St. Mary's University

The Faculty of Informatics

The Role of a Detection of vehicles' Speed and Net Overload in Road Traffic Accident Analysis with Data Mining Approach: Case of Addis Ababa

A Thesis Presented
by
Kaleab Yirga

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
in
Computer Science

June, 2022

ACCEPTANCE

The Role of a Detection of vehicles' Speed and Net Overload in Road Traffic Accident Analysis with Data Mining Approach: Case of Addis Ababa

By

Kaleab Yirga Assayea

Accepted by the Faculty of Informatics, St. Mary's University, in partial fulfillment of the requirements for the degree of Master of Science in Computer Science

Thesis Examination Committee:

_____

Internal Examiner
Michael Melese (PhD)

_____

External Examiner
Minale Ashagrie (PhD)

_____

Dean, Faculty of Informatics
Alembante Mulu (PhD)

Tuesday, June 21, 2022

DECLARATION

I, the undersigned, declare that this thesis work is my original work, has not been presented for a degree in this or any other universities, and all sources of materials used for the thesis work have been duly acknowledged.

Kaleab Yirga Assayea
Full Name of Student

_____

Signature

Addis Ababa

Ethiopia

This thesis has been submitted for examination with my approval as advisor.

Alembante Mulu (PhD)
Full Name of Advisor

_____

Signature

Addis Ababa

Ethiopia

June, 2022

ACKNOWLEDGMENTS

This thesis would not have been feasible without the direction and assistance of several people who contributed and extended their important help in the preparation and completion of this thesis in various ways.

First and foremost, I thank Almaty God for blessing me with tranquility, knowledge, and mental clarity, which has enabled me to complete this study project. This research project has progressed due to His abundant grace.

Second, I want to express my gratitude to Dr. Alembante Mulu, my thesis advisor. He has been quite helpful in the design of research topics and the determination of objectives that are relevant to the studies as my supervisor he also went over the document to make sure it was of good quality.

Also, I'd want to convey my heartfelt gratitude to Dr. Getahun, who was a tremendous help to me during the seminar course and the project title selection.

Finally, I would like to thank my family and friends for creating positive working environment and the provision of encouragement throughout the study.

DEDICATION

To: Anteneh Yirga (**Ant'ye**) was my younger brother who died Three years ago in a vehicle accident.

Contents

Acronyms

ADS	Accident Detection Server
ADU	Accident Detection Unit
AATP	According To Addis Ababa Traffic Police
BS	Base Station
CO	Carbon Dioxide
CCTV	Closed Circuit Television
DSRC	Dedicated Short Range Communication
DVSS	Dopplers Vehicular Speed Sensor
DC	Electrical Energy to Mechanical Energy
GPS	Global Positioning System
GSM	Global System for Mobile Communication
IR	Infrared Sensor
ITS	Intelligent Transport System
LCD	Light Limiting Diode
LPG	Liquefied Petroleum Gas
LCD	Liquid Cristal Display
UNO	Micro Controller
MANET	Mobile Ad-Hoc Network
OERD	Object And Event Detection and Response
OSN	Online Social Network
ODD	Operational Design Domain

RSU	Road Side Unit
RTA	Road Traffic Accident
SMS	Short Message Service
V2V	Vehicle To Vehicle
VSN	Vehicular Social Network
V2 I	Vehicular To Infrastructure
VSSN	Veteran Supplemental Support Network
WAVE	Wireless Access in Vehicular Environment
WHO	World Health Organization

List of Tables

List of Figures

Abstract

From the advent of increased transportation, overloading and over speeding of vehicles has become the major causes for accidents and killing many lives. Transport authorities are employing advanced traffic management system (ATMS) to improve vehicular traffic management efficiency. ATMS currently uses intelligent traffic lights and sensors distributed along the roads to achieve its goals. Furthermore, there are other promising technologies that can be applied more efficiently in place of the above-mentioned ones, such as vehicular networks.

In this study, the researcher tried to assess road traffic accident causes and control mechanisms undertaken by authorities, in Addis Ababa traffic police bureau and Addis Ababa Road Authority. As findings revealed, overload and over speed has immensely contributed to the incidence of road traffic accident in Addis Ababa.

This research demonstrates that the Ethiopian traffic management system has been using very old systems which have very limited capacity. In this proposed work, monitoring driving behavior with the help of wireless sensor technology is the target. So, the proposed research work focused on developing a model for integrating wireless sensor network and vehicular social network.

This paper presents a model, that can classify accidents well with a better accuracy as fatal, serious, and slight or property-damage was selected and evaluated. Experiment results reveal that the use of logistic regression is helpful in detecting causes of the accident. In this work it has been proved that driving over speed and over net load of vehicles are the major causes of traffic accidents in Addis Ababa and also the research proved that automation provides better performance than a human handled system.

***Keywords**: Sensor network, Data mining, Traffic Accident, Models*

Chapter One

1.1 Introduction

Vehicular transportation is an essential part of the modern human life development. However, the ever-increasing number of road accidents, traffic congestion and other such issues become obstacles for the realization of necessity vehicle transportation [9]. As the integration of the Internet of Vehicles and social networks, vehicular social networks (VSNs) are promising to solve the above-mentioned problems by enabling vehicle-to-vehicle and vehicle-to-infrastructure communication [9].

In this busy life schedule, people always prefer to drive at very speed rather than low speed to reach their respective places in time. Thus, it is necessary to understand the need of a technology which would be used as a speed limit enforcement system. A system which helps to limit the speed of the vehicles and the owner would be punished under the law due to over speed and this is the best method for making people to drive at normal designated speeds. From the advent of increased transportation, over speeding of vehicles has become one of the major causes for accidents and killing many lives [5].

With the rapid development of current logistics industry, the question of transport vehicles overloading is becoming more and more serious. Overloaded vehicles will cause extensive damage to the road and huge economic losses to society, which seriously threaten traffic safety. According to statistics, 70% of traffic accidents are caused by vehicles overloading [10]. Therefore, the overloading of transport vehicles has become one of the crucial issues of traffic management.

In addition to creating a sustainable development and promoting vehicle transportation efficiency Car traffic accidents have been extremely normal in the current world especially on our country Ethiopia. The need to check this has been exceptionally fundamental and various strategies have been utilized up until now. With the headway in innovation, distinctive administering bodies are requesting a type of mechanized innovation to control this issue of over speed driving and net overloading. In this situation, proposing a model to recognize the vehicle which is being driven over the given most extreme speed limit and net overload that the individual streets or expressway limits.

Nowadays, Intelligent Transport System (ITS) is one of the hottest research areas due to the high number of problems in transportation. A number of people dies, a lot of properties are lost due to traffic

problem in the world. So, researchers investigate VANET as one of the solutions to decrease the problem. Vehicular Ad-hoc Networks (VANETs) is a wireless type of network that allow opportunistic vehicular communications among vehicles i.e., vehicle-to-vehicle (V2V), and from vehicle-to infrastructure (V2I), then objective to improve the quality of vehicle transportation system in addition to preserve human life that victims and injures by the cause of vehicle accidents. Vehicular Ad hoc Networks (VANETs) are special kind of Mobile Ad Hoc Networks (MANETs) that are formed between moving vehicles and road side unit [4]. By this time there is no such technology in our country Ethiopia but, for the coming future the government should start deploying and working on it.

A Vehicular Social Network (VSN) is an emerging field of communication where relevant concepts are being borrowed from two different disciplines, i.e., vehicular ad-hoc networks (VANETs) and mobile social networks. This emerging paradigm presents new research fields for content sharing, data dissemination, and delivery services [12].

The majority of modern vehicles are equipped with an intelligent network of sensors, a Global Positioning System (GPS), and transceivers for signal transmission and reception. The vehicle that is closest to the accident site receives emergency messages and responds by either slowing down or changing lanes, depending on the circumstances. A higher likelihood of receiving emergency communications and making decisions exists in vehicles away from the accident scene [1]. The most recent automobiles are being imported and built in Ethiopia, particularly in Addis Ababa. In an effort to avoid and control vehicle accidents in the city of Addis Ababa, this study attempts to analyze the primary causes of traffic accidents and design a model that can help road traffic accident

1.2 Motivation of the Study

According to the World Health Organization (WHO) report, every year more than 1.3 million people now die on the world's road and about 50 million people are injured or disabled as a result of road traffic crashes. There are only about 5,000 automobiles for every 100,000 people in Africa. Around 48,000 people live in Europe. Despite this, the World Health Organization (WHO) reports that the number of people killed in road accidents in Africa is about three times greater than in Europe.

According to research conducted by Ethiopia's Federal Transport Authority, 62 people died in 10,000 traffic incidents in the first eight months of 2019, with 6,886 others suffering serious injuries and 7,071 others suffering minor injuries. There are 700,000 automobiles in the nation. Per capita, Ethiopians

possess three autos per 1,000 people [37]. To help tackle the issues outlined above, vehicles must engage with ITS. As a result, developing and accessing a message awareness forwarding mechanism that is both effective and efficient is critical. This is an experimental study that will be carried out with the help of a powerful reusable simulation tool that is compatible with both network and traffic simulators.

1.3 Statement of the Problem

Many vehicle accidents occurred due to high speed and net overload driving of the vehicle [10], 11]. Driver which has misbehaviors, or careless and some other suddenly with lack of awareness drive the vehicle within high speed and carry net overload. A vehicle which travels through highest speed it is very difficult in order to control with little bit meter per second. Also, a vehicle that carry a passenger over its' availability amount and above its' performance is difficult to manage that vehicle well through journey. So, in this case many are killed and injured due to vehicle accidents each and every day in addition to the property.

Even if penalizing the driver who gives service to the society all of the time does not teach them anything. As a result, message-based vehicular communication is a possible solution to the problem. As a consequence, this research will be beneficial in overcoming such challenges and will make a significant result in the field of vehicular social networks.

Since the analysis made by using traditional methods focus on problems with much more manageable number of variables and cases than may be encountered in real world, they have limited capacity to discover new and unanticipated patterns and relationships that are hidden in conventional databases (Plate et ai, 1997). The absence of significant attempt that has been made so far to this level in identifying the major determinants of car accidents and establishing the most important factors influencing the severity of an injury in Addis Ababa Region justify the importance of this research. This research work will be groundwork for the effort of reducing vehicle accident in particular and improving the quality of life in general.

Hence, to this end, the study attempts to explore, investigate and answer the following main research questions.

1.4 Research Question

- How driving over speed limit and over loading vehicles contribute to a traffic accident?
- What is the maturity level of Addis Ababa traffic Management in implementing technologies for road traffic management?
- Which model is appropriate to predict the causes of traffic accident dataset and which algorithm best classify the KPI data?

1.5 General Objective

The study will be conducted with main objective to make assessment on the role of A Detection of vehicles' Speed and Net Overload in Road Traffic Accident Analysis with Data Mining Approach in the effort of preventing and controlling vehicle accident at the city of Addis Ababa

1.5.1 Specific Objectives

The specific objectives of the research are:

- To review literatures so as to understand the current sciences, technologies and standards related to this study.
- To examine the current car traffic accident monitoring methods of Addis Ababa traffic Management Authority.
- To conduct experiment and build a model.
- To evaluate the model using the selected tools and technologies.
- To draw recommendations based on the findings.

1.6 Scope and Limitation of the Study

This study will be conducted to detect high speed and net overload driving of the vehicles and forward awareness massage to vehicular social network and make an experiment using traffic accident dataset. Due to limitation computational resources the study covers road traffics around Addis Ababa only.

The scope of the research is limited to assessing the roles of driving over speed and net overload on the causes of traffic accidents with literatures and the possible application of data mining technology at Addis Ababa Traffic Office. And also, it's limited to examine the potentials of data mining techniques in

developing classification model by the use of three algorithms (KNN, SVM and Logistic regression) In support of traffic control activities.

1.7 Significance of the Study

This thesis work focuses on studding and proposing a model that are significance for the implementation of detection of vehicles excessive speed and net overload and road traffic accidents. This study will be conducted to detect high speed and net overload driving of the vehicles and forward awareness massage to vehicular social network and make an experiment using traffic accident dataset.

1.8 Methodology Adopted

The study employed a hybrid strategy to analyze data collected through questioning and interviews (qualitative and quantitative).

Appropriate mining is carried out based on the established objectives and an assessment of the available data. The CRISP-DM Industry type of model was chosen and run on the available data for the goal of producing models and spotting trends.

The objective component confusion matrix is created using precision, recall, TP, FP, f-measure, and the area under the ROC curve. During the subjective step, selected interesting rules from the chosen model were debated with carefully selected domain experts to evaluate how closely they agreed with the final extracted interesting rules.

1.9 Organization of the Study

This research report is organized into five chapters. The first chapter is discussed about the introduction part that introduces key points about the research including the background, the problem which the research paper wants to address, the general and specific objective of the study, the scope and others. The second chapter is devoted to review literatures on the major factors of road traffic accidents (RTAs), the effects of RTAs on human security, assessment of measures being undertaken to reduce and/or prevent the occurrences of road crashes in the road of Addis Ababa, and examines individual's perception towards RTA as a security threat along with the literature in the third chapter reviewed. Chapter three contains detail description research methodology. The fourth chapter deals with proposing a model. Finally, in the fifth chapter conclusions and recommendations were presented.

Chapter Two

2.1 Literature Review

This chapter presents some background information of the available literature related to the development of traffic accident rates, analysis on the vehicles speed limit and overloading as causing factors for traffic accident. And also tries to assess role of the vehicular social network and vehicular ad hoc networks on reducing traffic accidents. In general, in this chapter the study tries to review literatures so as to understand the current sciences, technologies and standards related to this study.

