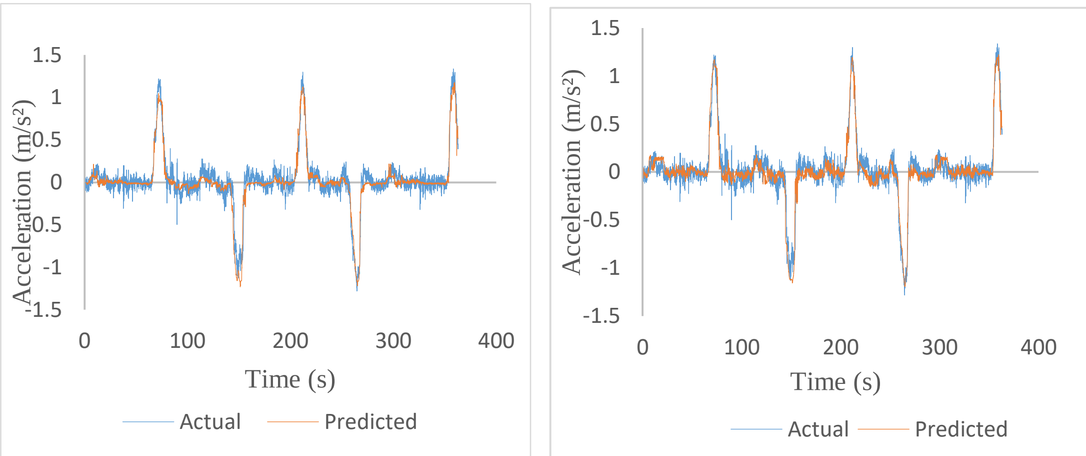


Figure 3

Comparison of Actual vs. Predicted Acceleration Over Time for a one Platoon Train Using(a) Gradient Boosting and (b) Random forest

(a) ***(b)***

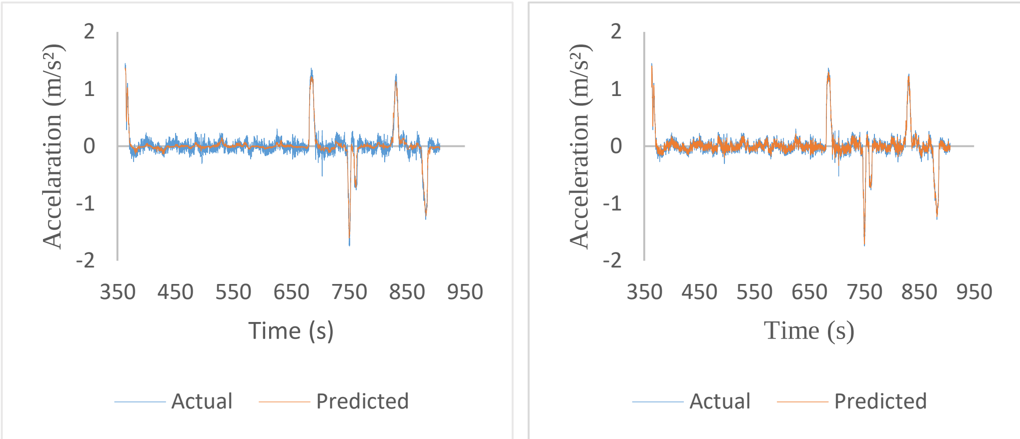


Figure 4

Comparison of Actual vs. Predicted Acceleration Over Time for two platoon Test Using (a) Gradient Boosting and (b)Random forest

(a) ***(b)***

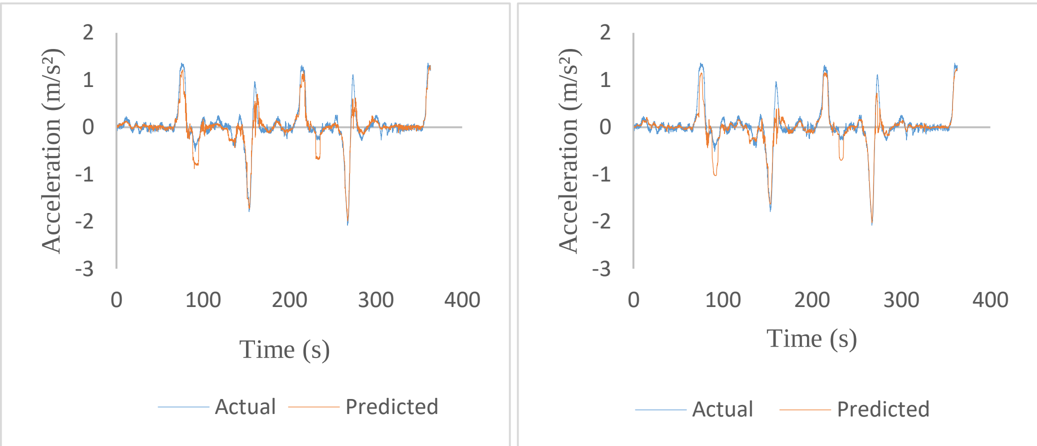
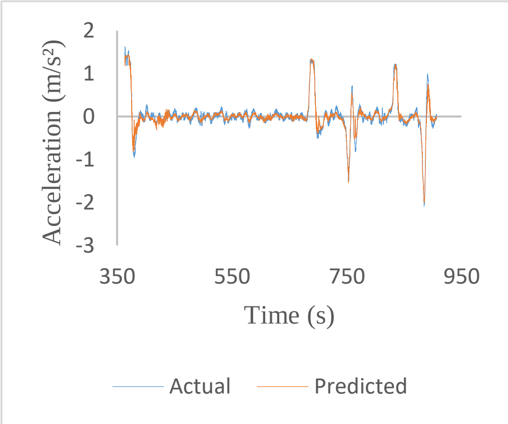


Figure 5

Comparison of Actual vs. Predicted Acceleration Over Time for a two Platoon Train Using(a) Gradient Boosting and (b) Random forest

(a)



(b)

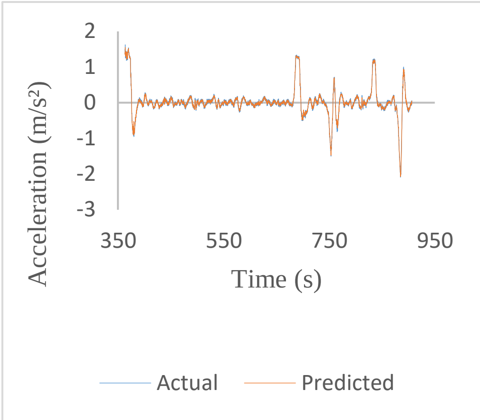
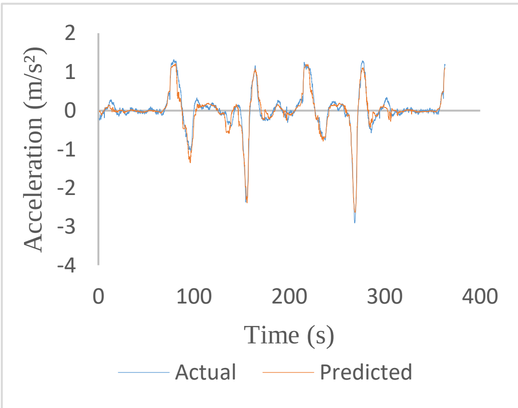


Figure 6

Comparison of Actual vs. Predicted Acceleration Over Time for three platoon Test Using (a) Gradient Boosting and (b) Random forest

(a)



(b)

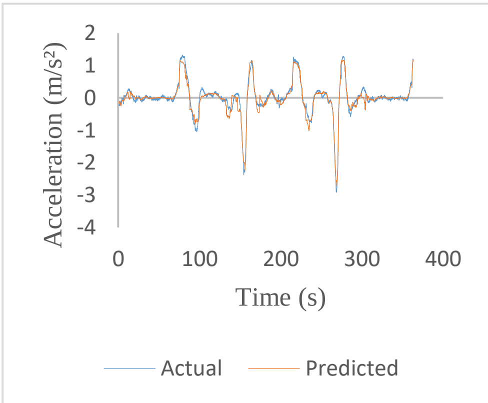


Figure 7

Comparison of Actual vs. Predicted Acceleration Over Time for a three Platoon Train Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

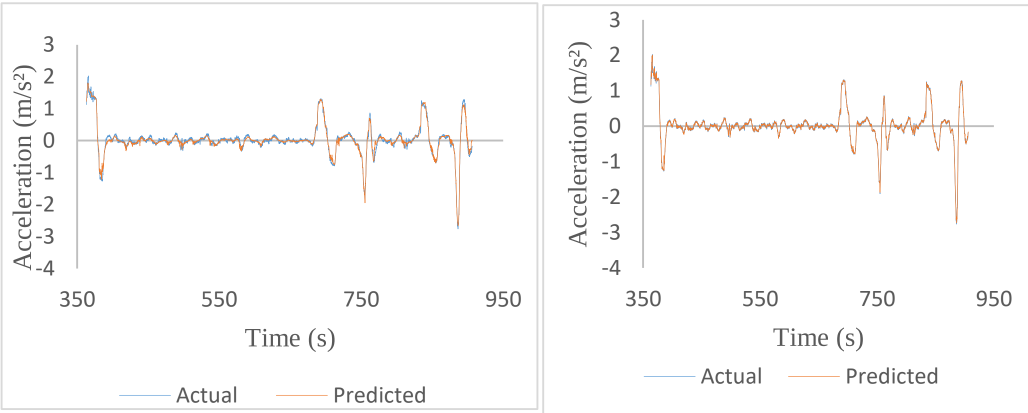


Figure 8

Comparison of Actual vs. Predicted Acceleration Over Time for four platoon Test Using (a) Gradient Boosting and (b)Random forest

