Proceedings of Second Shri Chhatrapati Shivaji Maharaj QIP Conference on Engineering Innovations
Organized by Shri Chhatrapati Shivaji Maharaj College of Engineering, Ahmednagar
In Association with Novateur Publications JournalNX-ISSN No: 2581-4230
February, 22nd and 23rd, 2019

Accident prevention systeam using android application

Shinde Aman Pandurang
Computer Department
SSMAP, Chass, Ahmednagar, India

Gaikwad Anjali Navnath
Computer Department
SSMAP, Chass, Ahmednagar, India

Ingale Komal
Computer Department
SSMAP, Chass, Ahmednagar, India

Ligade Mayuri Balasaheb
Computer Department
SSMAP, Chass, Ahmednagar, India

***Abstract*—Mostly, it is found that road accident happening are more frequent at certain specific locations i.e. black spot. The analysis of these black spot can help in identifying certain road accident factor that makes a road accident to occur frequently in that location. In this project we apply statistics analysis and data mining algorithms on the Fatal Accident dataset as an attempt to address this problem. Association rule mining is one of the popular data mining techniques that identify the causes of accident of road accident. In this project, we first applied k-means algorithm to group the accident locations into four levels, zero level, first level, second level and third level accident location. K-means algorithm takes accident level count as a factor to cluster the locations. Then we will use association rule mining to identify these locations. The rules show different factors associated with road accidents at different locations. For all this we will provide accident data that are issue from Nasik city commissioner office. Safety driving suggestions will be making based on accident data, association rules, classification model, and clusters obtained.**

I. INTRODUCTION

1.1 Project Idea

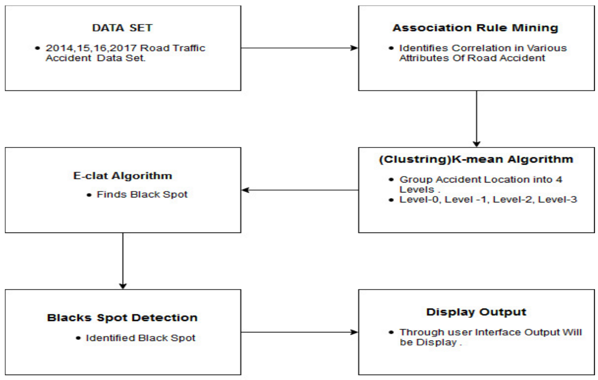
To identify important factors to road accidents in Nasik we have obtained a large data set every accident recorded in the Nasik district commissioner office in the Year2014-2017. The data is currently in an unsorted and scatter format and is stored in a Microsoft excel sheet database table. Unfortunately with the data in its current format, no relevant points or conclusions can be drawn. It is hoped that by applying data mining processes and techniques to the data set, relevant attributes and patterns can be established. And scientific study will also done that will helpful to government authorities and citizen. The main achievements of this project is to greater awareness of the conditions affecting road traffic accidents ,Establishing which individuals are most likely to be involved in a road traffic accident.

Problem Statement:
Road and traffic is most important issue not only for Indian government but also for common people. Mostly, it is found that road accident happening are more frequent at certain specific locations.

Objectives:

- When the vehicle is going towards accident spot, user will get notification through an android application.
- User will get voice notification for an black spot.
- Low cost solution for getting information about black spot.
- Voice notification will get before user reach to that place.

System Architecture:



For dataset purpose we required previous happened accident location. So we will take help of records of road traffic accident with reason how they happened.
In association rule mining from dataset we will collect various attribute points. By using clustering algorithm we will find black spots clusters and their clusters according to reasons. We finally we will generate level1, level2, level3 output according to accident reasons.

Motivation of the Project
Road accidents and traffic is most important issue not only for Indian government but also for common people. Road safety becomes a major public health concern. Everyday lots of vehicles driving on the road, and traffic accidents happen at anytime and anywhere. Some people die in accident also. As human being we all want to avoid accident and stay safe. To find out how to drive safer, data mining technique could be applied on the traffic accident dataset to find out some valuable information, thus give driving suggestion

II. SYSTEM REQUIREMENT

Hardware Requirement
Hardwar and software requirements for the system are stated below
Computer:

- Processor –Core i3.

- Bus speed **-** 2.5 GT/s DMI.
- Hard disk **-** 160 GB and
- Memory size – 1GB RAM.

Software Requirement

- jdk 8 or above
- Android studio v2.3.3
- Xamp server
- Apache tomcat server 7.0

Development Tool:

Developer Tools	**Description**
JDK 1.8	For JAVA Platform
Xamp	For My Sql database
Android studio v2.3.3	For android code editing
Apache Tomcat 7.0.56	For database servlets
Latex	For report generation

Testing Environment:

Software Required (Client Browser)	**Description**
OS	Windows , Linux
Browsers	Chrome , Mozilla Firefox etc.
MODEM Drivers	For internet connections

III. Conclusion

A thorough literature review revealed a gap in published studies on the relationship between road characteristics and RTA severity in Ethiopia. In this paper, we collected and cleaned traffic accident data, attempted to construct novel attributes, and tested a number of predictive models. The outputs of the models were presented for analysis to domain experts for feedback. The RTA is eager to continue the study to identify areas of interest that should be given resources for traffic safety. Finally, knowledge was presented in the form of rules using the PART algorithm of WEKA.

In contrast with the previously published work of the authors, which focused on driver characteristics, here we focused on the contribution that various road-related factors have on the accident severity. The results of this study could be used by the respective stakeholders to promote road safety. While the methods are simple, the results of this work could have tremendous impact on the well-being of Ethiopian civilians. The next step in the modeling will be to combine road-related factors with driver information for better predictions, and to find interactions between the different attributes. We also plan to develop a decision support tool for the Ethiopian Traffic Office.