2.1.1 Traffic Accident

Most investigations have revealed that 70% to 80% of all traffic accidents are due to human error. The term human error however is often controversial for. It doesn't satisfactory describe that large number of injuries and deaths that occurs on the road as the result of driving errors while abilities to do so are impaired by alcohol or drugs, lack of experience, lack at distribution of attention etc. (2002 WHO statistic). Car accidents are accidental collisions between automobiles. Car accidents can damage one or more autos, people, or structures. Car accidents also called traffic accidents, auto accidents, road accidents, and motor vehicle accidents cause thousands of deaths and hundreds of thousands of disabilities each year. Worldwide, car accidents kill an estimated one million people each year [13].

Road traffic accidents currently kill 1,800 Ethiopians a year and injure another 7,000. Alarmed by the increasing carnage, Shell Ethiopia, the largest fuel distributor in Ethiopia with a 43% market share, launched an awareness campaign: "Drive to Live". The campaign is intended to promote the value of safety rules and the benefit of implementing "defensive driving" for drivers employed by the transport companies [13].

2.1.3 Contributory Factors to the Accidents

There are many contributory factors for traffic accidents; some are discussed briefly in the following table [25].

Reasons	Total accidents in each year					Total	Percentage
	2009	2010	2011	2012	2013		
Drunk driving	3	2	696	130	45	826	1.54
Not give way for vehicles	1453	1185	1515	2838	4131	24758	43.57
Not give way for pedestrians	1115	1169	1692	2159	2514	8649	15.22
Coming too close	1984	1492	1580	1843	3431	10330	18.17
Improper turning	970	644	402	564	1149	3729	6.56
Over Speeding	35	65	317	515	499	1431	2.50
Over loading	50	20	38	67	67	242	0.42
Sleep driving	-	-	49	1	15	65	0.11
No use of right side	804	605	786	710	1274	4237	7.45
Improper passing	365	295	271	881	701	2504	4.40

Table 2.1, Factors That Causes Drivers to Inflict Traffic Accidents (Adopted from [25])

The impact of the accident is more sever on pedestrians than passengers and drivers. Pedestrians alone accounts 83.6 percent of accidents and followed by passengers which accounts 12.72 percent. From the populations, it is workers, students and jobless peoples more victimized by the accidents. Most of the accidents are occurred at market centers, churches, schools, offices and at houses centers. Although transportation service is vital for the day-to-day activities of human beings and base for development, in Addis Ababa it is not safe and become the source of loss of human and economic resources. There for in order to reduce its impact and if possible, to eliminate it and to maintain safety in the city the following options are suggested to implement by responsible and concerned bodies [25].

• Public education and awareness of road safety through public media, magazines, newspapers, posture, and screens.

• Licensing and training for drivers through ongoing assessment and training.

• Creating simple, easy-to-implement laws and rules that benefit both implementers and road users.

• Implementing agency must take rigorous action against unsafe road users by using adequate enforcement.

• In locations with a higher rate of traffic accidents, the government must send traffic police.

• Clearly designating pedestrian pathways, zebra crossings, and parks.

• Road construction should be thoroughly studied and based on certain technical specifications.

2.1.4 Road Traffic Accident at Addis Ababa

According to Tibebe 2005, Road traffic accidents are among the top leading causes of deaths and injuries of various levels. Ethiopia is one of the countries of the world experiencing highest rate of such accidents resulting in fatalities and various levels of injuries. Addis Ababa, the capital city of Ethiopia, takes the lion's share of the risk having higher number of vehicles and traffic. And the cost of these fatalities and injuries due to such road traffic accidents has a great impact on the socio-economic development of a society.

In developing African nations, an increasingly large proportion of financial resources are being diverted to the importation of vehicles and new road construction. The health impact of the transition has not been well documented in Africa and motor vehicle injuries have not figured prominently in the past as a priority public health problem (Taddele and Larson, 1991).

Apart from being serious health problem, road traffic accidents also create various socio-economic problems specifically on developing nations like Ethiopia. Especially if these accidents are accompanied by a greater volume of traffic system the problem associated with them will be more serious. In connection to this, volume and composition of traffic on the Ethiopian roads was reviewed by Ethiopian Road Authority, ERA (2002). The study provides analysis of traffic information of the year which is a primary input to the maintenance, construction, improvement, planning and management needs of

Ethiopian Road Authority. The report revealed that the need for better traffic system management and maintenance capacity strengthening is evident from simple consideration of volume and composition of vehicles on the road network of the country especially urban cities like Addis Ababa.

Abebe (2004), tried to assess points that needs immediate attention with respect to road safety. He stated that 81 % of the accident all over the county is due to drivers' fault and the other is due to vehicle, pedestrian and road faults. He also listed out that the main road safety problems are:

- Drivers not respecting pedestrian priority
- Over speeding
- Unsafe utilization of freight vehicles for passenger transportation.
- Poor skill and undisciplined behavior of drivers
- Less engineering effort in road design to consider safety
- Poor vehicle conditions
- Pedestrian not taking proper precautions
- Week traffic law enforcement
- Lack of proper emergency medical services

2.1.5 Vehicle Speed Detection

Different projects have focused on detecting vehicle speeds from a single camera lens. (Ferrier et al., 1994) introduced the application of background approximation and difference to vehicle speed detection (Rad et al., 2010, p.2555). (Ginsburg et al., 2015) measured vehicle speeds using a computer and a mounted camera as a low-cost alternative with less development time as compared to a typical traffic surveillance system.

Another vehicle speed detecting system was proposed by (Rad et al., 2010). They used a video camera, a computer and MatLab software to implement their system, still as a low-cost alternative and a more reliable system to radar surveillance traffic systems.

For speed estimation, [2] used Doppler's Vehicle Speed Sensor (DVSS) which is a fully self-contained, roadside mounted, vehicle speed measurement sensor. This non-intrusive, high performance speed sensor shatters existing sensor performance and cost points. In addition to the low unit cost, the sensor is extremely robust and performs without maintenance for years. The sensor is battery-powered, solar

charged, and mounts quickly on existing poles or overpasses. The Doppler Vehicle Speed Sensor - 102 uses 24 GHz Doppler microwave transceiver systems coupled to a Digital Signal Processor to measure and calculate the vehicle's speed. The DVSS-102 is capable of determining the average or composite vehicle speed for a multiple-lane freeway or highway [2].

2.1.6 Detection of Vehicles Overload

With the rapid development of current logistics industry, the question of transport vehicles overloading is becoming more and more serious. Overloaded vehicles will cause extensive damage to the road and huge economic losses to society, which seriously threaten traffic safety. According to statistics, 70% of traffic accidents are caused by vehicles overloading. Therefore, the overloading of transport vehicles has become one of the crucial issues of traffic management [10]. The increment in traffic accidents which are caused by vehicle overloading has become one of the major issues for traffic management [27].

Vibration sensor is used to collect vibration data, according to the varied sizes of vibration data to detect overload. The whole detection system consists of detection nodes, transmission networks and the back-end servers. Detection nodes are used for sampling, the data is transmitted to the PC through a wireless router so that the waveform generated by ground vibration can be observed in real time. Data is stored in the database and processed by the algorithm in the server. When the overload is detected, the server will send out a warning [10].

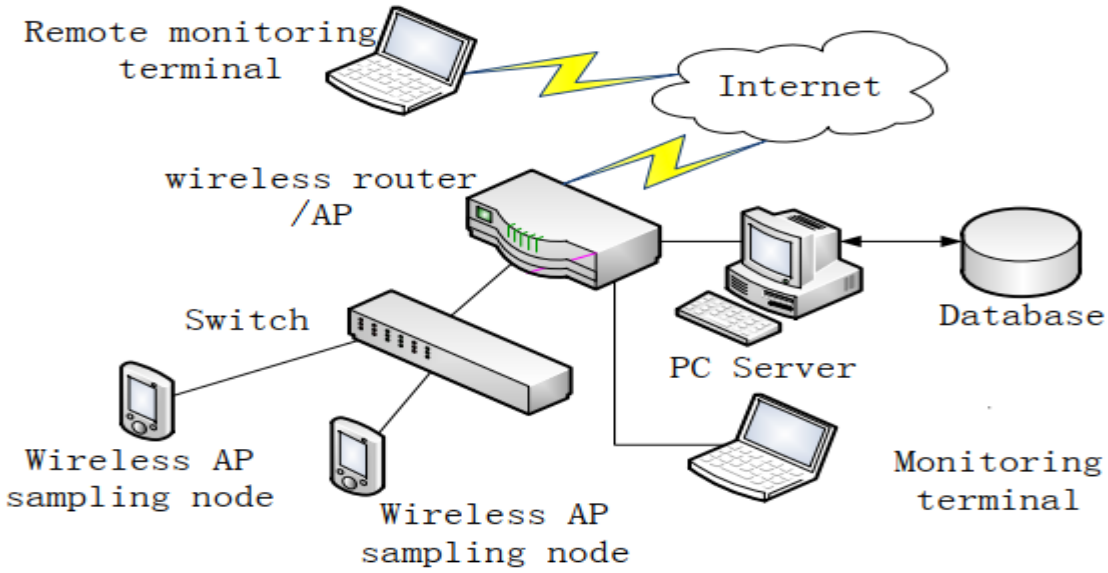

Figure 2.1, System architecture vehicle overload detection system [10]

2.1.7 Ad hoc Network

Being ad-hoc in nature, VANET is a type of networks that is created from the concept of establishing a network of cars for a specific need or situation. VANETs have now been established as reliable networks that vehicles use for communication purpose on highways or urban environments [17]. Being ad-hoc in nature, VANET is a type of networks that is created from the concept of establishing a network of cars for a specific need or situation. VANETs have now been established as reliable networks that vehicles use for communication purpose on highways or urban environments.

Vehicular ad-hoc networks are responsible for the communication between moving vehicles in a certain environment. A vehicle can communicate with another vehicle directly which is called Vehicle to Vehicle (V2V) communication, or a vehicle can communicate to an infrastructure such as a Road Side Unit (RSU), known as Vehicle-to-Infrastructure (V2I). The following Figure shows a typical VANET scenario.

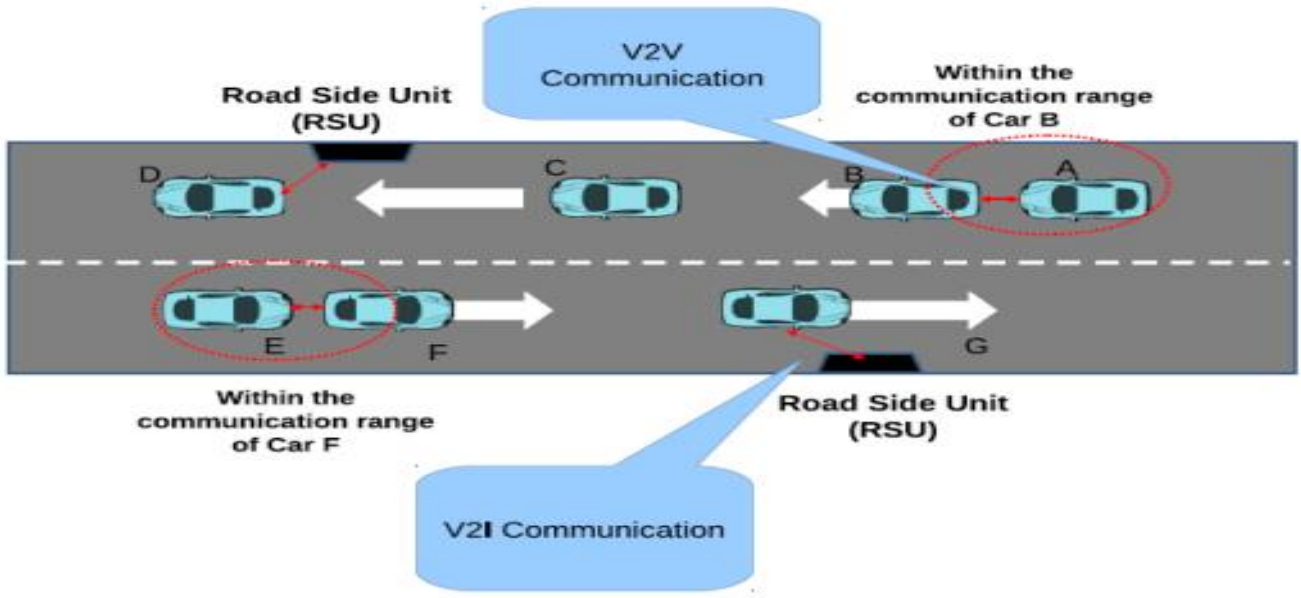

Figure 2.2, creating an Ad-hoc Network using Vehicles (Adopted from [17])

In VANETs, vehicles are equipped with wireless communication capabilities that allow them to connect together and to communicate with each other using distributed communications, called *Vehicle to Vehicle* (V2V). They can also communicate indirectly through a *Base Station* (BS), such communication is called *Vehicle to Infrastructure* (V2I) (Holzer et al., 2010). The Wireless Access in Vehicular Environment (WAVE) and IEEE 802.11p are at the core of the *Dedicated Short-Range Communication* (DSRC) standards that allow both V2V and V2I communication. Three types of networks architectures can be built using these technologies, as illustrated in the following Figure [26].

