

Kennesaw State University

DigitalCommons@Kennesaw State University

African Conference on Information Systems
and Technology

The 6th Annual ACIST Proceedings (2020)

Jul 2nd, 4:45 PM - 5:00 PM

Predicting Factors Contributing to Road Traffic Accident and Implying Driver's Driving Behavior in Addis Ababa City,

shimelis Shiferaw Deneke Mr.

Updated - AIS, shimlishshi@yahoo.com

SHIMELIS SHIFERAW DENEKE MR

AIS, shimlishshi@yahoo.com

Abdurahman Ali endalkachew Mr

AIS, alliendalkachew@gmail.com

Shimelis shiferaw Deneke

Jimma university

Abdurehman Ali

Addis Ababa University

Follow this and additional works at: <https://digitalcommons.kennesaw.edu/acist>



Part of the [Business Intelligence Commons](#), and the [Management Information Systems Commons](#)

Deneke, shimelis Shiferaw Mr.; DENEKE, SHIMELIS SHIFERAW MR; endalkachew, Abdurahman Ali Mr; Deneke, Shimelis shiferaw; and Ali, Abdurehman, "Predicting Factors Contributing to Road Traffic Accident and Implying Driver's Driving Behavior in Addis Ababa City," (2020). *African Conference on Information Systems and Technology*. 1.

<https://digitalcommons.kennesaw.edu/acist/2020/allpapers/1>

This Event is brought to you for free and open access by the Conferences, Workshops, and Lectures at DigitalCommons@Kennesaw State University. It has been accepted for inclusion in African Conference on Information Systems and Technology by an authorized administrator of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.

Kennesaw State University

DigitalCommons@Kennesaw State University

AFRICAN CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGY

Title: Predicting Factors Contributing to Road Traffic Accident and Implying Driver's Driving Behavior in Addis Ababa City

Author #1

Author #2

Date May 18, 2020
Addis Ababa / Ethiopia



The African Conference on Information Systems and Technology (ACIST)

ABSTRACT

This study tries to uncover which factor could be associated with Addis Ababa City road traffic accident and implicate drivers driving behaviors based on the data collected from post-accident. Which is the capital city of Ethiopia. The main objective of this research is Applying data mining technique to explore on factors that could associated with drivers related factors with road traffic accident to provide an insight way to reduce road traffic accident. The data was collected from road traffic accident from the real-world data collected at accident time in Addis Ababa City/Ethiopia. Totally 4285 datasets with 20 attributes were used. Which includes relevant attributes contains information about road, driver, vehicles and accident and other related variables. Finally, in the result implies novel association rules were identified, especially the male drivers' behaviors has indicated associated with speed driving behaviors and this implies need further investigated with behavioral science approach for more specific result.

Keywords: Data mining on road traffic accident, factors on road traffic accident, A prior algorithm on road traffic accidents, Addis Ababa City road traffic accidents.

INTRODUCTION

In this study, we have tried to explore the association of driver's behavior among various possible factors using one of machine learning techniques. Machine learning approach is believed to be used to uncover hidden patterns from real world road traffic accident data and based on injury severity level to get insight about what is really related to the driver's behavior that cause accident will be predicted. Currently road traffic accident is major cause of mortality, morbidity and also disability in the entire globe (Bigham, 2016).

This will help any stakeholders responsible in reducing road traffic accident. Policy makers can also use this as an input in considering different factors while working on traffic accident reduction.

This work attempts to answer: what hidden patterns that attributes to the cause of accident that could be associated with the driver's behavior? "Envisaging speeding behavior of driver from none behavioral data"

STATEMENT OF THE PROBLEM

In relation to the severity of road traffic accident in Ethiopia it is well articulated in various literature as one of the major health and development challenge. According to (Jima, 2019) in support of this notion in his finding states that, more than 276491 road traffic accidents, 912956km road network and 681000 number of motorized vehicle were newly introduced since 2007/08 until 2018 and also (Jima, 2019) stated that, due to Road traffic accident Ethiopia loses around 36.3 billion birr (estimated 1.3 billion \$ in current exchange rate of 28 birr for 1\$) in the past eleven (11) year in Ethiopia. In average annually Ethiopia loses around 0.9% of budget due to traffic accident in the past eleven year. Similarly in Increase on the number of motor vehicle, truck and other form of transportation machineries vital for the economic development of every level of economy (Adnan et al., 2018). However; as the number machineries increases the road traffic safety demand comprehensive traffic accident factors that could contribute in reducing the incident of traffic accident. And this will give better justification to do this research which help wise decision-making act from all perspectives of stockholders.

