A REAL TIME APPLICATION TO REDUCE TRAFFIC ACCIDENTS

Abhishek R Dalawai, Ashish Kumar Mahto, Dastageer Khan S, Dr. A B Rajendra

[1] *Abhishek R Dalawai, Student, Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru , Karnataka, India*
[2] *Ashish Kumar Mahto, Student, Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India*
[3] *Dastageer Khan S, Student, Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India*
[4] *Dr. A B Rajendra, Professor & HOD, Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India*

ABSTRACT

In this paper, we have examined the findings of studies on the elements that influence road accidents. As a result, it is a major worry for road safety departments as well as individuals. To find the factors that are closely associated to road accidents, road traffic data must be extensively evaluated. Factors such as the sort of collision, the road condition, the light effect, the weather, and the presence of a drunk driver must all be taken into account. Accidents are most common in locations near residential areas, zebra crossings, and school zones. It is required to mine the connection rules between significant risk factors from the statistics on these incidents in order to discover the major causes of these accidents.

This work extends the Apriori or Elcat algorithm to mine the association rules between risk factors and probes deep into the causes of traffic accidents on urban roads, taking into account the numerous layers and dimensions of accident data. The findings of the study enable the traffic department to develop appropriate accident prevention measures and increase traffic safety on city streets.

Keyword: - *traffic accidents, Apriori algorithm, factors, urban roads, causes of traffic accidents, precautions to be taken.*

1. INTRODUCTION

There are currently 1.4 billion vehicles on the road in the whole world, with a projected increase to 2.5 billion by 2040. The worrisome increase in vehicle numbers has resulted in an increase in the number of accidents day by day. Road accidents claim the lives of over 1.4 million people every year.

Accidents are something that both traffic departments and residents seek to avoid. The road traffic information is enormous, and data mining approaches are being employed to uncover the Accidents are something that both traffic departments and residents seek to avoid. The road traffic information is enormous, and data mining approaches are being employed to uncover the relationships in it. The accuracy of data, how effectively records are kept, and how well data is evaluated all play a role in the effective use of accident data.

In a big data set, association rules aid in the discovery of linkages such as patterns, correlations, and incidental structures between seemingly unrelated data.

The study's significance lies in its examination of traffic accident data factors on State Highways (SH) and Ordinary District Roads (ODRs). Recognizing the gravity of the situation, remedial efforts are made to reduce the danger posed by road accidents.

The goal of the project is to use data mining techniques to examine unintentional data from many types of roads, including plan roads, link roads, and central government-funded roads (PMGSY Roads). Data mining is a technique for extracting intriguing patterns and illustrated, understandable models from large databases. The knowledge-driven version is concerned with data harvesting and analysis, demand- driven collection of information sources, security and privacy considerations, and modelling based on user interests. Data dredging is used on data from the transportation and construction industries.

Data mining is an important tool for detecting suspicious activity. Road crashes are to blame for our country's higher severity levels.

However, data analysis on the statistical distribution level, let alone mining out the main accident chains from accident data, makes it impossible to detect the correlations between numerous risk variables of traffic accidents.

Due to the number of increases in the traffic accidents on urban areas/highways the traditional Apriori algorithm was upgraded from the perspectives of data structure measuring indices, and subjective constraints to address these flaws. The R programming language was used to optimize the algorithm flow. The association rules were mined in four steps in the modified algorithm: data processing, multidimensional data modelling, algorithm mining, and rule interpretation. Then, using multidimensional data, the improved Apriori algorithm was used to identify the primary causes of traffic accidents on metropolitan roadways.