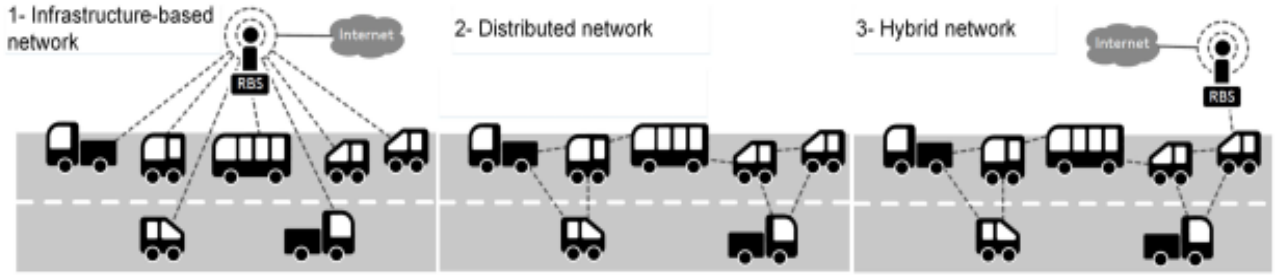

Figure 2.3 VANET architectures (Adopted from [26])

VANETs Applications

Vehicular applications have recently emerged with the widespread diffusion of smart mobile devices with network capabilities and access to user location. These applications aim at improving the users' traveling experience by providing real time traffic information and also aim at ensuring comfortable and safe trips to the travelers. There are two main classes of VANET applications: safety and non-safety applications (Kaveh, 2012). The first class (i.e., safety applications) aims at increasing the drivers' safety awareness to accidents and emergency conditions located in close proximity. Safety applications include collision and security distance warning, driver assistance, cooperative driving, cruise control, road information dissemination, automatic parking and automated driving (Holzer et al., 2010).

These applications are used to help drivers avoid collisions and coordinate among themselves at highway entries and at intersections by broadcasting warning messages over the considered network. Such applications aim at saving lives by reducing the number of car accidents and the associated damages. Non safety applications aim at providing comfortable and entertaining services to the drivers.

The latter range from real-time interactive games, weather information, restaurant locations, mobile e-commerce, and music downloads and content delivery.

Messaging Service for VANETs

Several research works advocate the design of a dynamic messaging system able to satisfy the online commuters' requirement in terms of enabling social interaction on the road by promoting the development of user oriented vehicular social applications and services. In fact, commuters are unknown and anonymous to each other and are sometimes hesitant about revealing their private information to an unknown public. The latter is one of many other reasons that constrains and limits their social interaction on the road (Luan et al., 2015b, a).

2.1.8 Vehicular Social Network (VSN)

With the rapid growth of smart devices and advent of advanced communication and computing capabilities, it is possible for drivers and passengers to communicate and socialize with other commuters on the roads. Not only limited to interest, but these commuters may also interact with each other passengers facing similar traffic condition, environmental factors, the same mobility pattern, or belong to the same community. Integration of advanced intelligent computing and social networking perspective into the vehicular environment has emerged as a new paradigm of Vehicular Social Networks (VSNs) with high potential for diverse applications; not only for future Intelligent Transportation System (ITS) but also for entertainment and drivers' comfort along the roads [12].

A Vehicular Social Network (VSN) is a mobile communication system formed by the combination of relevant concepts and features from the vehicular ad hoc networks (VANETs) and social networks. The idea of social cars emerges from the supposition that every driver can transfer information to different neighbors in light of normal hobbies; for example, Ford's idea auto Evos can straightforwardly shape an informal community with driver's companions. Beginning from essential components of VANETs, our point is to show how human social conduct can change the way autos are driven in the next couple of years. Today, person-to-person communication is a reality, and presenting social angles in VANETs permits vehicles to convey, as well as select, comparative neighboring in light of social measurements [1].

According to [1], essentially, a VSN is included two key parts:

1) A vehicular specially appointed system that speaks to the physical layer, and

2) an informal community system running on top of such a physical vehicular system. Subsequently, a VSN needs solid participation between social angles and physical system operational instruments.

Social Vehicles

Through vehicular communication, we allow vehicles to communicate with other cars within their proximity. However, so far, we have only looked at communicating vehicles as hollow entities. Taking into consideration the driver (and potential passengers) within the mobile units unlocks yet another dimension of the VANET. With this view, a driver can share data with his or her neighbors, that is, drivers and passengers in neighboring vehicles, based for example, on common interests, laying out the foundation for what is known as a VSN. Evidently, human behavior, for example, route preferences or individual selfishness, can have a large impact on the vehicular network, for example, its connectivity. With these considerations, researchers have shifted their focus in social aspects in order to improve the vehicular domain in both safety applications and infotainment.

Vehicular Social Sensor Networks

Social sensing is the estimation of events through online social networking. Hence, the sensing capabilities of online social networks (OSN) users have resulted in different networking applications. The application of sensor networking applications to mobile networks brings new challenges and opportunities. For instance, discusses the integration of the sensor networking paradigm with vehicular networks. The authors in [12] explain that VSNs use the advantages of the mobility of vehicles and the sensing capabilities of sensor networks. Furthermore, it also proposes to bridge vehicular sensor networks with social networks.

Vehicular sensor networks collect real-time data, and social networks provide the ability to share information among the users in different locations. This bridging enables the combination of data coming from the vehicular sensor networks and social networks. Application-specific information can also be from the collected data via estimation [1].

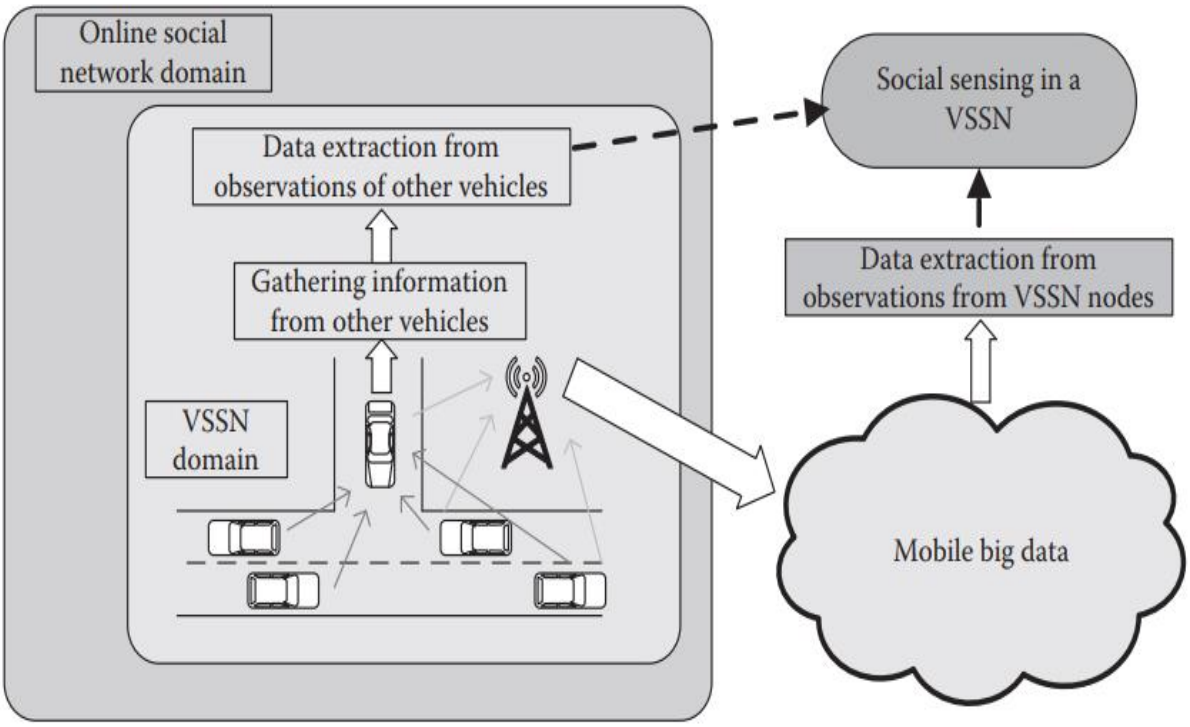

Figure 2.4, VSSN architecture and its relation to OSNs and VSNs (Adopted from [1])

Possible Application Areas

Vehicular Social Sensor Network (VSSNs) provides dynamic social sensing mechanisms. Movement of the users causes changes in the number of social nodes that participate in social sensing frequently. Furthermore, the link conditions and neighboring relations change frequently due to the mobility of social sensors. VSSN nodes are also in direct interaction with other VSSN nodes and the environment. Hence, the dynamic network architecture and direct interaction with the environment result in diverse application areas. Possible applications of VSSNs are as follows [1]:

- Up-to-date traffic information via vehicular social sensing and the determination of the traffic paths according to the traffic and road conditions.
- Real-time weather condition tracking based on information exchange in the VSSNs formed within a local region
- Real-time spot (on roads) measurements via reporting of the social sensors in vehicles to avoid traffic accidents beforehand

- Real-time event detection via social sensors; for example, drivers in the opposite lane may inform each other that there is an incident (this can be regarded as an example of a distributed social sensing mechanism with peer-to-peer communications)
- Real-time information sharing via social sensors; for example, drivers may share the possible parking spots in a crowded area to help other drivers
- Peer-to-peer vehicular social sensing on the road in case of no Internet connection

2.1.9 Accident Detection and Reporting System

The authors of [27] came up with a framework for automated driving system testable cases and scenarios. They develop an example of a preliminary test framework for Automated driving systems (ADS) that are in development and may come to market in the near to mid future. The following steps were conducted to support the development of the sample test framework.

- Identify concept ADS
- Identify attributes that define the operational design domain (ODD)
- Identify object and event detection and response (OEDR) capabilities
- Identify and assess failure modes and failure mitigation strategies

According to [2] the Three Researchers introduced a new system in automobile technology which is about how to keep a 10-meter distance between one vehicle and another vehicle so that the vehicle doesn't crash or cause any traffic problem. The system aims to prevent accidents mainly due to not knowing the following distance (i.e., 10m) between one vehicle and another vehicle.

The proposed system comprises an idea of having safety while reversing a vehicle, detects any object within the following distance, and displays the distance between one vehicle and another vehicle to the driver using LCD. They have used ultrasonic sensors to detect any vehicle on both the front and backside of our vehicle. This system is also used in a large crane which is mainly operated in the harbor area. If the car reaches 10 meters, the green color light glows. At an 8 Meter distance, yellow color light glows showing the warning. When it reaches a 5-meter distance red color light glows and which indicated that the driver is not following the indicators of the roadside.

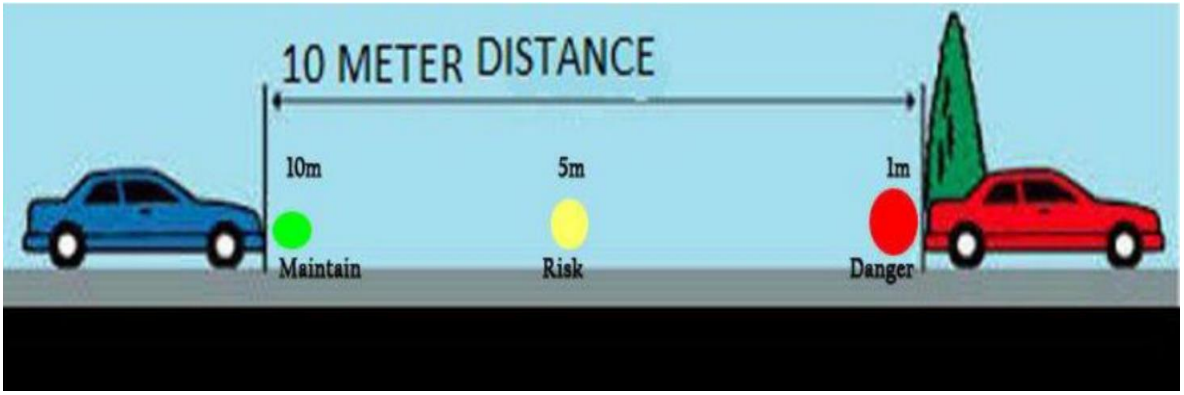

Figure 2.5 Proposed model for vehicle-to-vehicle proximity detection [2]

Many people around the world use their private vehicles as the main method of their daily transportation. Although, the use of such a method has its advantages, road accidents are very common and is considered as one of the leading causes of fatalities. Difficulties and delay in getting help on time after a road accident, is one of the factors that ranks road accidents as a leading cause of death. One way to minimize the severity of this is to equip vehicles with an integrated Accident detection & reporting system. Hence, the paper [31] presents an intelligent framework that can find the nearest location and report an accident place.

[31] Propose accident detection and reporting system by using GPS and GSM. The objective of this system is to minimize the delay of reaching the rescue in the accident location. This scheme is fully automated. Thus, it locates the accident spot accurately, calculates the distance, selects the nearest emergency station to the accident location and then sends alerting message to the selected location. Many kinds of research are heading to develop the features of the airbag. However, the success of any safety equipment depends on its correct implementation and specific rules to be followed.

The system detects accident from the vehicle and sends a message through GSM module. Another GSM module receives the message. A link is attached with the SMS so that the emergency station can display the exact location of the accident and its details Via Google map. Moreover, it gets detail SMS from the location of the accident. Hence there is a small variation in the coordinates, the initial value of latitude and longitude are same but fractional value changes with a small difference. Within 10 seconds if the reset button not pressed, GSM will send a message to the emergency station the example of text received when the accident occurred. It is seen that the text message consists of the coordinates in the form of Latitude and Longitude can be displayed in the Google map link [31].