On top of that In order to understand the severity of the problem examining some important reports on the traffic road accident related death and injuries and looking in to some of related literature will help in understanding the problem better. According to the latest WHO report (WHO 2015), shows that 1.25 million people are killed each year on the world related to traffic accidents and a 50 Million injuries per year (WHO, 2013). Besides that according to (Bener et al. (2003), (WHO, 2009) cited on (Beshah Tesema et al, July 2013) And also It has estimated that 90% of road traffic deaths occur in low and middle income countries (WHO, 2009).

Road Traffic Accidents Deaths in Ethiopia reached 15,015 or 2.50% of total deaths ranks Ethiopia #68 in the world. WHO works with partners—promote good practices to avoid traffic accidents which are related to helmet and seat-belt wearing, not drinking and driving, not speeding and being visible in traffic (WHO,2015).But, the accidents are still unavoidable and the mortality, morbidity disability caused by accident and cost of the accidents can be a series problem for both government and the people. Even though the Government of Ethiopia putting an effort to alleviate or mitigate the accident incidence nonetheless; the problem of care accident have becoming day to day phenomena as a result it is taking peoples life , increasing disability and suffering from the incident is both economical and socially painful to the nation in general.

Off curs different researchers are working on the road traffic accident safety from different perspectives moreover considering the magnitude and the impact of the problem attacking and approaching the problem from different direction will provide meaning full impact on the solution.

Besides According to different literature stated in the world, more than 90% reports of road traffic accident attributes to driver errors, to support this according to different results in Ethiopia reported that 76-81% accident happen because of driver related factor (Gidey, 2010, Beshah Tesema et al, July 2013).

In line with the above discussion, it semis the problem is found i.e. the driver related factors are the major one, but when and in what why do the factors contribute to accident? Or which factors come together could probably because accident is relevant question? These are question which could not be addressed in the traditional research methods and yet not answered, therefore the applying data mining techniques to predict driver's behavior which facto really attribute to accident will be examined using association rule.

RELATED WORKS

Data Mining is an interactive process of discovering valid and novel, useful and understandable patterns or models in large database (Han, Mannila and Smyth, 2001). Data Mining, according to (Han, Mannila and Symth ,2001) is a process that uses a variety of data analysis tools to discover patterns and relationships in data that may be used to make a valid prediction. Data mining uses advances in the field of Artificial Intelligence (AI) and Statistical techniques. Therefore, association rule analysis is being used in this research. Hidden patterns can be discovered from real world road traffic accident data. Different methods in analytics of road traffic data have been used and researched in different way. But, association rule analysis in specific application is less researched. (Moradkhani,F.,et al .,2014).

Similarly , in supporting the relevance data mining as research techniques for this research other researchers like (Bigham, 2016) has used similar approaches and had used the following attributes or

recorded details include time, date and location, type, location and maneuver of the vehicle, driver details, details of injury and injury severity. Just like this paper which has used secondary data source that was collected at accident event by government authority offices. Similarly (Bigham, 2016) had used secondary dataset which contains 4382 records of accident characteristics that was occurred in the study region

In selecting the specific data mining algorithms to deliver similar pattern related factors discovery. The association rule discovery commonly known as market basket analysis that attempts to discover hidden linkage between data items. Therefore, in this study, we will apply A priori association rule discovery algorithm to produce dependency rules which will predict occurrence of accidents factors based on occurrence of another accident factor and effective decisions can be taken to reduce accidents.

