MUSTAPHA NUHU YAHYA

ARTIFICIAL INTELLIGENCE BASED HYBRID MODELS

MASTER THESIS

YEAR 2020

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF CIVIL ENGINEERING

ARTIFICIAL INTELLIGENCE BASED HYBRID MODELS FOR CLASSIFICATION OF ROAD ACCIDENT SEVERITY

MASTER THESIS

Mustapha Nuhu YAHYA

Nicosia

September, 2020

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF CIVIL ENGINEERING

ARTIFICIAL INTELLIGENCE BASED HYBRID MODELS FOR CLASSIFICATION OF ROAD ACCIDENT SEVERITY

MASTER THESIS

Mustapha Nuhu YAHYA

Supervisor

Assoc. Prof. Dr. Beste ÇUBUKÇUOĞLU

Co- supervisor

Dr. Shaban Ismael ALBRKA

Nicosia

September, 2020

Mustapha Nuhu Yahya: ARTIFICIAL INTELLIGENCE BASED HYBRID MODELS FOR CLASSIFICATION OF ROAD ACCIDENT SEVERITY

Approval of Director of Institute of Graduate Sciences

Prof. Dr. KEMIL HUSNU CAN BASER

We certify that, this thesis is satisfactory for the award of the degree of Master of Science in Civil Engineering

Examining Committee in Charge:

Assoc. Prof. Dr. Beste ÇUBUKÇUOĞLU	Supervisor, Department of Civil and Environmental Engineering, NEU
Dr. Shaban Ismael ALBRKA	Co-supervisor, Department of Civil and Environmental Engineering, NEU
Prof. Dr. Huseyin GOKCEKUS	Department of Civil and Environmental Engineering, NEU
Dr. Nasradeen Ali KHALIFA	Department of civil Engineering,
Asst. Prof. Dr. Hussin Yahia	Middle East Collage, Oman Department of civil Engineering

I hereby declare that all information in this document has been obtained and presented by academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I had fully cited and referenced all material and results that are not original to this work.

Name, Last name: Mustapha Nuhu Yahya

Signature:

Date:

ACKNOWLEDGMENTS

Firstly, I will like to express my special thanks and appreciation to my co-supervisor, Dr. Shaban Ismael Albrak, my supervisor, Assoc. Prof. Dr. Beste Çubukçuoğlu and Engr Ibrahim Khalil Umar for their guidance and endless support during my studies here in Cyprus. They have always supported me academically and have given me the best guide ever in my academic life. Working with such respected and inspirational personalities has been a privilege and has ignited my interest in research.

My sincere appreciation goes to my Dear parents Hon Justice Nuhu Yahya Galadanchi and Hajiya Fatima Aminu Ibrahim and to my entire my entire family members especially Engr. MN Yahya, for their prayers, love, care, protection and courage they have given me throughout my life. Also, to my beloved wife Asmaa Alkasim Habib, I say a big thank you.

I would like to express my profound gratitude to all my colleagues and friends who have always been a source of strength and encouragement especially Engr. Bashir Aminu, Engr. Mubarak Diso, Engr Yasir Ibrahim Khalil, Hussein Mustapha Gaya, and many more whose names are too numerous to be mentioned in this context.

To my parents…

ABSTRACT

Road traffic accident (RTA) is a global event placing a serious social threat resulting to the death of millions of people every year and causing substantial economic damages to people, their communities and nations as a whole, thereby affecting the gross domestic product (GDP) of most countries.Increase in the number of fatalities is recorded in many countries and a significant percentage of road traffic deaths happens among pedestrians, cyclists and motorcyclists who are the most vulnerable road users. In this study a Nigerian highway was selected where lack of regular maintenance of the installed temporary traffic control devices, non-compliance with standards, human behaviors like over speeding, traffic rules violation in the work zones of highways were identified to be the major causes of RTA in Nigeria. This research aims to categorize the injury severity in RTA with both sensitivity rates and high accuracy. The set of data used covers 608 numbers of individual road traffic incidents of vehicle crashes involving a large number of vehicles involved, obtained from the Federal Road Safety Corps (FRSC) dataset in Nigeria, from October 2008 to October 2013. Using three non-linear classification learning paradigms as linear classification model all trained using hybrid learning approaches, the injury severity levels was classified into fatal and non-fatal classes. Experiment results reveal that among the machine learning paradigms considered, the hybrid ANN approach outperformed the three other individual approaches.

Keywords: Road traffic accident; Accident severity; Artificial Neural Network (ANN); Support Vector Machine (SVM); Boosted Regression Model (BRT); logistic regression model (LRM)

ÖZET

Karayolu trafik kazası (RTA), her yıl milyonlarca insanın ölümüyle sonuçlanan ciddi bir sosyal tehdit oluşturan ve bir bütün olarak insanlara, topluluklarına ve milletlerine önemli ekonomik zararlara neden olarak gayri safi yurtiçi hasılayı (GSYİH) etkileyen küresel bir olaydır. Ölümlerin sayısında artış birçok ülkede kaydedilmektedir ve karayolu trafik ölümlerinin önemli bir yüzdesi, en savunmasız yol kullanıcıları olan yayalar, bisikletliler ve motosikletliler arasında meydana gelmektedir. Bu çalışmada, kurulan geçici trafik kontrol cihazlarının düzenli bakımlarının yapılmaması, standartlara uyulmaması, aşırı hız gibi insan davranışlarının, otoyolların çalışma alanlarında trafik kurallarının ihlalinin başlıca nedenleri olduğu tespit edilen bir Nijerya karayolu seçilmiştir. Nijerya'da RTA. Bu araştırma, RTA'daki yaralanma şiddetini hem duyarlılık oranları hem de yüksek doğruluk ile kategorize etmeyi amaçlamaktadır. Kullanılan veri seti, Ekim 2008'den Ekim 2013'e kadar Nijerya'daki Federal Yol Güvenlik Birlikleri (FRSC) veri setinden elde edilen, çok sayıda aracın karıştığı araç kazalarına ilişkin 608 ayrı karayolu trafik olayını kapsamaktadır. hepsi hibrit öğrenme yaklaşımları kullanılarak eğitilmiş lineer sınıflandırma modeli olarak sınıflandırma öğrenme paradigmaları, yaralanma şiddeti seviyeleri ölümcül ve ölümcül olmayan sınıflar olarak sınıflandırıldı. Deney sonuçları, makine öğrenimi paradigmaları arasında, hibrit YSA yaklaşımının diğer üç ayrı yaklaşımdan daha iyi performans gösterdiğini ortaya koymaktadır.

Anahtar Kelimeler: Karayolu trafik kazası; Kaza şiddeti; Yapay Sinir Ağı (YSA); Destek Vektör Makinesi (SVM); Artırılmış Regresyon Modeli (BRT); lojistik regresyon modeli (LRM)

TABLE OF CONTENTS

Contents

List of Tables

Lıst of Fıgures

LIST OF ABBREVIATIONS

AI:	Artificial Intelligence
ANN:	Artificial Neural Network
BRT:	Boosted Regression Tree
FA:	Factor Analysis
FRSA:	Featute Removal Sensitivity Analysis
KMO:	Kaiser-Mayer-Olkin
LRM:	Logistic Regression Model
PCA:	Principal Component Analysis
RTA:	Road Traffic Accident
SA:	Sensitivity Analysis
SISO:	Single-Input Single-Output
SVM:	Support Vector Machine
WHO:	World Health Organisation

CHAPTER I

Introduction

General Overview

Road traffic accident (RTA) is a global event placing a serious social threat. RTA results to death of approximately 1.35 million people every year (Akkaş, sokullu, and Ertürk 2020). RTA is causing substantial economic damages to people, their communities and nations as a whole. Such problems occur from medical costs as well as loss of employment for the disabled and the dead due to the accidents, and for victims who need to take a break off school or work to take care of their patients, thereby affecting the gross domestic product (GDP) of most countries by 3%. 20 to 50 million people suffer non-severe injuries with many resulting to a disability. Increase in the number of fatalities is recorded in many countries and if appropriate measures were not adopted by the year 2030, RTA will be listed as the fifth cause to death in the world, resulting to death of 2.4 million people every year (WHO 2015). A significant percentage of the road traffic deaths happens among pedestrians, cyclists and motorcyclists who are the most vulnerable road users. Studies shows that countries with low- and middle-income experiences more than 90% of world's road accident fatalities, though having approximately 60% world's vehicles, and covers 70% of road mortality. Also, from higher income regions, people of lower economic backgrounds happen to be more exposed to RTA and in countries like America, 65% of fatalities happens among vehicle occupants. High Speed was identified as the primary factor causing injury to vulnerable road users, but only 29% of worlds countries met up with the speed reduction policies with only 10% reported effective (WHO 2009). Studies revealed that persons from low income setting are more vulnerable to road accidents by far than those from affluent families regardless of the country's economic status. Accident injuries are the major cause to death for children and youth aged 5-29.

Many studies revealed that the roads of African region are the worst in the world. The region having approximately 2% of the worlds motor vehicle which is the least in the world. Despite the low number of vehicles in the region roads, studies show that it contributes 16% of the world's road mortality having a higher rate than any other region of the world. Vehicles in the Africa region stand a chance of being involved accidents more than 100 times a car in the

United Kingdom or the United State. The African region has an average mortality rate of 24.1 per 100,000 populations as compared to the world average of 18.0 per 100,000 populations. Youths aging 15-44 in the region are the most vulnerable road users (pedestrians and 2-3-wheel riders) and account for approximately 52% of the total fatalities (WHO 2013).

Nigeria, with an average death rate of 33.7 per 100,000 populations has recorded the highest number of fatal accidents in the region. Studies also revealed that, in every 4 fatal accidents recorded in the African region, one happens to be on Nigeria's highway (WHO 2013).

The causes of accidents are related to road, vehicle and human factors. The main problem with road safety in Africa is the lack of implementation of countermeasures which were in existence for decades (Assum, 1998). Politics have a great influence on the implementation of these countermeasures. This is because of the scarcity of resources available for road safety. The most effective countermeasures are costly and limit the road user's freedom to choose traveling speed. Politicians in the region are reluctant to overcome these side effects (Assum, 1998).

Table 1

Regional population, road facilities and registered vehicles (%) (WHO 2015)

Region	Population (%)	Vehicles (%)	Fatality (%)
Africa region	12	2	25
America region	14	27	15
Eastern Mediterranean region	9	4	19
European region	13	26	9
South-East Asian region	26	15	16
Western Pacific region	26	26	16
*World Population (2014)	7.18 billion		
*Number of World registered Vehicles (2014)	1.18 billion		
*Annual fatality rate (2014)	1.25 million		

Problem Statement

The conditions of Highways are becoming worst every day and at the same time, the number of registered vehicles is increasing worldwide. Studies show a 15% increase in the number of registered vehicles from 2007 to 2010 (WHO, 2020) which shows a rapid increase in the number of road users fleeing the road as well an increase in the fatality recorded, putting lives of motorists, road workers and other vulnerable road users at risk. Kano is a very large state in northern Nigeria being the most populous and the second most populous city in the country after Lagos. It is a business-based center with large numbers of industries that attracts a huge number of people from within and outside the state which leads to an increase in the number of traffic beyond roads capacity. Kano- Wudil road is a major and busy connecting road linking Kano state (north-west) with 5 north-eastern states (Bauchi, Yobe, Gombe, Adamawa, Taraba), plateau state from north-central and some parts of Jigawa state. As such, heavy traffic is experienced on the route and hence, an outcome the unfortunate RTA.

Need for Study

The rapid increase in the number of vehicles using the road daily and the increase in the number of accidents recorded necessitate the study. Certain locations along the road experience regular accidents which indicate problems instigating accidents at that blackspots. Therefore, this study will map out the accident-prone zones along with the case study so that possible preventive and mitigation measures can be adopted to enhance the safety of the road users and minimizes passenger/ driver risks. Also, predicting accidents using previous data along the route will help in minimizing accidents since accident prediction has a significant role in the improvement of traffic safety.

Research Aim

- To investigate the severity of road traffic accidents (RTA) on a rural highway in Nigeria using AI-based models.

Objectives

- To assess the severity level of the RTA on rural highways
- To perform non-linear sensitivity analyses to identify factors resulting in higher accident severity.
- To develop 3 AI-based models (ANN, BRT, SVM) for the classification of RTA severity
- To compare the accuracy of AI models and statistical models
- To propose mitigation measures to enhance safety on rural highways.

Scope of the Study

The study focuses on classifying the severity level of RTA on undivided rural highways in Nigeria by identifying parameters related to road traffic accidents.

Limitations

- The analysis will be based on the accident data obtained from the federal road safety corps (FRSC) and only the parameters recorded by FRSC
- Data obtained did not contain accident property damage

Chapter I summary

Road traffic accident is a global event posing a serious threat to everybody resulting to huge loss in lives, properties, economic damages to people, countries and the nation as a whole. Such problems necessitate the study with the aim of investigating the severity level of (RTA) on rural highway of Nigeria using AI based models by comparing the accuracy of AI models and statistical model, thereby proposing mitigation measures to enhance safety on rural highway

CHAPTER II

Literature review

Introduction

In order to come up with measures to avoid and minimize the incidence of road accidents severity, a study of the frequency of road accidents should be carried out. Such an investigation will include integrating the reports of each accident with the characteristics of the road, the environment, the conditions of the vehicle and the driver. These features can be used in the prediction and assessment of the severity of road accidents when connected to accident reports (Mannering & Bhat, 2014).

Currently, significant progress has now been made on vehicles and urban roads for safety purposes, but incidents of high severity are still occurring on urban roads regularly. Continuous study of significant factors contributing to RTA enables researchers to perform simulations to predict the extent or severity of the accident (Moghaddam, Afandizadeh, and Ziyadi 2011). As such, this chapter aims to review previous research on the RTA severity investigation, because of the presence of several variables leading to an assessable amount of accident severity, the association between RTA severity and risk factor cannot be adequately recognized. Those factors are broadly categorized into Human-Related Factors, Vehicle-Related factors, Environmental Factors and Geometric design-related Factors (Chang & Mannering, 1999). Further information on the socio-economic implications of the RTA was discussed, and performance analyses of the RTA severity using different classification models were also given in the section.

Overview on RTA occurrences in Nigeria

Transport is a significant aspect of the Nigerian economy, the catalytic impact of which cannot be overemphasized, especially on socio-economic growth. In Nigeria, over 80% of road transport is performed by road (Afolabi & Gbadamosi, 2017).

Transportation safety means avoiding accidents and minimizing accident losses. There is a more important factor in the likelihood of accidents resulting in death of people and property loss on our highway, as Nigeria becomes more mobile every day. On our highways, the effects of accidents are enormous. Accidents cause substantial damage to our country's current and potential active workforce, as well as, in many cases, significant social issues, deaths or severe injuries. This often leads to the loss of breadwinners, driving the affected family into poverty, and hampering children's education. This often leads to the loss of breadwinners, driving the affected family into poverty, and hampering children's education. Accidents placed high costs on healthcare facilities (Afolabi, 2009).

The population growth and increased degree of motorization have been attributed to the rising magnitude of fatal road traffic accidents globally. The major cause of mortality in teenagers and people of the peak age is motor vehicle collisions (Atunbi, 2009). The proportion and an actual number of traffic deaths encountered in the number of developing countries have increased, while developed nations are seeing a downward trend in the incidence of accidents by more than 20% (Emenike & Ogbole, 2008).

Road accidents are not only a dynamic phenomenon in terms of their multiple causes but also in the form of their effect on lives and property. In addition to the humanitarian dimensions of road safety, injuries and deaths resulting from road collisions have significant social and economic implications, contributing to the creation of spatial contact phobia for prospective travelers. In normal circumstances, this would have stopped all business ventures that would have taken place and nicked them in the bud due to differences in location and distance of business prone areas and buyers given the fear of incidents that may occur with regards to the risk of being involved in road traffic collisions.

There are physical, social, and economic consequences of road traffic accidents. The worldwide economic cost of road accidents in 2003 was valued at $518 billion annually, of which $100 billion occurred in poor developing countries (WHO 2003). Every year, Nigeria loses around 80 billion nairas to road accidents. Of all subjects involved in road accidents in Nigeria, 29.1% are badly hurt and 13.5% unable to return to work (Atubi & Gbadamosi, 2015).

Since the first car crash was registered, road traffic accidents in Nigeria have been troubling and very disturbing. In 1913, with the enactment of the first transport law, the Highway (Motor

Traffic) Ordinance, the main objective of which was to "reduce road traffic incidents to the lowest minimum" in the Southern Protectorate, Nigeria's attempt to resolve the safety problems on our roads began. With the amalgamation of the protectorates in 1914, a nation-wide decree followed in 1916. Following the United Kingdom Road Traffic Act of 1930, the country-wide statute was subsequently checked and adapted in 1940 and 1945. The Road Traffic Act, the Law of Carriage, the Federal Highway Act, and the 1988 Federal Road Safety Commission Decree, which was later revised in 2007, were eventually included in other legislation. However, despite many revisions to transport laws and the wide array of agencies, with states having departments for traffic control, road traffic crashes have continued to increase throughout the country.

Traffic accidents in Nigeria vary by states. Regarding known causes of accidents worldwide, which include very bad roads resulting from poor maintenance culture and poor road management, Nigeria has been consistently ranked as having the highest road traffic accident incidents in the world for obvious reasons.

Road Traffic Accident

The accident is defined as an unexpected event that happens by chance, something that happened that often injures someone or damaged something. Hence, road traffic accidents are an unpredictable occurrence arising from the activity of automobiles (Onakomaiya, 1988). Accidents may be fatal, leading to the death of a road user or a minor one. Accidents do not only happen; they're triggered by certain factors. In other words, any transport accident is not only a mere event but has been formed as a consequence of one cause or another. Proper knowledge and a good awareness of such causes will help significantly in avoiding them. This would ultimately contribute to the desired objective of safety awareness for road users in our community.

Causes of Accident

The causes of RTC are multi-factorial and include the involvement of a variety of pre-crash factors, including human, vehicles and the road and the environment (Haddon 1980). The contributing factors or causes of road traffic accidents can be broadly classified into three (3) major categories viz– Human factors, Vehicle/Mechanical factors and the Road/Environmental factors (Ihueze & Onwurah, 2018). In these three categories, more than 80% of all traffic collisions are said to be caused by human factors because the operational capacity of the drivers is very important to the causes and prevention of the accidents.

To provide countermeasures, several studies have been performed to examine and consider the factors that lead to RTC. For instance, using linear regression analysis, (Ojo, 2014) analyzed the factors leading to road traffic accidents in Ekiti State, Nigeria and found that distraction of drivers, over speeding (speed violation) and unsafe overtaking significantly contributed to road traffic accidents in the state.

A study in Oyo State, Nigeria shows that speed violation, poor roads, wrong overtaking, sudden mechanical defects, use of drugs, heavy rainfall and tyre burst led to a growing rate of motorcycle road crashes (Gboyega et al., 2012).

Between 2005 and 2012, Olawole studied the effect of weather (rainfall and temperature) on road accidents in Ondo State, Nigeria, and found that the associations between road accidents and weather components were generally low and never exceeded 0.41 (Olawole, 2016).

In Northern Region of Ghana, Nyamuame and others used binary logistic regression to examine the factors that led to traffic accident severity. The study found that overloading and obstruction in Ghana were the two most important factors leading to the occurrence of road crashes (Nyamuame et al., 2015). Also, Previous studies have shown that brake failure is one of the factors leading to road accidents in developing countries.

Stages of accident

As we all know, accidents are triggered and not only do they happen, the crucial assessment of the phenomenon of incidents clearly shows three unique stages. At any point when it is registered, the three stages are the cumulative consumption of an accident. The stages are interwoven and occur sequentially after each other. These stages are the Pre-Accident stage, the Accident stage, and the Post-Accident stage of highway safety (Afolabi & Gbadamosi, 2017).

Pre-accident stage: -

The Pre-Accident stage brings together all phases of preventive or precautionary steps to monitor or mitigate road accidents. All contributory factors such as the vehicle, environment, the road users and the preventive or precautionary steps taken to usually avoid accidents come under this process. This is an example of many situations that may cause accidents. In other words, all conditions and instances that precede the occurrence of an accident are presumed. We may also determine those circumstances before they are reported that are capable of causing an accident. This stage, in short, deals with accident prevention.

The accident phase:

The crash stage is the actual occurrence of the accident when the mechanical device is involved in an actual collision resulting in an accident. That is when the pre-crash stage cannot be averted. The consequence of the accident to the victim also belongs to this stage. Similarly, all the key measures of the crash process are the place at which the accident happened and the time of day.

Studies have shown that it is possible to reduce the deaths of drivers and passengers by up to 80 percent by using safety belts alone. Therefore, the focus of this second stage is on the prevention of injuries.

During road traffic collisions, the following steps should be taken:

- Assess the Situation: -
- Locate the casualty
- Quickly examine the victims
- Prevent the possibility of more fires, explosions, road traffic
- Hold vehicle stationary
- Turn off connections to the engine, fuel and battery
- Display warning signs
- Send for assistance.

Care of the Victim: -

- Rescue the injured victims
- Check for breathing, heartbeat, and consciousness
- Care about unconscious circumstances first
- Take care of bleeding and fractures,
- Use the first aid kit for the vehicle if accessible
- Move the casualties to the nearest hospital immediately.

Care of the Vehicle: -

- Maintain the accident vehicles stuck and in safe possession.
- Safeguard property against damage
- Take assistance from the local community
- Notify the police.

Post-accident phase

It is possible to characterize the post-crash stage as the process of determining or analyzing the effects of road accidents. This review uses measurable and qualitative analytical methods and is focused on social, environmental and political impacts. At this point, it is our concern to save those who do not need to die, with hospitalization, permanent disabilities, and premature deaths being minimized.

Indeed, the emphasis is on the provision of appropriate and timely emergency communications, transport and medical treatment to assess the livelihood of the crash survivors' continued survival. Therefore, the problem at this point is the reduction of severity, which would include the availability and skill of ambulance crews and attendants in treating accident scene casualties and the receptivity of hospital personnel to accident victims not followed by police officers.

Table 0

Contributory factors causing RTA

PHASE		HUMAN	VEHICLE & EQUIPMENT	ENVIRONMENT
Pre-crash	Accident prevention	Information Attitude Impairment Police Enforcement	Road Worthiness Lighting Banking Handling Speed Management	Road design and road layout speed limits pedestrian facilities
Crash	Injury Prevention during the crash	Use of restraints Impairment	Occupant Restraints Other safety Devices Crash- Protective Design	Crash-protective Roadside objects.
Post-crash	Life sustenance	First-Aid-Skill Access Medicals	Ease of access Fire to Risk	Rescue facilities Congestion

Source: (Afolabi & Gbadamosi, 2017)

Contributing factors to road traffic accidents

The contributory factors causing can either be human factors, vehicle factor or environmental factor/road factor. These contributory factors are discussed detailly below

The human factor:

Human factors account for about 80 percent of the country's causes of road traffic accidents. Drivers, pedestrians, law enforcement officers, and engineers are the main components of the human factor (Gu-Chang & Li, 2010).

Most Nigerian road drivers are very rude, discourteous and have no respect for human life. Quite a number of them often acquired the desirable knowledge of road use and hardly obey the traffic rules/regulations. This has resulted in regular unnecessary devastation with many casualties of lives on Nigerian highways. The fact that, among virtually all the important factors leading to the troubling proportion among traffic accidents in Nigeria, the human factor is at the top of the list, is almost to the point of indisputability. Indicators for checking the argument are obvious:

- Prevalent neglect by road users of road transport signs
- Lack of adequate driver instruction/training
- Reckless driving habit especially among teenagers
- Inexperience and unqualified drivers
- Speeding, unsafe driving and violation of traffic laws, in particular concerning speed limits,
- Driving under the influence of herbal concoction and drugs
- Lack of regard for other road users and concern for them
- Impatience and carelessness
- Vehicle overloading
- Fatigue
- Poor vision

The vehicle factor:

The vehicle is also a significant factor in road traffic accidents. However, road safety goes beyond routine inspection or prompt vehicle repair. It should be an everyday routine of care and inspection of all parts of a vehicle. The key causes in the vehicle are faults in the tires, brakes and inputs, all resulting from inadequate vehicle maintenance. The global economic downturn has had a significant effect on the quality of goods in the Nigerian markets, with people now preferring the use of sub-standard items such as used tyre, spare parts and secondhand vehicles. These undermine the ideals of protection when viewed against the phenomenon of used cars, combined with reckless driving and over-speeding. A failure of any of these components will eventually impact smooth driving, which can inevitably lead to a serious accident. The various elements of the vehicular factor that contributed to the accident are:

- Failure of Brake
- Tyre burst
- Failure of engines
- Using counterfeit spare parts
- Faulty and Blinding lights
- Poor vehicle maintenance

Essentially, a faulty vehicle, an unserviceable vehicle, or a poorly maintained car are all hazards that are extremely and likely to cause road accidents.