2.2 Related works

Many systems have been proposed for vehicle accident detection and preventions by researchers. The accident detection methods were first based on real time traffic analysis to forecast traffic flow which deals with the change in the traffic before the occurrence of the accident. This detection technique is known as Traffic-incident detection-algorithm based on nonparametric regression which was proposed by Shuming Tang and Haijiun[30]. Similar model for traffic accident automatic detection, recording and reporting at intersection using metadata registry which was proposed by Yong-Kul Ki, Jin-Woo Kim and Doo-Kwon Baik [4]. An accident detection system on highway, proposed by In Jung Lee [5] makes use of CCTV which view flow of vehicle trace is like as level spacing distribution as Wigner distribution.

[30] Proposes a framework which utilizes the contextual state of two wheels for detecting the occurrence of accident. The architectural diagram as shown in the following figure 2.6 contains two units Accident Detection Unit (ADU) and Accident Detection Server (ADS).

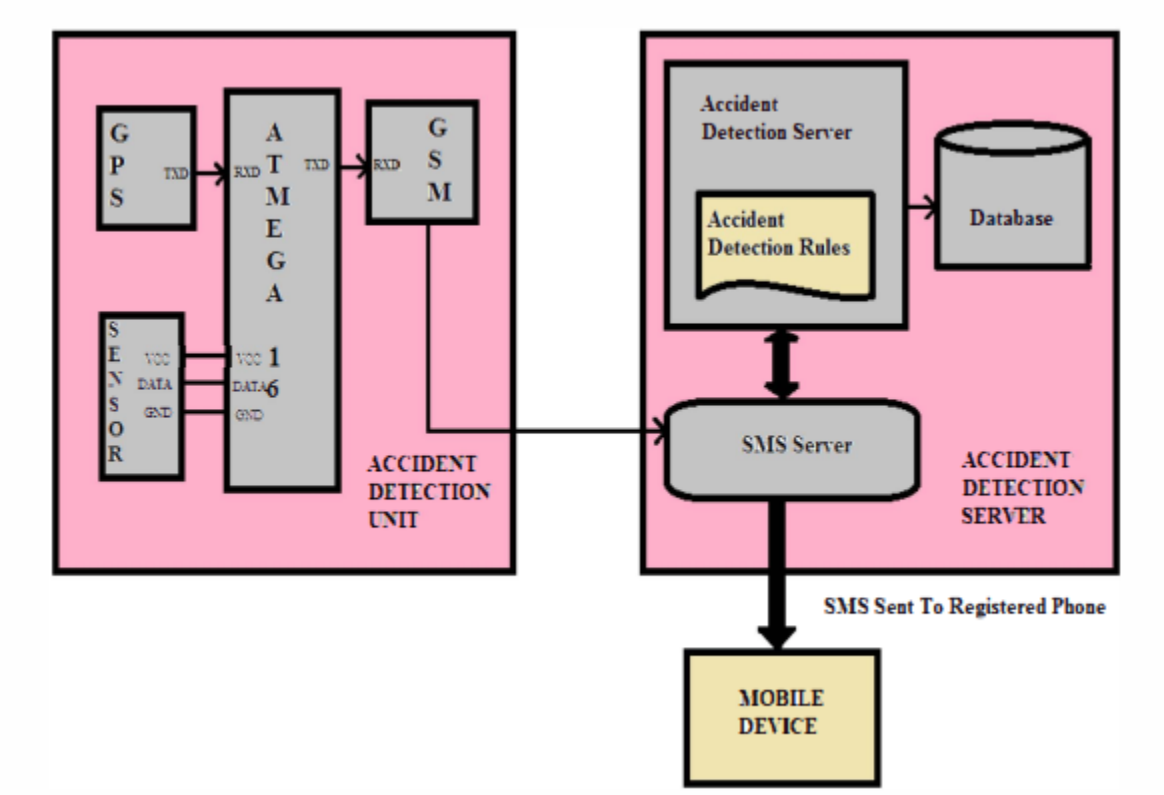

Figure 2.6 A proposed framework for accident detection and reporting system (Adopted from [30])

Lae Yin Mon1 and Khin Khin Saw, proposed vehicles detection systems for this system provides very effective in detection of over speed driving. This circuit is mainly consisting of Arduino UNO, two IR sensors, 1602A LCD and buzzer. When a vehicle is passing between the first and second sensors, the two sensors sense the object and then microcontroller program will start counting. When it passes cross the second sensors, the microcontroller will stop counting and calculate the speed as kilometer per hour. If the car's speed is over speed (100 km/h), the buzzer will be alarmed and the LED will be blinked. Then, LCD displays the speed of vehicles. Design a system of detection on highway road for vehicles. This system is to improve a device that detects over speeding of vehicle, gives warning using alarm and display vehicle's speed in LCD [23].

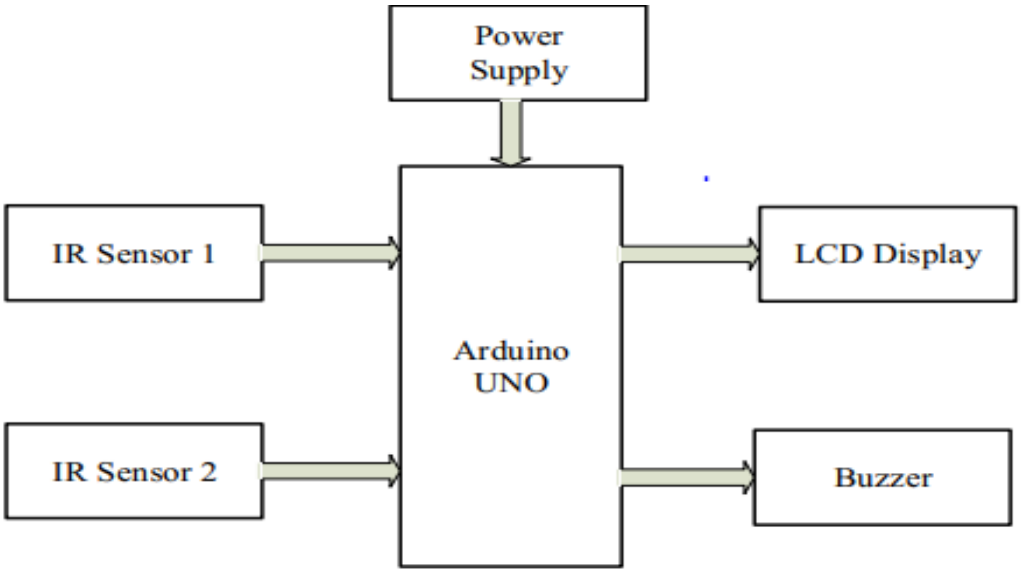

Figure 2.7 Block Diagram of the speed detection System [23]

Siquan, Min and Chuundong (2017), discussed the geophone sensor which was a new type of vehicle overloading detection system that had the ability to detect overloaded vehicles and spontaneously transfer the relevant data to a server or screen. However, it does not alert the vehicle owner from a distance.

Ramya, Palaniappan and Karthick (2012), examined a system that provides vehicle cabin safety based on an embedded system that monitors the level of toxic gases such as carbon monoxide (CO), liquefied petroleum gas (LPG) and alcohol within vehicles. Their system has the ability to send alert information

through an alarm system during dangerous situations, and also send an SMS to an authorized person through the GSM technology [24].

Ebenezer Narh Odonkor and Willie K. Ofosu, proposed vehicle overloading monitoring system with an automatic engine lock and a call alert. The accident notification was done through GPS and GSM technologies. It consists of a microcontroller, GSM module, power supply, keypad, LCD display, DC motor (car engine), load sensor, buzzer and an engine control block.

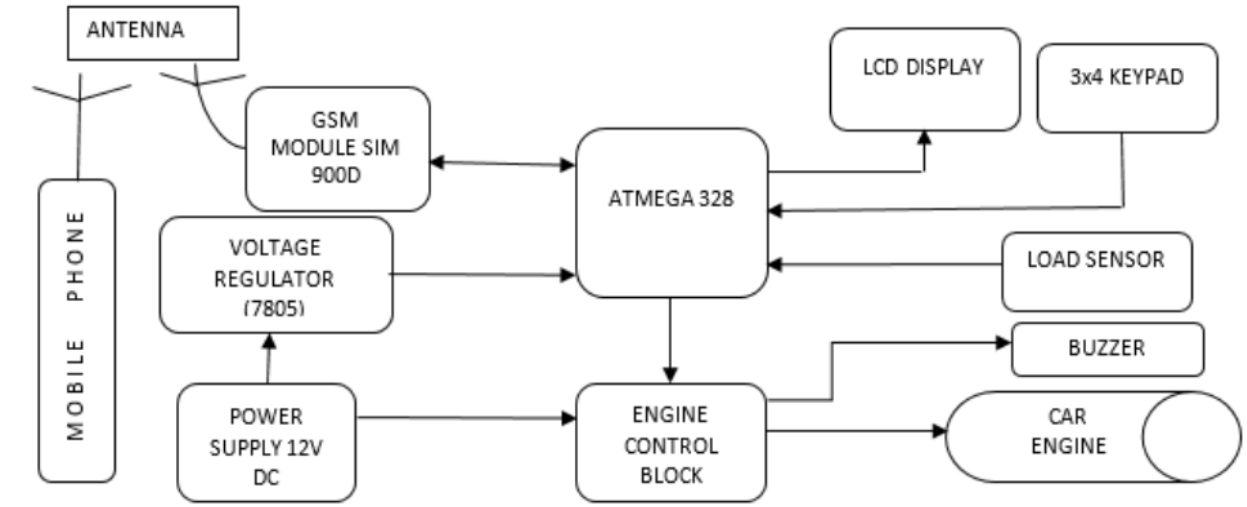

Figure 2.8 Block Diagram of the Vehicle Load Monitoring System with an automatic Engine Lock System based on GSM [24].

In vehicular networks, a vehicle sends a warning message to its neighbors as soon as identification of a risky situation. Vehicle Ad Hoc networks are important because they can provide communication for moving vehicles without the need of infrastructural services being deployed on the road. To avoid drastic results of accidents the moving vehicles need to be aware of the traffic situation by sending emergency messages using the wireless broadcast channel [5].

2.2.1 Summary for Related Works

A number of researches and articles published in a detection of vehicles' speed and net overload VSN literature shown in the analysis of these service-related works. Here are some of them which reviewed along those lines.

Author & Year	Title	Methods/ Approaches Techniques	Key Findings
Birhanu Mesfin (2021)	A Framework for Integration of Wireless Sensors Network and Object Detection System to Monitor Careless Driving: The Case of Addis Ababa City	The research followed a design science research method.	A Framework for Integrating wireless sensor networks and object detection systems to Monitor Careless Driving
Getamesay, Dereje and Kebebew (2019)	Lane Id based Selective Emergency Message Forwarding Scheme for VANET	Analyzing both primary and secondary data	A novel emergency message forwarding scheme for VANET (Vehicular Ad-hoc Network).
Saleh Dahl and Mats I (2019)	Vehicle speed measurement model for video-based systems	The intrusion line technique to measure the movement pattern vector with low computational complexity.	Proposes A model
Eric, Shawn, Michelle (2018)	A Framework for Automated Driving System Testable Cases and Scenarios	A four-stage approach was followed to identify ADS features: (1) review the literature, (2) define a framework for discussing ADS features, (3) define features and behaviors, and (4) Categorize the features.	Develop an example of a preliminary test framework for ADS that are in development and may come to market in the near to mid future
Amit, Srikrishna, Monika ,Saket, Sachin &	Automatic Accident Detection and Reporting Framework for Two Wheelers	Uses the Global Positioning System (GPS) to get the location of the accident and a message is sent through a	Presents an inexpensive but intelligent framework that can identify and report an accident for two-wheelers.

Mujeeb (2014)		microcontroller or a mobile device	
Fatma AL, Mohammed & Suliaman (2019)	Integrated Accident Detection & Reporting System	equip vehicles with an integrated Accident detection & reporting system	propose accident detection and reporting system by using GPS and GSM

Table 2.2 Summery of related research works

Chapter Three

3.1 Research Methodologies

This chapter discusses the research design and methodology in detail that is intended to provide the mechanism to solve the research problem of the study. A research methodology is defined as "the general approach the researcher takes in carrying out the research project", (Leedy & Ormrod, 2016 p. 8). It will consist of the research philosophy together with the research method and techniques. The choice of a suitable methodology is based on the nature of the research problem and the researcher's philosophical orientation and assumptions. According to (Leedy & Ormrod, 2016 p. 74), research design means the strategy that helps to solve the problem and simply it means planning. In addition to this, it includes the procedures, data collection, and data analysis techniques that the study will use to conduct the research.

3.1.1 Research Design

The research design is the overall map for concerning the theoretical research problem to relevant and practicable empirical research which means it provides a plan or a framework for data collection and analysis. The following is a conceptual model indicating the research process that is consists of a number of consecutive and related activities.

3.1.2 Research Approach

The study employed a hybrid strategy to analyze data collected through questioning and interviews (qualitative and quantitative).

And also, appropriate mining is carried out based on the established objectives and an assessment of the available data. The CRISP-DM Industry type of model was chosen and run on the available data for the goal of producing models and spotting trends.