The other goal of this study is to predict or understand the cause of accident by injury severity as target class which has four classes namely:

“Property Damage”, ”Series Injury”, “Fatal Injury “and “Slight Injury”.

Due to the availability of huge amount of data, almost all companies are looking to exploit their historical data for taking competitive advantages. The traditional technique which is manual analysis of data become a challenge and exceeds statistician and analyst’s capacity due to the size and complexity of data. The availability of this data has led to instated interests in data mining techniques and methods for extracting useful information and knowledge from the data to find insights and new opportunities (Provost & Fawcett, 2013).

Data mining therefore has been proven as a useful tool to address the need of converting the collection of data to useful information.

Result from The Association

In recent years, researchers from computing fields have been trying to analyze the real-world traffic data and struggle to divert to meaningful information called knowledge. Most researches are in utilizing this traffic data and studying different aspects of traffic accident injuries. Very related relevant papers can be summarized below as follows with the gap involved. (T.B., Tesema, & et al.,2005), used adaptive regression trees to mining rules that predict accident severity. But, the nature of road traffic data records most of the attributes are not related to each other. Since Decision Tree algorithm consider or assume attributes are related to each other (Langley, P.,1994), there is a need for applying association rule algorithms that did not considered or less researched. In fact, traffic data are in nature different attributes which has not any relationships at all. For e.g. Vehicle related factors and road related factors, driver related factors. Therefore, it is necessary to evaluate traffic data by using other algorithm like Apriori algorithm since others like (T.B.,Tesema et al.,2005), applies DT,NB and K-means algorithm to predict the severity

level based on road related factors. But still the real world dataset consists of different factors and the result is still unknown by investigating driver and car related factors.(Chong.,M. et al.2004),apply data mining technique and tries to develop a classifier model of injury severity by using data collected in automobile accidents. The goal was to improve accuracy of prediction of previous work by (Abdelwehab, et al, (2001) They have investigate the performance of neural networks and decision trees applied to predict drivers' injury severity in head-on front impact point collisions. For all the five injury classes such as: No Injury, Possible Injury, Non-Incapacitating Injury, Incapacitating Injury and Fatal Injury, the decision tree approach outperformed neural networks. Martin,L., et al., (2014), used data mining techniques to analyze the relationships between road conditions and crashes in the dataset collected during 2008-2010 accident records of a complementary road network of Andalusia, Spain government. The goal of the study was to improve road safety condition of Andalusia, Spain by analyzing the characteristics of the roads such as typology of roads, daily traffic volume, and proximity of urban zones. Complementary road network project defines the elements which technically called susceptible elements of improvements (ESM).The ESM elements show worse road condition than the standard of road safety. Such elements have a greater influence in accidents. Association rule mining algorithms the researchers able to find rules for the factors causing crashes. The researchers claim that they found the main causes for occurrence of accident. But the experimental report fails to show the rules explicitly rather they summarized and report the result. So it is really difficult to understand and interpret different factors for any organization responsible to reduce road traffic accidents.

As the previous literature shows, there are various factors on accident severity and many research proposed have been proposed to investigate this factors. The other scholars (Bae et al. ,2012) investigated the factors on traffic severity by considering the data heterogeneity. (Depire et al., 2008) also presented that the traffic accident data generally include heterogeneity of accidents. Bae et al. made an assumption that factors on traffic accident can be different depending on accident types, to reduce data heterogeneity. Both in the traditional research method indicated above and with knowledge discovery methods lack showing deep down to the real factors. In this work, I made an effort to investigate the main factors of accidents by using a priori algorithm for predicting accidents causing factor in real world heterogeneous data.

Therefore this paper uses association rule to predict what case the driver to behave or action with respect to the environment will be predicted in relation to the accident severity ,does the factor like age, driving experience, educational level and factors drivers environment and traffic environment had an impact on the level of injury(severity), moreover the drivers vs car connection(hired or own) , vehicle ownership (private, government) have been considered as a factors that influences each other this study

METHODOLOGY

This section will present the road traffic accident mining process. The knowledge discovery process steps (KDD) is shown in Figure:1 below. KDD could be employed for enhancing communication between domain experts and data mining method developers.