(a) ***(b)***

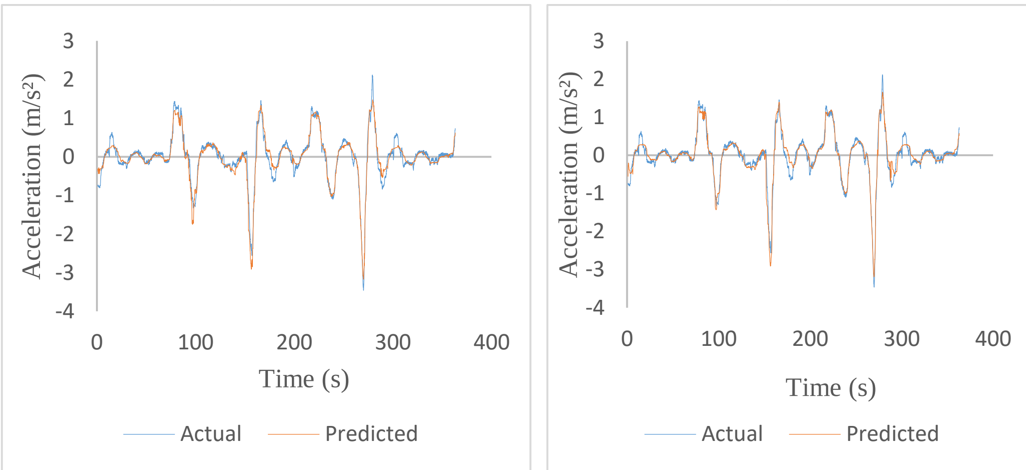


Figure 9

Comparison of Actual vs. Predicted Acceleration Over Time for a four Platoon Train Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

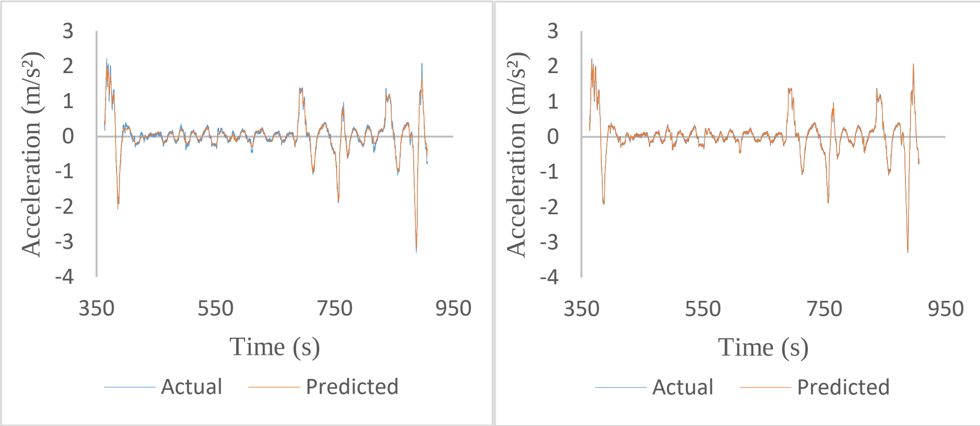


Table 2

Output train and test using vehicle Acceleration behavior for Gradient Boosting

Models	Train				Test			
	R^2	RMSE	MAE	NSE	R^2	RMSE	MAE	NSE
1 platoon GB	0.92	0.08	0.06	0.92	0.91	0.1	0.07	0.91
2 platoon GB	0.95	0.09	0.06	0.94	0.87	0.16	0.1	0.866
3 platoon GB	0.97	0.08	0.06	0.97	0.94	0.14	0.09	0.93
4 platoon GB	0.98	0.08	0.05	0.98	0.91	0.1	0.07	0.91

The performance evaluation of the Gradient Boosting (GB) model for vehicle acceleration behavior across four distinct platoon scenarios reveals a consistent trend of high accuracy and predictive reliability in the training phase. The model achieves R^2 values between 0.92 and 0.98 across the four platoons, indicating a strong correlation between predicted and actual values. The low RMSE (ranging from 0.08 to 0.09) and MAE (consistent at 0.05 to 0.06) in the training phase further attest to the model's precision. Notably, the Nash-Sutcliffe Efficiency (NSE) mirrors that of the R^2 values thereby emphasizing effectiveness of the model in capturing observed data variability. However, when applied to test data, there's a noticeable drop in the performance, more so with the 2nd platoon where R^2 reduces to 0.87 and RMSE increases to 0.16. This pattern suggests a possible overfitting problem common to complex models where the model seems to perform entirely well on training data but unimpressive in the case of unseen data. However, the test R^2 values remain above 0.87 and NSE values that are consistent with the R^2 such that the model still enjoys substantial forecasting power in new scenarios even as its accuracy is now slightly reduced. The analysis highlights the GB model potent learning and predictive trait of various vehicle acceleration behaviors in platoon scenarios with quarters re assessments over its likely adaptation in independent or more complex environments.

Table 3

Output train and test using vehicle Acceleration behavior for Random Forest

Models	Train				Test			
	R^2	RMSE	MAE	NSE	R^2	RMSE	MAE	NSE
1 platoon RF	0.99	0.03	0.02	0.99	0.91	0.1	0.08	0.91
2 platoon RF	0.99	0.02	0.01	1	0.83	0.19	0.1	0.82
3 platoon RF	1	0.02	0.01	1	0.94	0.13	0.09	0.94
4 platoon RF	1	0.02	0.01	1	0.94	0.16	0.11	0.94

The performance of the Random Forest (RF) model in predicting vehicle acceleration behavior across four platoon scenarios shows remarkable results in the training phase but reveals variability in the testing phase. In training, the model achieves near-perfect performance with R^2 and NSE values reaching 1.0 in the 3rd and 4th platoons, and 0.99 in the 1st and 2nd platoons, indicating an almost perfect fit to the training data. The RMSE and MAE are exceptionally low (as low as 0.01 and 0.02), further demonstrating the model's accuracy in this phase. However, the testing phase performance exhibits a notable decline, especially in the 2nd platoon, where the R^2 drops to 0.83 and RMSE increases to 0.19. While the R2 values in the testing phase for the 1st, 3rd, and 4th platoons remain high (above 0.90), indicating good predictive ability, the increase in RMSE and MAE suggests a decrease in prediction accuracy for new data. This pattern is indicative of overfitting in the training phase, where the model is highly tuned to the training data, potentially at the expense of its ability to generalize to unseen data. Despite this, the overall high R^2 values in the test phase demonstrate the RF model's robustness in predicting vehicle acceleration behaviors across different platoon scenarios, though caution is advised when applying the model in practical situations due to potential overfitting concerns.

b. Vehicle speed behavior

Figure 10

Comparison of Actual vs. Predicted Speed Over Time for a one Platoon Test Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

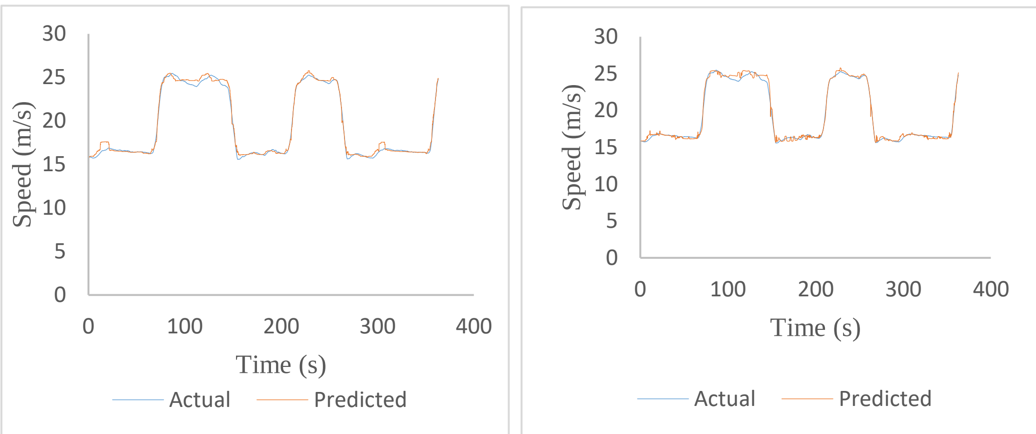


Figure 11

Comparison of Actual vs. Predicted Speed Over Time for a one Platoon Train Using Gradient Boosting and Random forest

(a) ***(b)***

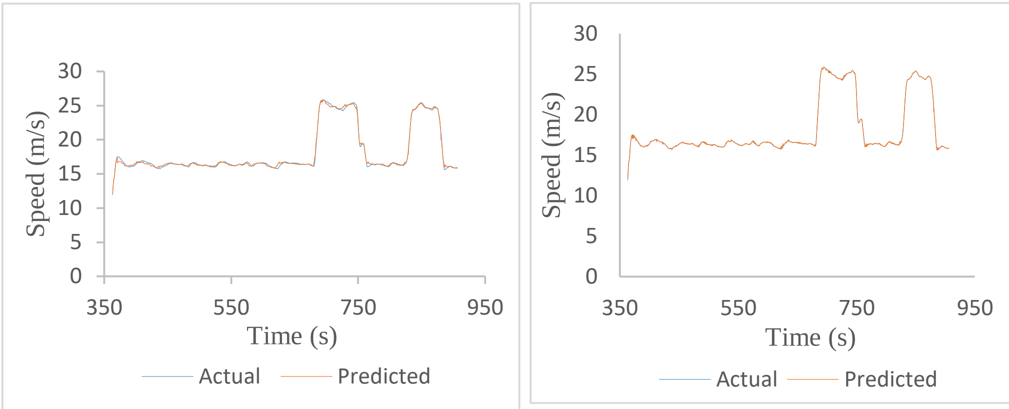


Figure 12

Comparison of Actual vs. Predicted Speed Over Time for a two Platoon Test Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

