

Perusal of Road Accident Parameters and Forecasting of Accident Types for New Roads

¹Suraj D, ²Sandeep Kumar S

¹ M. Tech student, ² Associate Professor

¹ Computer Science and Engineering,

¹ Cambridge Institute of Technology, Bangalore, Karnataka

Abstract: Road accidents is one of the dangerous events that results in major loss to human life. Road accidents can cause injuries for life time. There is a sudden rise in the number of accidents from past few years. So it is an important issue for the road traffic safety authorities. Road traffic data must be carefully analyzed in order to determine the relation between various parameters of road accidents. This paper describes perusal of factors influencing road accidents for existing roads and proposes the design and implementation of a technique for forecasting the accident types for new roads. Data mining techniques and tools are used for the analyses and prediction. Using this technique, road safety management authorities can take up several measures to bring down the number of accidents.

IndexTerms - Data Mining, Road accidents, Association rule mining, Naïve based classifier, Analyze factors, Predict accident types.

I. INTRODUCTION

Currently in the world, road traffic safety is one of the major concerns. According to CNBC, the number of vehicles on road worldwide will be 1.7 billion. The growth in the number of vehicles very alarming and this indirectly results in the increased number of accidents. An effective help in this situation is carrying out a careful examination of the road traffic data and find the hidden relation between the various parameters of accidents. Factors like weather conditions and road conditions must be considered.

Studies so far have shown that many type of accidents are due to inappropriate interpretation of the road accident parameters that resulted in improper design of roads and road safety mechanisms. Finding out the causes of road accidents helps to reduce and mitigate the damage caused by them. By using the results of analyses proposed in this paper traffic departments can rectify the improper safety measures. Also by using the forecasting of accident types for new roads, traffic authorities can design the new roads in a better way so as to reduce the costs of later correcting them.

Data mining is an approach for identifying important trends and patterns in huge datasets. The knowledge gained by this can be used to construct models that are helpful in future and also to predict target attribute based on the past data. Data mining is used to automatically discovery patterns, predict likely outcomes, creating actionable information, focus on large data sets and databases. Once the project objectives and requirements are finalized, it is formulated as data mining problem and preliminary implementation plan is developed for it.

For example, the business goal might be: "How to sell more product to customers?" This can be formulated as data mining problem as: "Which customers are likely to purchase my product?" A model that predicts this must be built on data that describes the customers who have purchased the product in the past. Collected data must contain relationships between customers who have purchased the product and customers who have not purchased the product. Customer properties might include age, locality, and so on.

Association rules are used for identifying hidden relations among the variables in large data sets. Association rules are represented in the form of rules or implications. Support and confidence are used to identify hidden relationship among the items in the large data sets. Apriori algorithm is the most popular data mining approach for finding frequent item sets and discovering patterns. We adopt this approach to discover the relationship among the various factors of road accidents.

Classification is a data mining function that assigns items in a collection to target categories or classes. For example, a classification model could be used to identify loan applicants as low, medium, or high credit risks. Classification takes place in two steps: Building the classifier based on training data; Using this classifier for classifying the query instances. Classifiers are tested by comparing the forecasted values with the target values training examples. Past data is divided into two sets: One for building the classifier; the other for testing the classifier. Rating a classifier causes class assignments and probabilities for each case.

Characteristics of classification algorithms include: Learning method, Tasks handled, Expressive power, Training time, Testing time, classifier comprehensibility. Naïve based classifier is probabilistic classifier that allows us to use the statistical properties of the training data to predict the likely class for a query instance. It is naïve in that it assumes that the values for each attribute are independent. Naïve based classifier is quite powerful algorithm, can express complex concepts, and is fast to generate and to classify new examples. We use Naïve based classifier to forecast the accident types for new roads.

Proposed System

Road safety is an important part of our lives, so it is necessary to continuously improve it with all possible and available opportunities and resources. Descriptive or predictive mining models applied on past data about accidents along with other

important parameters such as weather and road conditions creates an interesting alternative with potentially useful results for road traffic departments.

Proposed system describes how data about road accidents is used to extract frequently occurring patterns and important factors causing different types of accidents. Proposed system discovers the associations among factors of road accidents. Also, it predicts the type of accidents for new roads. Proposed system is a real time application that is useful for government sector as well as public

Identification of relation between parameters of road accidents for existing roads is done using apriori algorithm and association rule mining technique. Prediction of accidents for new roads is done using Naïve Based algorithm and Classification technique.

II. IMPLEMENTATION

The task of finding relation between factors influencing road accidents is treated as a data mining problem where association rule mining is used to find hidden patterns and affinity among various parameters of road accidents. Association rule mining extracts interesting rules between various attributes in a large accident data set. It generates various set of rules that define the relation among different attributes in the accident data. Support and confidence are used to identify strong rules. A support value suggests the frequency of occurrence of a rule in the data set. On the other hand, confidence defines the reliability of a rule. Hence, association rules with high confidence and support value are of main interest in the analyses. Apriori algorithm is used to generate association rules in our project.