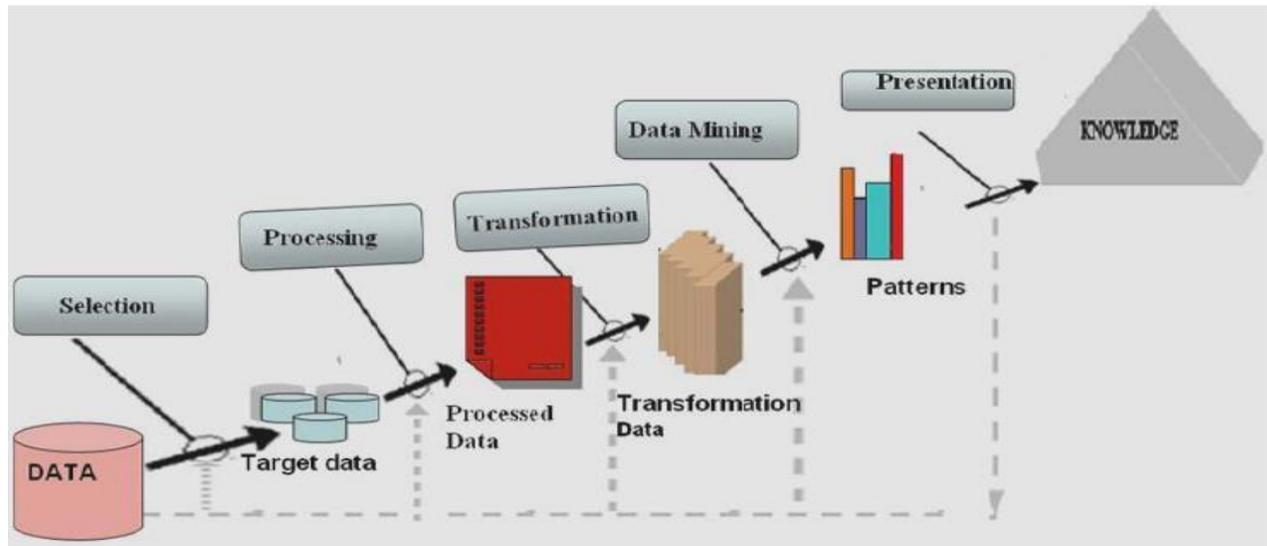


Figure 1: KDD Process [Adopted from (Han,j.&Kamber,M.,2006)]

KDD is often used as a synonym for Data Mining. KDD is the process of finding useful information and patterns in data. KDD can be defined as the whole process involving: data selection; data pre-processing: cleaning; data transformation; mining; result evaluation using the various techniques to extract useful information/pattern from the data. Therefore, DM is a step in KDD which is the and visualization. (Han, j.&Kamber, M.,2006). Data Mining, on the other hand, refer to the modeling step use of algorithms to extract hidden patterns & knowledge in data. Knowledge discovery in databases (KDD) process steps which are employed in this study are summarized as follow:

Selection

Selection is the first step in KDD process for obtaining data from various heterogeneous sources such as databases, data warehouses, files, non-electronic records, etc. Data mining requires collecting great amount of data to achieve the intended objective. Thus, the researcher in this study collects the data from Addis Ababa traffic accident department which is under federal police office.

Preprocessing

Before feeding the data to DM algorithms, we have to make sure the quality of data. Preprocessing in the second step in KDD in order to address quality related issues. In this research, after the data is collected from Addis Ababa Transport Authority the researcher made some preprocessing issues for assuring the quality of data which includes cleanse inconsistent & incorrect data; fills incomplete records; predict missing values; correct erroneous & anomalous data. Since this all preprocessing methods requires different algorithms and domain experts. Each of these preprocessing tasks are done in collaboration with traffic accident investigation officers. Unknown values were identified and labeled with code.