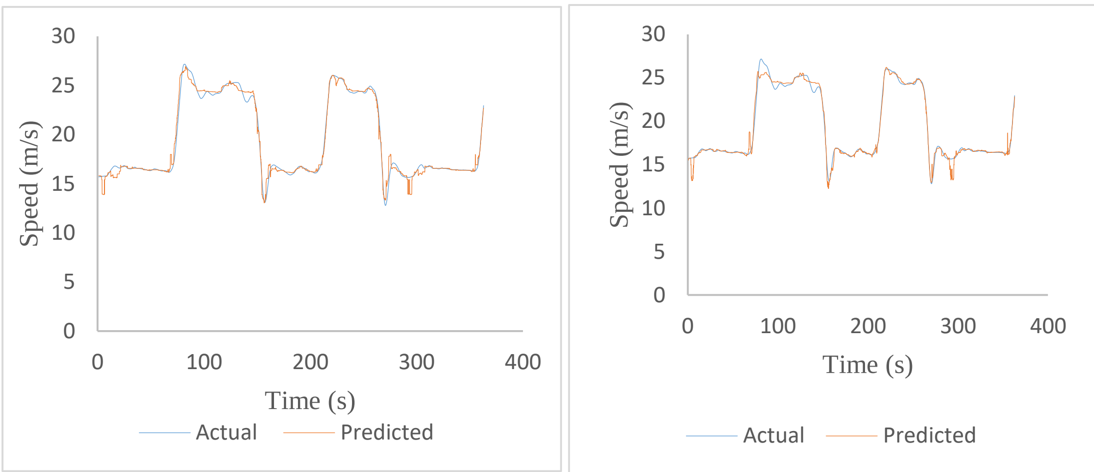


Figure 13

Comparison of Actual vs. Predicted Speed Over Time for a two Platoon Train Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

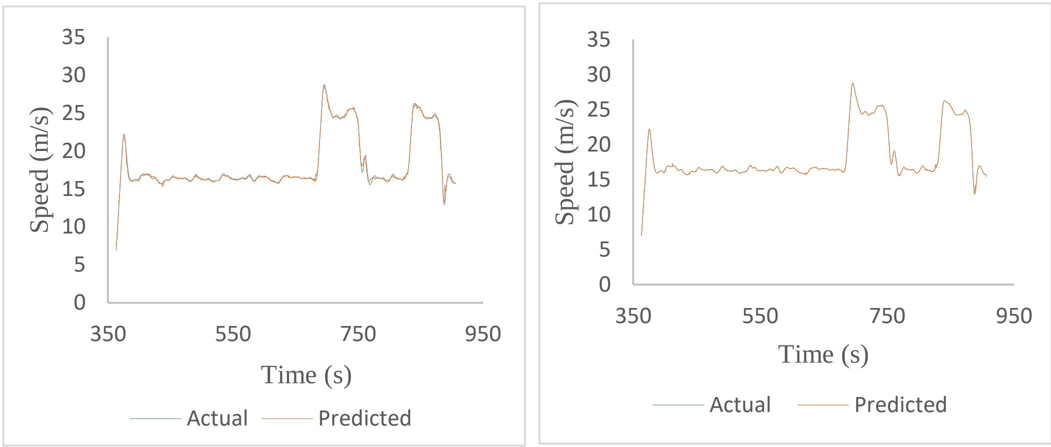


Figure 14

Comparison of Actual vs. Predicted Speed Over Time for a three Platoon Test Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

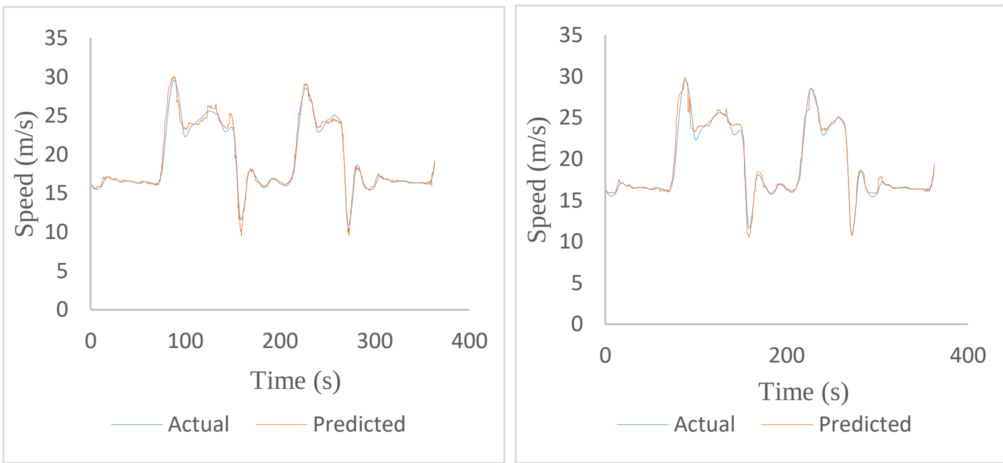


Figure 15

Comparison of Actual vs. Predicted Speed Over Time for a three Platoon Train Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

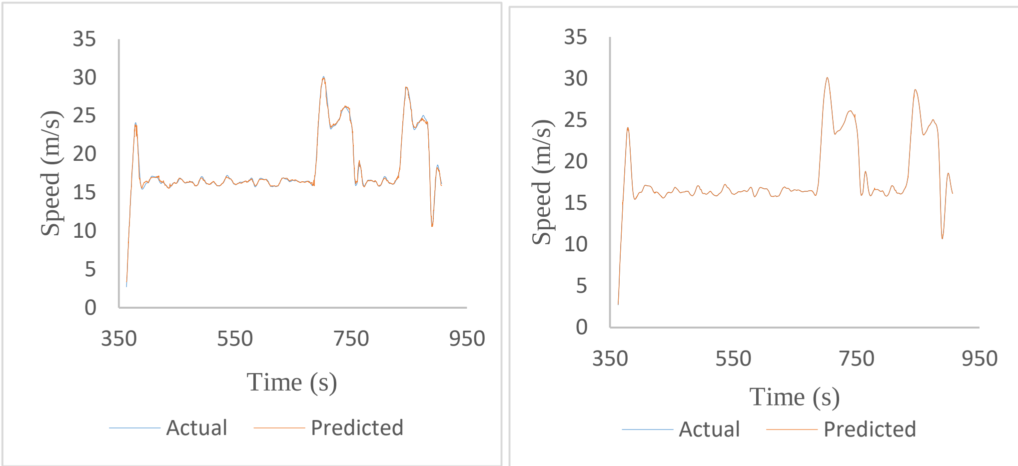


Figure 16

Comparison of Actual vs. Predicted Speed Over Time for a four Platoon Test Using (a) Gradient Boosting and (b) Random forest

(a) ***(b)***

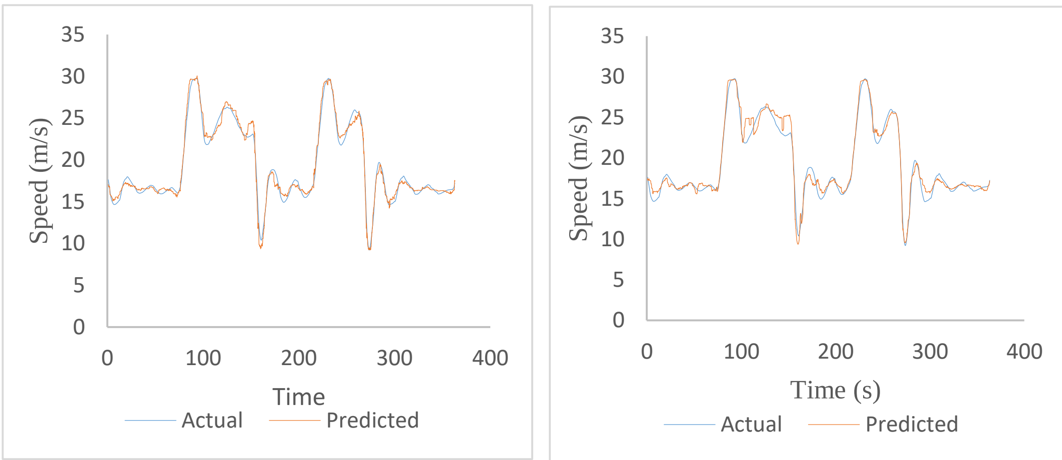


Figure 17

Comparison of Actual vs. Predicted Speed Over Time for a four Platoon Train Using (a) Gradient Boosting and (b) Random forest