IV. Results

All three classifiers performed similarly well with respect to the number of correctly classified instances (Table 2). Table 2: Summary of experiments conducted S.n Classification Models (classifiers) Number of correctly classified instances Accuracy in percentage 1 Decision Tree (J48) 14,666 80.221% 2 Naive Bayes 14625 79.9967% 3 K-Nearest Neighbors 14777 80.8182% The prior on the Property Loss class was approximately 75%, Slight Injury occurred approximately 10% of the time, Severe Injury occurred 8% of the time, and very few accidents were Fatal. Compared our prior (on Property Loss), we perform better than without having a model with respect to accuracy. However, accuracy alone does not completely describe the prediction efficiency, and other means of evaluating our predictive models are necessary. The receiver operating characteristics (ROC) curve, also known as the relative operating characteristic curve, is a comparison of two operating characteristics as the criterion changes. It can be represented by plotting the fraction of true positives (TPR = true positive rate) versus the fraction of false positives (FPR = false positive rate). An ROC analysis provides tools to select possibly optimal models and to discard suboptimal ones independent from (and prior to specifying) the cost context or class distribution. The ROC analysis is directly and naturally related to the cost/benefit analysis of diagnostic decision making. The area under the ROC curve (AUC) quantifies the overall discriminative ability of a test. An entirely random test (i.e., no better at identifying true positives than flipping a coin) has an AUC of 0.5, while a perfect test (i.e., one with zero false positives or negatives) has an AUC of 1.00. Since the accuracies of the above models were almost identical, we used ROC curves to further evaluate our models, using 20% (3,657) of the instance data. Some of the visualizations of the threshold curves are presented below, followed by a summary of the AUCs for each class value of the target class for each model. Table 3: Summary of the AUCs S.n Classification model (classifiers) Class values AUCs Property Loss 0.699 Slight Injury 0.608 Fatal 0.815 1 Decision Tree (J48) Severe Injury 0.736 Property Loss 0.752 Slight Injury 0.680 Fatal 0.855 2 Naive Bayes Severe Injury 0.761 Property Loss 0.884 Slight Injury 0.875 Fatal 0.965 3 K-Nearest Neighbors Severe Injury 0.918 In all cases, the AUCs were significantly >0.5, with the K-nearest neighbors model displaying AUCs closest to 1. These results indicate that all models predicted new instances well.

A predictive model is useless if it cannot represent knowledge in a way that end users can understand. Many learning techniques look for structural descriptions ("rules") of what is learned, which can become fairly complex. These descriptions are easily understood by the end user, and explain the bases for new predictions. Classification rules are a popular alternative to decision trees in representing the structures that learning methods produce, partly because each rule seems to represent an independent “nugget” of knowledge (Witten and Frank 2000). The antecedent (or precondition) of a rule is a series of tests, similar to those at decision tree nodes. The consequent (or conclusion) defines the class(es) that apply to instances covered by that rule, or perhaps provides a probability distribution over the class(es).

Acknowledgment

We acknowledge the technical support of SSMAP Chass, Ahmednagar

References

Abdalla, I. M., R. Robert, et al. (1997). "An investigation into the relationships between area social characteristics and road accident casualties." Accidents Analysis and Preventions 5: 583-593.

18 Beshah, T. (2005). Application of data mining technology to support RTA severity analysis at Addis Ababa traffic office. Addis Ababa, Addis Ababa University.

Beshah, T., A. Abraham, et al. (2005). "Rule Mining and Classification of Road Traffic Accidents Using Adaptive Regression Trees." Journal of Simulation 6(10-11).

Chang, L. and W. Chen (2005). "Data mining of treebased models to analyze freeway accident frequency." Journal of Safety Research 36: 365- 375.

Chang, L. and H. Wang (2006). "Analysis of traffic injury severity: An application of non-parametric classification tree techniques Accident analysis and prevention " Accident analysis and prevention 38(5): 1019-1027.

Chong, M., A. A., et al. (2005). "Traffic Accident Analysis Using Machine learning Paradigms." Informatica 29(1).

Getnet, M. (2009). Applying data mining with decision tree and rule induction techniques to identify determinant factors of drivers and vehicles in support of reducing and controlling road traffic accidents: the case of Addis Ababa city. Addis Ababa Addis Ababa University.

Kononov, J. and B. N. Janson (2002). "Diagnostic Methodology for the Detection of Safety Problems at Intersections " Journal Transportation Research Record 1784(2002): 51- 56

Mussone, L., A. Ferrari, et al. (1999). "An analysis of urban collisions using an artificial intelligence model." Accident Analysis andPrevention 31: 705-718.

Ng, K. S., W. T. Hung, et al. (2002). "An algorithm for assessing the risk of traffic accidents." Journal of Safety Research 33: 387-410.

Ossenbruggen, P. J., J. Pendharkar, et al. (2001). "Roadway safety in rural and small urbanized areas." Accidents Analysis and Prevention 33(4): 485-498.

Sohn, S. and S. Hyungwon (2001). "Pattern recognition for a road traffic accident severity in Korea." Ergonomics 44(1): 101-117.

Sohn, S. and S. Lee (2002). "Data fusion, ensemble and clustering to improve the classification accuracy for the severity of road traffic accidents in Korea. ." Safety Science 41(1): 1-14.

Srisuriyachai, S. (2007). Analysis of road traffic accidents in Nakhon Pathom province of Bangkok using data mining. Graduate Studies. Bangkok, Mahidol University..