3.1.2.1 Cross-Industry Standard Process for Data Mining (CRISP-M)

Data mining is a process which finds useful patterns from large amount of data [37]. Data mining is a process of extraction of useful information and patterns from huge data. It is also called as knowledge discovery process, knowledge mining from data, knowledge extraction or data /pattern analysis (Bharati M. Ramageri).

Accordingly, to achieve the objective of this study Cross-Industry Standard Process for Data Mining (CRISP-M) has been used. This is because CRISP-DM has been widely applied in data mining studies and also it is flexible for different data (Kantardzic, M. (2003)). CRISP-DM model is divided into six phases. These are Business understanding, Data understanding, and data preparation, Modeling, Evaluation and Deployment.

Business Understanding

In order to understand the business: structured and semi-structured interview conducted with two domain experts. In addition to this, the researcher conducted field measurements with network optimization team within how to operate equipment's. Direct observation of the work done is held by the researcher for basic practical understanding.

Data Understanding

the database of traffic accident data is the primary target data set of the study. In this phase, the original data collected for this research should be described briefly. Its description includes listing out attributes with their respective values, missing and outlier and evaluation of their importance to the research goal. Additionally, careful analysis of the data and its structure is done together with domain experts by evaluating the relationships of the data with the problem at hand and the DM tasks to be performed.

Data Preparation

Data preparation step concerns deciding which data will be used as input for DM methods for subsequent steps. It involves data integration, attribute subset selection, generalization, discretization and concept hierarchy and data cleaning are performed. To check the completeness of data records, to remove or correcting noise and missing values and finally to make more suitable for the data processing task. As a result, preprocessing and cleaning will be employed to fill missing values or to clean noisy data.

Data Mining Task

Based on the identified goals and the assessment of the available data, appropriate mining Model is chosen and run on the prepared data, having this purpose CRISP-DM Industry type of model has been used for the building models and identification of patterns. For the prediction purpose the researcher used K-nearest Neighbor, SVM support vector machine and Logistic regression algorithm based on its performance relative to other prediction algorithms.

Data mining tool selection: After the collections of data to make suitable for DM software the researcher used Orange. Orange is chosen because of its widespread use in different DM researches and familiarity of the researcher with the software.

Evaluation

to evaluate the performance of the classifiers both objective and subjective evaluation techniques are applied. For the objective part confusion matrix, precision, recall, TP, FP, f-measure and the area under the ROC curve are used. For the subjective part some interesting rules from the selected model with purposely selected domain experts to measure how much they are agree with the final interested rules extracted by the selected model.

3.1.2.2 Mixed approach

The study also employed a mixed approach (qualitative and quantitative) for the analysis of data gathered through questioner and interview through Microsoft excel tool. Nowadays, a number of researchers applied mixed approach to reduce the weakness of qualitative and quantitative approaches. For instance, the weakness of qualitative design includes knowledge produced may not be generalized to other people or other settings (i.e., findings may be unique to the relatively few people included in the research study); the results are more easily influenced by the researcher's personal biases and idiosyncrasies, etc. Quantitative design weakness also includes knowledge produced may be too abstract and general for direct application to specific local situations, contexts, and individuals (Kothari, 2004).

The evolution of mixed approach is associated with the field of social psychology. It intends to converge, i.e., to triangulate the different qualitative and quantitative data sources (Creswell, 2003). Among the mixed approach strategies, transformative strategy is employed in this inquiry. According to Creswell (2003), transformative strategy may or may not be sequential during data collection. Priority

can be given to either of the designs (qualitative or quantitative). The strategy has a theoretical perspective to guide the study. Accordingly, the researcher has given weight to qualitative design for the sake of closely looking participant's social situation, the way they make sense of their lives or experiences, and perception and/or attitude.

With all its weakness, this strategy is chosen owing to the fact that it gives the researcher discretion to decide on the issue of priority and enables to use theoretical frameworks that a researcher to rely on. Needless to say, the problem at hand benefits from the advantages of mixed approach. Hence, the approach is worthy of being adopted for this study. Because, the research is grounded on qualitative data gathered through interview, questionnaire and other sources coupled with the quantitative statistical reports (Yared Debebe, 2014).

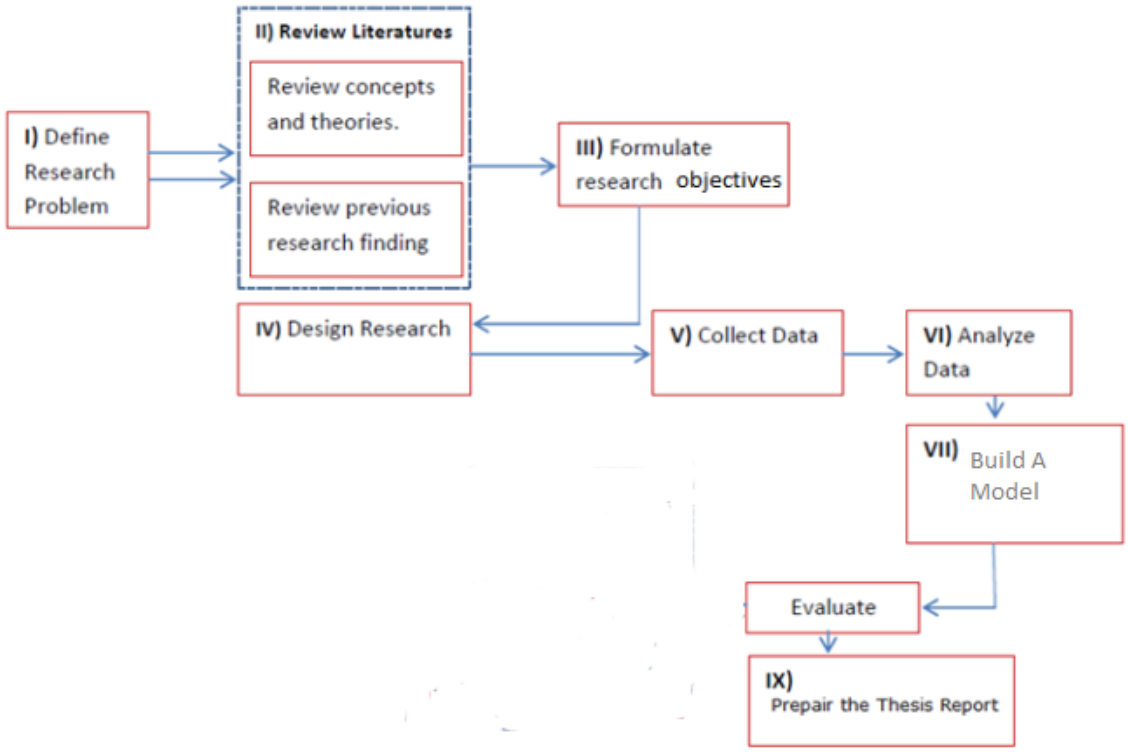

Figure 3.1 Research process flow chart (Adopted from [2])

The Overview of the process steps are described as follows:

1) Define Research Problem

The study defines the specific research problem and justifies the value of the solution which is done via interviews of concerned stakeholders, distributing questionnaires and with continuous analysis of literature reviews of frameworks on related areas and others.

2)Review literatures

The study tries to review literatures so as to understand the current sciences, technologies and standards related to this study. And also presents some background information of the available literature related to the development of traffic accident rates, analysis on the vehicles speed limit and overloading as causing factors for traffic accident.

3)Formulate Research Objective

Define the objectives the researchers infer the objectives of a solution from the definition of the problem and knowledge of what is possible and feasible. The objective of the solution is derived from the problem statement.

4) Design A Research

Research design is a plan that guides the investigator in the process of collecting, analyzing, and interpreting observations (Yin, 2014:28). It articulates what data is required, what methods are going to be used to collect and analyze this data, and how all of this is going to answer the research question. Hence this research is qualitative and quantitative which intended to explore and design vehicle speed and net overload detection and message forwarding model. This section discusses the processes and techniques used in carrying out the study. It also gives a description of the respondents including information on the study population, the number of respondents and how they were selected. It also provides an outline of research design and the instruments for data collection.

5)Data collection

Regarding the data collection method, the researcher used semi-structure interview, various documents, observations, questionnaires and self-administration. The interviews were conducted in Amharic

language and the questionnaires are prepared in English language. And for data mining process the researcher took the original data (traffic accident data) collected by Addis Ababa traffic office.

6)Analyze Data

This portion of the survey is concerned with the respondents to understand respondents who participate in filling the questionnaire for this research. Respondents are requested to fill their information and answered the prepared questions. And also, a data mining tool Orange is used to make a model and analyze the traffic accident data.

8) Designing A Model

Modeling is one of the major tasks which are undertaken under the phase of data mining in CRISP methodology.

8)Evaluation

Observation and measuring how well, the model supports the solution to the problem. The evaluation is aimed at comparing the objectives of the solution to the actual results derived in the design and development phase.

3.1.2 Data Collection and Analysis

This study will be going to apply data collection techniques interview of drivers, foot-travelers, institution how gives driver license and traffic polices mainly. The data collection sampling technique is a mixed type of sampling techniques, means it will be going to apply simple random sampling from a probability sampling technique to interview passengers and pedestrians randomly because peoples almost have the same understanding about drivers and car accidents. But, to interview the drivers the research will apply the purposive or judgmental sampling from non-probability sampling technique. Because purposive sampling is useful to collect data from purposely selected drivers that drives for many years and drivers that committed vehicle accidents so far while driving.

Secondary data source like books, reports, journal and conference articles and white paper from websites of reliable authors and organizations have been used to get information about detection of vehicles speed and net overload-based message forwarding framework for transportation sector.

3.1.3 Population and Sample

The study consisted of the selected traffic police of the Addis Ababa Traffic Management Agency, selected vehicle drivers and pedestrians/passengers as a sample of the population. The sample is done on 20 traffic police, 30 vehicle drivers and 50 pedestrians & passengers, the questionnaire is distributed randomly individually of those who are engaged in Addis Ababa city to gather a good response rate considering their situations well.

3.1.4 Sampling

For the purpose of this study, purposeful sampling technique is selected. It is widely used in qualitative research for the identification and selection of information-rich cases for the most effective use of limited resources (Patton, 2002). This involves identifying and selecting individuals or groups of individuals that are especially knowledgeable about or experienced with a phenomenon of interest (Cresswell & Plano Clark, 2011). The target population is drawn from of drivers, foot-travelers, institution how gives driver license and traffic polices mainly. And the sample in this research consists of 20 traffic police 30 vehicle drivers and 50 pedestrians & passengers.

Interview

In qualitative research approach interview is the most used data collection method. As a result, in this study the researcher has used interview as a data collection method. Patton (2002) pointed out that there are three main kinds of qualitative collections these are:

- **Interviews**: Open-ended questions and probes yield in-depth responses about people's experiences, perceptions, opinions, feelings and knowledge. Data consist of verbatim quotations with sufficient context to be interpretable.
- **Observations**: Fieldwork descriptions of activities, behaviors, actions, conversations, interpersonal interactions, organizational or community processes, or any other aspect of observable human experience. Data consist of field notes rich detailed descriptions, including the context within the observations were made.
- **Documents**: Written materials and other documents from organizational, clinical, or program records, memoranda and correspondence, official publication and reports; personal diaries, letters, artistic works, photographs, and memorabilia, and written responses to open-ended

surveys. Data consist of excerpts from documents captured in a way that records and preserves context.

Questionnaires

In Quantitative research approach survey questionnaire is used as a primary data collection method. Questioner is very popular in survey research because they provide a greater uniformity of responses and more easily processed (Gay, Mills and Airasian 2012). In this study, we use a quantitative research approach for data collection through questionnaires, questionnaire data collection method applied for traffic police of the Addis Ababa Traffic Management Agency and organizations who provide vehicle services.

Ethical Consideration

The purpose of the study has been disclosed to the interviewees before the interview and asked their willingness to participate in this academic study. And during the interview they have been asked their permission to record the interview. In order to assure the confidentiality, the interviewee names and details were not disclosed.

Sampling Techniques

The study area was in Addis Ababa where frequent traffic accident occurs. It was identified through the data collected from the Addis Ababa City Traffic Police office archive. To select informant quota, systematic and purposive sampling were employed. Quota and systematic sampling are used for the sake of obtaining diverse information from the target black spot areas. Because, the problem prevails in one place for example, Meskel square may not be similar with ayat. Purposive sampling also gives discretion to choose informants based on their familiarity to the issue.

Accordingly, ten traffic policemen's respondent from Addis Ababa Road Transport Authority and Traffic Police Traffic Accident Control and Inspection Office, were selected purposively. Thirty pedestrians were involved in the unstructured interview. Five black spot areas, namely Uraile, Bambis, Bete-megist, Kasanchis and Behirawi, were purposively selected each having six quota distributions. And also, ten driver respondents included.

3.1.5 Data Collection Techniques

The required data were gathered from primary and secondary sources. Thus, the primary sources of data were collected through the following instruments.

A. Interview

Semi-structured and unstructured interviews were conducted with informants. Informants from Addis Ababa Traffic Police and Traffic Accidents Control and Inspection Office

B. Questionnaire

In addition to in-depth interviews, the researcher utilized open ended questionnaire for the sake of obtaining, giving the opportunity for respondents to vent and explain their thoughts towards the subject of inquiry. To check the validity of the questionnaire a pilot test was conducted. Finally, ten questionnaires were distributed to respondents.

Secondary sources were collected from Addis Ababa City Police Commission. For the purpose of analyzing what is gathered through primary sources and in an effort of making it reliable, reviews of relevant books, journals, articles, and reports (extracted from international organizations such as WHO, World Bank and others) were conducted. The secondary sources were then integrated with the primary sources so that the research will be comprehensive enough to capture elements of the phenomenon under study.