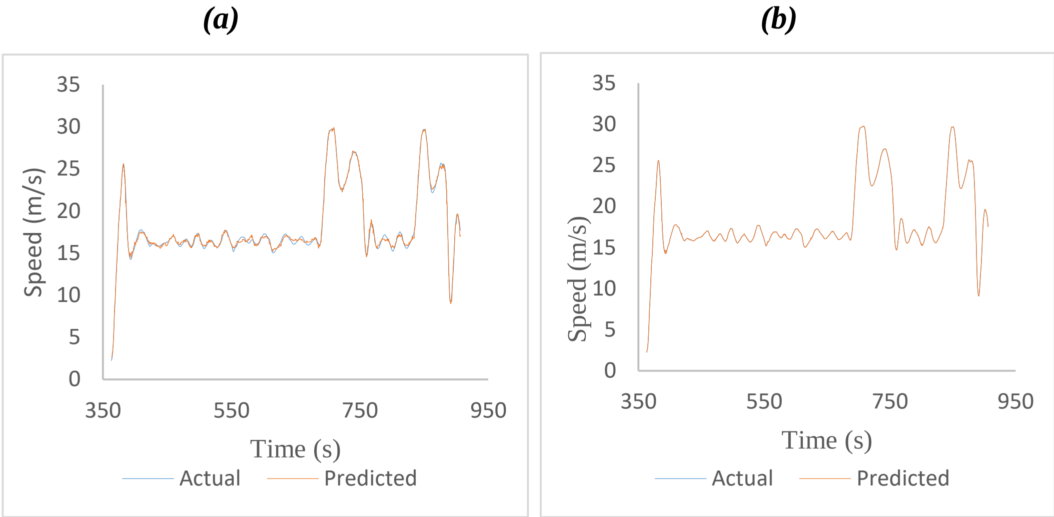


Table 4

Output train and test using vehicle Speed behavior for Gradient Boosting

Models	Train				Test			
	R^2	RMSE	MAE	NSE	R^2	RMSE	MAE	NSE
1 platoon GB	0.997	0.2	0.14	1	0.99	0.35	0.25	0.99
2 platoon GB	0.997	0.21	0.13	1	0.98	0.52	0.27	0.98
3 platoon GB	0.997	0.22	0.14	1	0.98	0.57	0.33	0.98
4 platoon GB	0.996	0.27	0.21	0.99	0.98	0.61	0.48	0.98

The Gradient Boosting (GB) model's performance in modeling vehicle speed behavior across four platoon scenarios demonstrates exceptional accuracy in the training phase, with a slight decrease in precision during the testing phase. In the training phase, the model achieves nearly perfect R^2 values, ranging from 0.996 to 0.997, and Nash-Sutcliffe Efficiency (NSE) scores are either perfect or nearly so, indicating an almost flawless fit to the training data. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), although relatively low, exhibit a gradual increase as the platoon scenario complexity escalates.

In contrast, during the testing phase, although the model maintains high R^2 values (ranging from 0.98 to 0.99), which indicates a strong predictive ability, there is a noticeable increase in RMSE and MAE. This trend is more apparent in more complex platoon scenarios, particularly in the 4th platoon, where the highest RMSE and MAE are recorded. The rise of error metrics over the test phase may suggest that this model, though being highly accurate over training data, does not generalize as well to new, unseen data. This potential overfitting issue means the model is wonderfully well tuned to training data losing some accuracy in a new scenario prediction. On the other hand, this result demonstrated the robustness of the GB model in capturing vehicle speed behavior across all platoons with constantly high R^2 and NSE values in the test phase.

Table 5

Output train and test using vehicle Speed behavior for random forest

Models	Train				Test			
	R^2	RMSE	MAE	NSE	R^2	RMSE	MAE	NSE
1 platoon RF	1	0.05	0.02	1	0.99	0.35	0.25	0.99
2 platoon RF	1	0.04	0.01	1	0.98	0.53	0.29	0.98

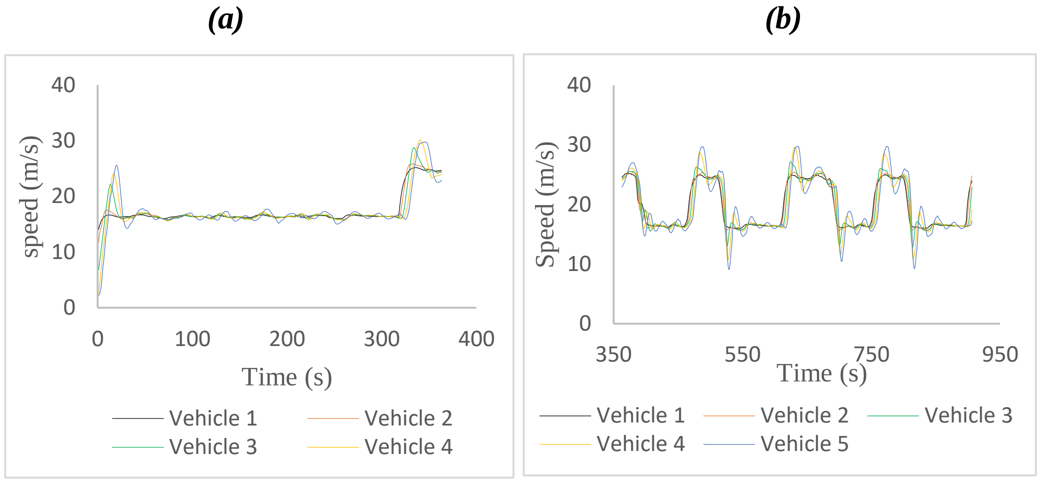
3 platoon RF	1	0.03	0.01	1	0.98	0.61	0.35	0.98
platoon RF	1	0.02	0.01	1	0.97	0.84	0.61	0.97

The results in terms of the vehicle's speed behavior prediction performance of Random Forest (RF) model across four different platoon scenarios are exemplary as showcased in case of training phase but exhibits some variability in the testing phase. In-training phase., the model achieves optimal performance, with R^2 and Nash-Sutcliffe Efficiency (NSE) values reaching a perfect score of 1.0, and exceptionally low RMSE and MAE, reflecting an almost ideal alignment with the training data. This level of accuracy is consistent across all platoons, showcasing the model's precision in the training context.

However, when evaluated in the testing phase, there is a noticeable increase in RMSE and MAE, indicating a decline in prediction accuracy for new data. While the R^2 values remain high (between 0.97 and 0.99) across all platoons, suggesting a strong ability to predict vehicle speed behavior, the increased error metrics in the test phase, especially in more complex scenarios like the 4th platoon, point to potential overfitting during training. This overfitting implies that the model, while highly tuned to the training data, may not generalize as effectively to unseen data. Despite this, the consistently high R^2 values in the test phase affirm the RF model's robust predictive capabilities across different platoon scenarios.

Figure 18

Speed profile for a platoon of five vehicles over time with original on (a) test and (b) train data



The figure shows the speed profiles for a platoon of five vehicles equipped with Adaptive Cruise Control (ACC) even before facing slope-induced changes, the platoon was not in a steady-state condition. The black line indicates the platoon leader, who applies cruise control to sustain a consistent pace. However, the following vehicles show significant fluctuations in speed, failing to achieve a constant velocity. This implies that alterations in slope have a substantial influence on the dynamics of the vehicles. The fluctuations in velocity also indicate the influence of string instability, whereby the variations in speed of the leader are magnified in the subsequent vehicles.

The observed dynamics arise from the combined influence of road geometry and string instability, posing difficulties in maintaining stable vehicle velocities inside an ACC system.

Figure 19

Speed Profile of a Platoon of Five Vehicles Over Time with (a) Gradient Boosting and (b) random forest Predictions on Test Data

(a) ***(b)***

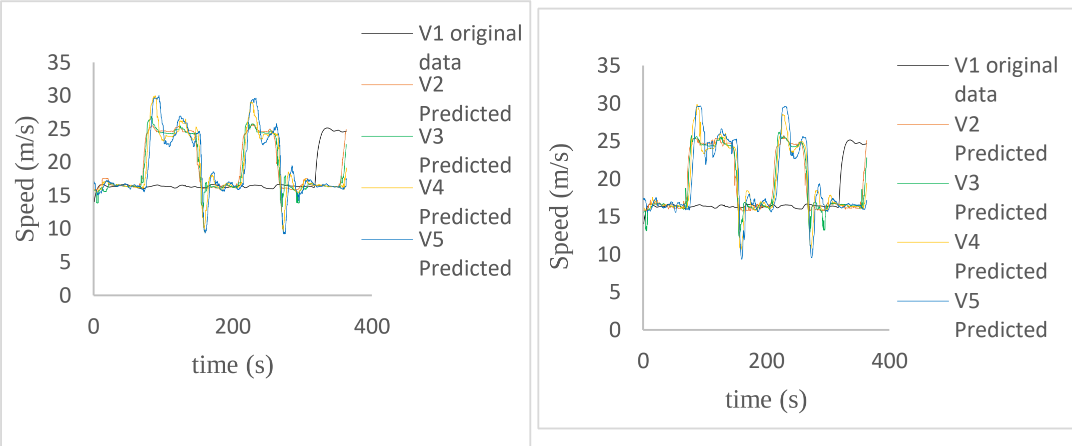
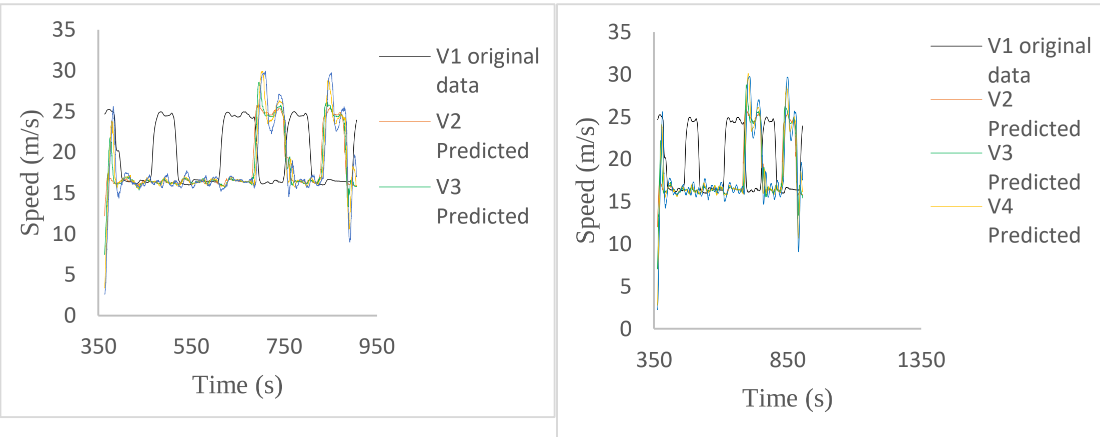