The environment factor:

Within the background of Nigeria, there is a clear debate as to whether poor roads can be related to the high rate of road accidents. Or, if the good and improved highways on which the country has spent so much are not a contradictory feature. The statement is against the backdrop that significant decreases in accident rates have not been seen despite the development of new roads in the region, but rather seem to be growing. In other words, other factors, particularly the human elements contributing to the tragedy, need to be concentrated on.

Environmental factor includes:

- Bad/poor roads
- Weather conditions
- Hazardous bend/curve
- Abandoned vehicles/ broken down
- Non-controlled animals
- Intrusion on the highway

Socio-Economic Consequences of RTA in Nigeria

The impact of road traffic collision cannot be overemphasized. Nigeria has a poor road traffic crash record. In social and economic terms, Nigeria has had a fair share of fatalities from road traffic collisions. Today, road accidents in Nigeria have taken away so many lives that hardly any single disease approaches its mortality prowess. As a result of road traffic collisions, people died prematurely and properties worth several millions of Naira were destroyed. In comparison, damages resulting from misery, deprivation and social disturbances, which may be difficult to quantify in monetary terms, are considered to be part of the basic social cost of an accident. Social costs in form of distress, i.e. loss of friends, close relatives, colleagues, parents, etc., which ultimately contribute to psychological depression. Other victims who may not have died may bear remnants of disabilities, such as blindness, loss of limbs, or even lifelong wheelchair bindings. Such victims and families experience significant psychological distress, often from stigmatization or emotional imbalances.

There are numerous effects of the negative chain reaction of road traffic collisions. For example, because of the death of a breadwinner, an individual may drop out of school due to a lack of fulfillment of essential needs, the resulting reaction may contribute to the negative social life of that person by becoming a dupe, armed robber, assassin, hoodlum or even a fraudster (Emenike & Ogbole, 2008).

The effects of road traffic collisions range from the physical, social and economic effect on people to the economic impact on the national economy and the impact on the vehicles itself. Over the years, road transport has contributed modestly to the economy's Gross Domestic Product (GDP). It usually accounts for no less than 80% of the share of GDP produced by the transport industry as a whole (Agbonkhese et al., 2013).

Road collisions also often contribute to the loss of traffic structures, such as bridges, and hence the loss of public transport infrastructure. Road traffic incidents have also adversely impacted the country's manpower capabilities. Approximately one-quarter of those involved in road traffic incidents are killed by police accident cases reports, while the remaining three-quarters suffer injuries.

Given the general effect of human and material loss on road injuries, it is tragic that the number continues to grow despite the government's attempt to curb the RTA exercise. Solving the traffic accident injury issue involves a multi-dimensional approach that will, in reality, require major stakeholders in the transport sector with a key role played by the government.

Previous studies on RTA using machine learning models

By reviewing various papers, it has been found that researchers in different countries use different methods to analyze accidents and find reasons for reducing the safety of rural roads. Below are some previous studies discussing road accident analysis and classifications using different approaches and models.

ANN model

Driver injury patterns patterns (fatal injury, possible injury, evident injury and no injury) was investigated with multi-class classification in ROR crashes, based on ML analysis (binary ANN and RF models), with data obtained from 2011 to 2013 in Washington State (Zhu, Li, and Wang 2018). The variables used for the investigation were time variables, demographic variables, vehicles (car, truck, pickup and others) and environmental variables. The findings indicate that in fatal accidents or serious injuries, the main factors instigating the incidents are truck usage, lack of restraint, driver impairment, being female, distractions, overtaking manoeuvres, rollover accident type and dawn/dusk conditions

A study by (Delen et al., 2006) came up with a research that uses an ANN to model the relationship between the severity levels of accidents (four levels: possible injury, no injury, fatal injury and non-incapacitating injury) and the causative factors related to the accidents that happened in the United States. Factors considered were vehicle types (SUV, passenger cars, vans and pickups/ light trucks), collision type: multiple vehicles collision (stuck/striking, back/ front/ side crash, head-on, rear-end) or single-vehicle accident (roll-over), personal information and environmental information. The seatbelt violation, being under the influence of drugs and alcohol, as well as passengers' gender and age and their vehicle type were important factors in accidents. Among their findings, the authors noted that no single factor is a key cause, rather the combination of them (seatbelt violation, being under the influence of drugs and alcohol, passengers' gender and age and the vehicle type) may be a key determinant.

Another research by (Casado-Sanz et al., 2020), driver factor as a cause of the accident was analyzed on a crosstown road of Spain to investigate the risk factors increasing the severity of accidents. It was observed from the results that factor increasing the severity of the accidents are lateral crosstown roads, wider lanes, low traffic volume, a high number of heavy vehicles on the road, infraction and absence of road markings.

using death registry data from 2012 to 2016, (Abdous & Mahmoudabadi, 2018) attempted to examine factors leading to increased exposure to the risk of rural residents in Iran. Conducted test for the selection of most dominant input parameters. Results found that, relative to those crossing urban roads, female pedestrians around intercity roadways were less vulnerable to fatal accidents, rather moving around minor/rural roads resulted in higher mortality exposure. There were no substantial variations in the vulnerability of female pedestrians to crashes on different days, but the wearing of dark outfit increases their vulnerability significantly.

Regression tree model

A study in Canada by (Lee & Li, 2015) established a research to predict driver accident severity (severe and non-severe) with one or two vehicles involved. The vehicles studied included heavy-trucks, cars and light trucks. Accidents, drivers, the climate, cars, facilities and traffic characteristics were also taken into account. Due to their elevated severity outcomes, ejection from a vehicle and head-to-head collisions were highlighted. Results show that there are distinctions between the drivers of heavy trucks and the other drivers: the probability of severity increases with regular traffic, and with the driver's age, the number of trucks increases.

Another study used CART model to institute a relationship between driver severity in accidents (three levels: injury, no injury, fatality) with trucks involved (above 10,000lb) in Taiwan. The research includes parameters related to a driver, environmental conditions, road, vehicle type (car, trailer, light truck), collision type (head-on, rear-end, sideswipe, collision with rail, overturn) and the accident characteristics (location, time). Results showed that driving under the influence of alcohol, seatbelt violation, head-on collision and light truck use as contrary are among the most important variables that increase driver severity (Chang & Chien, 2013).

Using CART algorithms with Conditional Inference-Forest, the severity (two levels: incapacitating injuries and non-incapacitating injuries) was analyzed in various types of accidents (head-on, rear-end, single-vehicle crash, sideswipe) on urban arterial roads of Florida. Identifying the influence of road type, traffic, driver and vehicles was the major objective. Passenger cars, light trucks, light slow-moving vehicles and heavy vehicles were studied. The result showed that alcohol/drug use, non-use of the seatbelt, speed limit violation and driver exceeding the age of 55 years are among the most significant parameters that worsen accident severity (Das et al., 2009).

Another research applied CART to investigate the extent of accident injury severity (three levels: injury, no injury and fatal) using 2001 report data of Taiwan National Traffic Research. Parameters considered were: vehicle involved (car, bus, pickup, truck, bicycle, motorcycle and pedestrians), collision type (head-on, pedestrian-vehicle, sideswipe, fixed object, rear-end), weather conditions and road types. The result showed that the most important factor in assessing the seriousness of accident driver injuries is the type of vehicle. Under the name 'vehicle type', the authors included motorized vehicles and vulnerable users, and the tree divided them into two divisions. The most prone are pedestrians, cyclists and motorcyclists when they are hit by motorized vehicles because of the degree of severity displayed in the tree's right branch (Chang & Wang, 2006).

Using the logit model investigated the crash pattern and factors contributing to accidents on two-lane rural highways to suggest safety measures to tackle the effects of the incidents. Results show that important factors contributing to severe crashes were pedestrian crashes on rural roads. In the rainy season, which typically had poorer weather conditions, pedestrian accidents have higher severity especially when the involved vehicle was a car (López et al., 2014).

Another research by (Zhou et al., 2019) came up with a study that analyzes the effects of various factors on the accident severity by adjusting the CART models real-time ridesharing vehicles. Data used for the study is a monthly accident data from Chicago police records. The initial data was resampled due to the high disparity (only 60 out of 2624 crashes were the most severe crashes), and the authors verified that the outcome of the prediction improved. Also, the performance indicators of the model such as G-mean and ROC area were better. Many parameters from the accident data were recognized as important indicators for accident severity.

A study to investigate the pattern of severity for drivers involved in heavy and light truck accidents by (Chen et al. 2016) in the United State. CART model was used for the selection of most dominant input parameters (crash-level, driver and vehicle-level). SVM model was used to evaluate the influence of variables on severity (no injuries, incapacitating and non-incapacitating injuries). Results show that significant and dangerous accidents are associated with drug use/alcohol, driving conditions and seatbelt usage.

An analysis of pedestrian crashes to understand the dissimilarities and interdependence between crash patterns and to develop strategies that will enhance pedestrian safety (Montella et al., 2011). Association rules and classification trees are the data-mining techniques used on the data related to pedestrian accidents that occurred in Italy. The results of the study show that higher crash severity of pedestrians was exhibited on rural roads. Also, at the night, the association between fatal crashes and older pedestrians was stronger in rural areas.

Support vector machine

A study by (Theofilatos et al., 2019) compares the real-time predictive power of Deep Learning (DL) and Machine Learning (ML) models, considering Naive Bayes, k-nearest neighbor, SVM, DT, RF, deep neural network and shallow neural network models. Sensitivity, Accuracy, Specificity and Area Under Curve (AUC) were the performance metrics used. The result showed a good performance of the Naive Bayes model as compared with the others because of its minor complexity.

A study in Spain by (Casado-Sanz et al., 2019) discussed the causative factors of severity of accidents on rural roads. Results showed that poor visibility by glare or weather conditions, traffic rules violations by drivers and pedestrians and speed limit violations were the major risk factors increasing the severity of pedestrian accidents.

Logistic regression model;

(Kononen et al., 2011) uses Logit designed models to estimate the likelihood of an accident in which at least one or more passengers experience severe or debilitating injuries (with Injury Severity Score (ISS) of more than or equal to 15). Parameters used for the study were vehicle type, change in speed, accidents involving a single vehicle vs. multiple vehicles accident, use of seatbelt and the direction of the collision. Result of this study shows that the use of a seatbelt, change in speed and direction of collision are the most important factors in accident severity prediction.

(Dadashova et al., 2014) uses time series analysis to study the severity and frequency of van accidents on roads: linear regression with parameters modified from Box-Cox and their autoregressive faults (Unobserved Components Model (UCM) as well as Demand for Road use, Accidents and their Gravity (DRAG). Factors related to drivers, fleet, economic factors, exposure variables and also legislative actions using the macroeconomic models were assessed as significant on the selected outcomes. The most important variables were driver behaviors, economic conditions, and categories of road infrastructure for greater injury severity in accidents.

A research to investigate the effects of passengers on driver accident severity using a random parameter logit model (LM) to obtain the differences in three accident situations (one, two and three passengers with driver included) together with several parameters for the roadway, environment driver attributes and the vehicle characteristics. The findings indicate that passenger(s) gender and age are both important factors, Confirming the difficulty of the relationships that need to be examined (Behnood & Mannering, 2017).

A study investigating accident injury severity (four levels: no injury, evident injury, possible injury, fatal injury) and the dissimilarities in accidents involving one or two light vehicles according to the driver gender. Also, different types of accidents (overturned, stuck an object, run-off roadway, others) were examined in the research by adopting a MLM model. Twenty-two thousand accident data record was collected from the state of Washington for the analysis. Their findings indicate major gender disparities in injury severity, even in the same type of accident. The authors concluded that further studies are required to better explain their findings, such as naturalistic studies. They also indicated that risk compensation should be present in the case of certain types of vehicles such as LTVs so they could have the impression of a self-protected driver. The risk of high accident severity rises for both sexes when the seat belt is not used (Ulfarsson & Mannering, 2004).

Using MLMs to examine the factors that can have a major effect on driver-injury severity categories (four levels: no injury, pain complaint, noticeable injury, and serious /fatal injury), in incidents involving large trucks and occurring in rural and urban areas of California by (Khorashadi et al., 2005). The study focuses on collisions (broadside, rear, another type) of one or multiple vehicles (opponent: passenger cars, tractor, trailer). The findings indicate that many variables affect the severity of the injuries to the driver: vehicle (type, number of vehicles involved, occupancy), environment (rain, road lighting, snow and fog), road geometry (concrete median barrier, number of lanes), and traffic characteristics (travel time, stop and go, location of collision and type).

Using the Logistic Regression Model approach, (Toy & Hammitt, 2003) uses an accident data in the United States to analyze the severity of accident (serious injury or death) for a driver in a two-vehicle accident (SUVs, pickup, vans, cars and trucks), with drivers age, gender, vehicle body type, restraint used, and the configuration of the crash as independent variables. The result of the studies shows that pickups, vans and SUVs are more assertive and pose more risk to smaller vehicles and have more protection to damage than cars.

Using mixed logit model approach to investigate and identify important factors contributing to the pedestrian accident injury severities in both rural and urban areas by (Chen and Fan 2019). The results showed that factors such as dark light condition and heavy trucks significantly increased pedestrian injury severities in both urban and rural areas.

A study by (Sahebi et al. 2015) predicts the pedestrian crash severities and the affecting factors on the rural traffic. Logit model was used for the analysis of factors affecting injury or death of pedestrian accident on rural highways of Tehran area of Iran. More than 50 variables were used for this study describing dependent variables. Results of this research showed that accident time and the type of vehicle involved in the accident had an important role in pedestrian accident severity.

Other related studies using other models;

Latent Classes Model (LCM) and MLM to study the severity of driver injury in single-vehicle crash with road factor on rural roads (slopes, straight and curved portions, number of lanes and traffic signs) and dangerous driver attitude due to drugs and seatbelts violation. The result shows a higher severity when both conditions were met at the moment of the incident (Li et al., 2019).

Using RF model to study the driver contributory behaviors to accidents on the rural roads of the United State (dual lane and two-way roads). Probit models were used for studying the contributing factors data for the collision. Results show that driving mistakes were more among young drivers and the concrete curbs on roads increase the risky behavior of drivers (Wu & Xu, 2018).

A research on signalized intersections in Miami, Florida studying driver injury severity in two (2) passenger car collisions with driver, roadway, environment, vehicle characteristics as well as crash identification as variables. Within the chosen data mining models, when studying prediction capacity and cost, RF was superior to C4.5 and IB. the result shows very substantial variations between them in driver severity by age and gender (Mafi et al., 2018).

Another study applied OPM to research the important explanatory variables of accident injury severity (three levels: fatal injury, incapacitating and non-incapacitating injury) of big truck collisions, such as collision type (rollover, sideswipe, angle, rear-end, multi-impact, head-on, others), type of vehicles involved (light vehicles (car, bus, pickup) or big truck), and characteristics of the driver (age, seat belt use, distraction, alcohol use, vision). Results of the studies show that severity of injuries is most related to truck drivers use of alcohol, driver distractions and drivers' emotional condition (such as being upset, in a hurry or clinically depressed). The vehicle type is also an important factor. Vans are involved in accidents that are more serious than cars. A higher mortality rate is recorded among the occupants of cars or vans when a collision involves a truck (Zhu and Srinivasan 2011).

A research carried out investigating the expanse of the road traffic problems. The result showed the RTA injuries and fatalities occurred most on rural roads than urban areas. Although the higher frequency of pedestrian accidents happened on urban roads, the result shows that the severity of injury to pedestrians is more on rural roads (Afukaar et al., 2003).

New approach to measure the probability of incidents related to darkness in terms of the odds ratio. Results show an increase in the risk of an injury by approximately 50% in rural areas (Johansson et al., 2009).

A research carried out on the Dutch road to look into the effects of road lighting on accidents. Result of the study shows a 140% average increase in risk on rural roads with lights and 360% on rural roads without lights (Wanvik, 2009).

A study in Iran using clustering analysis to analyze patterns of pedestrian crashes. Results show that in general, van/car/pickup crashes on rural highways, as well as heavy vehicle accidents on both urban and rural roads, were fairly less frequent and more likely to be fatal as compared to other pedestrian crashes (Kashani & Besharati, 2017).

Chapter II summary

Many relevant studies of the frequency of road accidents were analysed to come up with the measures to avoid and minimize the unwanted incident. Studies showed that RTA is not only a dynamic phenomenon in terms of their multiple causes but also in the form of their effects on lives and properties. These causes are multifactoral and includes the involvement of a variety of pre-crash factors.

CHAPTER III

methodology

Introduction

The Kano Maiduguri Road was selected as a thesis case study for the classification of accident severity along the route (Figure 3.1). The road section is a rural highway connecting the Kano city with many states in the North-Eastern and North Western part of Nigeria. The road is undergoing reconstruction in which it is been changed from the 2-lane single carriageway to a 4-lane divided highway. The first part of the contract has been given to Dantata & Sawoe construction company (Nigeria) Limited, which covers 101.365km out of the total distance of 591km, that is from Kano to Shuarin. For this thesis, only the first 38km of the road was considered, (Kano – Wudil) which is known as a dangerous road section from the traffic safety point of view. In this particular road section, a 15 km of a new road was constructed adjacent to the existing roadway. Traffic was diverted to this newly constructed carriageway giving the contractor room to reconstruct the existing road which will become one carriageway of the road. This chapter provides the step by step procedures adopted in this study to achieved the desired results.

Location of study area

Kano is a very large state in northern Nigeria being the most populous in the country and third (3^{rd}) largest city after Lagos and Ibadan. It is a business-based center with large numbers of industries that attracts a huge number of people from within and outside the state It was recorded to have a population of 9,401,288 as at 2006 census and 13,076,892 published by the national bureau of statistics in 2016, which gives 6.34% of Nigeria population with a growth rate of 3.36% annually. The road is the commonest mode of transportation linking Kano with other states of the country and most neighboring African cities (Barau et al., 2013).

Kano is a big state and a capital of same name province. It is situated at the northern part and recognized as the third (3^{rd}) largest city in the country. It lies along with geographical coordinates of 12. 00 22N and 8. 53 20E. located 840km from the Sahara Desert and 1140km away of the Atlantic Ocean. As such, the climate is quite hot and dry. Kano is a commercial based state, popularly known for its agricultural practices, manufacturing and transportation, has 44 local government areas, eight (8) of which formed the metropolitan with a total area of 499 square km (193 square meters). These local governments are; Kano Municipal, Dala, Gwale, Tarauni, Nassarawa, Fagge, Ungogo and Kunbotso (see Figure.2) (Okunola et al., 2012). Moreover, the schematic representation of the work methodology is represented in Figure 3.