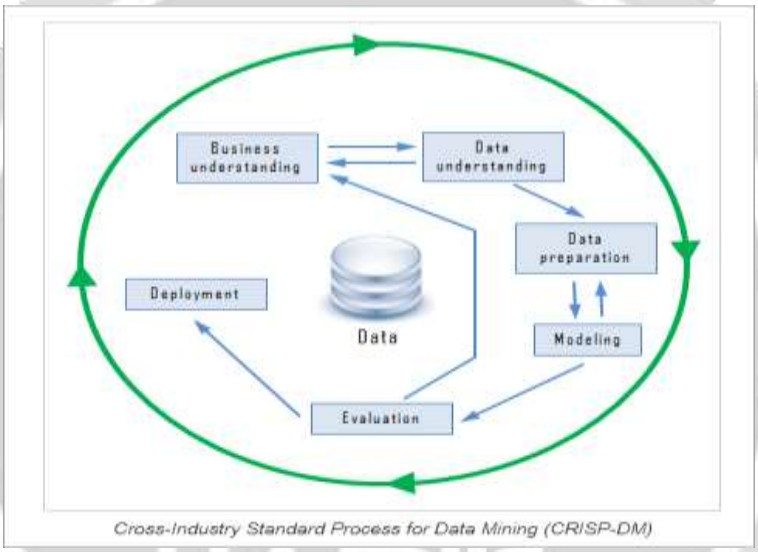


Fig: Data Mining(process) Technique

The main objectives of the paper are:

- Proposed system is an online browser based application.
- It is an real time application where we can access the application from different location.
- The major objective of the system is to reduce the traffic accidents.

It make uses of the factors such as speed limit, weather condition, vehicle condition, location and other parameters to reduce the traffic accidents.

2. LITERATURE SURVEY

Below Table shows the comparative study of

Sl.No	Research Paper	Focus	Limitation
1	TIBEBE BESHAH, SHAWNDRA HILL: Mining Road Traffic Accident Data to Improve Safety: Role of Road-related Factors on Accident Severity, IEEE, 2018	Find correlations and co- occurrences between data sets and help to show the probability of relationships between data items, within large data sets. It identifies frequent if-then associations.	Aims at injury severity prediction Takes more time for prediction. Not implemented as real time application
2	PRIYANKA A. NANDURGE, NAGARAJ V, DHARWADKAR: Analyzing Road Accident Data using Machine Learning Paradigms, IEEE, 2017	Each cluster is associated with a centroid. The main aim algorithm isto minimize the sum of distances between the data point and their corresponding clusters. Combined with association rule.	Only clustering done. Does not predicts the traffic accident patterns.
3	PRAJAKTA S. KASBE, APEKSHA V. SAKHARE: A Review on Road Accident Data Analysis Using Data Mining Techniques, IEEE, 2017	Goal of the SVM algorithm is to create the best line or decision boundary that can segregate n- dimensional space into classes so that we can easily put the new datapoint in the correct category in thefuture.	Using tools like Wekaatool the results can be easily obtained but the testing of these is not possible. Small Data-set used forprediction. Less accurate results.
4	MS. GAGANDEEP KAUR, ER. HARPREET KAUR: Prediction of the cause of accident- and accident- prone location on roads using data mining techniques, IEEE, 2017	RStudio allows users to develop and edit programs in R by supporting a large number of statistical packages, higher qualitygraphics, and the ability to manageyour workspace	Data Mining used. Huge data required Tools used for prediction Not implemented as realtime application.
5	SURAJ D, SANDEEP KUMAR S: A Survey on Analyses of FactorsRelated to Road Accidents Using Data Mining Techniques, 2018	Objective of discovering patterns/trends/groupings amonglarge sets of data and transforming data into more refined information	Uses data miningtechniques. Huge data required. Less number of acciden types used.