Data Analysis

The data gathered from different sources is analyzed using mixed approach. The data gathered through interview, questionnaire, recorded data from traffic authorities and other institutions are transcribed into themes and analyzed from the perspective of human security approach and literatures. Tables, charts and figures are used to clarify and substantiate explanations.

CHAPTER FOUR

EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Introduction

In this chapter, the researcher describes the techniques that have been used in developing a model to predict the occurrence of traffic accident. This study incorporated the typical stages which have been organized according to CRISP Industry processing model. Here the researcher discusses the experimentation process by relating the steps followed, the choice made, the task accomplished, the result obtained and evaluation of the model.

4.2 Data Collection

Collecting, analyzing and understanding the content and structure of the data available is one of the most important tasks that need close attention. With respect to this specific research the sole source of data about an accident is the daily accident report form to be filled and reported by the traffic police officers. It consists of full details about a given accident. Through successive update, the office keeps this data in excel file at file format. In line with this, the raw data was initially collected regarding vehicle accidents from the office's flat database and from daily accident record file.

4.3 Dataset Description

The study used data from the Addis Ababa Traffic Office The initial data source for the study contained traffic accident records from May 13, 2008 E.C up to August 30, 2014 E.C. The original dataset consists a total number of 5,545 instances and 18 attributes. Because of the unavailability of important attributes in the accident data the data set selected for this specific research covers the time from May 13, 2008 E.C up to August 30, 2014 E.C, which is a total of 5,207 records.

The following table shows attributes of the dataset with its description.

S.No	Attribute name	Type	Description
1	Accident ID	Number	A number to identify a ggiven accident uniquely
2	Accident Date	Date/time	Date of an accident
3	Accident Day	Text	The day of an accident in week days
4	Accident Time	Date/time	The time of the accident
5	Driver Age	Number	Age of the driver
6	Driver Sex	Text	Gender of the driver
7	Driver Educational Level	Text	Educational level of the driver
8	Ownership of the car	Text	Driver-vehicle ownership Driver Vehicle Relation
9	Driver Experience	Text	Driving experience of the driver
10	Driver License Level	Text	Status of the driver's license
11	Vehicle Type	Text	The type of the vehicle
12	Vehicle Plate Code	Number	Plate Number, Home, Government, Private company, NGO
13	Car Defect	Text	Problem or defect of the vehicle
14	Defendant's vehicle ownership	Text	Private, corporate, Government
15	Weather Condition	Text	The weather condition at the time of acidulate
16	The type of road	Text	
17	The condition of the road	Text	
18	The cause of the accident	Text	**Causes for the accident**

Table 4.1 List of attributes with description

4.4 Preparing Data for Analysis

And in order to get maximum out of the data using data mining tools and techniques, the data need to be cleaned and organized. Among others, data collection, data cleaning, attribute se lection, data formatting and transformation, and dimensionality reduction are the most important activities under data preparation, which finally resulted in creating target data set.

In this research, original dataset was collected and organized in Amharic language by Addis Ababa traffic police office. For data preprocessing the researcher translated into English version and prepared for analysis. The data was in an excel file format with 18 attributes, after attribute selection is done the attributes become 15. In line with this, the raw data was initially collected regarding vehicle accidents from the office's flat database and from daily accident record file. After keying the manual format of data, which took considerable time, careful analysis of the data and its structure was done together with the domain experts by evaluating the relationship of the data with the problem at hand.

4.5 Model Building

Modeling is one of the major tasks which are undertaken under the phase of data mining in CRISP methodology. In this phase several data mining techniques are applied and their parameters are adjusted to optimal values. Typically, different techniques can be employed for similar data mining problems. Some of the tasks include: - selecting the modeling technique, experimental setup or design, building a model and evaluating the model [33].

4.6 Selecting Modeling Technique

Selecting appropriate model depends on data mining goals. Consequently, to attain the objectives of these research three classification techniques has been selected for model building. The analysis was performed using orange environment. Among the different available classification algorithms in Orange, KNN, SVM and logistic Regression are used for experimentation of this study. The researcher selected the above algorithms, easy of understanding and interpretation of the result of the model.

4.7 Experiment Design

Before building a model, we need to generate a procedure or mechanism to test the model's quality and validity. For instance, in supervised data mining tasks such as classification, it is common to use classification accuracy measure or error rates as quality measures for data mining models. Besides, other standard measure including precision, recall, sensitivity and specificity are available. Therefore, the test design specifies that the dataset should be separated into training and test set and builds the model on the training set and estimates its quality on the separate test set. The process of building predictive models requires a well-defined training and validation protocol in order to ensure that most accurate and robust prediction. In this research 5545 datasets are used for training and testing. Orange 3.6.0 software has used to set up and measure the quality, validity and test of the selected model. For purpose of this study k-fold (10-folds) cross validation and percentage split test options are used because of its relatively low bias and variations. Accordingly, the datasets are randomly partitioned equally into ten parts. Hence, 90% of the dataset is for training and 10 % for testing for former and the dataset are partitioned in to percentages (70-30) splits option meaning 70% of the dataset for training and remaining for testing). Moreover, Witten and Frank reported that 10-fold cross validation has been proved to be statistically good enough in evaluating the performance of the classifier. To build the model of this research 14 independent and 1 dependent variables or attributes are used. The details of the selected attribute are discussed in above table 4.1.

4.8 Running Experiments

As discussed in the above section 4.2., the researcher applied two methods for running different kinds of experiments namely, 10-fold cross validation and percentage splits. Based on the above methods establishing scenario for model to be developed is very important to see the model result and analysis of each result, to compare the result of one model with the previous one and finally help us to find out the outperforming model based on criteria of evaluation. Consequently, for both of the methods following scenario has been done for each of three selected model with default parameter value of Orange 3.32.0.

4.9 Model building using K-Nearest Neighbor (KNN)

K-nearest neighbor is a supervised learning algorithm where the result of new instance query is classified based on majority of K-nearest neighbor category. The K-Nearest Neighbor (KNN) [33], is

one of the very powerful algorithms of machine-learning algorithms is used in this research. There are various other machine learning algorithms like Decision trees, Support vector machine (SVM), K-nearest neighbor, K-means clustering, Naive Bayes, Random Forest etc. The performance of these algorithms is depending on the data set provided. So before deploying any model selection of particular algorithm play very important role in the performance of the model.

Advantages of KNN

- KNN is called Lazy Learner (Instance based learning) as it does not learn anything in the training period. It does not derive any discriminative function from the training data. In other words, there is no training period for it. The training dataset is stored which is used to learns which helps in making real time predictions. As compared to other algorithms which requires training e.g., SVM, Linear Regression etc, KNN algorithm much faster.
- The KNN algorithm is one of the unsupervised machine learning algorithms, so training is required before making any predictions. The new data item is added in to the corresponding cluster and does not influence the accuracy.
- To implement KNN, only two parameters required i.e., the value of K and the distance function used to compute distance to put the data items in to the appropriate cluster.

Disadvantages of KNN

- Since in KNN the distance of new data item is computed with all other existing data points which is very time consuming in case of large data sets and degrades the performance of the algorithm.
- Does not work well with high dimensions. The KNN algorithm does not work well with high dimensional data because with large number of dimensions, it becomes difficult for the algorithm to calculate the distance in each dimension.
- Need feature scaling: We need to do feature scaling (standardization and normalization) before applying KNN algorithm to any dataset. If we do not do so, KNN may generate wrong predictions.
- Sensitive to noisy data, missing values and outliers: KNN is sensitive to noise in the dataset. We need to manually impute missing values and remove outliers

In my experiment, I have simulated this work using orange and use dataset to collect the data and optimize such data. An analytical task has done using KNN algorithm. This helps to find a prediction analysis to get better results. The objective of our work to find Couse of accidents in a city based on different parameters. After experiment done through KNN, result shows as the following table.

Algorithm	**Area Under Curve (AUC)**	**Classification Accuracy (CA)**	**F1 Score**	**Precision**	**Recall**
K-nearest Neighbor	0.616	0.246	0.238	0.238	0.246

Table 4.2 experimental result in KNN

4.10 Model building using Support Vector Machine (SVM)

SVMs were introduced first by Cortes and Vapnik for training linear machines efficiently (Cortes, et al., 1995). One of the simplest tasks for such models is the linear separation of various, multidimensional vectors representing two classes. SVMs find the hyper plane having one dimension less than the original dimensionality of the vectors separating the two classes. The target of the separation is to maximizing the distance of the elements of the classes from the hyper plane on the different sites of it. The closest class elements from the two classes are called support vectors (Hamel, et al., 2009).

One of the most effective tools of the SVMs is using kernel functions. The idea to ensure higher class separation capability is to transform the input space into another space having usually higher dimensionality. This space is called as feature space. When an appropriate transformation is found for the problem analyzed, typically it results better modeling accuracy and usually it results no significant increase in computational time. Another very important feature of SVMs is that the target function of their training for building up its kernel is quadratic and convex having no local but a global extreme (Cristianini, et al., 2000). These features and their promising applications result that SVMs are very popular in machine learning applications. SVMs were further developed and extended to handle much more complex assignments, e.g., multiclass classification even if when the classes are not linearly separable. In this case the target of the SVM is to minimize the number of misclassified class elements

together with the maximization of the distance between the separating hyper plane and the support vectors (Hamel, et al., 2009).

Based on the promising results in classification assignments SVMs were extended to realize also estimation tasks. Similar to the classification their estimation capabilities are considered also successful even if the dimensionality of the input space is very high (Hamel, et al., 2009). Variety of further SVM improvement research activities were done and are under development also today, one of such a very important result is the introduction of the SVM type called Least Squares SVM that is able to solve the model fitting through solving a linear equation set instead of the original extreme search algorithm included in SVMs (Suykens, et al., 2002, Valyon, et al., 2005). The input-output configuration strongly influences the accuracy of the developed model especially if dependencies between parameters are non-invertible [34]. After experiments are done through SVM, result shows as the following table.

Algorithm	**Area Under Curve (AUC)**	**Classification Accuracy (CA)**	**F1 Score**	**Precision**	**Recall**
Support Vector Machine (SVM)	0.636	0.229	0.217	0.212	0.229

Table 4.3 experimental result in SVM

4.11 Model building using Logistic Regression

Logistic regression model is one of the most widely used models to investigate independent effect of a variable on binomial outcomes in medical literature. However, the model building strategy is not explicitly stated in many studies, compromising the reliability and reproducibility of the results. There are varieties of model building strategies reported in the literature, such as purposeful selection of variables, stepwise selection and best subsets. However, the principal of model building is to select as less variables as possible, but the model (parsimonious model) still reflects the true outcomes of the data [34]. In this research, I will show the experimental results of accident dataset through logistic regression.

Algorithm	Area Under Curve (AUC)	Classification Accuracy (CA)	F1 Score	Precision	Recall
Logistic Regression	0.680	0.300	0.224	0.223	0.300

Table 4.4 Experimental result in Logistic Regression

4.12 Results and Discussion

I created predictive models for detecting causes of traffic accidents using Accident dataset from the Addis Ababa traffic police office with different machine-learning algorithms. The best results were achieved through logistic regression. Generally, most of the algorithms used achieved mean AUCs greater than 60%. The best algorithm (logistic regression) produced an AUC of 68%.

The methods and concepts for building predictive models for use in traffic accidents, as well as the challenges and difficulties faced when analyzing traffic accidents data, have been well described. These models are generally constructed using conventional statistical techniques such as logistic regression. The main studies have compared the results obtained through using a specific technique with the results obtained through logistic regression.

In comparison with results from logistic regression on the same dataset, the models created using support-vector machines and k-nearest neighbor produced slightly better results. In the first, models created using artificial logistic regression on data from traffic accidents produced better results (AUC = 68%) than models created using support-vector machines (AUC = 63.6%) and k-nearest neighbor (AUC = 61.6%).

Algorithm	Area Under Curve (AUC)	Classification Accuracy (CA)	F1 Score	Precision	Recall
K-nearest Neighbor	0.616	0.246	0.238	0.238	0.246
Support Vector Machine (SVM)	0.636	0.229	0.217	0.212	0.229
Logistic Regression	0.680	0.300	0.224	0.223	0.300

Table 4.5 Summery on Experimental results

4.13 Sensors

4.13.1 Speed Sensors

Sensors are installed in the form of an electric pole vertically in the order of object detector, speed sensor, proximity sensor from up. The camera and LED can take the upper positions of the poll. This form is chosen due to the idea of saving space, good lookup of the system components, and easy management of in one place.

The Number of sensor requirement can be calculated by considering things which are described under the following formula [2].

$$\mathbf{N=}\ 2.A.\ \pi$$
$$r2\sqrt{27}$$

Where N is the number of sensor Nodes to calculate, A is the area that is supposed to be under monitoring and the effective sensing radius of the sensor node is represented by r.

4.13.2 Overload Detector

The whole detection system consists of detection nodes and transmission networks. Detection nodes are used for sampling, the data is transmitted to the vehicular Adhoc network (VANET) so that the waveform generated by ground vibration can be observed in real time. Data is stored, processed by the

algorithm and disseminate into VSN. When the overload is detected, the server will send out a warning massage to the nearest vehicles within a fixed diameter (100m).