Figure 1

Map of Kano showing case study

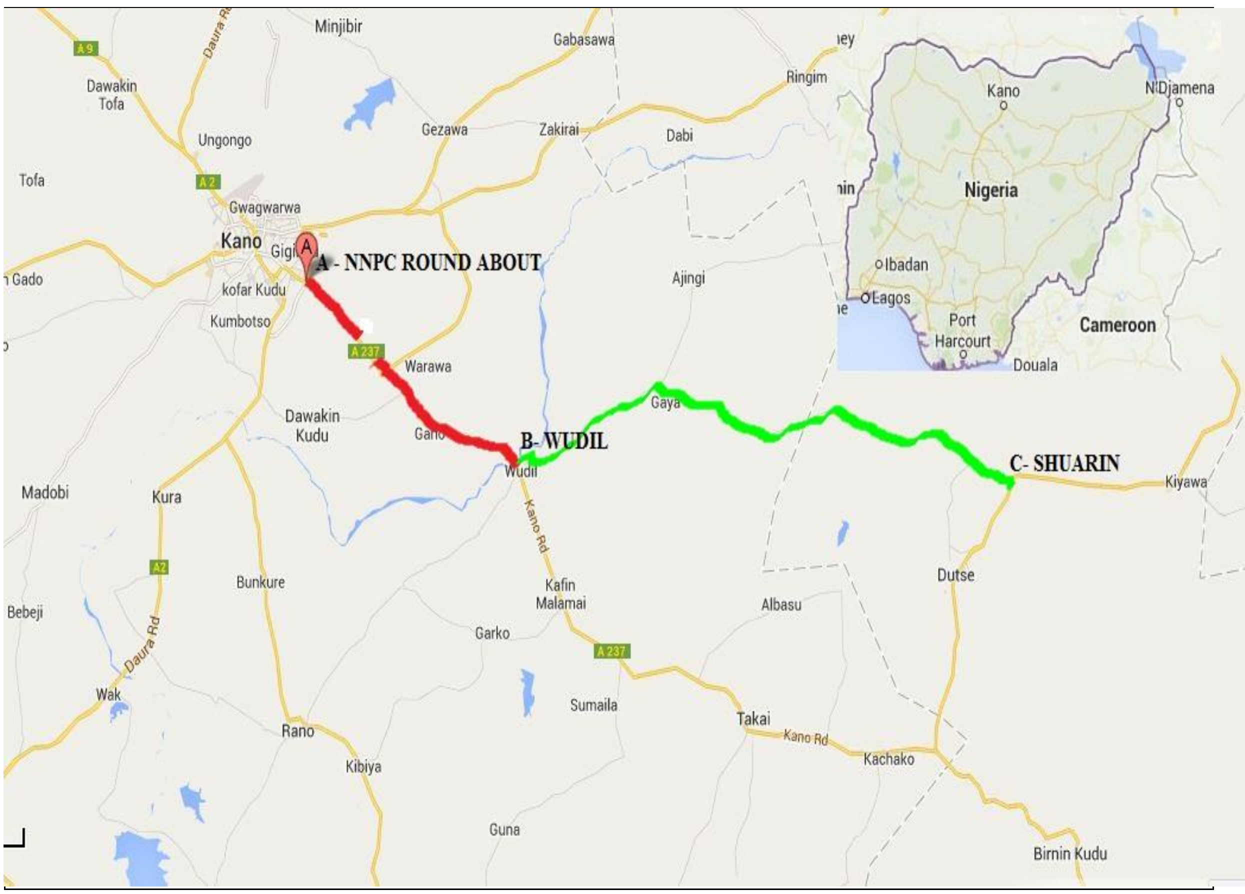


Figure 2

Map of Nigeria showing case study state

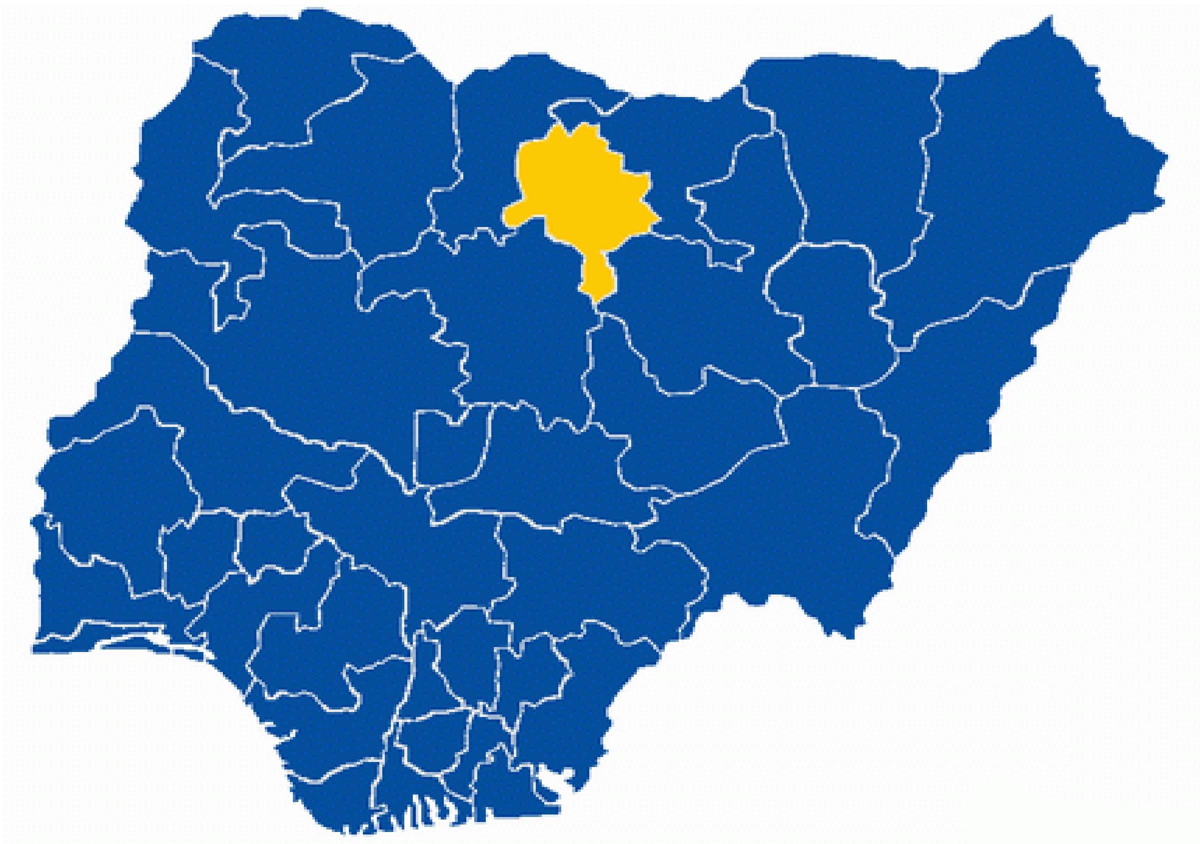
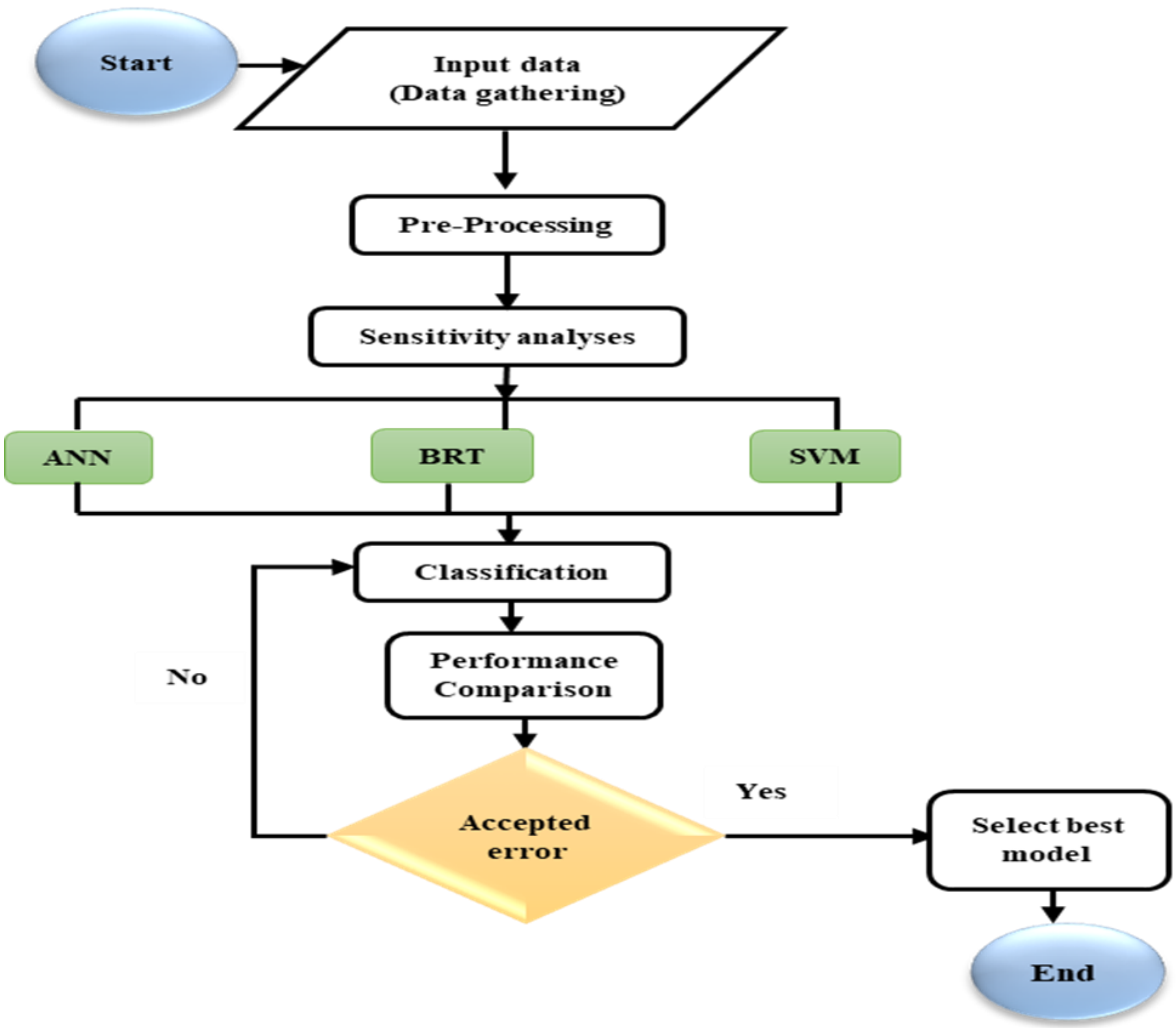


Figure 0

Schematic representation of the general methodology



Accident data (input data)

Accident records for 5years were obtained from the Federal Road Safety Commission (FRSC). The records give the total number of accidents, fatalities, injuries, number of vehicles involved, accident cause, number of people involved. The accidents were reviewed by their causes, fatalities and injuries per year. A major accident was then identified in the analysis. Also, locations having the highest number of accidents were identified from the records. Major accidents that occurred during the construction stage were not documented by the contractors, but pictures of the accidents involving both the construction equipment and other road users were supplied by contractors. Records from the FRSC include accidents that were not part of the road section under study. Only relevant accidents were extracted from the record.

Data pre-processing

Data pre-processing assists the neural network in learning the relevant patterns, which subsequently improves the data fitting and prediction accuracy. This includes data cleaning to handle missing values, noisy data and outliers. Data integration and transformation and lastly data reduction (Nourani, Gökçekuş, et al., 2020). The sigmoid activation function was used in the study's neural network with an upper bound of one and a lower bound of zero.

Data normalization is used to bring all the inputs and outputs variables into same range before feeding them to the AI models to avoid overshadowing of the data in lower numeric range by those in the higher numeric range (Nourani et al., 2019). Another benefit of the data normalization is the simplification if the numerical calculations in the model which in turns increases the model's accuracy and reduces the time taken to obtain the global/ local minimum. In this study, the data were normalized between 0 and 1 using the equation below

$$N_{norm} = \frac{N - N_{min}}{N_{max} - Nmin} \qquad (1)$$

Where N_{norm} is the normalized data values, $N,\ N_{min},\ and\ N_{max}$ are the observed, minimum and maximum values of the traffic accident data respectively. For model development, the normalized data is divided into three; 70% for training, 15% for validation and 15% for testing purposes.

Sensitivity analysis

The technique used to determine how independent variable values will impact a particular dependent variable under a given set of assumptions is defined as a sensitivity analysis. Is a model that determines how to target variables are affected based on changes in other variables known as input variables. This model is also referred to as what-if or simulation analysis with the primary objective of selecting the most dominant input parameters. It is a way to predict the outcome of a decision given a certain range of variables. By creating a given set of variables, an analyst can determine how changes in one variable affect the outcome.

The efficiency of any AI model depends on the imposed input parameters. Too much input parameters lead to overfitting and increase the complexity of the model (Tien Bui et al., 2019). On the other hand, a model with insufficient input parameters cannot model the process accurately. As such, a nonlinear Quadratic SVM kernel sensitivity analysis was used in the study to select the most significant and relevant input parameters to be used in the models. In the sensitivity analysis, each of the input parameters is used to predict the accident severity level for the study area individually via SVM model. The performance of each model is evaluated for both calibration and verification phases of the modelling. The average value of the performance criterion obtained in the calibration and verification steps is then used to rank the contribution of parameters to the accident severity and only the significant parameters are then considered for the model development. Three (3) types of sensitivity analysis were adopted for the study and they are;

- Feature Removal Sensitivity Analysis (FRSA)
- Single-input Single-output sensitivity analysis (SISO) and
- Principal Component Analysis (PCA).

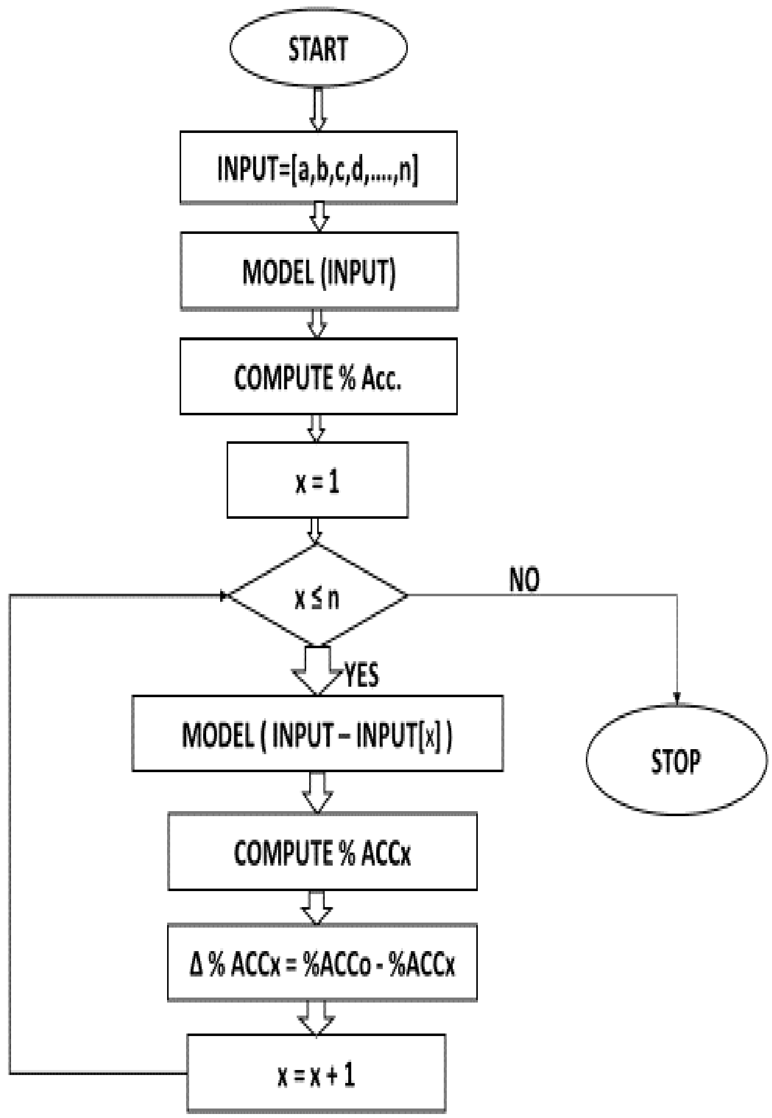
Feature Removal Sensitivity Analysis (FRSA)

Here in FRSA, the relative importance of input parameters is determined. It involves a four steps procedure as follows. Figure 4 shows the schematic representation.

- All parameters trained and tested as inputs to predict outputs
- Performance computed (Percentage Accuracy)
- One parameter removed from already trained and tested models and the new model was trained and tested without that parameter and corresponding percentage Accuracy value computed
- The procedure repeated for all parameters

Figure 4

Schematic representation of Feature Removal Sensitivity Analysis (FRSA)



Single-input Single-output (SISO)

This is a control system of simple single variable with one predictor variable as input and a single output as response. Is a non-linear sensitivity analysis where the relative importance of predictor variables is determined for the selection of most dominant input parameters. It involves a four steps procedure as follows;

- All parameters trained and tested as inputs to predict outputs
- Performance computed (Percentage Accuracy)
- All predictor parameters removed from already trained and tested model living single input parameter and a new model was trained and tested. Corresponding percentage Accuracy value was computed.
- Procedure repeated for all predictor variables individually.

For both FRSA and SISO, corresponding decrease in percentage Accuracy values obtained from the models are used to rank relative importance of parameters. Higher reduction in percentage Accuracy value indicates high relative importance and vice-versa.

Kernel Principal Component Analysis (KPCA)

PCA has been one of the popular multivariate statistical methods used to reduce the size of high-volume data. The dimensionality reduction is typically accomplished by arbitrarily determining the linear association between the variables (Wang, Zhao, and Kuang 2016). As mentioned above, however, standard PCA allows for linear dimensional reduction, while KPCA is a more efficient algorithm for mapping a non-linear data set process. The ability to work without any non-linear optimization, compared to other non-linear approaches, is the key importance of the kernel algorithm (Zhu et al. 2019). Input variables are modified and used as independent PCA variables by implementing this method (Noori et al., 2009). Kaiser-Meyer -Olkin (KMO) is one of the most frequently used statistics in any factor analysis (FA) to assess the suitability of data (Solgi et al., 2017). It is possible to explain the classification of the KMO coefficient as follows: Excellent ≥ 0.9, Very good = 0.8-0.89, Good = 0.7-0.79, Moderate = 0.6-0.69, Poor = 0.5-0.59 and Unacceptable < 0.5. The coefficients of KMO and the index of KMO are presented in equations (2). Further clarification of the PCA can be obtained in these studies (Solgi et al. 2017, Zhang 2009).

$$KMO = \frac{\sum\sum \gamma_{ij}^2}{\sum\sum r_{ij}^2 + \sum\sum r_{ij}^2} \qquad (2)$$

where r_{ij} is the correlation coefficient between the variable of i and j, and a_{ij} is the partial correlation coefficient between them.

Assuming the formula in equation (4) should match a non-linear transformation $\emptyset(x)$ from the original sample covariance matrix C in F space, the predicted non-features have zero mean:

$$\frac{1}{N}\sum_{i=1}^{N} \emptyset\ (X_t) = 0 \qquad (3)$$

$$C = \frac{1}{N}\sum_{i=1}^{N} \emptyset\,(X_i)^{\emptyset}\,(X_i)^T \qquad (4)$$

If a function of the kernel is defined as:

$$k\left(X_i\,, X_j\right) = \emptyset\ (X_i)^T\,\emptyset\, X_j \qquad (5)$$

The notation of the matrix can be used as:

$$K_{ak}^2 = \lambda_k\, N k_{ak}, \qquad (6)$$

Where,

$$K_{ij} = k\left(X_i, X_j\right) \qquad (7)$$

and a_k is the N-dimensional column vector of a_{ki} as:

$$a_k = \left[a_{k_1}, a_{k2,\ldots,}a_{kN}\right]T \qquad (8)$$

$a_k$can be solved by

$$K_{ak} = \lambda_k N a_k \qquad (9)$$

and the resulting KPCA can be calculated using

$$yk(X) = \emptyset(X)^T \; VK = \sum_{i=1}^{N} \; a_{ki} \, k \; (X_1 \; X_i) \qquad (10)$$

If there is no zero mean for the predicted dataset $\{\emptyset(x_i)\}$, the Gram matrix $\widetilde{K}$can be used to replace the kernel matrix K. The matrix of Gram is given by:

$$\widetilde{K} = K - 1_N K - K 1_N + 1_N KN \qquad (11)$$

where 1_N is the $N \times N$ matrix with all elements equal to $1/\, N$.

To specifically measure $\emptyset(x_i)$, so the kernel matrix can be constructed directly from the training data set $\{x_i\}$ (Q. Wang, 2012). The standard steps for the reduction of kernel PCA dimensionality can be summarized as(i) create the kernel matrix k from a collection of training data $\{x_i\}$ using equation (3.7); (ii) using equation (11), compute the Gram matrix $\widetilde{K}$ (iii) substituting K with $\widetilde{K}$, solve for vector $a_i$using equation 14 (iv) compute $yk(x)$ using equation (10). The polynomial kernel and the Gaussian kernel are two widely used kernels. This study employs the function of the Gaussian kernel as:

$$k(X, Y) = \exp(-1 \; \|1X - Y\|^2) \; / \; 2\sigma^2 \qquad (12)$$

Classification models

Artificial Neural Network (ANN)

ANN is a modelling tool focused on a series of linked units and nodes called artificial neurons that are built to imitate a suitable architecture of the biological neural system. It gains knowledge through training and generates outputs based on the knowledge of the relationship within the training data between the input parameters. The ANN structure consists of three different and interconnected layers (i.e. input layers, hidden layers and the output layers) in that order (Yap & Karri, 2011). As shown in Figure 3.5, each layer consists of interconnected weighted neurons (processing elements) that measure the desired output based on the input data relationship using an internal transfer function (Akğngör & Doğan, 2009). The output of each neuron is calculated by the sum of its inputs by some non-linear function. The connections are called edges. Usually, neurons and edges have a weight that changes as learning progresses.

The weight increases the signal intensity at a link or reduces it. Neurons may have a threshold such that a signal is only transmitted if that threshold is exceeded by the aggregate signal. Neurons are usually aggregated into layers. On their inputs, different layers can perform various transformations. In each neuron, the transfer function converts input signals into output signals. There are various kinds of transfer functions, among which linear, threshold and sigmoid functions are the most common (Sharma et al., 2005). A model of a neuron consists of three fundamental components, synapses (connecting links) distinguished by their power, a linear combiner that adds weighted input signals and an activation function to limit the output of the neurons' amplitude range to certain finite values. A neuron p can be represented mathematically as follows:

$$U_P = \sum_{j=1}^{m} w_{pj}\, x_j \qquad (13)$$

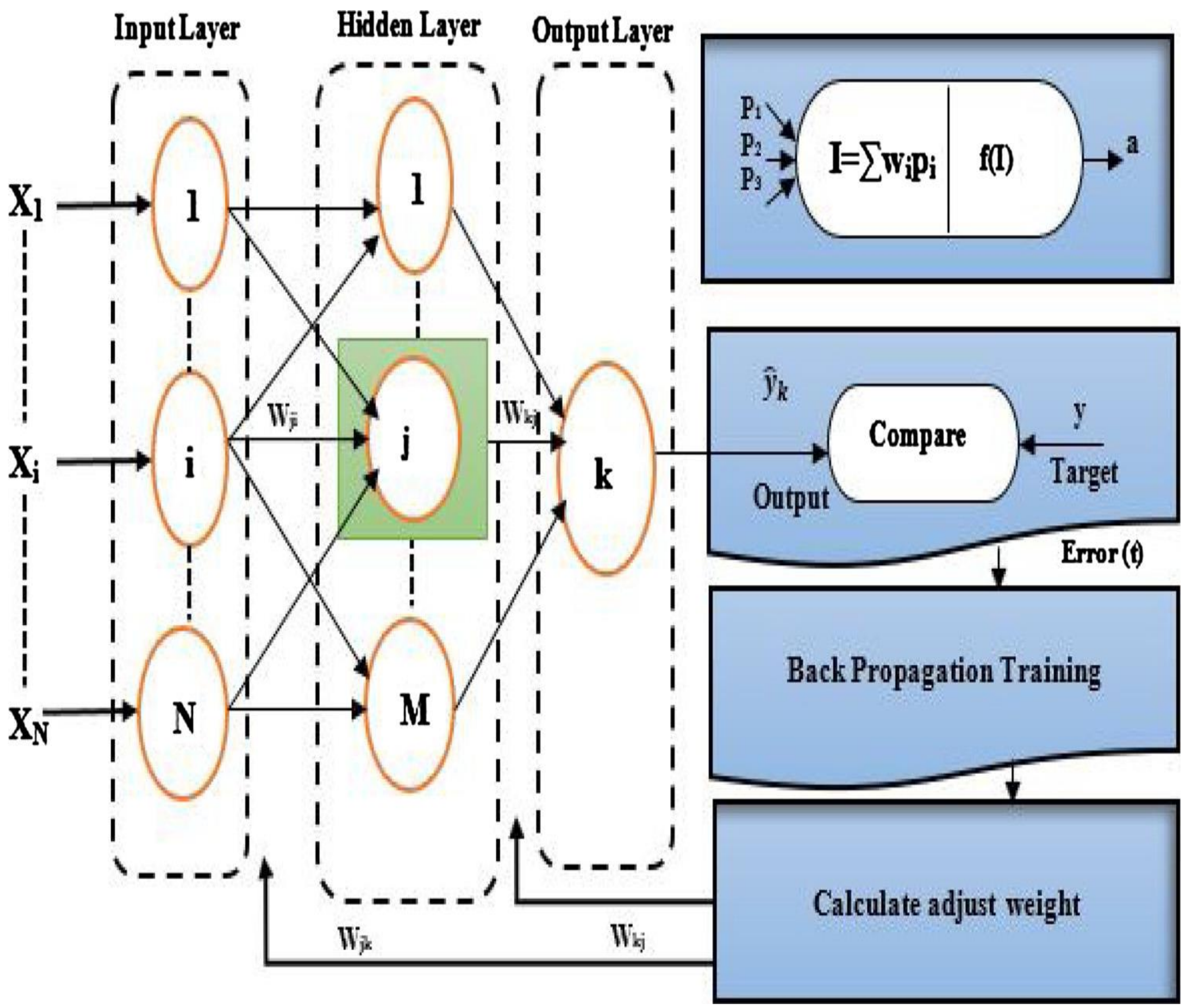
J = 1,2,3, …, m

$$yp = f(U_P + b_p) \qquad (14)$$

Where $x_1, \dots, x_m$ denotes the input signals, $w_{p1}, \dots, w_{pm}$ represents the synaptic weights of neuron p, U_P is the linear combiner, b_p the bias, *f*(.) is the activation function and y_p is the output signal of the neuron p. Information processing within the architecture of the neural network is carried out in the order in which the layers are organized, (i.e. from input layers to hidden layers and finally to the output layers). Each layer's output is an input to the next layer (Mlakar & Boznar, 2011).

Figure 5

A three-layered feedforward neural network with backpropagation algorithm.



a) Number of Hidden Layers

In a neural network, the number of hidden layers gives a network an ability to generalize. Increasing the number of neurons increases the time of computation and increases the probability of over-fitting since this can cause the network to memorize rather than generalize.

It is necessary to decide on the number of nodes in the hidden layer as it helps determine the architecture of the neural network. The research compared the corresponding goodness of fit value with the different number of nodes. In calculating the number of hidden nodes, the following equation was used (Yuen & Lam, 2006).