6	EMI JOHNSON, JUBY MARY ABRAHAM, SAMEERA SULAIMAN, PADMA SURESH L, DEEPA RAJAN S: Study on road accidents using data mining technology, IEEE, 2018	The data set used for implementation is only static data available on the UCI Machine Learning Repository. Data Mining tools used, discovering patterns/trends/groupings and transforming data into more refined information.	Uses data mining techniques. Huge data required. More time required for prediction.
7	DAGHAN DOGAN, SETA BOGOSYAN: Performance Analysis of SVM, ANN and KNN Methods for Acoustic Road-Type Classification, IEEE, 2019	Stores all the available data and classifies a new data point based on the similarity. ANNs for nonlinear systems dynamic modeling and identification with SVM	Concept used to classify road types Can't predict traffic accidents patterns Less accurate results
8	ZHENGAO, RUIFENGPAN, XUESONGWANG, RONGJIE YU: Research on Automated Modeling Algorithm Using Association Rules for Traffic Accidents, IEEE, 2018	Finding correlations and co-occurrences between data sets and help to show the probability of relationships between data items, within large data sets and identifying frequent if-then associations.	Less parameters used. Takes more time for prediction. Not implemented as real time application
9	K. GEETHA, C. VAISHNAVI: Analysis on Traffic Accident Injury Level Using Classification, IEEE, 2015	Discovering patterns/trends/groupings andtransforming data into more refined information.	Huge data required. Only classification is done. Aims to find the injuryseverity.

3. METHODOLOGY

Step 1: Data Collection

We are working on a real-time application, and we are developing a new application that includes data servers (used to store data). Data collection entails gathering information from various sources. Year, Speed Limit, Weather Conditions, School Zone, Humps, Hospital Zone, Road Type, Men at Work, Accident Types are all included in the data.

Step 2: Data Preparation

Data from servers is extracted and analyzed here. Complete data extraction and analysis in which we remove irrelevant data while retaining data needed for processing. According to the project, only certain types of accidents are required to generate outputs.

Step 3: Specify Constraints

SUPPORT COUNT

The proportion of transactions containing that item (A) compared to the total number of transactions in the data set.

CONFIDENCE

The confidence of an item set is defined as the ratio of the total number of transactions containing the item set to the total number of transactions containing the LHS.

Step 4: Association Rules Mining (Eclat Algorithm)

The most well-known and straightforward data mining technique is probably association (or relation). To identify patterns, we make a simple correlation between two or more items, often of the same type.
For example, using market-basket analysis, we might discover that a customer always buys cream when they buy strawberries, and thus suggest that the next time they buy strawberries, they might also want to buy cream. To process e-commerce data and identify patterns, we employ the eclat algorithm or the apriori algorithm. In this section, we generate patterns related to traffic accidents.

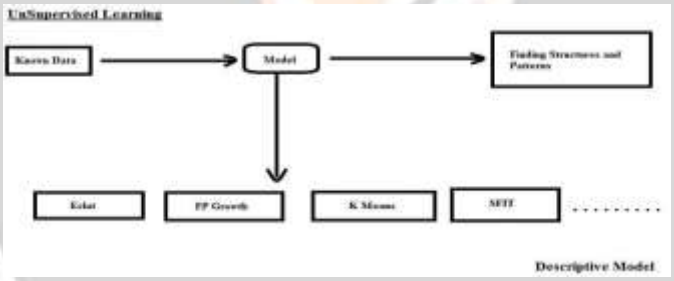
Step 5: Patterns Prediction

The system predicts the relationship between the types of frequent traffic accidents in this case. Machine learning is the study of a system using data. Machine learning is a subset of data science in which machine learning algorithms are used to process data.

Unsupervised Learning

A descriptive model is used for tasks that would benefit from the knowledge gained from summarizing data in novel and interesting ways. In the unsupervised learning technique, there are no predefined labels. The goal is to investigate the data and discover some structure within it. On transactional data, unsupervised learning works well.

Clustering and association learning techniques were used to create a descriptive model. We have many efficient algorithms, such as the "eclat algorithm," the "AIT algorithm," the "SFIT algorithm," the "STEM Algorithm," the "FP Growth algorithm," the "K Means algorithm," the "Fuzzy C Means algorithm," and so on.