Figure 20

Speed Profile of a Platoon of Five Vehicles Over Time with (a) Gradient Boosting and (b) random forest Predictions on Train Data

(a) ***(b)***



CHAPTER VI

Discussion

5.1 Introduction

In this regard, it has been noted that the present study has paved a great pathway to insights regarding the application of advanced machine learning models through the associated connected vehicles in improving road safety and traffic dynamics within diverse platoon scenarios. By categorizing data into one, two, three and four-platoon scenarios, the research thus successfully analyzed vehicle acceleration and speed motion behaviors and pointed out the complexity nature of vehicle interaction in platoon formations. In addition, the application of Gradient Boosting (GB) and Random Forest (RF) models to predicting acceleration and speed behavior of a vehicle had its merits but also came with several shortcomings in the application of these models. In the area of vehicle acceleration under GB model training cases, remarkably high accuracies were observed in all platoon scenarios with R^2 within the range 0.92-0.98. However, a decrease in performance observed in testing with respect to the 2nd platoon scenario for which can be said that there might be an overfitting problem. Similarly, RF model indicated almost perfect performance in training mode but showed a certain extent of variance from testing mode especially in 2nd platoon scenario where the R^2 values suddenly dropped. Again, for vehicle speed behavior, the GB model elicited exceptional accuracy during training phase where R^2 values almost achieved 1.

However, during testing phase, minimal increments in error metrics were realized albeit maintaining a strong predictive ability. For behavior to speed, best characterized by training phase perfect R^2 and NSE score RF model. Though testing phase error metrics increased, it could be inferred that its predictive ability deteriorated for new data -that is mostly more complex platoon scenarios, when analyzing a case study of five vehicles, the platoon found that there were significant speed fluctuations among these following vehicles. In other words, this indicates the impact of string instability and road geometry in maintaining stable velocities within an Adaptive Cruise Control system before changes caused by slope. These fluctuations indicate the challenges which a steady state condition has in platoon dynamics. This is what led to the highlights on the importance of connected vehicles in increasing road safety. The posed accuracy levels by the vehicle behavior prediction machine learning models, particularly in a platoon extremely, indicate that through connected vehicles, there would be a remarkable reduction in accidents. Vehicles fitted with such sensors and communication technologies to facilitate information exchange in real-time for adequate decision-making become very critical in pre-emptive adjustments to traffic conditions and accident-reducing modalities. Summing it up, this research successfully demonstrated a complete and nuanced view of vehicle behaviors in platoon scenarios with the help of some very effective advanced machine learning models. On one hand, it showcases the efficacy of these very models, and on the other, it points out the thorough validation of these models along with minor adjustments required in practical application. This clearly establishes the promise of connected vehicles, upheld through advanced machine learning technologies, to enhance traffic safety and management strategies, thereby creating meaningful areas for research consideration as well as implementation possibility in intelligent transportation systems.

CHAPTER VI

Conclusion and recommendation

6.1 Conclusion

The research undertook an explorative pathway to enrich car-following models with advanced machine learning techniques that provide profound insight into their application, as well as connected vehicles, to enhance road safety and traffic dynamics within a major urban arterial. The research analysis sought to look at the acceleration and speed motion behaviors among the vehicles within diverse platoon scenarios thus, making it possible to identify the vehicle interaction behavior within the underlying platoon formation.

The critical time-series of simulation provided information in detail relating to the dynamic behavior. While the specific nature of the data (e.g., speed, acceleration) isn't explicitly specified, this relatively structured format perhaps is meant to indicate that the models were successful at capturing the temporal dynamics associated with multiple vehicles. The results indicate in detail and comprehensiveness that all the tested and training phases of the acceleration and speed prediction model running over different conditions and scenarios suggest extensive exploration. Intricate graphs and figures on results signified great comprehension and representation of the vehicle dynamics involved. However, complexity and depth of data posed some difficulties in extracting specific results and figures.

Gradient Boosting (GB) as well as Random Forest (RF) model was applied to an acceleration of prediction as well as the speed behavior of vehicles, exhibiting their merits and also some flaw. All platoon scenarios with GB model training cases showed a remarkably high accuracy with R^2 values in the range of 0.92-0.98. However, discouraging trends were observed as far as testing is concerned since they exhibited a declining trend through all the platoon scenarios and that too from a 1st platoon scenario to 2nd one which his highly indicative of overfitting. The RF model, though ensured a certain extent of variance from testing mode, also practically exhibited very high performance in training mode especially in the case of 2nd platoon where R^2 values suddenly took a drop back.

In the training phase, where R^2 values almost approached 1, the GB model proved for great accuracy for vehicle speed behavior. However, in the test phase, little increases in error metrics were observed while maintaining an excellent predictive capability. The RF model best characterized by perfect R^2 and NSE scores was the error metrics of the training phase who increased in the testing phase indicating the prediction capability of the RF may deteriorate for new data, mostly in the more complex platoon scenarios.

this study for five vehicles in a platoon showed that diminution of progressive declines in speed was noticeable while the lead vehicle was able to sustain normal speed with significant fluctuations in the following vehicles indicating the impact of string instability and road geometry to maintain stable velocities within an Adaptive Cruise Control system before changes caused by slope. These fluctuations underscore challenges in achieving a steady-state condition in platoon dynamics.

6.2 Recommendation

- Enhanced Data Interpretation and Model Validation: This should involve intensification as well as deepening of data analysis on specific metrics and model comparisons. Cross-validation, regularization, or diversifying the dataset can enhance the models' generalizability and reliability.

- Inclusion of Contextual Data: Incorporation of other data in these models could highly improve them, as the prediction accuracy would raise highly. For example, weather conditions, driver behaviors or even types of roads are some contextual data that can improve these predictions. Including or understanding the factors in context gives a clear understanding of how adaptable the models can be to almost real-world situations.

- Substantial Real-world Testing: Before being applied to practical use, these models must undergo substantial testing in various environmental settings. This is necessary so that the model can demonstrate and prove its robustness, coherence, and adaptability to real-world complexities and uncertainties.

- On Going Research and Development: Machine learning has gained so much pace and its applications in intelligent transportation systems that even today, many open research questions exist in the area. In the future, new models, algorithms, and optimization techniques will have to be considered by the researchers for handling emerging challenges within the same domain. These recommendations may provide valuable insights for future research to contribute to developing safer, more efficient and intelligent transportation systems in efforts towards integrating connected vehicles, advanced machine learning technologies, and the potential that both fields yield.

REFERENCES

Abbas, M. M., Trani, A. A., & Wang, L. (2015). *Assessment of Vehicle-to-Vehicle Communication based Applications in an Urban Network Taehyoung Kim Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree .*

Addas, A. (2021). *The Impact of COVID-19 Lockdowns on Air Quality — A Global Review*.

Aftabuzzaman, M. (2007). Measuring Traffic Congestion- A Critical Review 30 th Australasian Transport Research Forum 30 th Australasian Transport Research Forum. *30th Australasian Transport Research Forum*, *February*, 1–16.

Albayrak, G. (2023). Optimizing of Discrete Time-cost Trade-off Problem in Construction Projects Using Advanced Jaya Algorithm. *Periodica Polytechnica Civil Engineering*, *67*(3), 806–818. https://doi.org/10.3311/PPci.22156

Aramrattana, M., Habibovic, A., & Englund, C. (2021). Safety and experience of other drivers while interacting with automated vehicle platoons. *Transportation Research Interdisciplinary Perspectives*, *10*, 100381. https://doi.org/10.1016/j.trip.2021.100381

Arvin, R., Khattak, A. J., Kamrani, M., & Rio-torres, J. (2020). Safety evaluation of connected and automated vehicles in mixed traffic with conventional vehicles at intersections. *Journal of Intelligent Transportation Systems*, *0*(0), 1–18. https://doi.org/10.1080/15472450.2020.1834392

Asghar, A., Mirjalili, S., Faris, H., & Aljarah, I. (2019). Harris hawks optimization : Algorithm and applications. *Future Generation Computer Systems*, *97*, 849–872. https://doi.org/10.1016/j.future.2019.02.028

Aslani, M., Seipel, S., Mesgari, M. S., & Wiering, M. (2018). Traffic signal optimization through discrete and continuous reinforcement learning with robustness analysis in downtown Tehran. *Advanced Engineering Informatics*, *38*(August), 639–655. https://doi.org/10.1016/j.aei.2018.08.002

Atasoy, B., & Akkinepally, A. (2019). *MIT Open Access Articles Dynamic Toll*

Pricing using Dynamic Traffic Assignment System with Online Calibration. *10*(October).