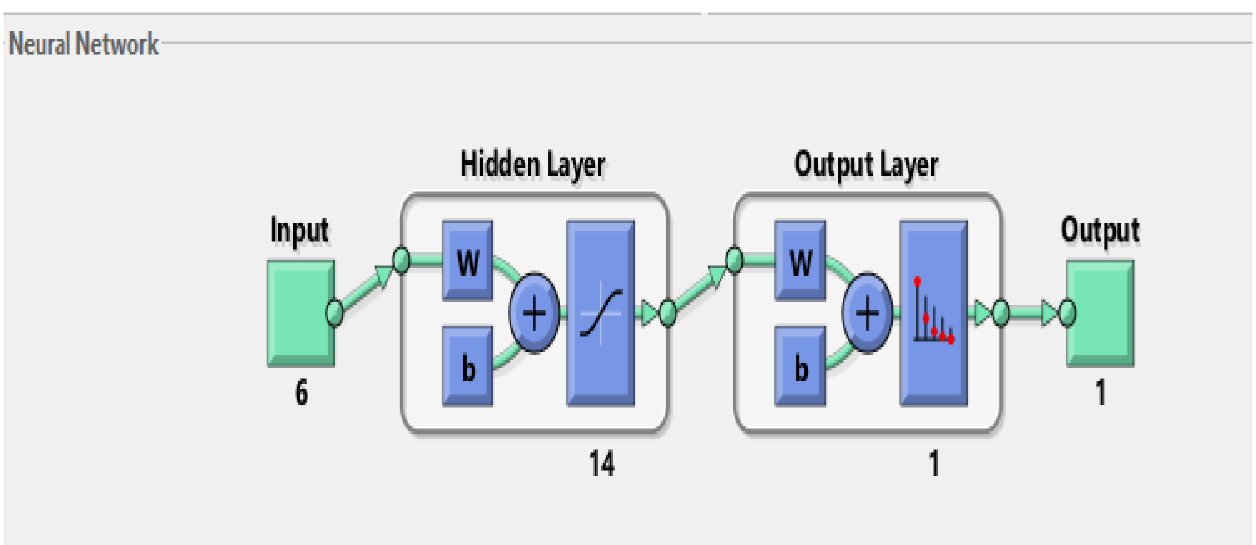
$$n = \frac{N_i + N_0}{2} + \alpha \qquad (15)$$

where n denotes the number of hidden nodes, N_i is the number of input neurons, N_0 is the number of output neurons and α was arbitrarily taken to be 2.

The value of the determination coefficient was used, using the dataset, to calculate the optimum number of hidden neurons in our neural network. As seen in Figure 3.6, this research,, therefore, settled on fourteen (14) hidden neurons.

Figure 6

ANN model showing the number of hidden neurons



Support vector machines (SVM)

SVR is a machine learning technique and a regression method used for the modelling of complex linear and non-linear processes which is being developed based on support vector machine (SVM) concept. It works by identifying the optimal decision boundaries that separate data points from different groups or classes and then predict the class of new observations based on this separation boundary. Like other SVM based methods, minimizing the operational risk is the major objective of the SVR which is different from other black box models where the main objective is to minimize the error between measured and predicted values. The SVR involved two stages, at first, the data are fitted into linear regression, then the output passes through a nonlinear kernel which takes the nonlinear form of the data. Given a set of training data $\{(x_{i,,}{}_{d_i})\}_i^N$ (where $x_{i,,}{}_{d_i}$ and N represents input vector, actual value and a total number of data patterns). The general expression of the SVR function can be written as (Xu et al., 2015):

$$y = f(x) = \square\ \emptyset(x_i) + b \qquad (16)$$

Where ω, $\emptyset(x_i)$, $x\ and\ b$ represent m-dimensional weight vector, feature spaces, non-linearly mapped from an input vector and bias respectively (Gunn & others, 1998). Parameters b and w can be computed by giving positive values for the slack parameters of $\xi and\ \xi^*$and minimization of the objective function as (wang et al 2015):

$$Minimize: \frac{1}{2}\|w\|^2 + C\left[\sum_i^N(\xi_i + \xi_i^*)\right] \qquad (17)$$

$$Subject\ to: \{\ w_i\emptyset(x_i) +\ b_i - d_i \le \varepsilon + \xi_i^* \qquad (18)$$

where $\frac{1}{2}\|w\|^2$ is the weights vector norm, C is the regularized constant that defines the trade-off between the empirical error and the regularized expression, and ε is the size of the tube that corresponds to the precision of the approximation placed within the training data points. By defining Lagrange multipliers $\alpha i\ and\ \alpha_i^*$, the above optimization issue can be modified to a dual quadratic optimization issue. By solving the quadratic optimization problem, the vector w can be determined as (Xu et al., 2015):

$$w^* = \Sigma_{i=1}^N(\alpha_i - \alpha_i^*)\varphi(x_i) \qquad (19)$$

The SVR's final expression can be written as (Xu et al., 2015):

$$f(x, \alpha_i, \alpha_i^*) = \Sigma_{i=1}^{N}(\alpha_i - \alpha_i^*)K(x, x_i) + b \qquad (20)$$

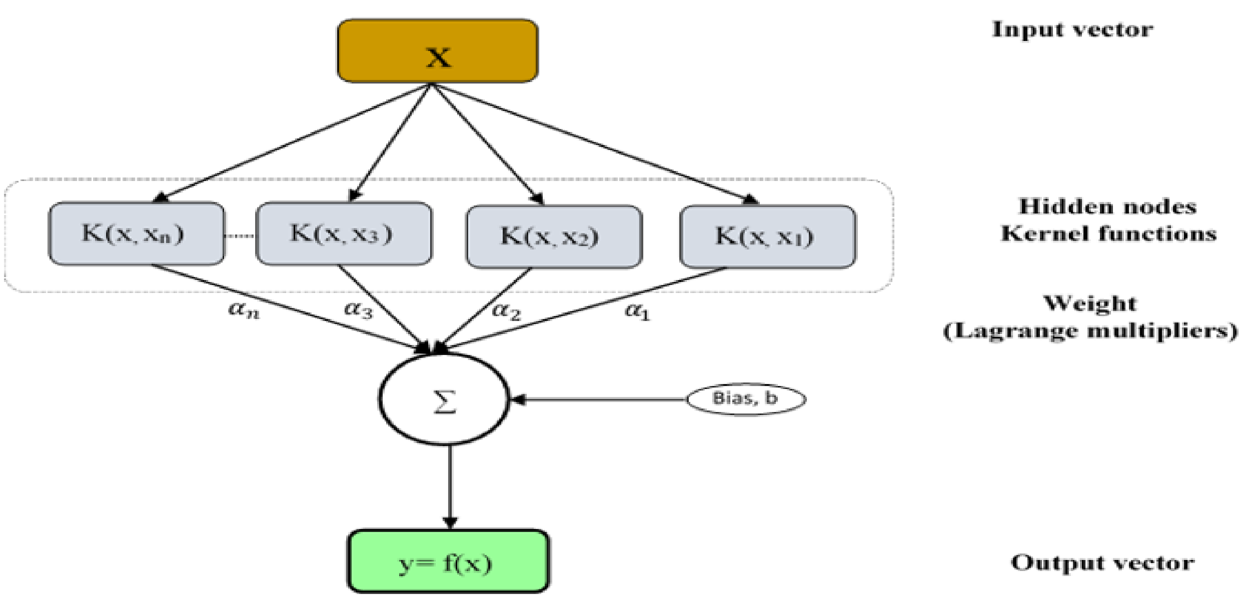
$k(x_i, x_j)$ is a kernel function that carries out nonlinear mapping into the space of the feature, and b is the bias term. The Gaussian Radial Basis Function (RBF) kernel is a widely used kernel feature represented as:

$$k(\, x_{i,} x_2) = \exp(-\gamma \|x_1 - x_2\|^2) \qquad (21)$$

where γ is the kernel parameter. The general conceptual structure of SVR is presented in Figure 7 below.

Figure 7

Schematic representation of the SVM algorithm



Boosted Regression Tree (BRT), model

Boosted Regression Trees (BRT) are well-known algorithms for the problem of classification. The Classification and Regression Trees (CART) model is a univariate binary decision hierarchy with each node (internal) in the tree specifying a binary test on a single parameter, a branch signifying an outcome and the leaf node representing the class distribution or class labels. CART works by selecting the best parameter at the root node and divide the data into two classes, dividing the data into two disjoint sections in such a way that the class labels in each branch are as uniform as possible, and then dividing each section is applied successively, and so on. If there are examples from n classes in the dataset A, $gini$ index, $gini\ (A)$ is defined as;

$$gini\ (A) = 1 - \Sigma_{j=1\ to\ n}\, p_j\ \hat{}2 \qquad (22)$$

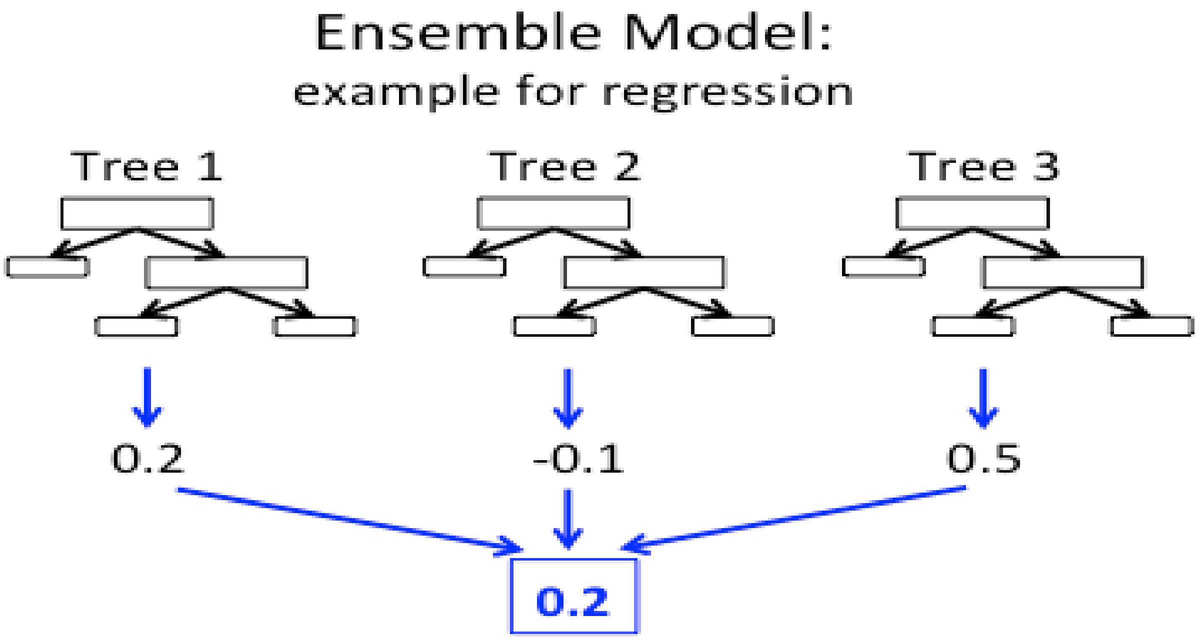
where the relative frequency of class j in A is p_j (Chong et al., 2005). Datasets A is separated into two smaller subsets A1 and A2 with sizes N1 and N2, the split data $gini$ index contains examples of n groups, then the $gini$ index, $gini\ (A)$ is defined as:

$$gini_{split}(A) = \frac{N_1}{N}\ gini\ (A_1) + \frac{N_2}{N} gini\ (A_2) \qquad (23)$$

BRT comprehensively checks for univariate splits. The attribute gives the smallest $\boldsymbol{gini_{split}(A)}$ to split the node. BRT extends the tree from a root node recursively and then prunes the big tree down gradually. The benefit of a decision tree is that it is very easy to derive classification rules from trees. More specifically, information can be expressed by a decision tree in the form of if-then rules; for each path from the root to the leaf node, one rule is made. Figure 8 shows the schematic representation of the model.

Figure 1

Schematic representation of BRT model



Logistic regression model

Logistic regression is a linear relationship model used generally to model discrete output parameters, especially for binary output variables. The basic concept is to assess event likelihood by observing the association between input and out variables (Dong et al., 2009). When the expected likelihood of an occurrence is greater than or equal to 0.5, it is considered the first group, i.e., event; otherwise, the second group, i.e., no event (Johnson, Wichern, and Johns 2002)

This is a special case of the Generalized Linear Model (GLM) that generalizes the ordinary linear regression by allowing an output variable that follows the exponential family through an appropriate relation function to be connected to the linear model. When the output variable is binomially distributed with parameter P_i, the logit function is used for the link function. The LR decides the relationship between the output variable $Y = (y_1, \dots, y_n)$, given $P = (p_1, \dots, p_n)$, and a set of k predictors, $X = (x_1, \dots, x_k)$:

$$y_1|P_1 \sim Bernoulli\ (p_i)\ , i = 1, \dots , n \qquad (24)$$

$$logit(p_i) = \log(\frac{p_i}{1-p_i}) = \beta_0 + \beta_1\, x_{1i} + \cdots \beta_k x_{ki} \qquad (25)$$

The likelihood of occurrence of crash severity level *i* denoted by P_i, can be calculated by solving $\beta = (\beta_1, \dots , \beta_k)$ using the maximum probability process (Annette et al., 2008).

$$p_i = \frac{\exp\{\beta_0 + \beta_1\, x_{1i} + \dots, + B_k\, x_{ki}\}}{1 + \exp\{\beta_0 + \beta_1\, x_{1i} + \dots, + B_k\, x_{ki}\}} \qquad (26)$$

Based on the RTA's estimated, accident severity level can be predicted. LR can either be multinomial or binomial. Multinomial logistic regression (MLR) or multinomial logit is using the multi-class output variables. The multinomial regression with the multinomial output variables, $Y = (y_1, \dots , y_n)$ and multiple input variables, $X = (x_1, \dots , x_n)$, consists of *C*−1 non-overlapping logit models. The likelihood of occurrence of accident severity level *c* (i.e., c^{th}logistic model), P_{ic}, can be estimated as;

$$P_{ic} = P(y_i - c|x) = \frac{\exp\{\beta_{c0} + \beta_{c1}\, x_{1i} + \dots, + \ \beta_{ck}\, x_{ki}\}}{1 + \sum_{c=1}^{c=-1} \exp\{\ \beta_{c0} + \beta_{c1}\, x_{1i} + \cdots, + \beta_{ck}\, x_{ki}\}} \qquad (27)$$

The crash severity level can be predicted based on the estimated P_{ic}.

Model performance evaluation

In model classification analysis, a classifier is typically evaluated by a confusion matrix (Table 3.1). In the table, the columns are the predictions of the classifier and the rows are the real classes. The number of positive cases rightly listed are represented as True Positive (TP). The number of positive cases wrongly labelled as negative is the False Negative (FN). The number of negative cases which are wrongly categorized as positive cases is False Positive (FP) and the number of negative cases correctly reported as positive cases is True Negative (TN).

Table 0

Confusion matrix and the four measurements for 2-classes classification.

	Predicted Positive (fatal)	Predicted Negative (non-fatal)
Actual Positive (fatal)	True Positive (TP)	False Positive (FP)
Actual Negative (non-fatal)	False Negetive (FN)	True Negative (TN)

By convention, in imbalanced data modelling, we regard the minority class as the positive class, while the majority class is regarded as the negative class. In classification problems based on the confusion matrix, we derive most of the performance measures used. Equations 28-32 below outline some of these success steps.

In classifying RTA severity, it is not possible to simply capture the classification power of a model by the correct classification rate. Accident severity datasets are usually imbalanced, having disproportionately more data sets in the non-fatal class than in the fatal class. If untreated, such a structure could lead to overfitting or training models that look promising on the outside with high accuracy rates (accuracy is characterized as the model's ability to correctly predict classes of accident severity on a test range.; see Eq. (28)), but in reality, non-informative. A trivial model that predicts all accidents to be non-fatal is a severe example of a poor model in a two-class issue of fatal and non-fatal classes. Such a model will have a very high rate of accuracy, while the value of an accident classification model is primarily focused on accurate classifications of higher severity classes. (e.g., fatal crashes), usually known as "sensitivity" (i.e., the model's capacity to correctly identify the degree of severity as 'fatal' (Jeong et al., 2018), see Eq. (29)). In contrast, a model classifying all accidents as fatal would produce high sensitivity, but a lower accuracy score. There is also a strong trade-off between sensitivity and accuracy scores of accident severity models that can only be overcome by treating imbalanced data appropriately. Therefore, this imbalanced data structure requires extra steps in model evaluation and training: (i) using acceptable measures of assessment, and (ii) balancing before training the dataset.

In evaluating model efficiency, limitations of the classification accuracy rate could be addressed by using additional statistical measures, namely, true positive (TP), true negative (TN), false positive (FP), and false-negative (FN) (see Table 3.1 for a detailed description) to bring about more comprehensive metrics. Using these measures, accuracy, sensitivity, specificity, precision and G-mean as defined in equations 28-32 respectively, can easily be evaluated for 2-class classification problems. Collectively, these metrics help to show a more accurate image of the overall success of the model (Parikh et al., 2008). Ultimately, as a compact measurement measure, the geometric mean (or G-mean) can be used to compare the general performance of different models. The G-mean is measured as the square root of the sensitivity and specificity product and will have high values when both specificity and sensitivity are high and there is a slight difference between the two metrics. (Kubat et al. 1997). Finally, while it is important to disclose several metrics that can provide a detailed image of model efficiency, steps need to be taken in the first place to generate high-performance models. This can be accomplished by creating a balanced dataset based on the original imbalanced datasets on which it is possible to train models.

$$\text{Accuracy} = \frac{TP + TN}{TP+TN+FN+FP} \tag{28}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{29}$$

$$\text{Specificity}= \frac{TN}{FP+TN} \tag{30}$$

$$\text{Precision Factor}= \frac{TP}{TP+FP} \tag{31}$$

$$\text{G-Mean}= \sqrt{Sensitivity \times Specificity} \tag{32}$$

The accuracy of a classifier is the most significant and widely recorded measure. This metric tests an algorithm's overall performance. As previously shown, however, predictive accuracy can be a misleading indicator of measurement when dealing with imbalanced data. This is because more weights are imposed on the majority class in such situations than on the minority class, which makes it harder for a classifier to do well on the minority class. Sensitivity, another indicator of efficiency, measures the accuracy of positive cases, while specificity measures the precision of negative cases. Sensitivity evaluates the efficacy of the classifier in the positive / minority class, while specificity evaluates the efficacy of the classifier in the negative/majority class. There's generally a trade-off between the sensitivity and the specificity for any given study. On the other hand, precision measures the exactness of a model. For a classifier, a greater precision value is an indicator of a strong classifier.

Geometric Mean

The Geometric Mean (G-Mean) is a metric that calculates the balance between the success of classifications in the majority and minority groups. A low G-Mean is an indicator of bad results even though the negative cases are correctly listed as such in the classification of positive cases. To avoid overfitting the negative class and underfitting the positive class, this measure is necessary.

Chapter III summary

Significant data of accident records (input data) were obtained containing relevant information which was pre-processed to improve the data fitting and prediction accuracy. The data was normalised to bring all input and output variables to same range before feeding them to the AI model. Different sensitivity analysis techniques were tested in order to select the most dominant input parameters. The three AI models (ANN, BRT, SVM) and statistical model were tested for the best accuracy. The performance of the four models was compared to select the best model with the highest accuracy

CHAPTER IV

Findings and discussion

General Overview

The speedy growth in the number of vehicles using rural highways in Kano leads to considerable growth in the traffic congestion which was marked as a serious problem posing a threat to the road users, the economy, travelling behavior and the environment at large. These problems causing severe traffic incidents that lead to loss of lives and properties to individuals and the government at large.

Sensitivity analysis results of the variables selected for the study was shown as well the results of the comparisons of the road traffic accident severity classification models shows the performance measures (accuracy, sensitivity, specificity, precision, G-mean) of all the models. It was presented in this chapter the accident distribution by location to show the accident-prone zones and also the accident cause contribution.

Accident Records

The section of the road understudy was classified as a dangerous road section by the Federal Road Safety Commission (FRSC) due to its several numbers of black spots. Most drivers have also agreed with that classification. Accident records were collected from the Federal Road Safety Commission. Table 4 and Figure 9 below gives a summary of the accidents recorded by the commission on the route from the year 2009 to 2013.

Table 0

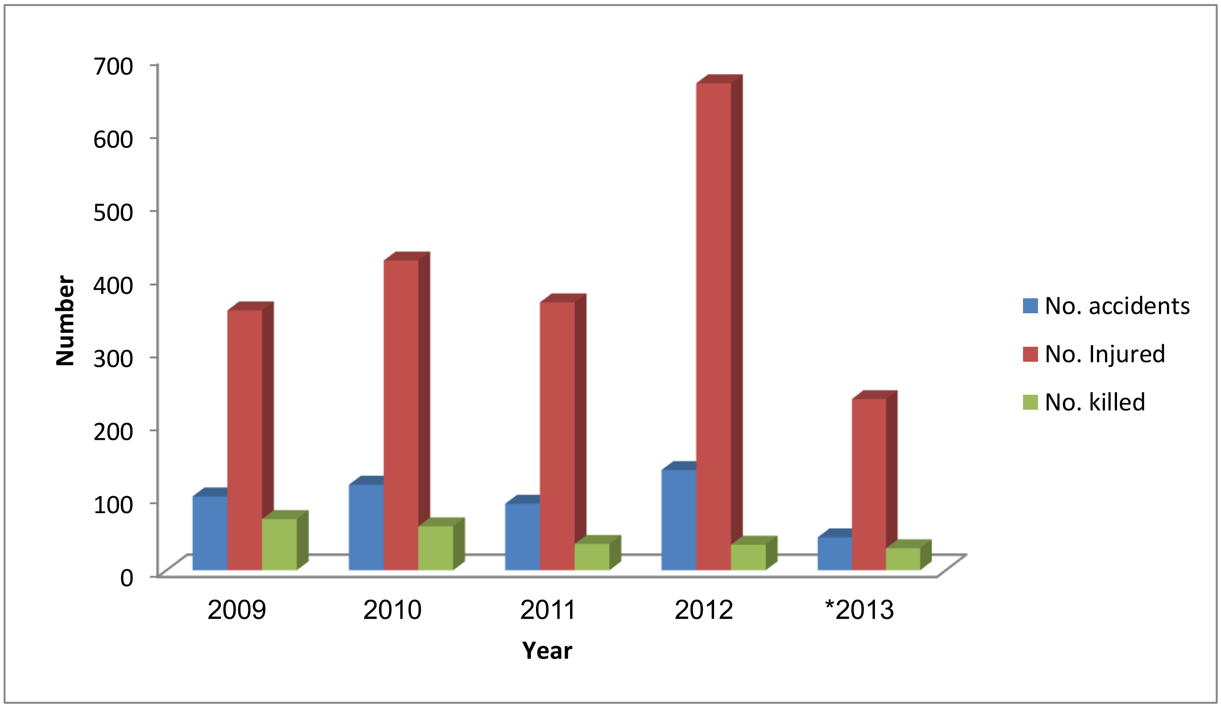
Accident Record

Year	No. accidents	No. Injured	No. killed	Total causalities
2009	79	313	54	367
2010	104	352	50	402
2011	78	327	33	360
2012	93	581	28	609
2013	45	234	30	264

Source: Federal Road Safety Corps (FRSC)

Figure 9

Accident Data for Road Section from January 2009 to November 2013



The average fatality on the road section is 42 deaths in 89 accidents annually translating to approximately 1- fatality in every 2 accidents, and leaving over 393 people with various degrees of injuries.

Accident distribution along the case study area

From the accident data collected from the FRSC, the accident distribution along the case study area was found to be in twenty (20) different locations of the highway. A significant number of accidents were recorded along the route within 5 years. Wudil-Bauchi road with 15.14 was found to have the highest percentage of accidents recorded. Wudil-Kano, Gano town, Polac and Garin-dau also having higher percentages of accidents with 12.81, 12.65, 8.49 and 6.16 respectively. Other locations with their accident distribution percentages were shown in Figure 4.1 below. The reasons behind the rampant incidents of RTA on this route varies with locations. Among the major reasons are unclear driver vision along sharp bends on the highway without proper traffic signs. Presence of market places that increases traffic volume also plays a vital role in escalating accident rate which attract a large number of people from within and outside the state, generating heavy traffic with vehicles exceeding the road provision. Others are the presence of work-zones along the road that necessitates road diversions to a single lane undivided highway. These work-zones increases vehicle congestion especially at entrance and exit phases.

However, the study is limited to eight (8) contributory causes of accidents which are WOV, SLV, TBT, DGD, MDV, >1, RTV and MOC. SLV was found to be having the highest percentage distribution with 42.76%. over speeding is the second most common cause of the accident (Ismail & Yahia, 2013). Table 5 shows the accident causes code. Other percentages of causes contributing to RTA are shown in Figure 10 and Figure 11 represent the graphical interpretation of accident distribution by location.