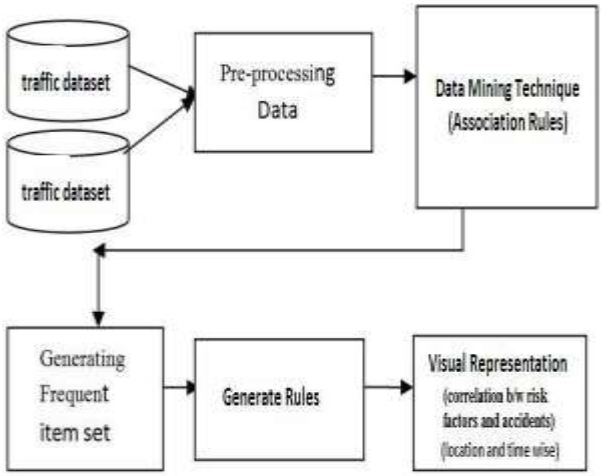
In the project, we use the "eclat algorithm or Apriori" to discover the relationship between different types of traffic accidents. The Eclat algorithm is one of the most efficient and takes the least amount of time to process data. This algorithm works well for both small and large data sets.

Step 6: Results

Generated patterns/trends (visualizing relationship b/w different accident types)

Step 7: Visual Representation

The final patterns are displayed on the GUI for the users. When a user logs in to the application, the system displays outputs on a graphical user interface (GUI).



Apriori or Eclat algorithm is selected because of the following advantages.

1. Faster Results (takes less time for Prediction) and it is efficient.
2. It works better for small data set as well as huge data set.
3. Scanning the whole database once is enough.
4. It works better for multiple constraints.

4. CONCLUSION

We discussed the existing techniques for reducing traffic accidents in this paper. We also briefed on some methods in the process of reducing accidents that have previously been attempted by some researchers. We saw various Data Mining techniques used in various papers, and after analyzing all of these papers, we concluded that the apriori or eclat algorithm is best suited for analyzing data related to traffic accidents. Because of the advantages it has over other algorithms. After analyzing all of the data and discovering frequent patterns, as well as the causes of traffic accidents. We provide precautions related to specific traffic accidents, which will be expanded by government authorities in intended to facilitate people from these accidents.

5. REFERENCES

1) "Mining Road Traffic Accident Data to Improve Safety: Role of Road-related Factors on Accident Severity" by TIBEBE BESHAH, SHAWNDRA HILL, IEEE, 2018

2) "Analyzing Road Accident Data using Machine Learning Paradigms", by PRIYANKA A. NANDURGE, NAGARAJ V, DHARWADKAR, IEEE, 2017

3) "A Review on Road Accident Data Analysis Using Data Mining Techniques", by PRAJAKTA S. KASBE, APEKSHA V. SAKHARE, IEEE, 2017

4) "Prediction of the cause of accident and accident-prone location on roads using data mining techniques", by MS. GAGANDEEP KAUR, ER. HARPREET KAUR, IEEE, 2017

5) "A survey on analysis of factors related to road accidents using data mining techniques", by SURAJ D, SANDEEP KUMAR S, IEEE, 2018

6) "Study on road accidents using data mining technology", by EMI JOHNSON, JUBY MARY ABRAHAM, SAMEERA SULAIMAN, PADMA SURESH L, DEEPA RAJAN S, IEEE, 2018

7) "Performance Analysis of SVM, ANN and KNN Methods for Acoustic Road-Type Classification", by DAGHAN DOGAN, SETA BOGOSYAN, IEEE, 2019

8) "Research on Automated Modeling Algorithm Using Association Rules for Traffic Accidents", by ZHEN GAO, RUIFENG PAN, XUESONG WANG, RONGJIE YU, IEEE, 2018

9) "Analysis on Traffic Accidents Injury Level Using Classification", by K. GEETHA, C. VAISHNAVI, IEEE, 2015

10) "Analysis of Injury Severity in Pedestrian Crashes using classification Regression Trees", by VICHIKA IRAGAVARAPU, DOMINIQUE LORD, IEEE, 2018