The task of predicting the possible accident type for new roads is treated as data mining problem where classification rules are used. Classification is a data mining technique that assigns items in a collection to classes that are predefined. A classification task begins with a training data set for accidents in which the accident types are known. Our classification is multiclass target as there are multiple accident types possible. In the model building process, classification algorithm finds relationships between the values of the factors influencing accidents like speed of the vehicle, weather, pedestrian distance, etc. and the type of accident like side impact, head on collision, etc. The relationships are conglomerated in a model, then it is used to apply to a different accident data set in which the type of accidents are unknown. Naïve Bayes algorithm is used in for finding relationships. Classification model is then tested by comparing predicted accident types and the known accident types in the test data. The historical data is divided into two sets: one for building the model and one for testing the model.

3.1 Architecture

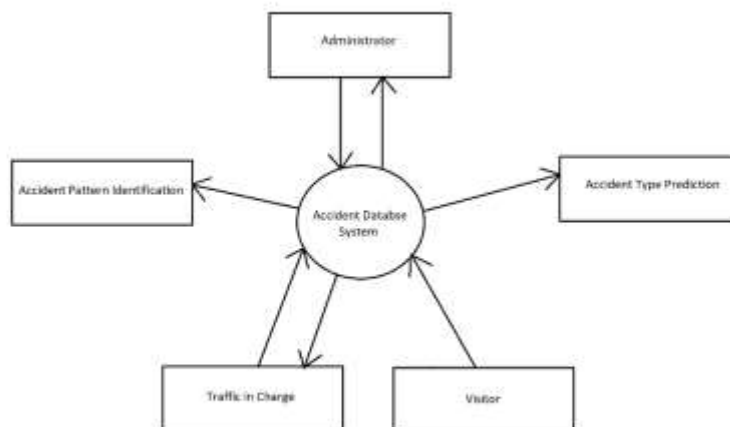


Figure 1: System Architecture

3.2 Administrator Module

Administrator manages the cities by uploading new cities, deleting the cities that are not applicable and updating the required cities. Administrator registers the City Traffic in charge and can add, delete and update the existing traffic in charge. City Traffic in Charge for a city is assigned by the admin. Admin sets the credentials for the City Traffic in Charge. Administrator adds, deletes and updates the parameters. Admin updates the values of these parameters.

3.3 Traffic in Charge

Traffic in Charge logs into the application providing the credentials set by the admin. Traffic in charger adds the road names in city and can edit, update and delete the existing road names and can view the existing road names. Traffic in Charge uploads the road accident details into the server. Based on this uploaded data, system predicts the patterns between road accidents.

3.3 Visitor Module

Visitor must be able to view the home page, contact us page. Visitor views the relation between the factors influencing accidents and the possible accident types for new roads.

3.4 Accident Pattern Identification Module

In this module, system uses “Association Rule mining” to analyze previous accidents data and to extract the accidents patterns. Patterns can be identified based on road wise, city wise, date wise, month wise etc. Apriori is a seminal algorithm for finding frequent itemsets using candidate generation. It is characterized as a level-wise complete search algorithm using anti-monotonicity of itemsets, “if an itemset is not frequent, any of its superset is never frequent”.

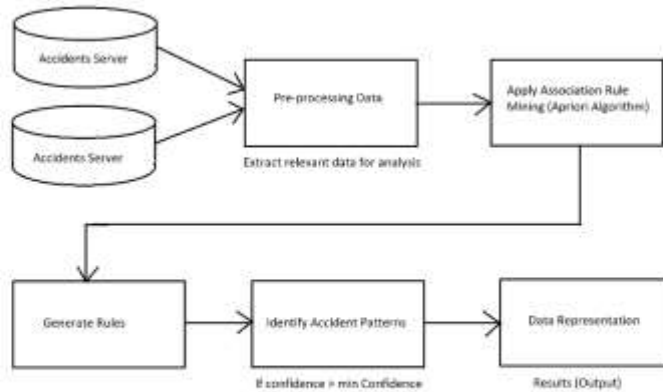


Figure 2: Accident Pattern Identification

3.4 Accident Type Prediction Module

In this module, system predicts the more possible accident type for new roads based on the past road accidents data. Here we make use of “Classification Rules” for the prediction. Given a set of training instances characterized by a set of attributes and associated with a class, Naïve based algorithm constructs a rule which will allow us to assign query instance to a class, given only the set of variables describing the query instance. In our case the classes represent the possible accident types for new roads.