Transformation

Transformation is the third steps in KDD for Converting the data collected from different sources into common new format. Apply data reduction & data categorization/binning to ease data mining. For example, the drivers age attribute values is in continues data type form so it is difficult for learning the data mining tool with this continues values. Thus, the researcher converts the age continues valued attribute to discrete attribute format for the ease of using the tool. The other issue was, the data are collected in excel data format. But, the employed DM tool requires the data in CSV OR ARFF file format. Therefore, the dataset format is changed to CSV data format. The csv file is attached in the folder given.

Mining Methods

The traffic accident data is collected from traffic office accident investigation department. Association rule discovery methodology used for detecting any relationships between road traffic accident factors. Specifically, A priori algorithm is employed for mining association rules. Finally, 4285 datasets were used.

Association Rule Mining (A Priori Algorithm)

Association rule mining is to find all rules on frequent item sets of the form $X \rightarrow Y$ that fulfills minimum confidence and support (Moradkhani, F., 2014). The general form of the rule is as follows: “IF event X occurs THEN event Y occurs as well, in M% of times, and this pattern occurs in N% of all events in the dataset”. Where, M is the Confidence, and N is the Support. Confidence, c, is the conditional probability that a transaction having X also contains Y which means conditional probability

(confidence) of Y given X greater than or equal to user defined threshold c. In this study A priori algorithm is used to mining interesting rules that fulfills the threshold values.

Support represents the probability that both events X and Y occurred simultaneously in the dataset. There are many rules could be mined by using this algorithm. So, it is really difficult for identifying which rules are interesting since it is very challenging difficult to set the threshold values. In order to solve this problem, the researcher used to consult traffic officers and Road and transportation office experts to identify interesting rules. So that we can identify different interesting rules which could implies the cause of road traffic accidents. By using expert's judgment as an input, I have set the minimum support value 0.25 and confidence 0.9 and maximum no. of rules 10.

The default minimum Support was 0.1 but in search of best rule I Have changed the following properties of a priori algorithm in weka

Car =True . this property, suggests that if enabling class association rule are mined instead of (the general) association rule. And the lowe bound mimSupport value were 0.25 , at 3 no association rule were produced in the experiment .

Result of A priori Association Rule



```

20:35:04 - Apriori
Number of large itemsets:

Number of large itemsets L(1): 18
Number of large itemsets L(2): 99
Number of large itemsets L(3): 270
Number of large itemsets L(4): 437
Number of large itemsets L(5): 435
Number of large itemsets L(6): 262
Number of large itemsets L(7): 90
Number of large itemsets L(8): 14

Found:

status=NoFault AccidentType=FrontToBackCollision 1099 ==> SeverityLevel=PropertyDamage 1073 conf:(0.98)
status=NoFault RoadType=Asphalt AccidentType=FrontToBackCollision 1099 ==> SeverityLevel=PropertyDamage 1073 conf:(0.98)
x=M AccidentType=FrontToBackCollision 1129 ==> SeverityLevel=PropertyDamage 1101 conf:(0.98)
x=M RoadType=Asphalt AccidentType=FrontToBackCollision 1129 ==> SeverityLevel=PropertyDamage 1101 conf:(0.98)
x=M WeatherCondition=GoodAire AccidentType=FrontToBackCollision 1124 ==> SeverityLevel=PropertyDamage 1086 conf:(0.98)
x=M RoadType=Asphalt WeatherCondition=GoodAire AccidentType=FrontToBackCollision 1124 ==> SeverityLevel=PropertyDamage 1096 conf:(0.98)
x=M RoadCondition=Dry AccidentType=FrontToBackCollision 1116 ==> SeverityLevel=PropertyDamage 1088 conf:(0.97)
x=M RoadType=Asphalt RoadCondition=Dry AccidentType=FrontToBackCollision 1116 ==> SeverityLevel=PropertyDamage 1088 conf:(0.97)
x=M RoadCondition=Dry WeatherCondition=GoodAire AccidentType=FrontToBackCollision 1114 ==> SeverityLevel=PropertyDamage 1086 conf:(0.97)
x=M RoadType=Asphalt RoadCondition=Dry WeatherCondition=GoodAire AccidentType=FrontToBackCollision 1114 ==> SeverityLevel=PropertyDamage 1086 conf:(0.97)

