A Machine Learning Approach for Prediction of Fatal Traffic Accidents

1Mrs. Soumya Dath G, 2Bhavana M B, 3Sahana R, 4Sharmila J
1Assistant Professor, 2,3,4Student
Department of Information Science and Engineering
GSSS Institute of Engineering and Technology for Women, Mysuru, Karnataka, India.

Abstract: The objective of this paper is to provide real time application which is an efficient tool for accident prediction. The system identifies relationships and possible hidden links between various factors that describe road accidents with fatal consequences. Using Association learning technique, the Eclat algorithm predicts the frequently occurring accidents in future and the relationship between different traffic accidents using the traffic dataset.

Index Terms: Machine Learning, Association Learning Technique, Eclat algorithm.

I. INTRODUCTION
With the advent of vehicle technology and road infrastructural development, the mobility of people from place to place increased in an exponential way. There are different kinds of transport systems that help people to do business or to enjoy or to study moving to different places. One of the most commonly used transport systems is a vehicle transport system. It is also the main cause of road traffic accidents in the world. Now, a number of road accidents are rising rapidly, leading to an increase in fatalities. Discovering the associations among the traffic accidents is the key factor in reducing the traffic accidents. In today's world, it is very difficult to keep track of road accidents and the causes of accidents. The current system is manual, where the public sector uses the general accounting data and analyzes the data manually, based on the analysis will take preventive measures to reduce the number of accidents. The system can be implemented as an application in real time. This theme can be implemented as government sector application. The system can be used by the public to find out accident patterns and the type of accident in a given area or city. The system identifies relationships and possible hidden links between various factors that describe road accidents with fatal consequences. Likewise, similar knowledge in the case of direct cooperation with all major stakeholders, such as the police, states and local authorities, can contribute to improving road safety.

II. RELATED WORKS
Zhen Gao, Ruifeng Pan, Xuesong Wang, Rongjie Yu [1] a method of computing minimum Support based on interested frequent item sets was proposed, and two powerful methods of automatic acquiring strong rules were proposed. Finally, designed an automatic modeling algorithm using association rules, which can promote the application of association rule mining in the existing intelligent transportation system.

Emi Johnson, Juby Mary Abraham and Sameera Sulaiman [2] the traffic accidents were predicted using data mining technology called Naïve Bayes technique is used which is proved supporting to resolve traffic accident severity problem and conclude which one could be optimal technique in road traffic accident scenario.

Suraj D and Sandeep Kumar [3] an analysis made with respect to factors influencing road accidents. Road accident causes major damages to humans. It can cause lifetime injuries. There is a spike in the number of accidents over recent years. This paper makes comparative study on the analysis done so far in this regard.

Liling Li, Sharad Shrestha and Gongzhu Hu [4] the statistics analysis and data mining algorithms on the FARS Fatal Accident dataset as an attempt to address this problem. The relationship between fatal rate and other attributes including collision manner, weather, surface condition, light condition, and drunk driver were investigated.

Pragath Tiwari, Sachin Kumar and Denis Kalitki [5] performed the classification of road accident on the basis of road user category. We have used Self Organizing map (SOM), K-modes clustering technique to group the data into homogeneous segments and then applied Support vector machine (SVM), Naive Bayes (NB) and Decision tree to classify the data.

III. METHODOLOGY
The Association (or relation) is probably the better known and most familiar and straightforward data science technique. A simple correlation between two or more items, often of the same type to identify patterns. For example, Market-basket analysis, where we track people's buying habits, we might identify that a customer always buys cream when they buy strawberries, and therefore, suggest that the next time that they buy strawberries they might also want to buy cream. The Association rule mining “Eclat Algorithm” is used to predict the relationship between different traffic accidents using the traffic dataset.

The steps in the Eclat algorithm are as follows:
1. Get TID list for each item (DB scan)
2. TID list of {a} is exactly the list of transactions containing {a}
3. Intersect TID list of {a} with the TID lists of all other items, resulting in TID lists of {a, b}, {a, c}, {a, d} ...={a}-conditional database (if {a} removed)
4. Repeat from 1 on {a}-conditional database
5. Repeat for all other items
The basic idea is to use Transaction Id Sets (tidsets) intersections to compute the support value of a candidate and avoiding the generation of subsets which do not exist in the prefix tree. In the first call of the function, all single items are used along with their tidsets. Then the function is called recursively and in each recursive call, each item-tidset pair is verified and combined with other item-tidset pairs. This process is continued until no candidate item-tidset pairs can be combined.

IV. PROPOSED WORK

The proposed system is a real time application which is useful for government sector to reduce the number of traffic accidents. The system describes one possibility of how to use the collected data about traffic accidents to mine frequent patterns and important factors causing different types of accidents. It discovers strong rules hidden in these frequent item sets, which often reveal the link between factors affecting accidents and that can be used to reduce the number of accidents by breaking them.

System discovers the associations among traffic accidents. System discovers strong rules hidden in these frequent item sets often uncover the association between influencing factors of accidents, which can be used to reduce the occurrence of accidents by breaking them. System uses traffic accidents data to mine frequent patterns and important factors causing different types of accidents.

V. SYSTEM DESIGN

**Fig 1: System Architecture Diagram**

Prediction is a statement about future events. Price Prediction for agricultural commodity has become the need of the hour for farmers. Although future events are uncertain, so accurate prediction is not possible. This includes a decision-making support model that can be helpful for farmers to predict prices. This model includes a portal in which farmers are required to login with the credentials (username and password) which can be their name and mobile number as it is easy for them to remember. As shown in fig, after login the dashboard appears. Farmers have to enter commodity name and previous selling price of the crop. Based on the previous prices, this model will be able to provide average prices for a particular crop which will be beneficial for farmers to make better decisions and predict prices. The raw data & weather statistics information will be extracted and transformed into understandable format during data preprocessing step. The train data is then trained using Supervised learning and KNN algorithm which provides data along with its year and location. The results will be thus displayed in the appropriate portal.

**Fig 2: Block Diagram**
Figure 2 depicts the block representation of the prediction of accidents. Accident data are preprocessed from the database where it filters the irrelevant data and extracts the relevant data for the analysis. The system make use of ECLAT Algorithm in association learning technique to predict future accidents.

VI. RESULTS

The ECLAT algorithm uses the dataset as shown in Fig 3 to train itself and predict the accidents based on previously occurred accidents.

Fig 3: Accident Datasets

Fig 4: Accident Pattern generated for the ECLAT Algorithm
4 illustrates the prediction of accidents for a particular road on giving the road name in the particular city, the system makes use of association learning techniques that predicts the future occurring accidents and also displays the accident with more confidence rate.

Figure 5 illustrates the accident predicted by Apriori algorithm.

Figure 6 depicts the graphical representation of the obtained accuracy between Apriori and ECLAT algorithm. This shows that the ECLAT Algorithm requires less time to generate accident patterns when compared to Apriori algorithm.

VII. CONCLUSION
Road safety represents an important part of our lives, so it is necessary to continuously improve within all possible and available opportunities and resources. Descriptive or predictive Learning applied on historical data about occurred accidents in combination with other important information as weather or road conditions creates an interesting alternative with potentially useful and helpful outcomes for all involved stakeholders. These factors motivated the creation of this project to analyse available data samples describing road accidents representing a quite large amount of data which required the use of relatively new method in-memory data processing in this domain.
REFERENCES