The load sensor is used to detect or sense the change in the weight of the vehicle. If the voltage signal from the sensor is more than the reference voltage, the sensor sends a signal to the microcontroller. The microcontroller receives a signal from the sensor and then provides an output with respect to the input signal given by the sensor. When overloading is detected, the microcontroller sends a signal to the Vehicular Social Network (VSN) [24].

4.13.3 Vehicular Ad hoc Network

A Vehicular Ad-Hoc Network (VANET) is a wireless system, which facilitates data exchange, security and solace among the vehicles within a broadcasting range by using some of the static equipment. Ad-Hoc Network is available for the travelers inside the mobile vehicle without any entangled association and server correspondence by using an electronic gadget. Every vehicle furnished with VANET equipment is a hub in the Ad-Hoc network that can receive and hands-off messages by the remote system. In vehicular Ad-Hoc system utilization of distinctive appointed administration advance must be indispensable among the vehicles. (e.g.: Accessibility of Wi-Fi for simple, exact, powerful and basic correspondence among vehicles on element versatility). In the wireless protocols Wireless Access in Vehicular Environments (WAVE) is meant for communication between vehicular nodes within the direct communication range (1000m). Emergency applications are essential in a large distance (e.g., traffic status warning). A technique to perform multi-hop packet delivery to one or more nodes is one of the main tasks [5].

Vehicular Ad-hoc Network (VANET) is the most studied data delivery methods in vehicular communication (Yousefi et al., 2006). VANET entails the creation of an ad-hoc based communication mesh network of highly dynamic vehicular nodes. Packets being forwarded in the mesh nodes (i.e., vehicles) need to know the structure of the mesh network, the direction to forward the data traffic based on the vehicle location and route to the destination nodes/ vehicles. Multi-hop routing protocols for VANET provide different options that could be used for CVT applications based on their specific latency requirements (Saleet et al., 2011; Al Rabayah and Malaney,

2012; Tee and Lee, 2010; Jarupan and Ekici, 2010; Sahu et al., 2013; Dey et al., 2015b).

Real time traffic data collection is one of the major functions of transportation agencies to inform travelers and provide reliable transportation services. Historically infrastructure-based sensors (e.g., inductive loop detector, video cameras, infrared sensors, blue-tooth) have been extensively used for traffic data collection (Leduc, 2008; Chowdhury and Wang, 2007). The coverage of the infrastructure-based traffic data collection system is very limited and resource intensive to cover all highways. Most transportation agencies deploy roadway sensors to monitor traffic condition on major corridors, however, a growing number of agencies have been using services from private traffic data providers to report real-time traffic speeds and travel times for large transportation network (Herrera et al., 2010). [9]

The transportation data service providers aggregate real time traffic data collected from multiple sources, including mobile phones, freight fleets, and GPS devices, that cover a broader highway network in comparison to limited roadway sensor coverage (INRIX-Iowa DOT, 2015) [9].

4.13.4 Vehicular Social Network

With the rapid growth of smart devices and advent of advanced communication and computing capabilities, it is possible for drivers and passengers to communicate and socialize with other commuters on the roads. A Vehicular Social Network (VSN) is an emerging field of communication where relevant concepts are being borrowed from two different disciplines, i.e., vehicular ad-hoc networks (VANETs) and mobile social networks [12].

VSNs enable commuters to socialize and exchange information with other commuters on the roads. The vehicles in VSNs communicate opportunistically in the distributed architecture and information sharing with neighbors merely happens in given circumstances, such as for a particular social relationship and geographic position. VSNs are constructed on-the-fly and have short life resulting in a highly dynamic network. Data dissemination in VSNs presents several challenges. It is hard to understand and exploit the social relationship and behaviors of nodes in VSNs to improve data dissemination in VSNs. Dynamic network topology and intermittent connectivity in the vehicular environment are two of the factors influencing information dissemination in VSNs. The applications of VSNs demand for reliable and efficient mechanism and protocols for data dissemination which may be achieved considering

different network parameters and features. Mezghani et al. present an overview of recent achievements to enhance data dissemination in VSNs [12].

4.13.5 Summary of Responses from the questioners

The questionnaire is designed Table 4.6 summary of responses from taxi drivers and long vehicle drivers which is gathered by the means of a questionnaire

Number	Questions	Strongly Agree	Agree	Uncertain	Strongly Disagree	Disagree
1	I always use sidewalk if they are available	52 %	13 %	7 %	5 %	23 %
2	Follow the commands or all traffic signs and signals	15 %	40 %	3 %	30 %	12 %
3	Have you ever witnessed any live accident	20 %	29 %	10 %	22 %	19 %
4	Do you believe driving over speed limit and net overload can be Couse for traffic accident in Addis Ababa?	29 %	25 %	22 %	15 %	9 %
7	Addis Ababa sidewalks roads used as market is one of the reasons for traffic Accident?	25 %	10 %	14 %	21 %	30 %
8	always cross the street with zebra line	35 %	22 %	13 %	18 %	12 %

Table 4.6 pedestrian's *response summary*

Number	Questions	Strongly Agree	Agree	Uncertain	Strongly Disagree	Disagree
1	I used to drink and drive	15 %	13 %	7 %	40 %	25 %
2	I always Respect and drive within the speed limit	25 %	53 %	2 %	8 %	12 %
3	I used to violate red light	20 %	29 %	10 %	22 %	19 %
4	Driving Monitoring tools are sufficient enough	29 %	25 %	22 %	15 %	9 %
7	I used to drive over speed limit sometimes.	25 %	10 %	14 %	21 %	30 %
8	Driving over speed is a common cause of traffic accidents	35 %	22 %	13 %	18 %	12 %
9	Sometimes I used to drive over net overload	10 %	12 %	32 %	29 %	17 %
10	Driving over vehicle net overload is a common cause of traffic accidents?	25 %	10 %	14 %	21 %	30 %
11	I believe I am concerned about ensuring traffic safety.	22 %	34 %	15 %	15 %	14 %
12	Always wear a seat belt.	32 %	17 %	19 %	15 %	17 %
13	Regularly maintain the vehicle	24 %	3 %	54 %	10 %	7 %

Table 4.7 traffic police's response summary

Number	Questions	Strongly Agree	Agree	Uncertain	Strongly Disagree	Disagree
1	Ethiopian transport Authority has efficient traffic monitoring system?	57%	23%	9%	4%	7%
2	Luck of road Traffic Monitoring System is the cause for traffic accidents in Addis Ababa?	7%	9%	10%	35%	32%
3	Vehicle speed and net overload detector is the best solution for traffic accident	26%	25%	5%	25%	19%
4	There is luck of intelligent traffic monitoring system at ETA.	25%	17%	7%	21%	30%
5	Using vehicular social network to aware closer vehicles speed and net overload can support in reducing traffic accident.	35%	22%	10%	18%	15%
6	Addis Ababa City Administration Traffic Management Agency has to apply automated systems to support traffic management processes.	36%	13%	5%	29%	17%
7	Addis Ababa City Administration Traffic Management Agency has its own vehicle speed and net overload detection systems.	30%	22%	12%	19%	17%
8	Traffic problems in Addis Ababa Streets (please tick one square along each line)	-	-	-	-	-

Table 4.8 Driver's response summary

Chapter Five

5.1 Conclusions and Recommendation

This chapter is about the general conclusions of the whole project work and future directions on the use of vehicular social networks in the intelligent road traffic system for monitoring vehicle speed and vehicle net overload.

5.1.1 Conclusions

Addis Ababa, as capital city of Ethiopia and a host for many National and International Organizations, is experiencing a rapid growth of population. Addis Ababa is characterized by poor traffic control systems and regulations, overcrowded residential patterns, lack of good engineering systems which can help reduce the high traffic congestion.

According to Addis Ababa traffic police reports and interview results, most of the accidents in Addis Ababa have been caused by human errors. Of these accidents caused by human errors, drivers were also indicated as responsible for the cases. Failure to give way for pedestrians, over speed and overload were also the main reported and observed errors of drivers.

In this project, I have designed a vehicle speed and overload detection and message forwarding system by using wireless sensors and vehicular ad hoc network. This paper presented a model that can easily detect the speed of vehicles/ human and overloaded vehicles by utilizing sensors so that it alerts the overloaded & over speed of vehicles/human. The system accumulates information of vehicles and if over speed and overload occurs it alerts by giving buzzer or awareness message to the neighbor cars through vehicular ad hoc network.

The output of the project is to build a system able to detect an accident and forward awareness message to the nearest vehicle through vehicular social network. The designed detection system continuously monitors the overload and speed of the approaching vehicle. It minimizes the difficulties of traffic department and makes ease to control the rash driving/over speed and overload vehicles on highways.

5.1.2 Recommendations for Future Work

This research work can contribute a lot towards a comprehensive study in this area in the future, in the context of our country. Hence, based on the findings of this study, the following recommendations are forwarded.

- Currently, the Addis Ababa City Transport Authority and Addis Ababa traffic bureau maintain a large volume on drivers, vehicles and other data on manual files. In order to improve traffic management, there is a need to automate the system.
- The system could be connected to a database or cloud system to have all the data of the emergency station and has a direct connection with them.
- Develop the system by interconnection the camera with microcontroller to take images of the accident spot. The system can be improvised by adding more components to know the state seatbelt, the speed of the vehicle and the braking system
- This research work is limited to monitoring over speed and overload vehicles within the specific area which is sensitive to traffic accident, but the scope needs to be widened to areas which might be an entire city and all express and high ways which are common to car accident due to high traffic of vehicle and traffic laws that allows high speed.

References

[1] Anna Maria Vegni, Valeria Loscrì And Athanasios V. Vasilakos: "Vehicular Social Networks", Msc.Thesis,Dept.Comp.Sci,Tylor And Francis Group, Newyork, 2017.

[2] Birhanu Mesfin, "A Framework For Integration Of Wireless Sensors Network And Object Detection System To Monitor Careless Driving" MSc. Thesis, Dept. Comp.Sci, St.Mary's University, Ethiopia, 2021.

[3] Benedict S.O.Bwire, "Comparative Analysis Of Machine Learning Algorithms On Weighbridge Data For Overloaded Truck Prediction", Research Project, School Of Computing And Informatics, University Of Nairobi, Nairobi, November 2019.

[4] Getamesay Haile, Dereje Yohannes and Kebebew Ababu, "Lane Id Based Selective Emergency Message Forwarding Scheme For VANET", Volume 6, Issue 4, July – August (2019)

[5] S. Karimulla Basha and T.N. Shankar, "An Efficient Emergency Message Forwarding Technique with Improved Rebroadcast Suppression For Vanets", International Journal Of Engineering, Volume 7, 1228-1232, 2018.

[6] Bilal Ahmed, Asad Waqar Malik, Taimur Hafeez and Nadeem, "Services And Simulation Frameworks For Vehicular Cloud Computing: A Contemporary Survey", EURASIP Journal On Wireless Communication And Networking, Volume 1, Issue 1, 2019.

[7] S. Sridharan And S. Ramprakash," Simulation Of Vehicle Adhoc NET Work (VANET) Using Network Simulator (NS-2) Tools", International Journal Of Engineering Research & Technology (IJERT), Volume 4, Issue 24, ISSN:2278-0181,2016

[8] Feng Xia, Noor Ullah And Xiangjie Kong,"Vehicular Social Networks: Enabling Smart Mobility", (May 2017),IEEE Communication Magazine, Volume 10 , Available:
Https://Www.Researchgate.Net/Publication/316894938

[9] Kakan Chandra Dey A Anjan Rayamajhi Mashrur Chowdhury Parth Bhavsar James Martin, "Vehicle-To-Vehicle (V2V) And Vehicle-To-Infrastructure (V2I) Communication In A Heterogeneous Wireless Network - Performance Evaluation", (2016) Transportation Research 68 168–184

[10] Siquan Hu Min Kong And Chundong She, “Design Of Vehicle Overload Detection System Based On Geophone”, *Conf. Ser.* 887 012021, 2017.

[11] Manipriya Vutturi1, Susmitha Palle2, Sathyam Bonala,” Automatic Speed Detection And Reporting System Using Arduino”, International Journal Of Engineering Research & Technology (IJERT), Volume 5, Issue 24, ISSN: 2347-3878**,**2018

[12] Azizur Rahima, Xiangjie Konga, Feng Xiaa, Zhaolong Ninga, Noor Ullaha, Jinzhong Wanga, Sajal K. Dasb “Vehicular Social Networks: A Survey”, Msc.Thesis, Msc.Thesis,Dept.Comp.Sci, Missouri

[13] Fanueal Samson, “Analysis Of Traffic Accident In Addis Ababa: Traffic Simulation” Msc. Thesis, Faculty Of Technology Mechanical Engineering Department, Addis Ababa University, Ethiopia,
, June 2006.

[14] Fekede Asefa, Demeke Assefa And Gezahegn Tesfaye. (2014) “Magnitude Of, Trends In, And Associated Factors of Road Traffic Collision In Central Ethiopia”, Volume14, Available: Http://Www.Biomedcentral.Com/1471-2458/14/1072

[15] Biru Mohammed Derese1 and Dumga Kassahun Trueha, “Modeling Frequency Of Injuries Per Vehicle Crash In Gurage Zone, Southern Ethiopia”, Ethiop J Health Sci, Vol. 31, No. 1 January 2021

[16] Itoro Michael Ikon, “Vehicle Speed Detecting App”, Msc Final Report, Dept Software Engineering, University of Greenwich, England, London,

[17] Sabih Ur Rehman, M. Arif Khan, Tanveer A. Zia, Lihong Zheng, “Vehicular Ad-Hoc Networks (Vanets) - An Overview and Challenges (January 2013),” EURASIP Journal On Wireless Communications And Networking, Avaliable: Https://Www.Researchgate.Net/Publication/258248007

[18] Kibret Gebremedhin, “Truck Overload Fine Rationalization On Ethiopian Primary Roads” Msc Final Report,Dept Civil Engineering,Addis Ababa University,Ethiopia, 2015

[19] Dastan Bamwesigye And Petra Hlavackova, “Analysis Of Sustainable Transport For Smart Cities”, MDPI,Volume 11, PP 2140,Swizerland 2019.