```

Figure: 2 Result of A priori association result

=== Run Information ===

Scheme: weka. associations .A priori -N 10 -T 0 -C 0.9 - D 0.05 -U 1.0 -M

0.25 -S -1.0 A -c -1 Relation:Accident Data shine 1vsc023

Instances: 4285, Attributes: 20

Driver related: Driver Sex, Driver Age, Driver Educ. Level, Driver Car Connection, and Driving Exp. Years

Vehicle related: Vehicle Status, Vehicle Movement, Vehicle Ownership, Vehicle Type, and Service Year

Road related: Road Separation, Road Junction, Road Type, Road Condition

Weather related: Light Condition, Weather Condition

Accident related: Accident Type, Accident Factor, License Grade, Severity Level (Class)

The buffer result of A priori association rule output is included in the project folder. For example, among the rule generated rule no. 10 is

DriverSex=M, RoadType=Asphalt, RoadCondition=Dry, and WeatherCondition=GoodAire AccidentType=FrontToBackCollision)1114 ==> SeverityLevel=PropertyDamage) 1086 [conf:\(0.97\)](#)

Which means that *“if the gender of the driver is male, the road Condition is Dry, Weather Condition is Good air, Accident Type is Front To Back Collision then Severity Level is property Damage 1086 , with confidence interval of [conf:\(0.97\)](#)”*

Taking the variable into their natural group we have vehicle related, Road related weather related and accident related variables

1. **VehicleStatus=NoFault**, AccidentType=FrontToBackCollision1099==>SeverityLevel=PropertyDamage 1073 [conf:\(0.98\)](#)
2. VehicleStatus=NoFault, **RoadType=Asphalt**, AccidentType=FrontToBackCollision, 1099 ==> SeverityLevel=PropertyDamage 1073 [conf:\(0.98\)](#)
3. **DriverSex=M**, AccidentType=FrontToBackCollision,1129==>SeverityLevel=PropertyDamage 1101 [conf:\(0.98\)](#)
4. DriverSex=M,**RoadType=Asphalt**,AccidentType=FrontToBackCollision,1129==> SeverityLevel=PropertyDamage1101 [conf:\(0.98\)](#)

5. DriverSex=M **WeatherCondition=GoodAire** AccidentType=FrontToBackCollision 1124 ==>
SeverityLevel=PropertyDamage 1096 [conf:\(0.98\)](#)

From the give rule developed from A prior algorithm result If we considered the above highlighted variables and considering all experiment results found under this study :

Vehicle status has no problem, the road type is asphalt and the driver sex is male weather condition is good And the collision type front to back, this variable could have explaining potential to describe the drivers behavior during the accident time.

Just like market basket analysis this variable have some association with drivers behavior towards speed of driving , since having good weather condition , good road condition(Asphalt) and good vehicle status could have association with speed and being male diver could also have hormonal influence towards speed even though this has to be justified for behavioral science. But it will indicate good factorial association with the occurrence of the accident in this work.

This association is novel since it implies an appropriate indication towards the cause of driving behavior which will be expressed in terms of speed.

CONCLUSIONS AND RECOMMENDATION

This study will investigate possible application of data mining technology at Addis Ababa Traffic Office, for analyzing different factors of road traffic accidents.

A priori algorithm can also discover interesting rules and relationships between factors affecting road traffic accidents focusing on drivers driving behavior especially.

This study can be an input for all stakeholders who are functioning to reduce road traffic accidents by understanding the drivers driving behavior towards speed, in different road, weather condition need to be considered. The limitation for this study is the training dataset classes are not equally labeled. This may create a bias in predicting different classes. In future, it is better to label the dataset in equal instances of classes which can help to avoid biases. And moreover behavioral studies are needed to determine exactly the speed driving behaviors of drivers by including the following variables driver sex, vehicle status, road type in different weather condition possibly give a clear cases of accident .Generally we can say as factor to traffic accident there are a lot of supporting evidences that

implies it is associated with drivers errors which is general but this work could show that more specifically from drivers error bringing speed viruses controlling the speed related issues on board could be the contribution of this work . besides the limitation other researchers might consider with more data set that comprise the entire country since this research had used dataset collected at the capital city Addis Ababa only.