Bartin, B., Ozbay, K., Yanmaz-tuzel, O., & List, G. (2005). *Modeling and Simulation of*. *October*. https://doi.org/10.1109/ITSC.2005.1520079

Bigazzi, A. Y., Figliozzi, M. A., Clifton, K. J., Bigazzi, A. Y., Figliozzi, M. A., Traffic, K. J. C., Bigazzi, A. Y., Figliozzi, M. A., & Clifton, K. J. (2015). Traffic Congestion and Air Pollution Exposure for Motorists : Comparing Exposure Duration and Intensity Traffic Congestion and Air Pollution Exposure for Motorists : Comparing Exposure Duration and Intensity. *UJST*, *9*(7), 443–456. https://doi.org/10.1080/15568318.2013.805345

Boluma Mangata, B., Landry Gilgen, M., Tene Patience Ryan, T., Makwem Hanse, T., & Oshasha Fiston, O. (2022). Road traffic analysis on the congestion problem using the Ford-Fulkerson algorithm. *Put i Saobraćaj*, *68*(4), 19–25. https://doi.org/10.31075/pis.68.04.03

Brackstone, M., & Mcdonald, M. (2000). *Car-following : a historical review*. *2*(1999), 181–196.

Breiman, L. (1984). *No TitleClassification and Regression Trees (1st ed.). Routledge.* https://doi.org/10.1201/9781315139470

Bureau, T., & Highway, O. S. (n.d.). *A study of t r a f f i c capacity*. *10*.

Cai, C., Jia, C., Nie, Y., Zhang, J., & Li, L. (2023). *A path planning method using modified harris hawks optimization algorithm for mobile robots*. *2018*. https://doi.org/10.7717/peerj-cs.1473

Campolongo, F., Saltelli, A., & Cariboni, J. (2011). From screening to quantitative sensitivity analysis . A unified approach. *Computer Physics Communications*, *182*(4), 978–988. https://doi.org/10.1016/j.cpc.2010.12.039

Catalin, C., Jorge, B., Genevieve, D., & Aziz, N. (2022). Bond Graph and Computational Fluid Dynamics in Traffic Flow. *2nd International Conference on Systems and Computer Science*, *8*, 246–251. https://doi.org/10.1109/IcConSCS.2013.6632055

Cattin, J., Leclercq, L., Pereyron, F., & Faouzi, N. El. (2018). *Calibration of Gipps ’* *car-following model for trucks and the impacts on fuel consumption estimation*. 367–375. https://doi.org/10.1049/iet-its.2018.5303

Celikoglu, H. B. (2013). *An Approach to Dynamic Classification of Traffic Flow Patterns*. *28*, 273–288. https://doi.org/10.1111/j.1467-8667.2012.00792.x

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250.

Chang, X., Li, H., Rong, J., Zhao, X., & Li, A. (2020). Analysis on traffic stability and capacity for mixed traffic flow with platoons of intelligent connected vehicles. *Physica A: Statistical Mechanics and Its Applications*, *557*, 124829. https://doi.org/10.1016/j.physa.2020.124829

Chen, D., Laval, J. A., Ahn, S., & Zheng, Z. (2012). Microscopic traffic hysteresis in traffic oscillations: A behavioral perspective. *Transportation Research Part B: Methodological*, *46*(10), 1440–1453. https://doi.org/10.1016/j.trb.2012.07.002

Chen, Y., Wenjuan, E., Wang, X., Wan, Q., Wang, C., & Yang, N. (2022). Multi-vehicle Cooperative Merging Control Strategy for Expressway under New Mixed Traffic Environment. *2022 IEEE 7th International Conference on Intelligent Transportation Engineering, ICITE 2022*, 603–608. https://doi.org/10.1109/ICITE56321.2022.10101440

Chu, L., Liu, H. X., Oh, J. S., & Recker, W. (2003). A calibration procedure for microscopic traffic simulation. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2*, 1574–1579. https://doi.org/10.1109/ITSC.2003.1252749

Ciuffo, B., Mattas, K., Makridis, M., Albano, G., Anesiadou, A., He, Y., Josvai, S., Komnos, D., Pataki, M., Vass, S., & Szalay, Z. (2021). Requiem on the positive effects of commercial adaptive cruise control on motorway traffic and recommendations for future automated driving systems. *Transportation Research Part C: Emerging Technologies*, *130*(July 2020). https://doi.org/10.1016/j.trc.2021.103305

Clark, H. (2020). Abusing language. *AIMS Journal*, *32*(1).

Colombaroni, C., & Fusco, G. (2014). *Artificial Neural Network Models for Car Following : Experimental Analysis and Calibration Issues for Car Following : Experimental*. *2450*. https://doi.org/10.1080/15472450.2013.801717

Constantinescu, Z., Marinoiu, C., & Vladoiu, M. (2010). Driving style analysis using data mining techniques. *International Journal of Computers, Communications and Control*, *5*(5), 654–663. https://doi.org/10.15837/ijccc.2010.5.2221

Daganzo, C. F. (2007). Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, *41*(1), 49–62. https://doi.org/10.1016/j.trb.2006.03.001

Dantas, G., Siciliano, B., Boscaro, B., Cleyton, M., & Arbilla, G. (2020). Science of the Total Environment The impact of COVID-19 partial lockdown on the air quality of the city of Rio de Janeiro , Brazil. *Science of the Total Environment*, *729*, 139085. https://doi.org/10.1016/j.scitotenv.2020.139085

Dawson, D. (2019). *Developing Emergency Preparedness Plans For Orlando International Airport (MCO) Using Microscopic Simulator WATSim*. *2006*, 2004–2019.

de Palma, A., & Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, *19*(6), 1377–1399. https://doi.org/10.1016/j.trc.2011.02.010

Delpiano, R., Herrera, J. C., Laval, J., & Coeymans, J. E. (2020). A two-dimensional car-following model for two-dimensional traffic flow problems. *Transportation Research Part C: Emerging Technologies*, *114*(February), 504–516. https://doi.org/10.1016/j.trc.2020.02.025

Demirtas, M. (2021). *Sizing optimization and design of an autonomous AC microgrid for commercial loads using Harris Hawks Optimization algorithm*. *245*. https://doi.org/10.1016/j.enconman.2021.114562

Derbel, O., Peter, T., Zebiri, H., Mourllion, B., & Basset, M. (2013). Modified Intelligent Driver Model for driver safety and traffic stability improvement. In *IFAC Proceedings Volumes* (Vol. 46, Issue 21). IFAC. https://doi.org/10.3182/20130904-4-JP-2042.00132

Ding, D. (2011). *Modeling and simulation of highway traffic using a cellular automaton approach*. *December*, 34. http://uu.diva-portal.org/smash/get/diva2:483914/FULLTEXT01.pdf

Du, J., Rakha, H. A., Filali, F., & Eldardiry, H. (2020). International Journal of Transportation COVID-19 pandemic impacts on traffic system delay , fuel consumption and emissions. *International Journal of Transportation Science and Technology*, *December*, 1–13. https://doi.org/10.1016/j.ijtst.2020.11.003

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors*, *37*(1), 65–84. https://doi.org/10.1518/001872095779049499

Fan, T., Wong, S. C., Zhang, Z., & Du, J. (2023). A dynamically bi-orthogonal solution method for a stochastic Lighthill-Whitham-Richards traffic flow model. *Computer-Aided Civil and Infrastructure Engineering*, *38*(11), 1447–1461. https://doi.org/10.1111/mice.12953

Fox, S., Ney, D., & Verrucci, E. (2018). Liberalisation , urban governance and gridlock : Diagnosing Yangon ' s mobility crisis ☆. *Cities*, *July*, 1–13. https://doi.org/10.1016/j.cities.2018.07.008

Francesco, M. Di, & Rosini, M. D. (2015). *Conservation Laws from Follow-the-Leader Type Models via Many Particle Limit*. *217*, 831–871. https://doi.org/10.1007/s00205-015-0843-4

Friedman, J. H. (2001). *Greedy Function Approximation : A Gradient Boosting Machine Author (s): Jerome H . Friedman Source : The Annals of Statistics , Oct ., 2001 , Vol . 29 , No . 5 (Oct ., 2001), pp . 1189-1232 Published by : Institute of Mathematical Statistics Stable UR*. *29*(5), 1189–1232.

Ge, Q., Ciuffo, B., & Menendez, M. (2014). *Comprehensive Approach for the Sensitivity Analysis of High-Dimensional and Computationally Expensive Traffic Simulation Models*. *February 2015*. https://doi.org/10.3141/2422-14

Greenberg, H. (1959). An Analysis of Traffic Flow. *Operations Research*, *7*(1), 79–85. https://doi.org/10.1287/opre.7.1.79

Hall, R., & Chin, C. (2005). Vehicle sorting for platoon formation: Impacts on highway entry and throughput. *Transportation Research Part C: Emerging*

Technologies, *13*(5–6), 405–420. https://doi.org/10.1016/j.trc.2004.09.001

Hamdar, S. H. (2008). Modeling Driver Behavior as a Stochastic Hazard-Based Risk- Taking Process. *Statistics*, *2000*(July), 1–24.

Hamdar, S. H., Mahmassani, H. S., & Treiber, M. (2015). From behavioral psychology to acceleration modeling: Calibration, validation, and exploration of drivers' cognitive and safety parameters in a risk-taking environment. *Transportation Research Part B: Methodological*, *78*, 32–53. https://doi.org/10.1016/j.trb.2015.03.011

Han, J., Shi, H., Chen, L., Li, H., & Wang, X. (2022). *The Car-Following Model and Its Applications in the V2X Environment : A Historical Review*. 1–34.

Han, J., Wang, X., & Wang, G. (2022). Modeling the Car-Following Behavior with Consideration of Driver, Vehicle, and Environment Factors: A Historical Review. *Sustainability (Switzerland)*, *14*(13), 1–27. https://doi.org/10.3390/su14138179

Hazelton, M. L. (2001). *Inference for origin ± destination matrices : estimation , prediction and reconstruction*. *35*.