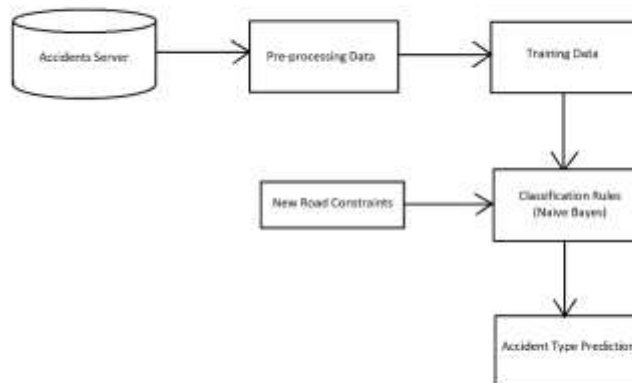


Figure 3: Accident Type Prediction

III. RESULTS AND DISCUSSION

The below figure shows the output of apriori algorithm used in identifying the relation between factors influencing road accidents.

Patterns

Road: kuvempunagar

Start Date: 1/1/2018

End Date: 04/27/2018

Generate Patterns

Accidents Dataset Found

Dataset (Road Accidents)!!!

Sl. No	AccidentId	Road Name	Description	Registered Date	Accident Type
1.	13	kuvempunagar	death	1/25/2018	hit-run
2.	15	kuvempunagar	major injuries	1/25/2018	collision
3.	16	kuvempunagar	fault of driver	1/25/2018	single car accident
4.	17	kuvempunagar	due to rain fall	3/10/2018	hit-run

Generated Patterns !!!

LHS		RHS	Confidence
collision	→	hit-run	100.00%
collision,sporting accident	→	hit-run	100.00%
hit-run	→	collision	100.00%
hit-run,sporting accident	→	collision	100.00%
Inexperience	→	sporting accident	100.00%

The below output shows the output of naïve based algorithm for predicting the accident type for new roads.

Accident Type Prediction (Naive Bayes)!!!

Enter New Road: Mb Road

Set Constraints for Road and Click Predict Button

Sl.No	Constraint Name	Value
1.	Speed	40KM
2.	WeatherCondition	Fog
3.	SchoolZone	Yes
4.	Humps	rumble strips
5.	hospital zone	35km
6.	age	20
7.	month	january
8.	gender	male
9.	camera	Yes
10.	vehicle type	2 wheeler
11.	vehicle no	kA 09 m 1964
12.	Number of riders	2
13.	injury	fracture
14.	accident index	Value Notset
15.	police force	No
16.	no of vehicles	3 vehicles
17.	no of casualties	3
18.	date	28/1/2018

19. day of week	sunday
20. time	7:30pm
21. highway no	NH 57
22. road type	mud road
23. junction details	4 junctions
24. light conditions	dim
25. spot lighting	yes
26. accident location	double road
27. age of vehicle	15 yrs.
28. driving experience	5
29. Nationality(A)	indian
30. station name	ashoka puram station
31. right hand pin	60
32. falling rocks	moderate
33. men at work	Yes
34. straight prohibited	Yes

Predict Accident (NB)

Road Name: Mb Road
Accident Type: drink n drive

IV. CONCLUSIONS

Interesting relationship between the factors of road accident can be found using Association Rule mining technique. As we have used a comprehensive dataset related to accident data, the results of the analysis is commendable. Also, possible accident types are predicted for new roads using classification. As we have used the Naïve based algorithm for the prediction of accident, the result is reliable.

The above analysis and prediction can be used by traffic authorities to take necessary actions to reduce the number of accidents on road. Results have potential and wide scope for application of data processing and analytical methods in the field of road safety. The discovered new knowledge can be used for proposing specific actions for concrete roads; effective police checks on critical road points; building safe roads; effectively targeted information campaigns reducing the number of road accidents.

REFERENCES

- [1] František Babi, Karin Zuskáová, "Descriptive and Predictive Mining on Road Accidents Data", in SAMI, 2016.
- [2] Gagandeep Kaur, Harpreet Kaur, "Prediction of the cause of accident and accident prone location on roads using data mining techniques", in ICCNT, 2017.
- [3] Liling Li, Sharad Shrestha, Gongzhu Hu: "Analysis of road traffic fatal accidents using data mining techniques", in SERA, 2017.
- [4] Rishi Sai Reddy Sudireddy, Uttam Mande: "Prediction of Road Accidents Using Correlation Based on Map Reducing", in CICN, 2016.
- [5] Nikhat Ikram, Shilpa Mahajan: "Road accidents: Overview of its causes, avoidance scheme and a new proposed technique for avoidance", in INDIACom, 2017.
- [6] Suwarna Gothane; M. V. Sarode: "Analyzing Factors, Construction of Dataset, Estimating Importance of Factor, and Generation of Association Rules for Indian Road Accident", in IACC 2017.
- [7] Sachin Kumar, Durga Toshniwal: "Analysing road accident data using association rule mining", in ICCCS, 2015