REFERENCES

- Adnan, S., Shah, R., Ahmad, N., Shen, Y., Pirdavani, A., Basheer, M. A., & Id, T. B. (2018). *Road Safety Risk Assessment : An Analysis of Transport Policy and Management for Low- , Middle- , and High-Income Asian Countries*. 30. <https://doi.org/10.3390/su10020389>
- Bigham, B. S. (2016). Road Accident Data Analysis : A Data Mining Approach. *Researchgate*, (September).
- Jima, D. D. (2019). Road Traffic Accident in Ethiopia from 2007/08-2017/18 (Since Ethiopia Millennium) (Vol. 18). <https://doi.org/10.1313/aijser.v2i2.90>
- Abdelwahab, H. T. & Abdel-Aty, M. A(2001)., Development of Artificial Neural Network Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections. Transportation Research Record 1746, Paper No. 01-2234. B.
- Tesema e tal , (July 2013), Integrated Information Architecture in Support of Road Safety Organizations: The Case of Ethiopia, A PhD Dissertation, Addis Ababa University, Ethiopia
- Bae, Y.,Kim,J.,Chung,J.,(2012),Traffic accident analysis for High way on weather condition and Time, International conference on Transport, Environment and civil engineering, Malaysia
- Chong ,M.,Abraham,A.,Paprzycki,M.,(2004),Traffic accident analysis using Decision Tree and Neural Networks, Oclahoma State university, USA.
- DEBELA DEME et al (October 2016) "TRAFFIC ACCIDENT CAUSES AND COUNTERMEASURES OF ADDIS ABABA-ADAMA EXPRESSWAY" Addis ababa University Msc thesis ,
- Depaire, B., Wets, G., Vanhoof, k.,(2008) ,Traffic accident segmentation by means of latent class clustering, Accident Analysis and Prevention 40, pp.1257-1266.
- Han,J.,Kamber,M.,(2006),Data Mining: Concepts and Techniques, second edition, Isevier.http://www.who.int/gho/publications/world_health_statistics/2016/en/,last accessed date, January
- Jacob P., (2010),Python Text processing with NLTK 2.0 cookbook",PACKT,
- Jian-feng Xi, et.al , (9 January 2014) "The Model of Severity Prediction of Traffic Crash on the Curve " Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2014, Article ID 832723, <http://dx.doi.org/10.1155/2014/832723>,

- Langley, P., & Sage, S. (1994, July). Induction of selective Bayesian classifiers. In Proceedings of the Tenth international conference on Uncertainty in artificial intelligence (pp. 399-406). Morgan Kaufmann Publishers Inc.
- Lenka Komackova and Milos Poliak, 2016 , “ Factors Affecting the Road Safety” Journal of Communication and Computer doi:10.17265/1548-7709/2016.03.006
- Martin,L.,et al J.,(2014),using data mining technique to road safety improvement in Spanish roads,Science direct,vol. 160(507-614)
- Minister of Transport, N. R. (February 2011). National Road Safety strategic plan of Ethiopia 2011 to 2020. A.A
- Ministry of Interior,(2013,) Ministry of Interior, General Traffic Department, (Annual Statistical Report, 2013), Kingdom of Saudi Arabia
- (Moradkhani,F.,et al ., 2014), “Road Accident Analysis: A data mining approach”, Provost,F, and Fawcett,T,(2013), “Data Science for Business”,O’Reilly,978-1-449-36132-7.
- WHO ,2013. Global status report on road safety 2013: Supporting a decade of action. Luxembourg:
- WHO 2015, Global status report on road safety 2015, Violence and Injury Prevention, http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/
- Xue, Y., et al. (2006). NBA-Palm: prediction of palmitoylation site implemented in Naive Bayes algorithm. BMC bioinformatics, 7(1), 4