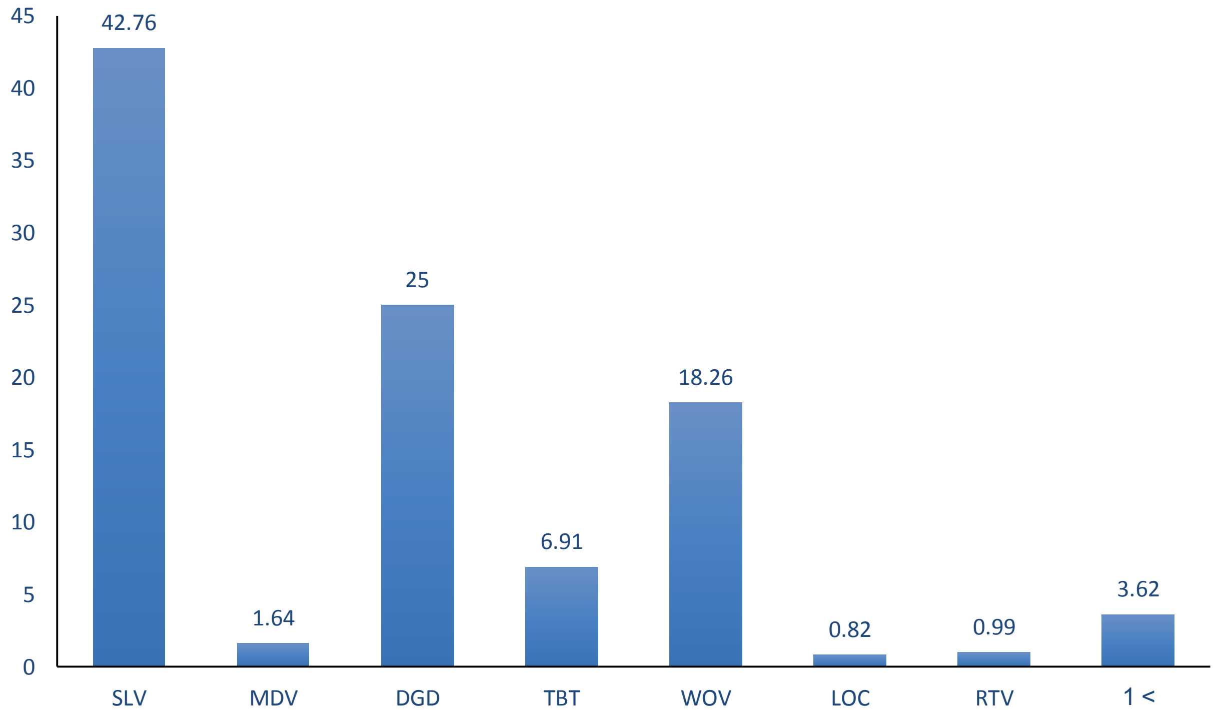
Table 5

Causes of Accidents Code

Causes of Accidents	Code
Wrong overtaking	WOV
Over speeding	SLV
Tire burst	TBT
Dangerous driving	DGD
Mechanically defective vehicle	MDV
Morethan 1 cause	>1
Fatigue	RTV
Run-off	LOC

Figure 10

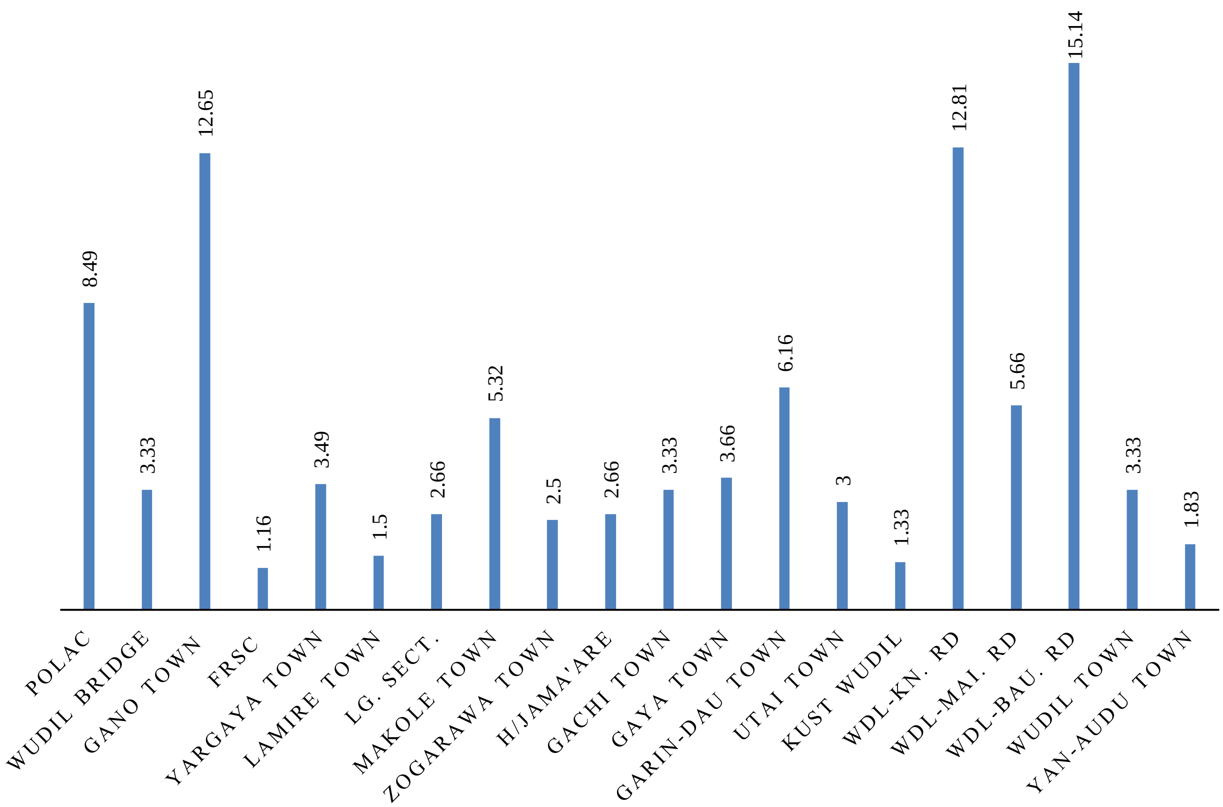
Accident cause contribution



Exploratory data analysis

Figure 11

Accident distribution by location



Result of sensitivity analysis

To obtain excellent and significant results, it is important to select the most dominant input parameters for any black-box model, i.e. input parameters having high relevance with the output variables. Several AI-based models using different kernel functions were employed to train many severity models and the best model was obtained using Quadratic SVM kernel. As such, a nonlinear Quadratic SVM kernel was used in the two types of the nonlinear sensitivity analysis of the accident severity predictor variables adopted in this study to determine the most dominant input parameters. This is because the linear forms of sensitivity analysis used for selection of model's inputs parameters such as the linear regressions and Pearson correlations have been criticized for the selection of relevant input parameters in many engineering studies since they mostly have a nonlinear relationship (Gan et al. 2012; Mansourkhaki et al. 2018). Some reports have questioned the use of such linear measures (Issac & Israr, 2014). In this study feature removal sensitivity analysis and the single input single, out sensitivity analysis was used.

Discussion of Feature Removal results

In the feature removal sensitivity analysis, A quadratic SVM was used to model the accident severity using all the inputs six parameters and the model accuracy was recorded as 94.8%. subsequently, one parameter from the six inputs was removed with replacement and each time a parameter is removed, a model is trained with the remaining five inputs parameter and the accuracy is recorded. The procedure was repeated until all the parameters were removed. The corresponding decrease in the accuracy upon removal of each parameter is recorded. The relevance of the parameters was ranked based on the resulting decrease in the model's accuracy. The most relevant parameter is that which resulted in the highest decrease in the model accuracy as compared with the model having all the six parameters (Nourani, Gökçekus, et al., 2020). As shown in Table 6, C is the most dominant parameter resulting in a decrease of 25.9% upon its removal from the six input parameters. The other parameters have almost similar importance resulting in a decrease accuracy of 10-11%. This shows that all the parameters have reasonable importance and should be used in the model.

Table 6

Result of feature removal sensitivity analysis

Parameters	Accuracy (%)	Rank
All without R	85.0	6
All without V	84.8	5
All without I	84.7	4
All without t	84.3	3
All without T	84.1	2
All without C	68.9	1

Discussion of Single-input-single output results

Here, a simple single variable system with one predictor and one output as the response was adopted for the selection of most dominant input parameters. Predictor variable with less accuracy shows less relevance to the output variable, thus, eliminated. Quadratic SVM was similarly used for training the models in the single-input single-output sensitivity analysis and the result presented in Table 7. C having the highest percentage shows a good relevance with the response variable and is the most dominant. The other five (5) predictor variables show a slight change in their accuracies, thus showing closely similar relevance with the output. The high accuracy of all models trained showed a significant relevance of all predictor variables with the response. As such, all input parameters can be used for the model's classification. Table 7 shows the performance accuracy and ranking of the six accident severity predictor variables.

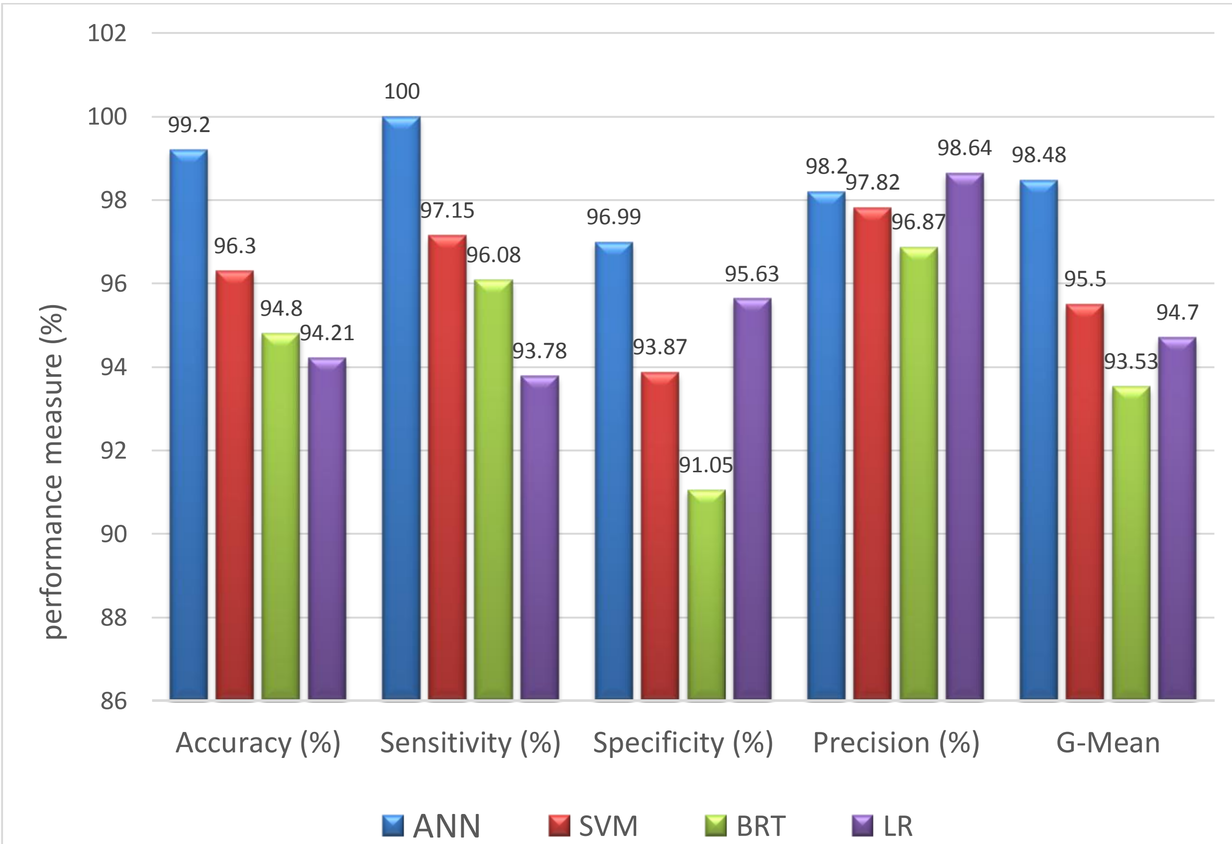
Table 7

Single input single output analysis results

Parameters	Accuracy (%)	Rank
C with output	94.8	1
T with output	72.5	2
I with output	71.9	3
t with output	71.3	4
V with output	70.1	5
R with output	69.8	6

Figure 12

Graphical representation of the model's performance



Discussion of results (confusion matrix and ROC curve)

The results of the severity classification models were graphically presented in the confusion matrix and ROC curve as shown in Figure 13 and Figure 14 respectively. The results further indicated that the ANN model has a higher classification accuracy than all other models. The ANN has classified the fatal accident with 98% accuracy and the non-fatal accident with 100% accuracy. The high classification accuracy of the fatal accident has been supported by a study conducted by (Sarkar & Sarkar, 2020) where ANN predicted fatality rate of 96.15% accuracy. Similarly, the SVM and BRT classified fatal accidents with 96 and 95% accuracy.

The ability of the SVM to fit the data to linear regression before finally fitting it into the nonlinear function and its ability minimize the operational risk in the process gives it the upper hand in its ability to classify the fatal accidents with high accuracy. The BRT models also obtain its high performance from an ensemble of different models which several studies such (Nourani, Gökçekuş, et al., 2020) proved it improves models accuracy. The result obtained by the LGR was also good, however not as high as the AI-based models. This indicates that there is a strong linear relationship between the input variables and the accident severity. The LGR as seen in Figure 13 classified the non-fatal accidents with higher error than all the AI-based models resulting in overall decrease accuracy of the model.

Figure 13

Models confusion matrix (a) SVM, (b) BRT, (c) ANN, (d) LR

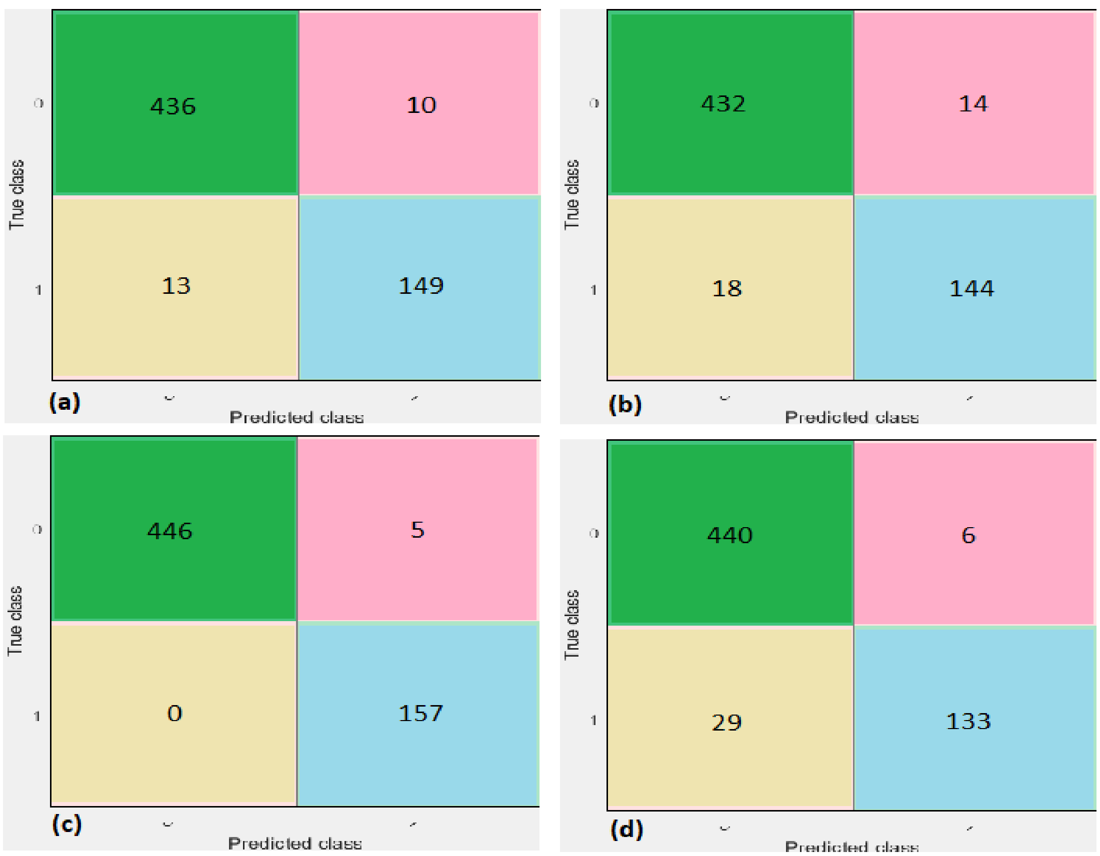
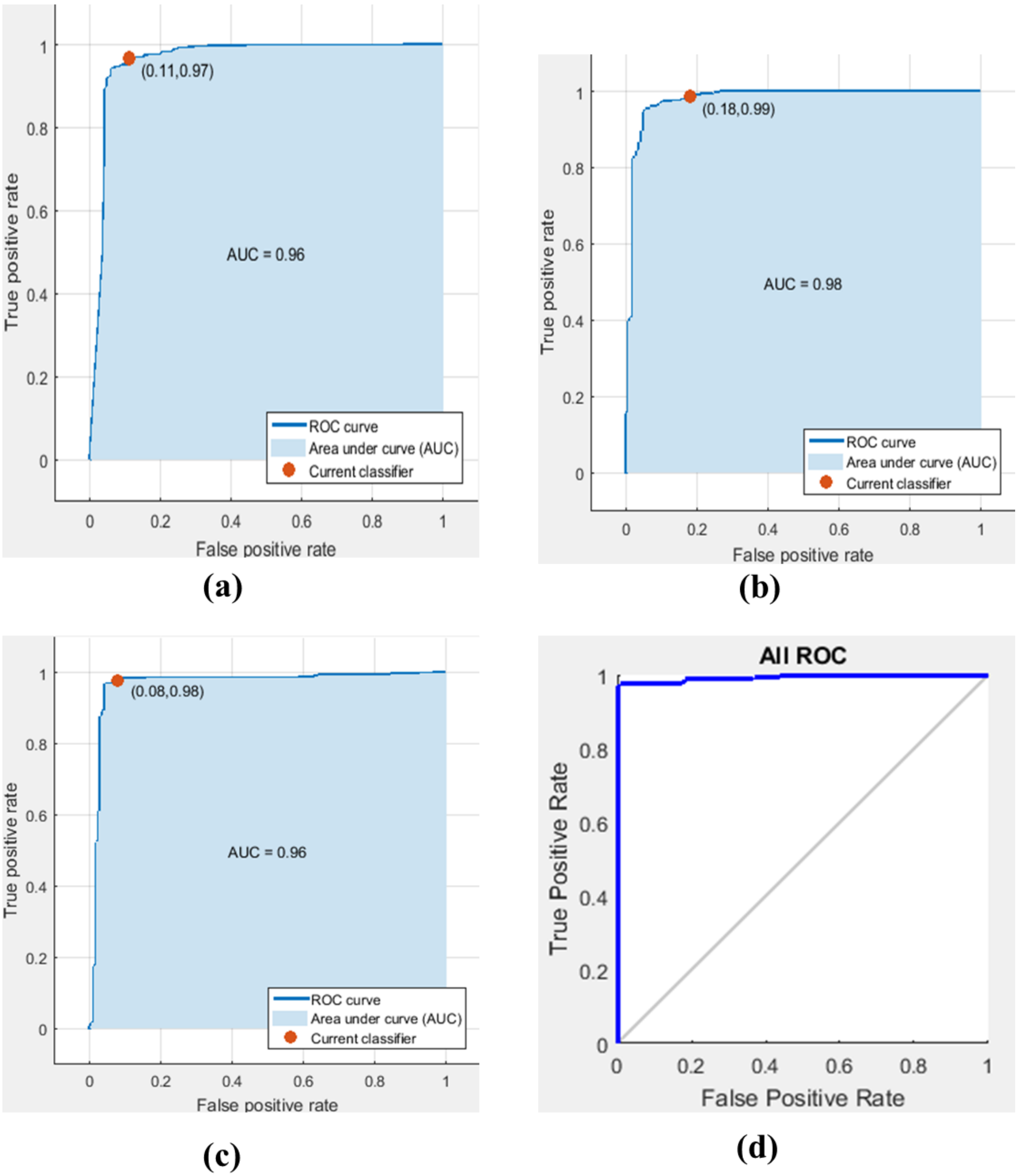


Figure 14

Models ROC curve (a) BRT, (b) LGR, (c) SVM, (d) ANN



Chapter IV summary

This chapter shows that single input-single output sensitivity analysis technique has the highest accuracy with much relevance with the output variables. Best model was achieved using quadratic SVM kernel, and also, ANN model has shown the highest sensitivity and outperform the other three (3) models. This indicates a good result and will give a balance between success of classification in the minority and majority groups and will surely handle the threat of underfitting in the positive class and overfitting in the negetive class.

CHAPTER V

Discussion

For the classification of the accident severity, three AI-based classification models (ANN, SVM and BRT) and a logistic regression model were used in the study. The data was divided into 75% training 25% testing for modelling purposes. The six potential inputs variables were all used as inputs to all the classification models. This due to the higher interaction obtained between the accident severity and each of the inputs parameters as seen in Table 6 and 7. For training the ANN model, Levenberg Marquardt backpropagation algorithm with tansig function were used to train many ANN models with different structure by changing the number of hidden neurons in the models. The optimum structure which gives the highest accuracy in the accident severity classification was obtained using 14 hidden neurons and six input parameters. For the SVM and BRT models, the best models with high accuracy were obtained using the quadratic kernel and least square boost algorithms respectively. All the AI-based models have resulted in high accuracy (Table 8) with accuracy value >90%. This is due to the versatility of the AI-based models in establishing a nonlinear relationship in complex engineering processes. The ANN model has demonstrated the highest accuracy with 99.2% accuracy level followed by SVM, BRT and lastly LGR.

Also, The ANN model has demonstrated the highest sensitivity with 100% sensitivity level followed by SVM, BRT and lastly LGR. ANN model having the highest percentage of sensitivity is a good indication of the efficiency of the model. It shows how good it measures the accuracy of positive cases and how it evaluates the efficacy of the classifier in the minority class. For G-mean, ANN model outperforms the other three model with 98.48%, followed by SVM, LR and lastly BRT with 95.50, 94.70 and 93.53 percentages respectively. All are an indication of good results as they will give a balance between the success of classification in the minority and majority groups and will tackle the treat of underfitting in the positive class and overfitting the negative class. Figure 4.4 shows the graphical representation of the models performance.

Table 8

Comparing models classification accuracy

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	G-Mean
ANN	99.20	100.0	96.99	98.20	98.48
SVM	96.30	97.15	93.87	97.82	95.50
BRT	94.80	96.08	91.05	96.87	93.53
LR	94.21	93.78	95.63	98.64	94.70

Road traffic accident (RTA) as a significant public health matter is an event leading to personal injuries or property damage taking place in an area proposed for public transport or places generally used for transportation. RTA is a serious scenario posing a great panic to both road users and pedestrians due to loss of lives and properties which need to be grasped by a multi-disciplinary approach. Accidents cannot be prevented, as such; measures have to be taken to reduce the likelihoods of occurrences. These mitigation measures includes every attempt possible to save lives, break free from injuries, lower the degrees of injuries severity, avoid property damages, treatment reduction and compensation costs and also avert the deprivation of productive time and morale. These measures are proposed to lower the chances of accidents happening by carefully looking into accident causes thereby reducing accident severity. These mitigation measures can be categorically classified into primary and secondary measures.

Primary measures: They are measures proposed to remove all circumstances leading to road accidents. E.g. speed reduction, correct traffic signs, underage driving policies, enforcing the correct attitude for driving, safe vehicle driving and use of alcohol on wheel etc.