Hidas, P. (2005). *Modelling vehicle interactions in microscopic simulation of merging and weaving*. *13*, 37–62. https://doi.org/10.1016/j.trc.2004.12.003

Hoogendoorn, S. P., & Bovy, P. H. L. (2001). State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, *215*(4), 283–303. https://doi.org/10.1177/095965180121500402

Horňák, I., & Přikryl, J. (2015). *Terms of use : ON TRAFFIC DATA n ´ Institute of Information Theory and Automation*.

Hou, Y. (2021). *Research and empirical Analysis of Traffic flow Modeling based on fluid Mechanics Research and empirical Analysis of Traffic flow Modeling based on fluid Mechanics*. 8–13. https://doi.org/10.1088/1755-1315/692/4/042102

Jia, S., Li, Y., & Fang, T. (2021). *Can Driving-Restriction Policies Alleviate Tra c Congestion ? A Case Study of Beijing , China Can driving*-restriction policies *alleviate traffic congestion ? A case study of Beijing , China*.

Jia, Y., Qu, D., Han, L., & Lin, L. (2021). *Research on car-following model based on molecular dynamics*. *13*(2), 1–10. https://doi.org/10.1177/1687814021993003

Kardani, N., Bardhan, A., Roy, B., Samui, P., Nazem, M., & Jahed, D. (2021). A novel improved Harris Hawks optimization algorithm coupled with ELM for predicting permeability of tight carbonates. *Engineering with Computers*, *0123456789*. https://doi.org/10.1007/s00366-021-01466-9

Kaysi, I., & Alam, G. (2000). D River B Ehavior and T Raffic S Tream I Nteractions. *Journal of Transportation Engineering*, *126*(6), 498–505.

Kerner, B. S. (2007). *Control of Spatiotemporal Congested Traffic Patterns at Highway Bottlenecks*. *8*(2), 308–320.

Kerner, B. S., & Klenov, S. L. (2004). Spatial-temporal patterns in heterogeneous traffic flow with a variety of driver behavioural characteristics and vehicle parameters. *Journal of Physics A: Mathematical and General*, *37*(37), 8753–8788. https://doi.org/10.1088/0305-4470/37/37/001

Khan, Z. H., & Gulliver, T. A. (2018). *A macroscopic traffic model for traffic flow harmonization*.

Khashayarfard, M., & Nassiri, H. (2021). *Studying the Simultaneous Effect of Autonomous Vehicles and Distracted Driving on Safety at Unsignalized Intersections*. *2021*.

Khodayari, A., Ghaffari, A., Kazemi, R., & Braunstingl, R. (2012). A Modified Car-Following Model Based on a Neural Network Model of the Human Driver Effects. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *42*(6), 1440–1449. https://doi.org/10.1109/TSMCA.2012.2192262

Kumar, M., Kumar, K., & Das, P. (2021). Study on road traffic congestion: A review. *Recent Trends in Communication and Electronics*, *June*, 230–240. https://doi.org/10.1201/9781003193838-43

Lahmar, I., Zaier, A., Yahia, M., & Boaullegue, R. (2023). *A Novel Improved Binary Harris Hawks Optimization For High dimensionality Feature Selection*. *171*, 170–176. https://doi.org/10.1016/j.patrec.2023.05.007

Lazar, H., Rhoulami, K., & Rahmani, D. (2016). *A Review Analysis of Optimal Velocity Models*. 123–131. https://doi.org/10.3311/PPtr.8753

Letters, T. (2013). *car-following and traffic stream model Validation of Van Aerde ' s simplified steady- state car-following and traffic stream model*. *7867*. https://doi.org/10.3328/TL.2009.01.03.227-244

Levi-civita, T., & Pura, M. (2022). *Seminario Dottorato 2022/23*. 1–189.

Li, Y., & Sun, D. (2012). Microscopic car-following model for the traffic flow: The state of the art. *Journal of Control Theory and Applications*, *10*(2), 133–143. https://doi.org/10.1007/s11768-012-9221-z

Liang, K. Y., Mårtensson, J., & Johansson, K. H. (2016). Heavy-Duty Vehicle Platoon Formation for Fuel Efficiency. *IEEE Transactions on Intelligent Transportation Systems*, *17*(4), 1051–1061. https://doi.org/10.1109/TITS.2015.2492243

Lighthill, M. J., & Whitham, G. B. (1955). *On Kinematic Waves . II . A Theory of Traffic Flow on Long Crowded Roads Article cited in :* https://doi.org/10.1098/rspa.1955.0089

Lin, P. W., Kang, K. P., & Chang, G. L. (2004). Exploring the effectiveness of variable speed limit controls on highway work-zone operations. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, *8*(3), 155–168. https://doi.org/10.1080/15472450490492851

Logghe, S., & Immers, L. H. (2008). *Multi-class kinematic wave theory of traffic flow*. *42*, 523–541. https://doi.org/10.1016/j.trb.2007.11.001

Lopez, P. A., Behrisch, M., Bieker-walz, L., Erdmann, J., Fl, Y., Hilbrich, R., Leonhard, L., Rummel, J., Wagner, P., & Wießner, E. (2018). *Microscopic Traffic Simulation using SUMO*. 2575–2582.

Lownes, N. E., & Machemehl, R. B. (2006). *Proceedings of the 2006 Winter Simulation Conference L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, eds. 2001*, 1406–1413.

Lykov, S., & Asakura, Y. (2018). *Anomalous Traffic Pattern Detection in Large Urban Areas : Tensor-Based Approach with Continuum Modeling of Traffic*

Flow.

Ma, J., Li, C., Liu, Z., Duan, Y., Lei, Y., & Xiong, L. (2016). On Traffic Bottleneck in Green ITS Navigation : An Identification Method. *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, 1–5. https://doi.org/10.1109/VTCSpring.2016.7504381

Maiti, S., Winter, S., Member, S., Kulik, L., & Sarkar, S. (2020). *The Impact of Flexible Platoon Formation Operations*. *5*(2), 229–239. https://doi.org/10.1109/TIV.2019.2955898

Martin, P. (1993). Autonomous intelligent cruise control incorporating automatic braking. *SAE Technical Papers*, *4*(4). https://doi.org/10.4271/930510

Mathew, T. V, & Radhakrishnan, P. (2010). *Calibration of Microsimulation Models for Nonlane-Based Heterogeneous Traffic at Signalized Intersections*. *March*, 59–66.

May, A. D. (1997). Introduction to traffic flow theory. In *Transport Planning and Traffic Engineering*. https://doi.org/10.1016/b978-034066279-3/50018-9

Mehmood, A., & Easa, S. (2017). *Modeling Reaction Time in Car-Following Behaviour Based on Human Factors*. *January 2009*.

Milanés, V., & Shladover, S. E. (2014). Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies*, *48*, 285–300. https://doi.org/10.1016/j.trc.2014.09.001

Minkin, L., & Whiting, P. (2018). *Road capacity with a steady flow traffic*.

Motamarri, R. (2021). JAYA Algorithm Based on Lévy Flight for Global MPPT Under Partial Shading in Photovoltaic System. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, *9*(4), 4979–4991. https://doi.org/10.1109/JESTPE.2020.3036405

Mutabazi, M. I. (1987). *Evaluation of Accuracy of U . S . DOT Rail-Highway Grade Crossing Accident Prediction Models*. 166–170.

Nagel, K., & Schreckenberg, M. (1992). A cellular automaton model for freeway traffic. *Journal de Physique I*, *2*(12), 2221–2229. https://doi.org/10.1051/jp1:1992277

Nagel, K., Schreckenberg, M., Nagel, K., & Schreckenberg, M. (1992). *A cellular automaton model for freeway traffic To cite this version : HAL Id : jpa-00246697 cellular*. *2*(12), 2221–2229.

Nelson, P. (1998). *THE PRIGOGINE-HERMAN KINETIC MODEL PREDICTS WIDELY SCATTERED TRAFFIC FLOW DATA AT HIGH CONCENTRATIONS*. *32*(8), 589–604.

Newell, G. F. (2002). A simplified car-following theory: A lower order model. *Transportation Research Part B: Methodological*, *36*(3), 195–205. https://doi.org/10.1016/S0191-2615(00)00044-8

Nu, D., Fuertes, W., Marrone, L., & Macas, M. (2022). *Machine Learning-Based Traffic Classification in Software-Defined Networking : A Systematic Literature Review , Challenges , and Future Research Directions*. *49*(4).

Peng, H. (n.d.). *Evaluation of Driver Assistance Systems — A Human Centered Approach*.

Pipes, L. A. (1953). *An Operational Analysis of Traffic Dynamics*. *274*. https://doi.org/10.1063/1.1721265

Punzo, V. (2016). *How Parameters of Microscopic Traffic Flow Models Relate to Traffic Dynamics in Simulation How parameters of microscopic traffic flow models relate to traffic dynamics in simulation : implications for model calibration*. *August*. https://doi.org/10.3141/2124-25

Qin, Y., Wang, H., & Ran, B. (2018). Stability Analysis of Connected and Automated Vehicles to Reduce Fuel Consumption and Emissions. *Journal of Transportation Engineering, Part A: Systems*, *144*(11). https://doi.org/10.1061/jtepbs.0000196

Radoni, J. R., & Davidovi, M. D. (2023). *TRAFFIC INTENSITY AND AIR POLLUTION BEFORE AND DURING LOCKDOWN IN NOVI SAD , SERBIA by*. *27*(3), 2333–2345.