[20] Road Safety Performance Review Ethiopia, United Nation Commit Ion For Africa, December 2020.

[21] Transport Policy Of Addis Ababa, Ministry Of Transport, August 2011.

[22] Betelhem Belete G/Hiwot, “Conceptual Security Framework For Mobile Banking Key Authentication And Message Exchange Protocols”, Msc Thesis, Dept Faculty Of Informatics, Addis Ababa, Ethiopia, June 2017

[23] Lae Yin Mon And Khin Khin Saw,(March-April 2019) “Design And Construction Of Speed Detection System For Vehicles” IJTSRD, Volume 3, Issue 3,Available: Http://Www.Ijtsrd.Com/

[24] Ebenezer Narh Odonkor And Willie K. Ofosu, (January 2020), “Design And Construction Of Vehicle Loading Monitoring System Using Load Sensor And GSM, Volume 10, Issue 1, Available: Https://Www.Researchgate.Net/Publication/340250768

[25] Endalkachew Mekonen, “Road Traffic Accident In Ethiopia”, (May 2020), Volume 2, Issue 2, Available: Https://Www.Researchgate.Net/Publication/341616346

[26] Saida Maaroufi, “A Dynamic Messaging System For Vehicular Social Networks” Msc Thesis, Dept Computer Sc, Canada, 2021

[27] Eric Thorn, Shawn Kimmel And Michelle Chaka, “A Framework For Automated Driving System Testable Cases And Scenarios”, US Department Of Transportation, 2018

[28] Ademar Takeo Akabane, Roger Immich, Richard Wenner Pazzi, Edmundo Roberto Mauro Madeira And Leandro Aparecido Villas Xploiting,(August 2019) “Vehicular Social Networks And Dynamic Clustering To Enhance Urban Mobility Management” IMPD, Avaliable: Www.Mdpi.Com/Journal/Sensors

[29] Anna Vegni, Valeria Loscrì, Abderrahim Benslimane, Anna Vegni, Valeria Loscrì, Abderrahim Benslimane, (July 9,2019) “A Framework For The Integration Of Online Social Networks With Vehicular Social Networks” , Hal Open Science ,Avaliable : Https://Hal.Archives-Ouvertes.Fr/Hal-02178424

[30] Amit Meena, Srikrishna Iyer, Monika Nimje,Saket Jogjekar, Sachin Jagtap,Mujeeb Rahman “Automatic Accident Detection And Reporting Framework For Two Wheelers” International Conference ICACCCT, 2014.

[31] Fatma AL Saadi1, Mohammed AL Jabri1, Suliaman AL Hasani Integrated “Accident Detection & Reporting System” , , May 2019 , Global Scientific Jornal, Volume 7, Issue 5, Avaliable: Www.Globalscientificjournal.Com

[32] Fikadu Mekasha Admassie, “Road Traffic Accident: Causes and Control Mechanisms: In Addis Ababa City”, Msc Thesis, Addis Ababa University, Dep Public Administration and Development Management, Addis Ababa, Ethiopia, June 2015.

[33] Devendra Prasad, Avinash Sharma, Sandip Kumar Goyal and Dr. Amit Kumar Bindal, “System Model for Prediction Analytics Using K-Nearest Neighbors Algorithm”, October 2019.

[34] Dr. Zsolt János Viharos And Krisztián Balázs Kis “Support Vector Machine (SVM) Based General Model Building Algorithm For Production Control” , September , 2011

[35] Zhongheng Zhang, “Model Building Strategy for Logistic Regression: Purposeful Selection”, (Jan 2016) [Site]: https://atm.amegroups.com/article/view/9400/html.

[36] André Rodrigues Olivera, Valter Roesler, Cirano Iochpe, Maria Inês Schmidt, Álvaro Vigo, Sandhi Maria Barreto Bruce Bartholow Duncan, Comparison of Machine-Learning Algorithms to Build a Predictive Model For Detecting Undiagnosed Diabetes - ELSA-Brasil: Accuracy Study, June 2017.

[37] *Bharati M. Ramageri*, Data Mining Techniques and Applications, available: https://www.researchgate.net/publication/49616224_Data_mining_techniques_and_applications 2015.

Appendix 1:

ident ID	Accident Day	Accident Date	Driver Age	Driver Gender	Driver Educational Level	Ownership of the car	Driving experience	Vehicle Type	Number plate	Defective car	Defendant's vehicle ownership	Weather	The type of road	The condition of the road	The cause of the accident
Acc-000001	Saturday	5/13/2008	20	Male	10	Owner	2	Lada	3	Without	Private	Good air	Asphalt	Dry	Crash and escape
Acc-000002	Sunday	1/1/2009	20	Male	9	Employee	2	Dolphin	2	Without	Private	Good air	Asphalt	Dry	Lead
Acc-000003	Sunday	1/1/2009	29	Male	10	Owner	10	Vitts	2	Without	Private	Good air	Asphalt	Dry	I fell
Acc-000004	Monday	2/1/2009	50	Male	Masters	Owner	12	Toyota	2	Without	Private	Good air	Asphalt	Dry	Not waiting distance
Acc-000005	Monday	2/1/2009	18	Male	7	Employee	Without	Hachbach	2	Without	Private	Good air	Asphalt	Dry	When the red-light breaks
Acc-000006	Thursday	2/13/2008	40	Male	10	Employee	13	Mercedes	35	Without	Help	Good air	Asphalt	Dry	I fell
Acc-000007	Sunday	1/1/2009	47	Male	12	Employee	18	Daph	4	Without	Government	Good air	Asphalt	Dry	I fell
Acc-000008	Tuesday	3/1/2009	39	Male	Masters	Employee	22	Kia	3	Without	Private	Good air	Asphalt	Dry	Not waiting in the distance
Acc-000009	Tuesday	3/1/2009	25	Male	10	Employee	8 months	Damascus	3	Without	Corporate	Good air	Asphalt	Dry	Not waiting in the distance

Acc-000010	Tuesday	3/1/2009	35	Male	12	Employee	15	Pickup	4	Without	Government	Rain	Asphalt	Moisture	Lead
Acc-000011	Tuesday	3/1/2009	50	Male	Degree	Owner	15	Toyota	2	Without	Private	Good air	Asphalt	Dry	Not waiting in the distance
Acc-000012	Sunday	1/1/2009	24	Male	10	Employee	3	Lada	1	Without	Private	Good air	Asphalt	Dry	I fell
Acc-000013	Wednesday	4/1/2009	65	Male	9	Employee	40	City bus	3	Without	Government	Good air	Asphalt	Dry	Not waiting in the distance
Acc-000014	Wednesday	4/1/2009	30	Male	8	Employee	1	FSR	3	Without	Private	Good air	Asphalt	Dry	I fell
Acc-000015	Wednesday	4/1/2009	35	Male	9	Employee	15	Minibus	1	Without	Private	Good air	Asphalt	Dry	Not waiting in the distance
Acc-000016	Thursday	5/1/2009	37	Male	Degree	Owner	3	Toyota	3	Without	Private	Good air	Asphalt	Dry	When he stands up
Acc-000017	Wednesday	4/1/2009	47	Male	12	Employee	17	Dolphin	3	Without	Corporate	Good air	Asphalt	Dry	By denying priority
Acc-000018	Thursday	5/1/2009	38	Male	Degree	Employee	7	D4D	3	Without	Corporate	Good air	Asphalt	Dry	When driving to Wala
Acc-000019	Saturday	5/13/2008	44	Male	6	Employee	14	Corolla	2	Without	Private	Good air	Asphalt	Dry	I fell
Acc-000020	Thursday	5/1/2009	28	Male	8	Employee	2	Don't worry	3	Without	Private	Good air	Asphalt	Dry	By denying priority
Acc-000021	Thursday	5/1/2009	30	Male	Diploma	Employee	14	Sport	3	Without	Corporate	Good air	Asphalt	Dry	When it turns upside down
Acc-000022	Thursday	5/1/2009	34	Male	Diploma	Owner	2	Corolla	2	Without	Private	Good air	Asphalt	Dry	When folded
Acc-000023	Thursday	5/1/2009	34	Male	12	Employee	8	Land from Rooster	3	Without	Corporate	Good air	Asphalt	Dry	When driving to Wala
Acc-	Thursday	5/1/2009	32	Male	12	Employee	5	Minibus	4	Without	Government	Good	Asphalt	Dry	Lead

000024												air			
Acc-000025	Thursday	5/1/2009	26	Male	10	Employee	1 month	Toyota	Police	Without	Police	Good air	Asphalt	Dry	Lead
Acc-000026	Friday	6/1/2009	34	Male	9	Owner	4	Toyota	3	Without	Private	Good air	Asphalt	Dry	Open to Grabs
Acc-000027	Saturday	5/1/2009	22	Male	10	Employee	5	Lada	1	Without	Private	Good air	Asphalt	Dry	By denying priority
Acc-000028	Saturday	7/1/2009	33	Male	10	Employee	1	Toyota	3	Without	Private	Good air	Asphalt	Dry	Not waiting in the distance
Acc-000029	Saturday	7/1/2009	37	Female	12	Employee	8 months	Daph	3	Without	Government	Good air	Asphalt	Dry	She opened the car door and fell

Partial View of original Sample Data

Appendix 2:

Questioners for purposely selected participants

Questionnaires

I) Personal Background (For ALL)

1) Full Name:
2) Address:
3) Age:
4) Gender:
5) Residential Address:
6) Education:
7) Marital Status:

II) Interview (For Pedestrians)

English

1) Do you have your own a car?
2) Have you seen any traffic accident in your life?
3) What are the causal factors of the accident?
4) How could this accident have been prevented?
5) What kinds of measures are being taken to reduce and/or prevent the effects of RTAs?
6) Do you believe driving over speed limit and net overload can be Couse for traffic accident in Addis Ababa?

□□□□

1). ለስራ ወይንም ለግል ትራንስፖርት አገልግሎት የሚጠቀሙበት ተሽከርካሪ (መኪና) አለዎት?

2). የመኪና አደጋ አጋጥመዎት ወይ ደግሞ አደጋ ሲደርስ አይተው ያውቃሉ?

3). ለመንገድ ትራፊክ አደጋ መፈጠር ምክንያት ሊሆኑ የሚችሉ ጉዳዮችን ቢነግሩን?

4). የትራፊክ አደጋን እንዴት መከላከል ይቻላል?

5). የትራፊክ አደጋ የሚያደርሰውን ጉዳቶች ለመቀነስ እና ለመከላከል ምን አይነት ዘዴዎችን መጠቀም ይቻላል?

6). የመኪኖች ከተፈቀደላቸው ጭነት በላይ ጭነው እንዲሁም ከሚፈቀድላቸው ፍጥነት አልፈው ማሽከርከር ለትራፊክ አደጋ መፈጠር ዋነኛ ምክንያት ናቸው ብለው ያምናሉ? ለምን?

III) Questionnaires (For Drivers)

1) I used to drink and drive.
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree
2) I always Respect and drive within the speed limit.
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree
3) I used to violate red light
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree
4) Driving Monitoring tools are sufficient enough.
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree
5) I used to drive over speed limit sometimes.
    - Yes
    - No
    - Maybe
6) Driving over speed is a common cause of traffic accidents?
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree
7) Sometimes i used to drive over net overload.
    - Yes
    - No
    - Maybe
8) Driving over vehicle net overload is a common cause of traffic accidents?
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree

9) Traffic accident cases are Justified Correctly.
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree
10) I believe i am concerned about ensuring traffic safety.
    - Strongly Agree
    - Agree
    - Neutral
    - Disagree
    - Strongly Agree

III) Questionnaires (For Trafic Polices)

1). ETA has efficient traffic monitoring system.

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Agree

2). Luck of road Traffic Monitoring System is the Couse for traffic accidents in Addis Ababa.

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Agree

3). Traffic problems in Addis Ababa Streets (please tick one square along each line)

	No Problem	**Minor Problem**	**Major Problem**
Traffic Speed	☐	☐	☐
Traffic Volume	☐	☐	☐
Heavy Vehicles	☐	☐	☐
Pedestrian Facilities	☐	☐	☐
Irresponsible Driving	☐	☐	☐
Safety Concerns	☐	☐	☐

4). Vehicle speed and net overload detector is the best solution for traffic accident.

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Agree

5). Vehicle speed and net overload detector is the best solution for traffic accident.

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Agree

6). There is luck of intelligent traffic monitoring system at ETA.

- Strongly Agree

- Agree
- Neutral
- Disagree
- Strongly Agree

7). Using vehicular social network to aware closer vehicles speed and net overload can support in reducing traffic accident.

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Agree

8). Addis Ababa City Administration Traffic Management Agency has to apply automated systems to support traffic management processes.

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Agree

9). Addis Ababa City Administration Traffic Management Agency has its own vehicle speed and net overload detection systems.

- Yes
- No
- Maybe

10). Do you have any suggestions to overcome the traffic problems?