Secondary measures: As earlier said, accidents cannot be prevented totally rather controlled, therefore, these measures reduce the injury severity in the cause of an accident. E.g. use of seat belts, use of helmets, smoke alarms, passengers limiting etc.

Measures to enhance Safety on Rural Highways

To enhance the safety of both road users and pedestrians on rural highways, excessive identification into the road users, vehicles and road infrastructures interactions will give a solution to the alarming RTA worldwide.

Road users

- Pedestrians, passengers and drivers should be educated about traffic signs and rules
- Only qualified persons should be issued driving license
- Limiting passenger number in a vehicle
- Driving on drug influence should be avoided

Vehicle

- Use of healthy vehicle will reduce the risk of accident occurrence

Road infrastructure / condition

- Pedestrian paths should be provided. Also, crossings for pedestrians at intersections should be adequate.
- Well maintained road with safety signs and proper signs should be provided
- Separate lanes should be provided for fast-moving and slow-moving vehicles
- Proper visibility should be maintained on junctions and work-zones

Legislative / Enforcement

- Strict traffic rules should be enforced by concerned authorities
- Haphazard parking should be prevented
- Stray of animals should be stopped near roads
- helmets and seat belts should be implemented

Chapter V summary

This chapter shows that all the AI-based models have resulted in high accuracy with accuracy value >90%. This is due to the versatility of the AI-based models in establishing a nonlinear relationship in complex engineering processes. Also, ANN model has demonstrated the highest sensitivity with 100% sensitivity level followed by SVM, BRT and lastly LGR.

CHAPTER VI

Conclusion and recommendations

The target of this research is to apply different classification models that able simultaneously to predict the degree of severity of RTA on rural highways with higher accurancy and sensitivity. When practising RTA severity classification models, giving an attention to sensitivity rates helps us to understand the severer effects of serious incidents. There are two reasons for which such models can be used: (a) Identifying variables that relate most to the RTA by sensitivity analysis and (b) real-time crash accident severity prediction in semi-autonomous vehicles. latterlly, if the models foresee serious accidents under human control, they can notify the autonomous entity of the decision to take control of the driver. Moroever, 4 hybrid machine learning algoriths and two procedures to balance the initially unbalanced accident data of motor vehicle crash severity level, with a large (608) crash datasetwere used in this research. Comparison of all trained models shows that FFBPNN combined with the over-sampling traetment give the best classification performance both in accuracy and sensitivity. Logistic regression model having the least classification performance among the models proves that the dataset is more of a non-linear relationship.

Recommendations for future work

- It is well known that the actual speed of a vehicle when an accident occurred is one of the most significant factors causing various injury levels. Unfortunately, there is no details in our dataset of the speed of the vehicles in an accident. It is highly likely that if the speed was known and used as a variable, it might have helped to boost the performance of the models analyzed in this paper.
- Future studies should focus on the spatial modeling of injuries to RTAs in Kano State and then scale up to Nigeria as a whole. Accident classification models should be generalized and the cumulative effects of the various explanatory variables should be taken into account.
- In similarly evaluating classification methods as proposed herein, further researches on accident severity classification models can take accident costs into account. When

making decisions, transportation agencies and policymakers may want to be mindful of the degree of uncertainty and inaccuracy of accident severity classification models. As reported, the accident costs were underestimated by all classification techniques, including the best model. Other statistical models and machine learning methods for classifying crash severity can therefore be investigated to reduce underestimation for future research, and their output can be correlated with the results of this analysis.

- Future research should focus on methods of database balancing as well as evaluating various functions of information transfer across model layers. In addition, it is recommended that databases with various dimensions be explored where possible, because a greater number of data will increase the generalization power and, thus, the adaptation of the data to the model. It is therefore proposed to investigate other methods based on network architectures, such as complex networks that are ideal for studying random phenomena of a complex nature, derived from multiple causes, irrespective of the structural versatility of models in the selection of variables and construction of stochastic models.

Chapter summary

In this chapter, the target of the research to apply different classification models to predict the degree of severity of RTA on rural highways with higher accurancy and sensitivity, ANN model happened to show highest accuracy while logistic regression model gave the least performance amongst the models. As such, the aim of the reseach is meet.

REFERENCES

Abdous, H., & Mahmoudabadi, A. (2018). *PW 0581 Burden of road traffic injuries on rural residents, a case-study for iran*. BMJ Publishing Group Ltd.

Afolabi, J. O. (2009). Topical Issues on Road Safety in Nigeria. *International Journal*, *17*(42), 1069–2406.

Afolabi, J. O., & Gbadamosi, K. T. (2017). Road traffic crashes in Nigeria: Causes and consequences. *Transport & Logistics: The International Journal*, *17*(42).

Afukaar, F. K., Antwi, P., & Ofosu-Amaah, S. (2003). Pattern of road traffic injuries in Ghana: implications for control. *Injury Control and Safety Promotion*, *10*(1–2), 69–76.

Agbonkhese, O., Yisa, G. L., Agbonkhese, E. G., Akanbi, D. O., Aka, E. O., & Mondigha, E. B. (2013). Road traffic accidents in Nigeria: Causes and preventive measures. *Civil and Environmental Research*, *3*(13), 90–99.

Akğngör, A. P., & Doğan, E. (2009). An artificial intelligent approach to traffic accident estimation: Model development and application. *Transport*, *24*(2), 135–142. https://doi.org/10.3846/1648-4142.2009.24.135-142

Akkaş, M. A., SOKULLU, R., & Ertürk Çetin, H. (2020). Healthcare and patient monitoring using IoT. *Internet of Things*, *11*, 100173. https://doi.org/10.1016/j.iot.2020.100173

Annette, J., Dobson, A. B., & Barnett, A. G. (2008). An introduction to generalized linear models 3rd edn. *Chapman Hall/CRC*, 149–163.

Assum, T. (1998). Road safety in Africa: appraisal of road safety initiatives in five African countries. *The World Bank and United Nations Economic Commission for Africa, 1998 (Working Paper No. 33). 2. THE GLOBAL IMPACT• 67*.

Atubi, A. O., & Gbadamosi, K. T. (2015). Global positioning and socio-economic impact of road traffic accidents in Nigeria: Matters arising. *Am Int J Contemp Res*, *5*(5), 136–146.

Atunbi, A. O. (2009). Urban Transportation. *An Appraisal of Features and Problems in Nigerin Sociaty International Journal of Geography and Regional Planning*, *1*.

Barau, A. S., Ludin, A. N. M., & Said, I. (2013). Socio-ecological systems and biodiversity conservation in African city: insights from Kano Emir's Palace gardens. *Urban Ecosystems*, *16*(4), 783–800.

Behnood, A., & Mannering, F. (2017). The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. *Analytic Methods in Accident Research*, *14*, 41–53.

Casado-Sanz, N., Guirao, B., & Attard, M. (2020). Analysis of the risk factors affecting the severity of traffic accidents on Spanish crosstown roads: The driver's perspective. *Sustainability*, *12*(6), 2237.

Casado-Sanz, N., Guirao, B., Lara Galera, A., & Attard, M. (2019). Investigating the risk factors associated with the severity of the pedestrians injured on Spanish crosstown roads. *Sustainability*, *11*(19), 5194.

Chang, L.-Y., & Chien, J.-T. (2013). Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. *Safety Science*, *51*(1), 17–22.

Chang, L.-Y., & Mannering, F. (1999). Analysis of injury severity and vehicle occupancy in truck-and non-truck-involved accidents. *Accident Analysis & Prevention*, *31*(5), 579–592.

Chang, L.-Y., & Wang, H.-W. (2006). Analysis of traffic injury severity: An application of non-parametric classification tree techniques. *Accident Analysis & Prevention*, *38*(5), 1019–1027.

Chen, C., Zhang, G., Qian, Z., Tarefder, R. A., & Tian, Z. (2016). Investigating driver injury severity patterns in rollover crashes using support vector machine models. *Accident Analysis & Prevention*, *90*, 128–139.

Chen, Z., & Fan, W. (2019). Modeling pedestrian injury severity in pedestrian-vehicle crashes in rural and urban areas: mixed logit model approach. *Transportation Research Record*, *2673*(4), 1023–1034.

Chong, M., Abraham, A., & Paprzycki, M. (2005). Traffic accident analysis using machine learning paradigms. *Informatica*, *29*(1).

Dadashova, B., Arenas, B. R., Mira, J. M., & Aparicio, F. I. (2014). Explanatory and prediction power of two macro models. An application to van-involved accidents in Spain. *Transport Policy*, *32*, 203–217.

Das, A., Abdel-Aty, M., & Pande, A. (2009). Using conditional inference forests to identify the factors affecting crash severity on arterial corridors. *Journal of Safety Research*, *40*(4), 317–327.

Delen, D., Sharda, R., & Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention*, *38*(3), 434–444.

Dong, L., Zhao, X., Wu, Q., & Yang, Y. (2009). Logistic regression for prediction and diagnosis of bacterial regrowth in water distribution system. *Transactions of Tianjin University*, *15*(5), 371.

Emenike, G. C., & Ogbole, A. (2008). Accidents and the Road Transport Industry in Nigeria. *Journal of the International Centre for Constructive Research (ICCR)*.

Gan, W. Q., McLean, K., Brauer, M., Chiarello, S. A., & Davies, H. W. (2012). Modeling population exposure to community noise and air pollution in a large metropolitan area. *Environmental Research*, *116*, 11–16. https://doi.org/10.1016/j.envres.2012.04.001

Gboyega, A., Ebijuwa, A. S., Oyetola, S. O., & Akinola, J. O. (2012). Factors influencing high rate of commercial motorcycle accidents in Nigeria. *American International Journal of Contemporary Research*, *2*(11), 130–140.

Gu-Chang, A., & Li, Y. (2010). Correlation Analysis on Driver's Human Factors and Accident Hazard in Road Traffic Accidents [J]. *Journal of Chongqing Jiaotong University (Natural Science)*, *1*.

Gunn, S. R., & others. (1998). Support vector machines for classification and regression. *ISIS Technical Report*, *14*(1), 5–16.

Ihueze, C. C., & Onwurah, U. O. (2018). Road traffic accidents prediction modelling: An analysis of Anambra State, Nigeria. *Accident Analysis & Prevention*, *112*, 21–29.

Ismail, A., & Yahia, H. A. (2013). Causes and effects of road traffic accidents in Tripoli--Libya. *Proceeding the 6th Civil Engineering Conference in Asia Region: Embracing the Future through Sustainability ISBN 978-602-8605-08*, *3*.

Issac, B., & Israr, N. (2014). *Case Studies in Intelligent Computing: Achievements and Trends*. CRC Press.

Jeong, H., Jang, Y., Bowman, P. J., & Masoud, N. (2018). Classification of motor vehicle crash injury severity: A hybrid approach for imbalanced data. *Accident Analysis & Prevention*, *120*, 250–261.

Johansson, Ö., Wanvik, P. O., & Elvik, R. (2009). A new method for assessing the risk of accident associated with darkness. *Accident Analysis & Prevention*, *41*(4), 809–815.

Johnson, R. A., Wichern, D. W., & others. (2002). *Applied multivariate statistical analysis* (Vol. 5, Issue 8). Prentice hall Upper Saddle River, NJ.

Kashani, A. T., & Besharati, M. M. (2017). Fatality rate of pedestrians and fatal crash involvement rate of drivers in pedestrian crashes: a case study of Iran. *International Journal of Injury Control and Safety Promotion*, *24*(2), 222–231.

Khorashadi, A., Niemeier, D., Shankar, V., & Mannering, F. (2005). Differences in rural and urban driver-injury severities in accidents involving large-trucks: an exploratory analysis. *Accident Analysis & Prevention*, *37*(5), 910–921.

Kononen, D. W., Flannagan, C. A. C., & Wang, S. C. (2011). Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes. *Accident Analysis & Prevention*, *43*(1), 112–122.

Kubat, M., Matwin, S., & others. (1997). Addressing the curse of imbalanced training sets: one-sided selection. *Icml*, *97*, 179–186.

Lee, C., & Li, X. (2015). Predicting driver injury severity in single-vehicle and two-vehicle crashes with boosted regression trees. *Transportation Research Record*, *2514*(1), 138–148.

Li, Z., Ci, Y., Chen, C., Zhang, G., Wu, Q., Qian, Z. S., Prevedouros, P. D., & Ma, D. T. (2019).

Investigation of driver injury severities in rural single-vehicle crashes under rain conditions using mixed logit and latent class models. *Accident Analysis & Prevention*, *124*, 219–229.

López, G., Abellán, J., Montella, A., & de Oña, J. (2014). Patterns of single-vehicle crashes on two-lane rural highways in granada province, Spain: in-depth analysis through decision rules. *Transportation Research Record*, *2432*(1), 133–141.

Mafi, S., Abdelrazig, Y., & Doczy, R. (2018). Machine learning methods to analyze injury severity of drivers from different age and gender groups. *Transportation Research Record*, *2672*(38), 171–183.

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research*, *1*, 1–22.

Mansourkhaki, A., Berangi, M., & Haghiri, M. (2018). A Neural network noise prediction model for Tehran urban roads. *Journal of Environmental Engineering and Landscape Management*, *26*(2), 88–97.

Mlakar, P., & Boznar, M. Z. (2011). Artificial neural networks-a useful tool in air pollution and meteorological modelling. In *Advanced Air Pollution*. IntechOpen.

Montella, A., Aria, M., D’Ambrosio, A., & Mauriello, F. (2011). Data-mining techniques for exploratory analysis of pedestrian crashes. *Transportation Research Record*, *2237*(1), 107–116.

Noori, R., Abdoli, M. A., Ghasrodashti, A. A., & Jalili Ghazizade, M. (2009). Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: a case study of Mashhad. *Environmental Progress & Sustainable Energy: An Official Publication of the American Institute of Chemical Engineers*, *28*(2), 249–258.

Nourani, V., Elkiran, G., & Abdullahi, J. (2019). Multi-station artificial intelligence based ensemble modeling of reference evapotranspiration using pan evaporation measurements. *Journal of Hydrology*, *577*(June), 123958. https://doi.org/10.1016/j.jhydrol.2019.123958

Nourani, V., Gökçekuş, H., & Umar, I. K. (2020). Artificial intelligence based ensemble model for prediction of vehicular traffic noise. *Environmental Research*, *180*, 108852. https://doi.org/10.1016/j.envres.2019.108852

Nourani, V., Gökçekus, H., Umar, I. K., & Najafi, H. (2020). An emotional artificial neural network for prediction of vehicular traffic noise. *Science of the Total Environment*, *707*, 136134. https://doi.org/10.1016/j.scitotenv.2019.136134

Nyamuame, G. Y., Aglina, M. K., Akple, M. S., Philip, A., & Klomegah, W. (2015). Analysis of road traffic accidents trend in Ghana: causing factors and preventive measures. *Int J Eng Sci Manag Res*, *2*(9), 127–132.

of Violence, W. H. O. D., Prevention, I., Violence, W. H. O., Prevention, I., & Organization, W. H. (2009). *Global status report on road safety: time for action*. World Health Organization.

Ojo, A. L. (2014). Predominant causes of road traffic accidents among commercial vehicle drivers in ekiti state, nigeria. *Int. J. Humanit. Soc. Stud*, 2321–9203.

Okunola, O. J., Uzairu, A., Gimba, C. E., & Ndukwe, G. I. (2012). Assessment of gaseous pollutants along high traffic roads in Kano, Nigeria. *International Journal of Environment and Sustainability*, *1*(1).

Olawole, M. O. (2016). Impact of weather on road traffic accidents in Ondo state, Nigeria: 2005-2012. *AnallelleUniiversiit{\u{a}}{\c{t}}iiiidiin Oradea--SeriiaGeografiie*, 44–53.

Onakomaiya, S. O. (1988). *Unsafe at any Speed: Towards Road Transportation for Survival*. University of Ilorin.

Organisation, W. H. (2015). Tables A2 & A10, data from 2013. *Global Status Report on Road Safety 2015*, *ISBN 978 92 4 156506 6*, 316–332.

Organization, W. H., & others. (2013). Road safety in the WHO African region: the facts 2013. *Geneva, Switzerland: World Health Organization*.

Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology*, *56*(1), 45.

Rezaie Moghaddam, F., Afandizadeh, S., & Ziyadi, M. (2011). Prediction of accident severity using artificial neural networks. *International Journal of Civil Engineering*, *9*(1), 41–48.

SAHEBI, S., MIRBAHA, B., MAHPOUR, A., & NORUZOLIAEE, M. H. (2015). *Predicting pedestrian accidents in rural roads using ordered logit model*.

Sarkar, A., & Sarkar, S. (2020). Comparative Assessment Between Statistical and Soft Computing Methods for Accident Severity Classification. *Journal of The Institution of Engineers (India): Series A*, *101*(1), 27–40. https://doi.org/10.1007/s40030-019-00422-7

Sharma, N., Chaudhry, K. K., & Rao, C. V. C. (2005). Vehicular pollution modelling in India. *Journal of the Institution of Engineers(India), Part EN, Environmental Engineering Division*, *85*(2), 46–53.

Solgi, A., Pourhaghi, A., Bahmani, R., & Zarei, H. (2017). Improving SVR and ANFIS performance using wavelet transform and PCA algorithm for modeling and predicting biochemical oxygen demand (BOD). *Ecohydrology & Hydrobiology*, *17*(2), 164–175.

Theofilatos, A., Chen, C., & Antoniou, C. (2019). Comparing machine learning and deep learning methods for real-time crash prediction. *Transportation Research Record*, *2673*(8), 169–178.

Tien Bui, D., Shirzadi, A., Shahabi, H., Chapi, K., Omidavr, E., Pham, B. T., Talebpour Asl, D., Khaledian, H., Pradhan, B., Panahi, M., & others. (2019). A novel ensemble artificial intelligence approach for gully erosion mapping in a semi-arid watershed (Iran). *Sensors*, *19*(11), 2444.

Toy, E. L., & Hammitt, J. K. (2003). Safety impacts of SUVs, vans, and pickup trucks in two-vehicle crashes. *Risk Analysis: An International Journal*, *23*(4), 641–650.

Ulfarsson, G. F., & Mannering, F. L. (2004). Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis & Prevention*, *36*(2), 135–147.

Wang, Q. (2012). Kernel principal component analysis and its applications in face recognition and active shape models. *ArXiv Preprint ArXiv:1207.3538*.

Wang, W., Zhao, D., & Kuang, Y. (2016). Decomposition analysis on influence factors of direct household energy-related carbon emission in G uangdong province—Based on extended K aya identity. *Environmental Progress & Sustainable Energy*, *35*(1), 298–307.

Wanvik, P. O. (2009). Effects of road lighting: an analysis based on Dutch accident statistics 1987--2006. *Accident Analysis & Prevention*, *41*(1), 123–128.

WHO. (2020). *No Title*.

Wu, J., & Xu, H. (2018). Driver behavior analysis on rural 2-lane, 2-way highways using SHRP 2 NDS data. *Traffic Injury Prevention*, *19*(8), 838–843.

Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. *ArXiv Preprint ArXiv:1505.00853*.

Yap, W. K., & Karri, V. (2011). ANN virtual sensors for emissions prediction and control. *Applied Energy*, *88*(12), 4505–4516.

Yuen, K.-V., & Lam, H.-F. (2006). On the complexity of artificial neural networks for smart structures monitoring. *Engineering Structures*, *28*(7), 977–984.

Zhang, Y. (2009). Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM. *Chemical Engineering Science*, *64*(5), 801–811.

Zhou, B., Zhang, X., Zhang, S., Li, Z., & Liu, X. (2019). Analysis of Factors Affecting Real-Time Ridesharing Vehicle Crash Severity. *Sustainability*, *11*(12), 3334.

Zhu, M., Li, Y., & Wang, Y. (2018). Design and experiment verification of a novel analysis framework for recognition of driver injury patterns: From a multi-class classification perspective. *Accident Analysis & Prevention*, *120*, 152–164.

Zhu, S., Heddam, S., Wu, S., Dai, J., & Jia, B. (2019). Extreme learning machine-based prediction of daily water temperature for rivers. *Environmental Earth Sciences*, *78*(6), 202.

Zhu, X., & Srinivasan, S. (2011). A comprehensive analysis of factors influencing the injury severity of large-truck crashes. *Accident Analysis & Prevention*, *43*(1), 49–57.

APPENDICES

APPENDIX A

Accident information for the Case study Route

Causes of accidents (speed limit violation, mechanical defective vehicle, etc.) are given code for easy use

Table A.1
Causes of Accidents

Causes of Accidents	Code
Wrong overtaking	WOV
Over speeding	SLV
Tire burst	TBT
Dangerous driving	DGD
Mechanically defective vehicle	MDV
On-street parking	SVO
Fatigue	RTV
More than 1 cause	LOC

Table A.2
Number of Accidents by Causes

YEAR	ACCIDENT CAUSES								
	WOV	SLV	TBT	DGD	MDV	SVO	RTV	LOC	TOTAL
2012	12	38	7	32	1	0	1	2	93
2011	17	20	3	37	0	0	1	0	78
2010	20	53	8	22	1	0	0	0	104
2009	12	41	8	16	1	0	0	0	78
TOTAL	61	152	26	107	3	0	2	2	353

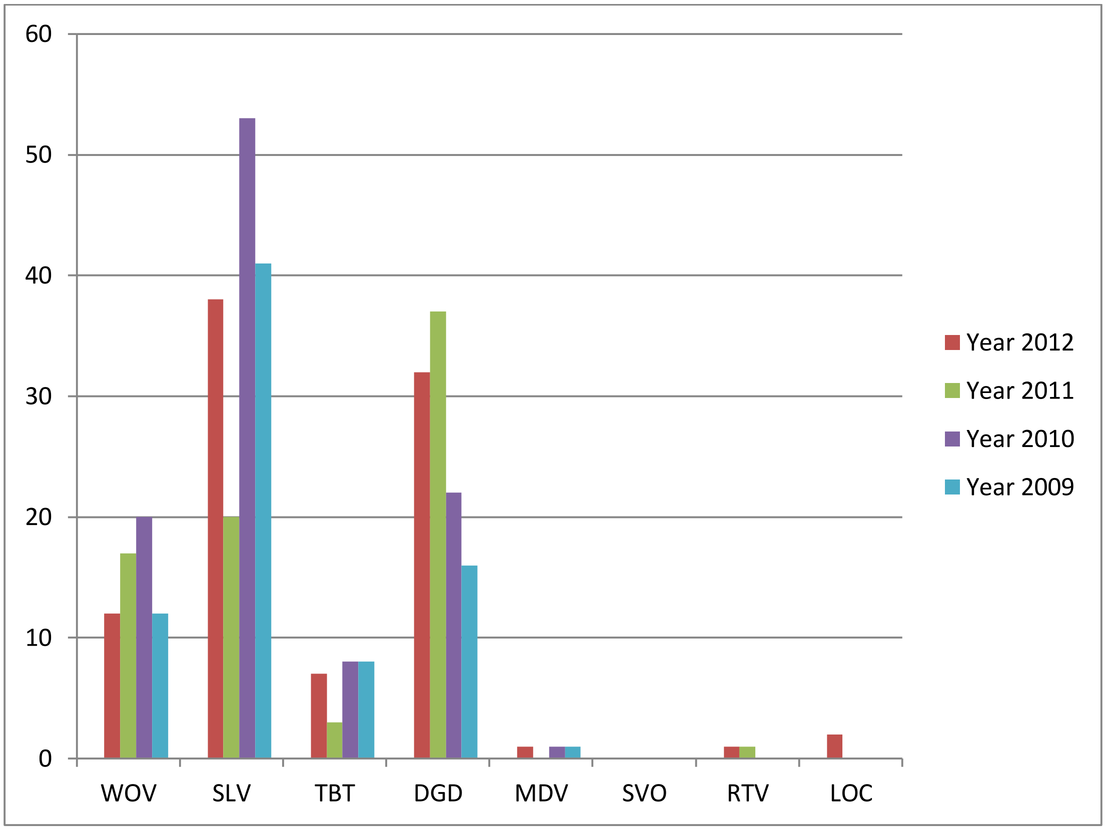


Figure AA.1

Cause Number of each Accidents

Table A.3
Number of Fatalities by Causes

YEAR	ACCIDENT CAUSES								
	WOV	SLV	TBT	DGD	MDV	SVO	RTV	LOC	TOTAL
2012	7	6	2	10	0	1	0	2	28
2011	12	11	2	7					33
2010	3	33	2	13	0	0	0	0	51
2009	10	17	12	15	0	0	0	0	54
TOTAL	32	67	18	45	0	1	0	2	166