Rajamani, R., & Zhu, C. (2002). Semi-autonomous adaptive cruise control systems.

IEEE Transactions on Vehicular Technology, *51*(5), 1186–1192. https://doi.org/10.1109/TVT.2002.800617

Rajé, F., Tight, M., & Pope, F. D. (2018). *Tra ffi c pollution : A search for solutions for a city like Nairobi*. *February*. https://doi.org/10.1016/j.cities.2018.05.008

Ren, X., Liu, J., & Wen, J. (2018). Congestion and Air Quality. *2018 5th International Conference on Industrial Economics System and Industrial Security Engineering (IEIS)*, 1–6.

Rodrigues, F., & Azevedo, C. L. (2019). Towards Robust Deep Reinforcement Learning for Traffic Signal Control: Demand Surges, Incidents and Sensor Failures. *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, 3559–3566. https://doi.org/10.1109/ITSC.2019.8917451

Samal, S. R., Gireesh Kumar, P., Cyril Santhosh, J., & Santhakumar, M. (2020). Analysis of Traffic Congestion Impacts of Urban Road Network under Indian Condition. *IOP Conference Series: Materials Science and Engineering*, *1006*(1). https://doi.org/10.1088/1757-899X/1006/1/012002

Sean, Z., Li, J., Li, X., Zhang, M., & Wang, H. (2017). *Modeling heterogeneous traffic flow : A pragmatic approach*. *99*, 183–204. https://doi.org/10.1016/j.trb.2017.01.011

Session, S. (2020). *Road safety*. *September 2017*.

Setiawan, E. B., Tarwidi, D., & Umbara, R. F. (2016). *Numerical Simulation of Traffic Flow via Fluid Dynamics Approach*. *3*(1), 93–104.

Shah, M. A. (2020). *Congestion Modelling and Level of Service Assesment of Urban Roads in*. *7*(8), 2230–2240.

Sherali, H. D. (2014). *Springer Optimization and Its Applications VOLUME 84*. *54*. http://www.springer.com/series/7393

Shladover, S. E. (2018). Connected and automated vehicle systems: Introduction and overview. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, *22*(3), 190–200. https://doi.org/10.1080/15472450.2017.1336053

Shladover, S. E., Su, D., & Lu, X. Y. (2012). Impacts of cooperative adaptive cruise control on freeway traffic flow. *Transportation Research Record*, *2324*(October 2014), 63–70. https://doi.org/10.3141/2324-08

Steven, E., & Thompson, D. (2015). *Cooperative Adaptive Cruise Control (CACC) For Truck Platooning : Operational Concept Alternatives*.

Sun, J., Zheng, Z., & Sun, J. (2023). *Stability Evolution of Car-Following Models Considering Asymmetric Driving Behavior*. *2677*(8), 361–371. https://doi.org/10.1177/03611981231156584

Taylor, P., & Castillo, J. M. (2012). *Transportmetrica Three new models for the flow – density relationship : derivation and testing for freeway and urban data*. *December 2014*, 37–41. https://doi.org/10.1080/18128602.2011.556680

Toledo, T., & Koutsopoulos, H. N. (2004). *Statistical Validation of Traffic Simulation Models*. *1876*, 142–150.

Traoré, M. K., Maiga, O., Traoré, M., Koné, Y., Traoré, K. M., Traoré, M. K., Maiga, O., Traoré, M., Koné, Y., Maïga, O., Koné, Y., Maïga, O., & Traoré, M. K. (2020). *Application of multi-perspective modeling and holistic simulation to Urban Transportation Systems To cite this version : HAL Id : hal-02977165 Application of Multi-Perspective Modeling and Holistic Simulation to Urban Transportation Systems*. https://doi.org/10.46354/i3m.2020.emss.013

Urban, P. I., Congestion, T., Qi, L., Member, S., Zhou, M., Luan, W., & Member, S. (2018). A Two-level Traffic Light Control Strategy for. *IEEE Transactions on Intelligent Transportation Systems*, *19*(1), 13–24. https://doi.org/10.1109/TITS.2016.2625324

Venkata Rao, R. (2016). Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *International Journal of Industrial Engineering Computations*, *7*(1), 19–34. https://doi.org/10.5267/j.ijiec.2015.8.004

Venkayya, D., & Devi, A. L. (2023). *Optimal Power Flow Analysis for Power Loss Reduction using Jaya Algorithm*. *5*(1), 1–10.

Wang, M., van Maarseveen, S., Happee, R., Tool, O., & van Arem, B. (2019).

Benefits and Risks of Truck Platooning on Freeway Operations Near Entrance Ramp. *Transportation Research Record*, *2673*(8), 588–602. https://doi.org/10.1177/0361198119842821

Wang, T., Hu, S., & Jiang, Y. (2021). Predicting shared-car use and examining nonlinear effects using gradient boosting regression trees. *International Journal of Sustainable Transportation*, *15*(12), 893–907. https://doi.org/10.1080/15568318.2020.1827316

Ward, J. a. (2009). Heterogeneity, Lane-Changing and Instability in Traffic: A Mathematical Approach. *Thesis*. http://www.personal.reading.ac.uk/~dj902857/publications/JAW_thesis.pdf

Wilhelm, F., & Lian, F. (2019). How speed and visibility influence preferred headway distances in highly automated driving. *Transportation Research Part F: Psychology and Behaviour*, *64*, 485–494. https://doi.org/10.1016/j.trf.2019.06.009

Wu, C., Ma, Z., & Kim, I. (2020). Multi-Agent Reinforcement Learning for Traffic Signal Control: Algorithms and Robustness Analysis. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*. https://doi.org/10.1109/ITSC45102.2020.9294623

Xie, R., Wei, D., Han, F., Lu, Y., Fang, J., Liu, Y., & Wang, J. (2018). *Technological Forecasting & Social Change The e ff ect of tra ffi c density on smog pollution : Evidence from Chinese cities*. *April*. https://doi.org/10.1016/j.techfore.2018.04.023

Xu, L., Wang, L. Y., Yin, G., & Zhang, H. (2014). Communication information structures and contents for enhanced safety of highway vehicle platoons. *IEEE Transactions on Vehicular Technology*, *63*(9), 4206–4220. https://doi.org/10.1109/TVT.2014.2311384

Yang, L., Fang, S., Wu, G., Sheng, H., Xu, Z., Zhang, M., & Zhao, X. (2022). Physical Model versus Artificial Neural Network (ANN) Model: A Comparative Study on Modeling Car-Following Behavior at Signalized Intersections. *Journal of Advanced Transportation*, *2022*(Idm). https://doi.org/10.1155/2022/8482846

Yang, L., Zhang, Y., & Zuo, J. (2021). An Attention-Based Spatial-Temporal Traffic Flow Prediction Method with Pattern Similarity Analysis. *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 3710–3717. https://doi.org/10.1109/ITSC48978.2021.9564947

Yang, Q. I., & Koutsopoulos, H. N. (1996). *A MICROSCOPIC TRAFFIC SIMULATOR FOR EVALUATION DYNAMIC TRAFFIC MANAGEMENT SYSTEMS*. *4*(3).

Yeo, H., & Skabardonis, A. (2009). Understanding Stop-and-go Traffic in View of Asymmetric Traffic Theory. *Transportation and Traffic Theory 2009: Golden Jubilee, July 2009*, 99–115. https://doi.org/10.1007/978-1-4419-0820-9_6

Yue, W., Li, C., Member, S., & Chen, Y. (2022). What Is the Root Cause of Congestion in Urban Traffic Networks : Road Infrastructure or Signal Control ? *IEEE Transactions on Intelligent Transportation Systems*, *23*(7), 8662–8679. https://doi.org/10.1109/TITS.2021.3085021

Zhang, H. M. (1999). *A mathematical theory of tra c hysteresis*. *33*.

Zhang, H. M., & Kim, T. (2005). A car-following theory for multiphase vehicular traffic flow. *Transportation Research Part B:Methodological*, *39*(5), 385–399. https://doi.org/10.1016/j.trb.2004.06.005

Zhao, F., Ma, R., & Wang, L. (2022). A Self-Learning Discrete Jaya Algorithm for Multiobjective Energy-Efficient Distributed No-Idle Flow-Shop Scheduling Problem in Heterogeneous Factory System. *IEEE Transactions on Cybernetics*, *52*(12), 12675–12686. https://doi.org/10.1109/TCYB.2021.3086181

Zhu, H. B., & Dai, S. Q. (2008). Analysis of car-following model considering driver’s physical delay in sensing headway. *Physica A: Statistical Mechanics and Its Applications*, *387*(13), 3290–3298. https://doi.org/10.1016/j.physa.2008.01.103

Appendix A Ethics Certificate

NEAR EAST UNIVERSITY

SCIENTIFIC RESEARCH ETHICS COMMITTEE

29.01.2024

Dear Abdınasır_Mohamed Yusuf

Your project **"Modelling And Predicting Car Following Behavior In Connected Vehicles: A Machine Learning Approach"** has been evaluated. Since only secondary data will be used the project does not need to go through the ethics committee. You can start your research on the condition that you will use only secondary data.

Prof. Dr. Aşkın KİRAZ

The Coordinator of the Scientific Research Ethics Committee

Appendix B Turnitin Similarity Report

Submit

	AUTHOR	TITLE	SIMILARITY			FILE	PAPER ID	DATE
	Abdinasir Mohamed Yu...	Full thesis	13%				2309721025	02-Mar-2024

Assist. Prof. Dr. Ikenna Uwanuakwa