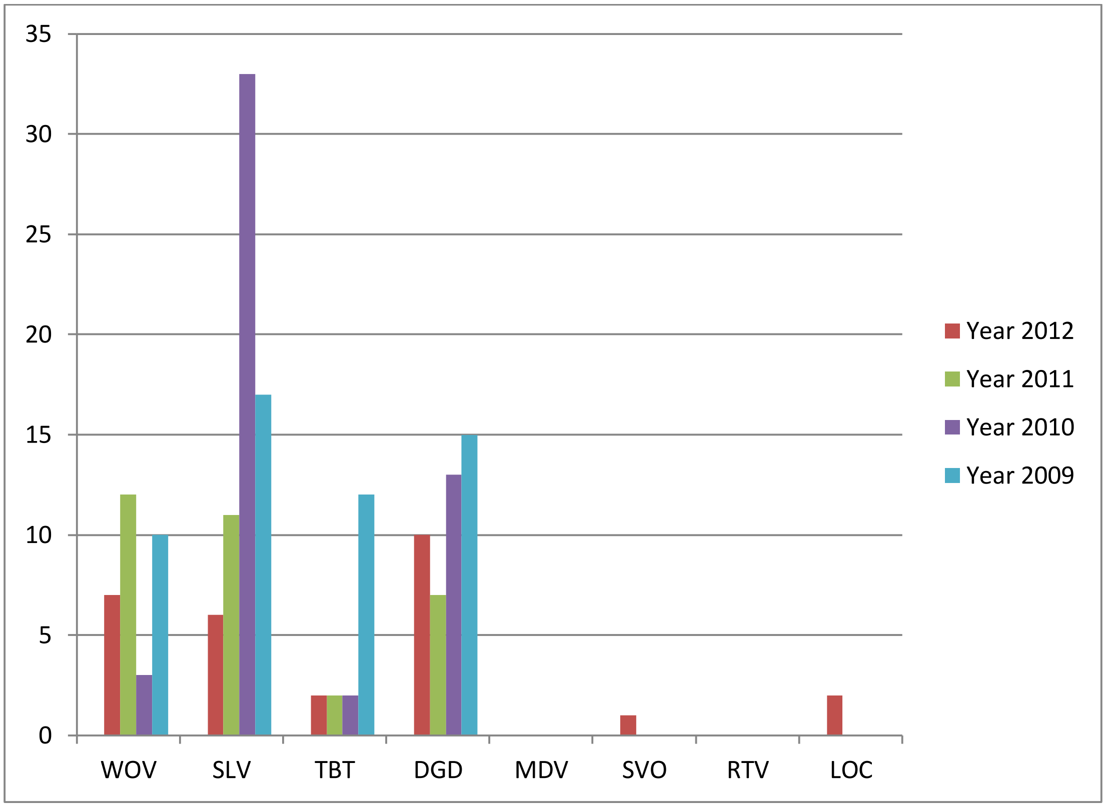


Figure AA.2

Number of fatalities for each Group

Table A.4
Summary for Number of Injuries

YEAR	ACCIDENT CAUSES								
	WOV	SLV	TBT	DGD	MDV	SVO	RTV	LOC	TOTAL
2012	95	207	62	185	15	6	0	11	581
2011	101	70	23	128	0	0	3	0	327
2010	89	172	24	67	0	0	0	0	352
2009	54	144	24	70	21	0	0	0	313
TOTAL	339	593	433	450	36	6	3	11	1573

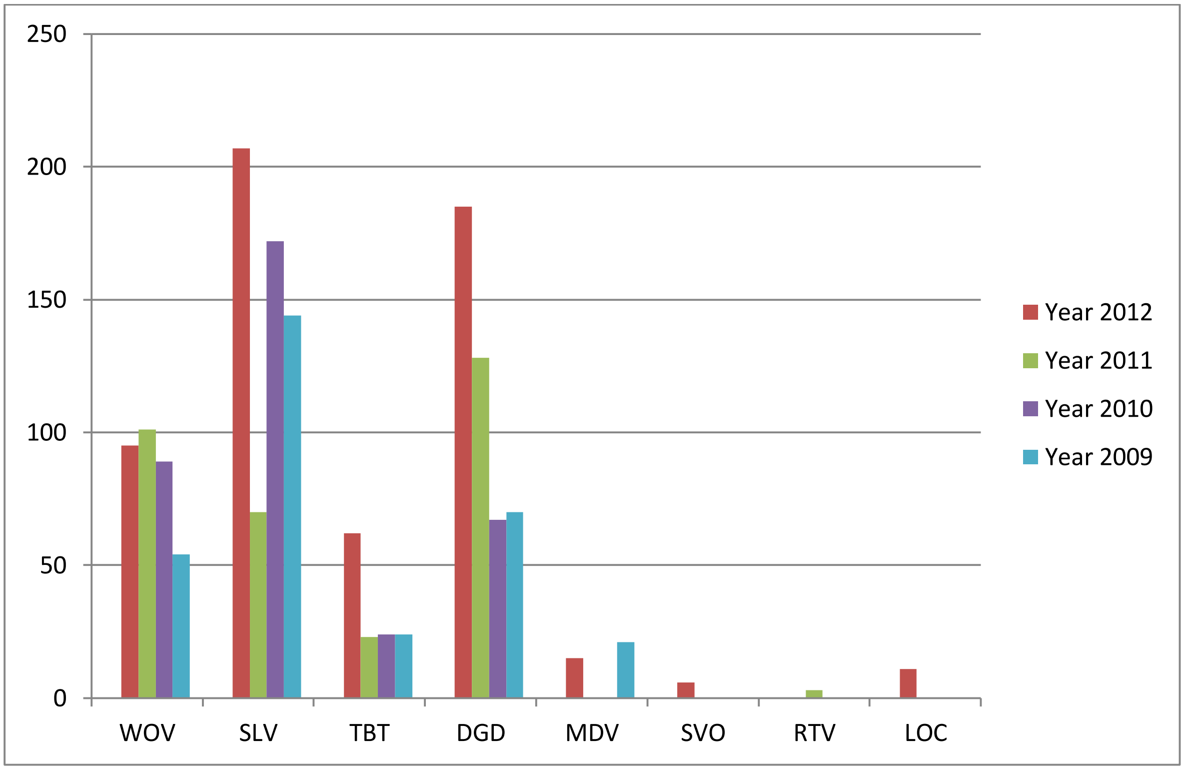


Figure AA.3

Chart Showing Summary for Injury

Table A.5
Accident Data for 2009

S/N	Date	Vehicle Type	Location	Total Injury	Total Fatality	Cause
1	04/01/2009	M/C	ZOGAWARA/WDL-KN	2	0	SLV
2	04/01/2009	CAR	K BABALE/WDL-KN	0	0	TBT
3	05/01/2009	CAR,BUS	G DAU/WDL-KN	4	0	GDG
4	05/01/2009	BUS	J GANO/WDL-KN	7	0	WOV

5	06/01/2009	BUS	POLAC/WDL-KN	7	0	TBT
6	08/01/2009	CAR	GANO/WDL-KN	16	0	WOV
7	09/01/2009	BUS, CAR	JEMAGU/WDL-KN	0	0	DGD,SLV
8	31/01/2009	M/C	MAKOLE/WDL-KN	2	0	SLV
9	01/02/2009	JEEP, M/C	G DAU/WDL-KN	2	0	DGD
10	09/02/2009	BUS	GANO/WDL-KN	3	0	DGD,SLV
11	14/02/2009	M/C	WUDIL/WDL-MAI	3	0	SLV
12	22/02/2009	CAR,TRUCK	GANO/WDL-KN	1	0	WOV,DGD
13	22/02/2009	CAR	G DAU/WDL-KN	0	0	DGD
14	03/03/2009	CAR,M/C	SHAGOGO/WDL-MAI	1	0	DGD
15	05/03/2009	BUS	GANO/WDL-KN	7	4	SLV
16	14/03/2009	M/C	J GANO/WDL-KN	2	0	SLV
17	16/03/2009	BUS	POLAC/WDL-KN	1	0	OLV
18	19/03/2009	CAR	MAKOLE/WDL-KN	2	0	SLV
19	23/03/2009	CAR	POLAC/WDL-KN	1	0	TBT
20	23/03/2009	CAR,M/C	MAKOLE/WDL-KN	3	0	DGD
21	27/03/2009	CAR,BUS	YAR GAYA,WDL-KN	13	0	WOV
22	02/04/2009	TRUCK	POLAC/WDL-KN	0	1	WOV

23	05/04/2009	M/C	WUDIL/WDL-MAI	2	1	WOV
24	11/04/2009	CAR	YAR GAYA/WDL-KN	14	0	SLV
25	16/04/2009	M/C	GANO/WDL-KN	1	0	SLV
26	18/04/2009	CAR,M/C	YAR GAYA/WDL-KN	2	0	SLV
27	29/04/2009	TRUCK	GANO/WDL-KN	1	0	SLV
28	04/05/2009	CAR,BUS	GANO/WDL-KN	12	3	SLV,DGD
29	06/05/2009	CAR	WUDIL/WDL-KN	0	1	SLV
30	08/05/2009	CAR	MAKOLE/WDL-KN	0	0	SLV
31	08/05/2009	CAR,M/C	POLAC/WDL-KN	2	0	SLV
32	13/05/2009	CAR	POLAC/WDL-KN	0	1	WOV
33	14/05/2009	TRUCK,M/C	WUDIL/WDL-MAI	1	0	WOV
34	18/05/2009	CAR,BUS	D FOREST/WDL-MAI	6	7	WOV
35	19/05/2009	N/A	G DAU/WDL-KN	4	0	WOV
36	08/06/2009	M/C	G DAU/WDL-KN	2	0	DGD
37	12/06/2009	BUS	KANYA/WDL-KN	1	1	SLV
38	14/06/2009	BUS	WUDIL/WDL-MAI	1	0	DGD
39	25/06/2009	M/C	DANBAGINA/WDL-KN	2	0	SLV

40	06/07/2009	M/C	WUDIL/WDL-MAI	2	0	SPV
41	18/07/2009	CAR	B/4 GANO	5	0	SLV
42	19/07/2009	BUS		10	1	SLV
43	19/07/2009		T/GUMSAU	5	0	SLV
44	22/07/2009	CAR,M/C	GANO/WUDL-KN	0	1	DGD,SLV
45	22/08/2009	M/C	WUDIL/WDL-MAI	2	0	SLV
46	22/08/2009	CAR,M/C	G DAU/WDL-KN	1	0	SLV
47	26/08/2009	M/C	WDL JUNCTION	2	0	SLV
48	07/09/2009	CAR	G DAU/WDL-KN	2	0	SLV
49	10/09/2009	CAR,TRUCK	KANYA/WDL-KN	8	0	DGD
50	10/09/2009	CAR,BUS,TRUCK	KANYA/WDL-KN	10	3	DGD
51	15/09/2009	CAR	T GUNSAU	5	0	SLV
52	16/09/2009	CAR.M/C	MAKOLE	6	2	SLV,GDG
53	18/09/2009	M/C	G DAU	1	0	SLV,DGD
54	18/09/2009	CAR,BUS	T GUNSAU	1	0	SLV
55	19/0/2009	BUS	GAYA/WDL-MAI	9	1	SLV
56	22/09/2009	CAR,M/C	D GAU/WDL-KN	1	0	SLV
57	28/09/2009	BUS	WUDIL/WDL-KN	10	0	TBT

58	28/09/2009	BUS,CAR	MARIRI/WDL-KN	11	3	SLV
59	29/09/2009	CART	K/LAMIRE	6	1	SLV
60	06/10/2009	JINCHENG	OPP.TOTAL	3	2	DGD
61	12/10/2009	J5	OPP. TOTAL	20	0	MDV
62	15/10/2009	BUS	DUN-DUN AJINGI	0	8	TBT
63	21/10/2009	CAR	ITALIYAR HAUSAWA WDL-BAU	2	0	TBT

64	31/10/2009	M/CYCLE	FRSC BASE	2	0	SLV
65	05/11/2009	CAR	GANO WDL-KN	2	2	SLV
66	14/11/2009	CAR	GANO WDL-KN	8	0	SLV
67	16/11/2009	CAR	POLAC WDL-KN	5	0	SLV
68	20/11/2009	P/UP	T/GUNSAU	5	1	SLV
69	23/11/2009	CAR,BUS	YAR GAYA	8	0	DGD
70	25/11/2009	CAR	GSSS WDL	2	0	SLV
71	27/11/2009	2 M/CYCLE	OPP. AMANA HOSPITAL	2	0	SLV
72	30/11/2009	CAR	ZUMBULAWA WDL-BAU	2	4	TBT
73	30/11/2009	M/CYCLE	OPP.AUDU MANAGER F/STATION WDL TOWN	2	0	SLV

74	03/12/2009	BUS	MAKOLE WDL-KN	2	0	WOV
75	05/12/2009	CAR,TYT BUS	R/GWANGWAN WDL-KN	13	6	DGD
76	07/12/2009	CAR,M/CYCLE	G/DAU WDL-KN	3	0	SLV
77	09/12/2009	BUS	J/ALH. ADO WDL-KN	3	0	SLV
78	10/12/2009	CAR,BUS	J/GANO WDL-KN	2	0	WOV
79	11/12/2009	CAR,TRAILER	WDL BRIDGE	2	0	TBT

Table A.6
Accident Data for the Year 2010

S/N	Date	Vehicle type	Location	Injury	Fatality	Cause
1	26/03/2010	P/UP	ZOGARAWA/WDL-KN	2	1	WOV
2	31/03/2010	M/C	G/DAU/WDL-KN	2	0	SLV
3	04/04/2010	CAR	ZOGARAWA/WDL-KN	10	0	SLV
4	05/04/2010	CAR/M/C	YAR GAYA/WDL-KN	1	1	SLV
5	10/04/2010	CAR	POLAC/WDL-KN	0	1	SLV
6	11/04/2010	CAR	GANO/WDL-KN	0	0	SLV
7	15/04/2010	M/C	J/GANO/WDL-KN	1	0	SLV
8	20/04/2010	BUS/CAR	R/GON-GON/WDL-KN	4	0	WOV
9	22/04/2010	BUS	JUNC/WDL-MAI	1	0	SLV
10	23/04/2010	BUS	MAKOLE/WDL-KN	1	1	SLV
11	24/04/2010	CAR	SCI.GAYA/WDL-MAI	1	0	SLV
12	01/05/2010	CAR	TOTAL/WDL-KN	1	0	DGD
13	02/05/2010	BUS	MARYAM/WDL-KN	29	2	SLV
14	03/05/2010	BUS	D/NA’ABBA	10	0	SLV
15	04/05/2010	M/C	WDL PARK/WDL-KN	1	0	DGD
16	06/05/2010	BUS/TRUCK	DANBAGINA/WDL-KN	2	0	DGD
17	13/05/2010	M/C	DANBAGINA/WDL-KN	3	0	TBT
18	15/05/2010	CARS	MAKOLE/WDL-KN	3	0	SLV
19	16/05/2010	BUS,TRUCK,BUS,BUS,	MEGA/WDL-KN	0	1	DGD
20	22/05/2010	CAR	BRIDGE/WDL-KN	1	0	DGD
21	24/05/2010	M/C	WDL-KN	3	0	SLV
22	20/06/2010	BUS/CAR	GANO/WDL-KN	16	1	SLV

23	29/06/2010	A CARS, M/C	R/GWAN-GWAN/WDL-KN	3	0	WOV
24	29/06/2010	BUS,CAR, TRUCK	YAR GAYA/WDL-KN	5	0	DGD
25	20/07/2010	M/C	TOTAL/WDL-KN	2	0	SLV
26	29/07/2010	TRUCK/BUS	YAR GAYA/WDL-KN	7	7	SLV
27	31/07/2010	BUS	KUST/WDL-MAI	8	0	TBT
28	01/08/2010	BUS	GANO/WDL-KN	4	0	SLV
29	01/08/2010	BUS	POLAC/WDL-KN	2	0	TBT
30	02/08/2010	M/C	POLAC/WDL-KN	1	0	SLV
31	02/08/2010	TRUCK/CAR	R/GWAN-GWAN/WDL-KN	4	1	SLV
32	03/08/2010	M/C	POLAC/WDL-KN	1	0	SLV
33	07/08/2010	P/UP	MAKOLE/WDL-KN	2	1	SLV
34	08/08/2010	CARS	ZOGARAWA/WDL-KN	7	0	WOV
35	09/08/2010	CAR	POLAC/WDL-KN	0	0	WOV
36	10/08/2010	CAR	D/BAGINA/WDL-KN	7	0	WOV
37	14/08/2010	M/C	BRIDGE/WDL-KN	2	0	SLV
38	16/08/2010	CAR	YAR GAYA/WDL-KN	5	0	WOV
39	17/08/2010	CAR	YAR GAYA	3	0	SLV
40	17/08/2010	CAR, BUS	MAKOLE/WDL-KN	3	0	WOV
41	20/08/2010	M/C	G/DAU/WDL-KN	3	0	SLV
42	21/08/2010	CAR	UTAI/WDL-KN	0	1	SLV
43	21/08/2010	M/C	POLAC/WDL-KN	1	0	DGD
44	23/08/2010	M/C, BUS	MAKOLE/WDL-KN	1	0	DGD
45	25/08/2010	CAR	JIDO/WDL-KN	5	0	DGD
46	27/08/2010	BUS	MAKOLE/WDL-KN	5	0	TBT

47	31/08/2010	P/UP, CAR	R/GWAN-GWAN/WDL-KN	6	0	WOV
48	31/08/2010	BUS, CAR,	ZOGARAWA/WDL-KN	9	3	DGD
49	01/09/2010	TRUCK	GAYA/WDL-KN	0	0	SLV
50	01/09/2010	BUS	GANO/WDL-KLN	0	0	SLV
51	01/09/2010	CAR	GANO/WDL-KN	0	1	SLV
52	02/09/2010	BUS, CAR	G/DAU/WDL-KN	3	1	TBT
53	03/09/2010	BUS, P/UP	D/NA'ABBA/WDL-KN	2	0	WOV
54	05/09/2010	BUS, CAR, TRUCK	R/GWAN-GWAN/WDL-KN	8	5	SLV
55	05/09/2010	CAR	POLAC/WDL-KN	1	0	TBT
56	13/09/2010	TRUCK, CAR	GWAN-GWAN/WDL-KN	8	1	DGD
57	13/09/2010	V/WAGEN	GANO/WDL-KN	7	0	SLV
58	17/09/2010	CAR	POLAC/WDL-KN	0	0	SLV
59	24/09/2010	CAR, M/C	YAR GAYA/WDL-KN	0	1	SLV
60	25/09/2010	BUS	GANO/WDL-KN	1	0	WOV
61	26/09/2010	CAR	GANO/WDL-KN	0	0	SLV
62	26/09/2010	CAR	ZOGARAWA	3	0	WOV
63	26/09/2010	BUS	GWAN-GWAN/WDL-KN	0	1	SLV
64	26/09/2010	CAR, M/C	TOTAL/WDL-KN	2	0	WOV
65	29/09/2010	CAR	G/DAU/WDL-KN	2	1	SLV
66	29/09/2010	M/C	POLAC/WDL-KN	1	0	SLV
67	03/10/2010	CAR	GWAN-GWAN/WDL-KN	3	0	SLV
68	03/10/2010	TRUCK	GANO/WDL-KN	1	1	DGD
69	07/10/2010	CAR,M/C	GARIN DAU/WDL-KN	1	1	SLV

70	12/10/2010	BUS,DAF,BUS	ZANGO/WDL-KN	6	1	DGD
71	16/10/2010	M/C	ZOGARAWA/WDL-KN	10	1	WOV
72	16/10/2010	CAR	MAKOLE/WDL-KN	2	0	SLV
73	18/10/2010	M/C	LAMIRE/WDL-BAU	2	0	SLV
74	30/10/2010	BUS	AMARAWA	2	0	TBT
75	02/11/2010	TRUCK	J/GANO/WDL-KN	2	0	SLV
76	05/11/2010	CAR	POLAC/WDL-KN	0	1	DGD
77	07/11/2010	M/C	MAKOLE/WDL-KN	0	2	DGD
78	08/11/2010	CAR	G/DAU/WDL-KN	1	0	DGD
79	11/11/2010	CAR	JAM'ARE/WDL-KN	4	0	SLV
80	15/11/2010	M/C	GWAN-GWAN/WDL-KN	1	0	SLV
81	19/11/2010	CAR	G/DAU/WDL-KN	2	0	SLV
82	23/11/2010	M/C	J/ADO/WDL-KN	3	0	DGD
83	26/11/2010	2M/C	BRIDGE/WDL-KN	1	0	SLV
84	27/11/2010	2 M/C	AMARAWA/WDL-MAI	5	2	DGD,SLV
85	29/11/2010	BUS	MAKOLE/WDL-KN	2	0	WOV
86	03/12/2010	TRUCK,CAR,M/C	BRIDGE/WDL-KN	2	0	DGD
87	05/12/2010	BUS,P/UP	GANO/WDL-KN	7	4	SLV
88	06/12/2010	CARS	POLAC/WDL-KN	11	0	WOV
89	09/12/2010	CARS	GANO/WDL-KN	1	1	DGD
90	11/12/2010	P/UP, M/C	MAKOLE/WDL-KN	11	1	WOV
91	13/12/2010	TRUCK	WDL TOWN	1	0	SLV
92	16/12/2010	BUS, CAR	MARIRI/WDL-KN	0	1	SLV
93	17/12/2010	M/C	G/DAU/WDL-KN	12	0	DGD
94	17/12/2010	CAR	OPP.0TEL/WDL-MAI	2	0	DGD

95	18/12/2010	CAR	SHAGOGO	1	1	SLV
96	21/12/2010	CAR	YAN AUDU/WDL-MAI	2	0	SLV
97	21/12/2010	TRUCK	SHAGOGO/WDL-MAI	0	1	SLV
98	22/12/2010	CAR, TRUCK	POLAC/WDL-KN	0	0	MDV
99	22/12/2010	CAR	MAKOLE/WDL-KN	5	0	WOV
100	23/12/2010	CAR	GANO/WDL-KN	4	0	WOV
101	24/12/2010	CAR	POLAC/WDL-KN	1	0	WOV
102	25/12/2010	CAR, BUS	YAR GAYA/WDL-KN	17	0	SLV
103	29/12/2010	CAR	GANO/WDL-KN	0	1	TBT
104	30/12/2010	CAR	POLAC/WDL-KN	1	0	SLV

Table A.7
Accident Data for Year 2011

S/N	Date	Vehicle type	Location	Injury	Fatality	Cause
1	04/01/2011	2 BUS, P/UP	GWAN-GWAN/WDL-KN	23	0	WOV
2	11/01/2011	CAR, M/C	HOTORO/WDL-KN	1	0	WOV
3	22/01/2011	CAR	GANO/WDL-KN	8	3	WOV
4	30/01/2011	P/UP	G/DAU/WDL-KN	1	0	SLV
5	09/02/2011	TRUCK, M/C	GWAN-GWAN/WDL-KN	4	2	TBT
6	12/02/2011	CAR	J/GANO/WDL-KN	0	1	SLV
7	12/02/2011	M/C	POLAC/WDL-KN	3	0	RTV
8	17/02/2011	BUS, P/UP	ZOGARAWA/WDL-KN	2	0	WOV
9	17/02/2011	P/UP	MAKOLE/WDL-KN	3	1	SLV

10	17/02/2011	TRUCK, M/C	J/GANO/WDL-KN	0	0	DGD
11	19/02/2011	TRUCK, BUS	GOGEL/WDL-KN	1	0	WOV
12	20/02/2011	M/C	MAKOLE	4	0	SLV
13	20/02/2011	CAR	J/ADO/WDL-KN	2	0	DGD
14	23/02/2011	M/C	J/GANO/WDL-KN	2	1	DGD
15	01/03/2011	JEEP	GANO/WDL-KN	1	0	SLV
16	01/03/2011	BUS, M/C	MAKOLE	1	0	DGD
17	04/03/2011	CAR,BUS, BUS	GANO/WDL-KN	6	0	DGD
18	05/03/2011	BUS CAR, BUS	GANO/WDL-KN	7	0	DGD
19	08/03/2011	BUS, CAR	GANO/WDL-KN	6	1	WOV
20	10/03/2011	CAR	UTAI/WDL-KN	2	0	TBT
21	11/03/2011	P/UP	J/ADO/WDL-KN	5	0	SLV
22	12/03/2011	TRUCK, CAR	GAYA/WDL-KN	4	5	SLV
23	12/03/2011	CAR	GWAN-GWAN/WDL-KN	7	0	WOV
24	12/03/2011	M/C	D/NA'ABBA/WDL-KN	2	0	DGD
25	13/03/2011	BUS,	GANO/WDL-KN	10	0	DGD
26	13/03/2011	TYT, M/C	G/DAU/WDL-KN	2	1	SLV
27	13/03/2011	TRUCK	GWAN-GWAN/WDL-KN	1	0	WOV
28	21/03/2011	CAR, BUS, TRUCK	J/GANO/WDL-KN	10	4	WOV
29	24/03/2011	BUS	NNPC/WDL-KN	6	0	WOV
30	26/03/2011	CAR	ZOGAEAWA/WDL-KN	6	1	WOV

31	26/03/2011	CAR	GWAN-GWAN/WDL-KN	3	2	SLV
32	26/03/2011	P/UP	J/GANO/WDL-KN	2	0	SLV
33	13/05/2011	CAR	G/DAU-WDL-KN	1	0	DGD

34	17/05/2011	TRUCK	GANO/WDL-KN	2	0	DGD
35	19/05/2011	TYT BUS/M/CYCLE	WDL R/ABOUT	1	0	DGD
36	19/05/2011	N/A	GAYA	8	0	WOV
37	23/05/2011	CAR	G/DAU-WDL-KN	1	1	SLV
38	03/06/2011	BUS	J/GANO/WDL-KN	5	0	WOV
39	06/06/2011	TRAILER/BUS/CAR	YARGAYA	8	3	WOV
40	06/06/2011	BUS/M/CYCLE	GARINDAU/WDL-KN	1	0	DGD
41	10/06/2011	M/CYCLE/	WDLTOWN	2	0	DGD
42	10/06/2011	M/CYCLE	U/WUDILAWA-WDL-M	2	0	DGD
43	11/06/2011	CAR/BUS	R/GONGONWDL-KN	6	0	DGD
44	11/06/2011	N/A	GANO-WDL-KN	2	0	DGD
45	14/06/2011	BUS	POLACWDL-KN	4	0	LSV
46	16/06/2011	BUS 2 M/CYCLE	YARGAYAWDL-KN	5	0	WOV
47	17/06/2011	3BUS,CAR	POLAC WDL-KN	20	0	DGD
48	21/06/2011	CAR,M/CYCLE	G/DAU WDL-KN	1	1	DGD
49	23/06/2011	BUS,CAR	J/GANO WDL-KN	14	2	DGDS
50	27/06/2011	M/CYCLE	POLAC WDL-KN	2	0	WOV
51	30/06/2011	BUS,M/CYCLE	G/DAU WDL-KN	2	0	DGD

52	05/07/2011	M/CYCLE	G/DAU WDL-KN	3	0	DGD
53	05/07/2011	BUS,CAR	DOGON MARKE-WDL-KN	6	2	DGD
54	08/07/2011	CAR,BUS	J/GANO WDL-KN	10	0	LSV
55	20/07/2011	CAR,M/CYCLE	DORAWA WDL-KN	2	0	DGD
56	29/07/2011	M/CYCLE	G/DAU WDL-KN	2	0	DGD
57	02/08/2011	CAR,M/CYCLE	POLAC WDL-KN	1	0	SLV
58	18/08/2011	BUS,M/CYCLE	WDLTOWN	2	0	SLV
59	20/08/2011	CAR	D/NAABBA WDL-KN	3	1	DGD
60	27/08/2011	CAR,P/UP	POLAC WDL-KN	3	1	DOV
61	01/09/2011	2 M/CYCLE	WDL-BRIDGE	2	0	SLV
62	13/09/2011	BUS,M/CYCLE	POLAC WDL-KN	3	0	DGD
63	19/09/2011	2 BUS	GANO WDL-KN	8	0	DGD
64	23/09/2011	CAR	GAYA WDL-KN	2	0	SLV
65	04/10/2011	CAR	G/DAU WDL-KN	2	0	DGD
66	28/10/2011	P/UP,CAR,P/UP,BUS	R/GWANGWAN WDL-KN	17	0	TBT
67	03/11/2011	CAR, ARTICULATED VEH.	G/DAU WDL-KN	1	0	DGD
68	07/11/2011	BUS	POLAC WDL-KN	5	0	DGD
69	09/11/2011	TANKER,M/CYCLE	J/GANO WDL-KN	1	0	DGD
70	09/11/2011	M/CYCLE	GANO WDL-KN	1	0	DGD

71	14/11/2011	BUS	POLAC WDL-KN	7	0	SLV
72	16/11/2011	CAR,M/CYCLE	WDL TOWN	3	0	DGD
73	16/11/2011	M/CYCLE	GANO WDL-KN	2	0	WOV
74	18/11/2011	BUS	GANO WDL-KN	6	0	SLV
75	23/11/2011	M/CYCLE	WDL TOWN	1	0	DGD
76	26/11/2011	N/A	J/ALI WDL-KN	1	0	DGD
77	28/12/2011	CAR,M/CYCLE	J/ALH. ADO WDL-KN	2	0	DGD
78	31/12/2011	CAR,BUS	J/GANO WDL-KN	10	0	SLV

Table A.8
Accident Data for the Year 2012

S/N	Date	Vehicle type	Location	Injury	Fatality	Cause
1	01/01/2012	CAR,M/CYCLE	GANO WDL-KN	0	1	DGD
2	03/02/2012	BUS	GANO WDL-KN	11	0	DGD
3	17/02/2012	CAR,BUS	J/GANO WDL-KN	13	0	SLV
4	08/03/2012	CAR,M/CYCLE	K/BABALE WDL-KN	5	0	SLV
5	16/03/2012	CAR,BUS	GANO WDL-KN	4	0	SLV
6	18/03/2012	CAR	G/DAU WDL-KN	1	0	DGD
7	22/03/2012	CAR,BUS	WDL BRIDGE	6	0	DGD
8	24/03/2012	2 CARS	J/GANO WDL-KN	0	0	DGD
9	25/03/2012	BUS,M/CYCLE	FRSC GATE	1	0	DGD
10	28/03/2012	TANKER,TRAILER	YARGAYA WDL-KN	3	0	DGD
11	08/04/2012	BUS	J/GANO WDL-KN	3	0	SLV
12	14/04/2012	2 CARS	ZOGARAWA WDL-KN	1	0	DGD

13	16/04/2012	CAR	HADEJIA RIVER BASIN WDL-KN	6	0	DGD
14	19/04/2012	N/A	WDL TOWN	3	0	DGD
15	20/04/2012	CAR	MAKOLE WDL-KN	3	0	SLV
16	21/04/2012	N/A	POLAC WDL-KN	0	1	SLV
17	21/04/2012	M/C	WDL BRIDGE	1	0	DGD
18	26/04/2012	CAR	MAKOLE/WDL-KN	0	1	WOV
19	04/05/2012	CAR	J/GANO/WDL-KN	7	0	SLV
20	07/05/2012	BUS, TRUCK	GWAN-GWAN/WDL-KN	0	1	SLV
21	13/05/2012	CAR	POLAC/WDL-KN	2	0	TBT
22	13/05/2012	BUS, M/C	MAKOLE/WDL-KN	2	0	SLV
23S	18/05/2012	BUS, TRUCK	POLAC	8	0	SLV
24	23/05/2012	CAR	N/A	2	0	SLV
25	24/05/2012	BUS,CAR	WDL JUNCTION	2	0	DGD
26	30/05/2012	CANTER	N/A	11	0	DGD
27	02/06/2012	M/CYCLE	G/DAU	2	0	SLV
28	05/06/2012	2 CARS,BUS	N/A	3	0	SLV
29	07/06/2012	TRAILER	POLAC WDL-KN	0	1	SLV
30	11/06/2012	M/CYCLE	MAKOLE	2	0	DGD
31	22/06/2012	P/UP,TRAILER	GANO WDL-KN	11	1	DGD

32	23/06/2012	2 CAR	T/COLLEGE WDL WDL TOWN	3	0	SLV
33	28/06/2012	2 CAR,BUS	G/DAU WDL-KN	10	0	SLV
34	29/06/2012	BUS,M/CYCLE	M/PARK WDL-KN	1	0	DGD
35	30/06/2012	BUS	ZOGARAWA WDL-KN	4	0	SLV
36	05/07/2012	BUS	WDL-BRIEDGE	1	0	SLV
37	07/07/2012	2 CAR	WDL-BRIEDGE	1	0	SLV
38	12/07/2012	CAR,M/CYCLE	GANO WDL-KN	1	0	SLV
39	15/07/2012	CAR	K/GARKO WDL-KN	0	0	DGD
40	20/07/2012	BUS,2 M/CYCLE	WDL-BRIEDGE	1	0	DGD
41	20/07/2012	CAR,M/CYCLE	GIDAN TAKARDA WDL-KN	0	0	DGD
42	23/07/2012	BUS,CAR	ZOGARAWA WDL-KN	8	0	WOV
43	24/07/2012	CAR,BUS	MAKOLE WDL-KN	4	0	SLV
44	25/07/2012	HILUX,JEEP	R/GWANGWAN WDL-KN	2	0	DGD
45	02/08/2012	M/C	POLAC	3	0	SLV
46	05/08/2012	CAR	MAKOLE	3	0	DGD
47	05/08/2012	3BUS, CAR	GANO/WDL-KN	5	0	WOV
48	08/08/2012	ACCORD	POLAC	6	0	WOV
49	11/08/2012	N/A	GANO	25	1	DGD
50	18/08/2012	N/A	GACHI	4	0	SLV
51	20/08/2012	TYT	KONAR G/ALI	7	5	WOV/SLV
52	24/08/2012	CANTER	GADAR JANNA	16	1	SLV

53	26/08/2012	CANTER	POLAC	5	0	SLV
54	02/09/2012	HIACE	KA0N	16	1	TBT
55	10/09/2012	HIACE	500M AFTER	11	0	SLV
56	22/09/2012	P/UP	DA0	2	1	SLV
57	22/09/2012	M/CYCLE	WUDIL BRIDGE	3	0	SLV
58	22/09/2012	M/CYCLE	H/JAMA'ARE	2	0	SLV
59	23/09/2012	GOLF	DARKI	8	1	LOC
60	23/09/2012	CAMRY	GANO	16	1	DGD
61	24/09/2012	N/A	POLAC	3	1	LOC
62	24/09/2012	CANTER	NEAR GAYA	6	1	WRONG PARKING
63	26/09/2012	N/A	UTAI BEND	5	0	TBT
64	29/09/2012	HIACE	KANYAR	12	1	SLV
65	30/09/2012	GOLF	KWANAR GOGORADO	6	0	SLV
66	06/10/2012	CIVIC	GADAR JANNA	11	0	SLV
67	14/10/2012	TYT	POLAC	3	2	DGD
68	17/10/2012	LITEACE	JIGAWAR GANO	3	1	TBT
69	18/10/2012	HIACE HIACE	KWANAR GARKO	32	0	SLV
70	19/10/2012	BUS	SHAGOGO	15	0	TBT
71	19/10/2012	LITEACE	OPP STADIUM WDL	30	3	DGD,SLV
72	23/10/2012	STARLET	POLAC	3	0	DGD
73	25/10/2012	M/CYCLE	FRSC GATE	3	0	DGD
74	01/11/2012	GOLF	KANYAR UTAI	6	0	DGD
75	01/11/2012	HIACE	GARIN DAU	12	0	SLV
76	02/11/2012	TANKER GOLF	WUDIL BRIDGE	7	0	TBT
77	08/11/2012	TIPPER	NEAR POLAC	3	0	SLV
78	08/11/2012	M/CYCLE	GARIN DAU	3	0	SLV

79	16/11/2012	LITEACE	JIGAWAR GANO	14	0	TBT
80	17/11/2012	CIVIC	H/JAMA'ARE	4	0	WOV
81	19/11/2012	SIENNA	KANYAR UTAI	10	0	WOV
82	22/11/2012	CANTER	KWANAR LAMIRE	15	0	MDV
83	27/11/2012	HIACE	NEAR POLAC	12	0	DGD
84	29/11/2012	CIVIC	GANO	6	1	DGD

85	06/12/2012	COROLLA	NEAR YAR'GAYA	9	0	WOV
86	08/12/2012	N/A	OPP WDL LG SEC	2	0	WOV
87	14/12/2012	MAZDA	MAKOLE	5	1	WOV
88	14/12/2012	FORD	GANO BEND	12	0	DGD
89	17/12/2012	M/CYCLE	FRSC GATE WDL	3	0	DGD
90	22/12/2012	VECTRA	MAKOLE	19	0	WOV
91	23/12/2012	ACCORD	MAKOLE	3	0	SLV
92	23/12/2012	SPACE RUNNER	GIDAN KAYA	3	0	OLV,SLV
93	26/12/2012	SHARON	'YAN TUKWANE	20	0	WOV

APPENDIX B

Traffic Volume Count

Table B.1
Manual Traffic Count for Kano- Maiduguri

MANUAL TRAFFIC COUNT SHEET				
Road section	Kano-Wudil			
Date	27-06-2013			
Weather	Partly cloudy			
Observer	Ibrahim Khalil			
Direction	Kano-Maiduguri			
Time	Vehicle	Tally	Total	Percentages
8-9	Cars		496	
	Trucks		74	
	Total		570	
9-10	Cars		536	
	Trucks		60	
	Total		596	
10-11	Cars		689	
	Trucks		97	
	Total		786	
11-12	Cars		528	
	Trucks		110	
	Total		638	
12-13	Cars		557	
	Trucks		82	
	Total		639	
13-14	Cars		605	
	Trucks		96	
	Total		701	

14-15	Cars		579	
	Trucks		85	
	Total		664	
15-16	Cars		750	
	Trucks		109	
	Total		89	
16-17	Cars		898	
	Trucks		18	
	Total		1006	
	Cars			
	Trucks			
	Total			

Table B.2
Manual Traffic Count for Maiduguri - Kano

MANUAL TRAFFIC COUNT SHEET				
Road section	Kano-Wudil			
Date	27-06-2013			
Weather	Partly cloudy			
Observer	Ibrahim Khalil			
Direction	Maiduguri-Kano			
Time	Vehicle	Tally	Total	Percentages
8-9	Cars		648	
	Trucks		34	
	Total		682	
9-10	Cars		677	
	Trucks		56	
	Total		733	
10-11	Cars		763	

	Trucks		48	
	Total		811	
11-12	Cars		622	
	Trucks		77	
	Total		699	
12-13	Cars		571	
	Trucks		78	
	Total		649	
13-14	Cars		56	
	Trucks		74	
	Total		730	
14-15	Cars		496	
	Trucks		45	
	Total		541	
15-16	Cars		573	
	Trucks		76	
	Total		636	
16-17	Cars		115	
	Trucks		751	
	Total			
	Cars			
	Trucks			
	Total			
	Cars			
	Trucks			
	Total			

Table B.3
Summary of Manual Traffic Count for Maiduguri - Kano Road

MANUAL TRAFFIC COUNT SHEET					
Road section	Kano-Wudil				
Date	27-06-2013				
Weather	Partly cloudy				
Observer	Ibrahim Khalil				
Direction					
TIME		KAN-MAID	MAID-KAN	TOTAL	% OF TRUCKS
8-9		682	570	1252	
	TRUCK	34	74	108	8.63
9-10		733	596	1329	
	TRUCK	56	60	116	8.73
10-11		811	786	1597	
	TRUCK	48	97	145	9.08
11-12		699	638	1337	
	TRUCK	77	110	187	13.98
12-13		649	639	1288	
	TRUCK	78	82	160	12.42
13-14		730	701	1431	
	TRUCK	74	96	170	11.88
14-15		541	664	1205	
	TRUCK	45	85	130	10.78
15-16		649	859	1508	

	TRUCK	76	109	185	12.27
16-17		751	1006	1757	
	TRUCK	115	108	223	12.69

APPENDIX C

Additional Case Study pictures

Figure C.1
Improper access of users into the work zone

Figure C.2
Unprotected working area Km 1+560

Figure C.3
Damaged traffic sign

Figure C.4
Missing buffer

Figure C.5
Safekeeping of equipment after work

Figure C.6
Posted speed limit through the work zone

Figure C.7
Dangerous pedestrian device

Figure C.8
Transition zone at Km 7+800

Figure C.9
Transition zone at 15+025

Figure C.10
Transition zone at 25+025

Figure C.11
Work equipment accident

Figure C.12
Tipper struck the median entering working area

Figure C.13
Truck left unattended after an accident

Figure C.14
Long truck accident

Figure C.15
Accident while off lording

Figure C.16
On-street parking Km 0+900

Figure C.17
Misuse of traffic control device Km 3+450

Figure C.18
Non-compliance to signs by road users Km 8+870

Figure C.19
Dangerous facility for pedestrians 3+000

Figure C.20
Equipment on road users’ path

Figure C.21
Improper sign placement

Figure C.22
Damaged warning signs Km 5+100

Figure C.23
Confusing traffic signs Km 7+800

Figure C.24
Non-reflective drums Km 0+000

Figure C.25
Missing buffer space Km 0+566

Figure C.26
Missing taper Km 9+500

Figure C.27
Dangerous flagging Km 8+890

Figure C.28
Workers exposed to moving traffic Km 18+480

NOW VIEWING: HOME > MY PAPERS > PAPER

About this page

This is your assignment inbox. To view a paper, select the paper's title. To view a Similarity Report, select the paper's Similarity Report icon in the similarity column. A ghosted icon indicates that the Similarity Report has not yet been generated.

paper

INBOX | NOW VIEWING: NEW PAPERS ▼

Submit File

Online Grading Report | Edit assignment settings | Email non-submitters

	AUTHOR	TITLE	SIMILARITY	GRADE	RESPONSE	FILE	PAPER ID	DATE
	Mustapha Nuhu Yahya	ABSTRACT	0%	--	--		1443161088	11-Nov-2020
	Mustapha Nuhu Yahya	CH1	0%	--	--		1443110900	11-Nov-2020
	Mustapha Nuhu Yahya	CH4	0%	--	--		1443113497	11-Nov-2020
	Mustapha Nuhu Yahya	CH5	0%	--	--		1443172678	11-Nov-2020
	Mustapha Nuhu Yahya	CH3	8%	--	--		1443712017	12-Nov-2020
	Mustapha Yahya Yahya	Thesis	10%	--	--		1443119319	11-Nov-2020
	Mustapha Nuhu Yahya	CH2	12%	--	--		1443111580	11-Nov-2020

Assoc. Prof. Dr. Shaban Ismael Albrka

NEAR EAST UNIVERSITY YAKIN DOĞU ÜNİVERSİTESİ

<u>ETHICS LETTER</u>

TO GRADUATE SCHOOL OF APPLİED SCİENCES

REFERENCE: MUSTAPHA NUHU YAHYA (20193835)

I would like to inform you that the above candidate is one of our postgraduate students in Civil Engineering department he is taking thesis under my supervision and the thesis entailed: ***Artificial Intelligence Based Hybrid Models for Classification of Road Accident Severity. The data used in our data collected from department of traffic engineering in Nigeria.*** The data used in the study was our data obtained from Traffic department in Nigeria.

Please do not hesitate to contact me if you have any further queries or questions.

Thank you very much indeed

Best regards

Assoc. Prof. Dr. Shaban Ismael Albrka
Faculty of Civil and Environmental Engineering, Near
East Boulevard, ZIP: 99138
Nicosia / TRNC, North Cyprus,
Mersin 10 – Turkey.

MUSTAPHA NUHU YAHYA

29B, Justice N Y Galadanchi Street, Kabuga Housing Estate, Janbulo, Gwarzo Road, Kano.

Phone; +2348063353538

Email; mustygaladanchi@gmail.com

Personal Data	
Date of birth:	30th January, 1992
Sex: Male:	Male
Marital Status:	Single
State of Origin:	Kano State
Local Govt. Area:	Gwale

CAREER OBJECTIVES

- To work and contribute to the best of my ability so as to improve the organizational objectives and accomplishment of overall mission statement.
- To improve an organizational setup through my academic training and technical skills by coordinating and supervising of various projects.

ACADEMIC QUALIFICATIONS WITH DATES

Near East University, Turkish Republic of Northern Cyprus 2019-2021

- M Sc Transportation and Highway Engineering

Nation Youth Service Corps, Kaduna 2017-2018

- Federal College of Education, Zaria

Bayero University Kano 2011-2017

- B. Eng Civil Engineering (2nd Class Honour)

Day Science College, Kano 2007-2010

- Senior Secondary School Certificate

Prime College, Kano 2004-2007

- Junior Secondary School Certificate

The Caliphate Nur. and Pri. School 1998-2004

- Primary Leaving Certificate

WORK EXPERIENCE

Duo Creed Global Limited, Kano. 2021- Date

Job Title: Project Engineer

Responsibilities:

- Work as site Engineer
- Supervision and coordination of all structural engineering work on site.

Federal College of Education, Zaria. 2017- 2018

Job Title: NYSC

Responsibilities:

- Worked with the Estate Department of the College
- Site visitations, supervision and coordination of projects ranging from structural, highway, hydro.
- Installations and general maintenance
- Observed the preparation, function and process of tender and contract documents in action

UYK NIG LTD, Kano. 2015

Job Title: Siwes II

Responsibilities:

- Worked with the Project Coordinator
- Supervision and coordination of structural engineering work.

Ministry of Works, Housing and Transport Kano State 2015

Job Title: Siwes I

Responsibilities:

- Worked with laboratory
- Testing soils and materials to determine the adequacy and strength of foundations, concrete, asphalt or steel.
- Suggestion and recommending the best quality material for work.
- Compiling checking and approving reports.

PROFESSIONAL CERTIFICATION

Bayero University Kano

- Professional diploma in computer studies 2012
- Certificate in computer studies 2010

PROFESSIONALL SKILLS

- Experience of working and managing within a changing environment
- Good commercial awareness including programming and delivery to budget/ timescales
- Detailed knowledge of construction principles and standards
- Good understanding of safety standards and CDM regulations

TECHNICAL SKILLS

- Ability to communicate effectively, clearly, fluently, interpersonally and in written forms
- Able to relate with people from various backgrounds in simple and understanding manner.

COMPUTER SKILLS

- Microsoft office tools
- AutoCAD

- MATLAB
- Arc GIS

OTHER SKILLS

- Ability to work under pressure and target solution
- Ability to work in a team and as an individual with minimum guide
- Creativity

HOBBIES

- Reading
- Travelling

REFEREES

- Dr. Kassim Muhammad
  Faculty of Engineering, Bayero
  University Kano
  Phone: +2348060598055
  kmohammed.civ@buk.edu.ng

- Engr. Mustapha Aminu Ibrahim
  Director works,
  Kano University of Science and Technology, Wudil.
  Phone: +2347063324352
  mustaphaaibrahim@yahoo.com

- Abdulhakim Ibrahim kiyawa
  Lecturer, Faculty of Environmental Management,
  Bayero University Kano.
  Phone: +2347039170868
  aikiyawa.em@buk.edu.ng