IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106119

"Cause Analysis of Traffic Accidents Using Data Science"

[1]Gowthami U, [2]Harsha T G, [3]Harshini B A, [4]Spandana H D, [5]Mrs. Divya B M

[1-4]Students, Department of Information Science and Engineering, B G S Institute of Technology B G Nagara Nagamangala (Tq), Mandya(D), India

[5]Asst. professor Department of Computer Science and Engineering B G S Institute of Technology B G Nagara Nagamangala (Tq), Mandya(D), India

Abstract: Traffic accidents on city streets are the consequence of coordinated activities by numerous elements, including humans, vehicles, roadways, and the environment. It is important to mine the connection rules between significant risk factors from the statistics on these incidents in order to determine the primary causes of these accidents. This method enhances the Apriori 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 many layers and dimensions of accident data. The parameters like support, confidence, and lift were modified according to the layer and dimension of certain characteristics in order to find qualifying association rules between risk factors. The findings were then filtered to provide a set of useful association rules. The system's data allow the traffic department to develop appropriate accident-prevention strategies and improve traffic safety on city streets. The major goal of the system is to discover the link between traffic accident risk variables and accident kinds. By analysing the key accident variables, the system was built as an automation to decrease road accidents

I. INTRODUCTION

The fast growth of urban expressways not only improves transportation ease, but also creates traffic safety issues. Over 6,000 traffic incidents occurred on the Shanghai Expressway between April and June 2014, indicating that the city need a safer transportation environment. Exploring the influence of accident-causing variables and implementing efficient methods to reduce the occurrence of accidents is a pressing necessity. Scholars have been studying the influence of influencing variables on traffic accidents in recent years, with a particular focus on people, vehicles, roads, and the environment. Some researchers[1-2] investigated driver behaviour and examined the features of the approach used to change lanes in order to detect risky driving habits. There are also studies[3] on the influence of road conditions on traffic accidents, which suggest that a prime and steep roadbed will compromise traffic safety to some level. Other studies[4] looked at how weather and dynamic traffic flow affected accidents. However, the majority of those research concentrated on the influence of one component (people, vehicles, roads, and the environment) on traffic accidents. Scholars employed a variety of knowledge science techniques in road safety studies after the advent of knowledge science technologies. Among these, association rule mining was frequently employed to investigate the relationship between the elements that influence road accidents. After all, powerful association criteria are prone to uncovering the network concealed inside accident data. To persuade them, we might use the two thresholds of Support and Confidence to evaluate the importance and believability of the principles, and Lift to type the validity of the principles.

The existing of association rule mining generally determined the model parameters (such as the minimum Support, etc.) by repeated experiments. For the huge results excavated, it's necessary for the experts to screen useful rules consistent with personal expertise manually. The tactic is inefficient and therefore the subjective screening process can't be translated into an objective algorithm, so it hinders the direct application of association rule mining in intelligent transportation system. In this system, we proposed a way to calculate the minimum Support within the modeling parameters, and suggests how to extract strong rules from the huge rules by a clustering method, or automatically filtering out the weak rules supported an expert experience related method. Finally, we built up an automatic modeling algorithm using association rules which might better promote the sensible application of association rule mining in existing intelligent transportation system.

Discovering the associations among the traffic accidents and causes for accidents is the key factor in reducing the traffic accidents. Reducing traffic accidents is a challenging task in the current traffic sector. We also get many tools and software to maintain traffic accidents, these tools just collects the data stores in sever but no analysis is done. There is no automation to find the accident risk factors which increases the traffic accidents.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106119

II. RELATED WORKS

1. IEEE PAPER TITLE: A Review On Road Accident Data Analysis Using Data Mining Techniques
YEAR OF PUBLICATION: 2017
AUTHORS: Prajakta S. kasbe, Apeksha V. Sakhare.
METHODOLOGY: SVM classifier, K means used.
LIMITATIONS:

- Using tools like Wekaa tool the results can be easily obtained but the testing of these is not possible.
- Small Data-set used for prediction.
- Less accurate results.

2. IEEE PAPER TITLE: Analyzing Road Accident Data using Machine Learning Paradigms.
YEAR OF PUBLICATION: 2017
AUTHORS: Priyanka A. Nandurge,Nagaraj V. Dharwadkar.
METHODOLOGY: combined result of kmeans clustering and association rule used.
LIMITATIONS:

- Only clustering done.
- Does not predicts the traffic accident patterns.

3. IEEE PAPER TITLE: STUDY ON ROAD ACCIDENTS USING DATA MINING TECHNOLOGY.
YEAR OF PUBLICATION: 2018
AUTHORS: Emi Johnson,Juby Mary Abraham, Sameera Sulaiman, Padma Suresh L, Deepa Rajan S,.
METHODOLOGY: The data set used for implementation is only static data available on the UCI Machine Learning Repository. Data Mining tools used.
LIMITATIONS:

- Uses data mining techniques.
- Huge data required.
- More time required for prediction.

4. IEEE PAPER TITLE: Performance Analysis of SVM, ANN and KNN Methods for Acoustic Road-Type Classification
YEAR OF PUBLICATION: 2019
AUTHORS: Daghan Dogan,Seta Bogosyan.
METHODOLOGY: SVM classifier, ANN and KNN methods used.
LIMITATIONS:

- **Concept used to classify road types**
- **Cant predict traffic accidents patterns**
- **Less accurate results.**

5. IEEE PAPER TITLE: Research on Automated Modeling Algorithm Using Association Rules for Traffic Accidents.
YEAR OF PUBLICATION: 2018
AUTHORS: Zhen Gao, Ruifeng Pan, Xuesong Wang, Rongjie Yu.
METHODOLOGY: association rule used.
LIMITATIONS:

- Less parameters used.
- Takes longer for prediction.
- Not implemented as real time application.

III. PROPOSED WORK

Proposed system may be a real time application which is beneficial for state sector to scale back the amount of traffic accidents. Traffic safety represents a crucial a part of our lives, so it's necessary to continuously improve within all possible and available opportunities and resources. Descriptive or predictive mining applied on historical data about occurred accidents together with other important information as weather or traffic conditions creates a stimulating alternative with potentially useful and helpful outcomes for all involved stakeholders.
Proposed system describes one possibility of the way to use the collected data about traffic accidents to mine frequent patterns and important factors causing different types of accidents.
Proposed system uses parameters like "speed limit", "weather condition", "Humps", "men at work" etc.. to predict the correlation between risk factors and accident types.

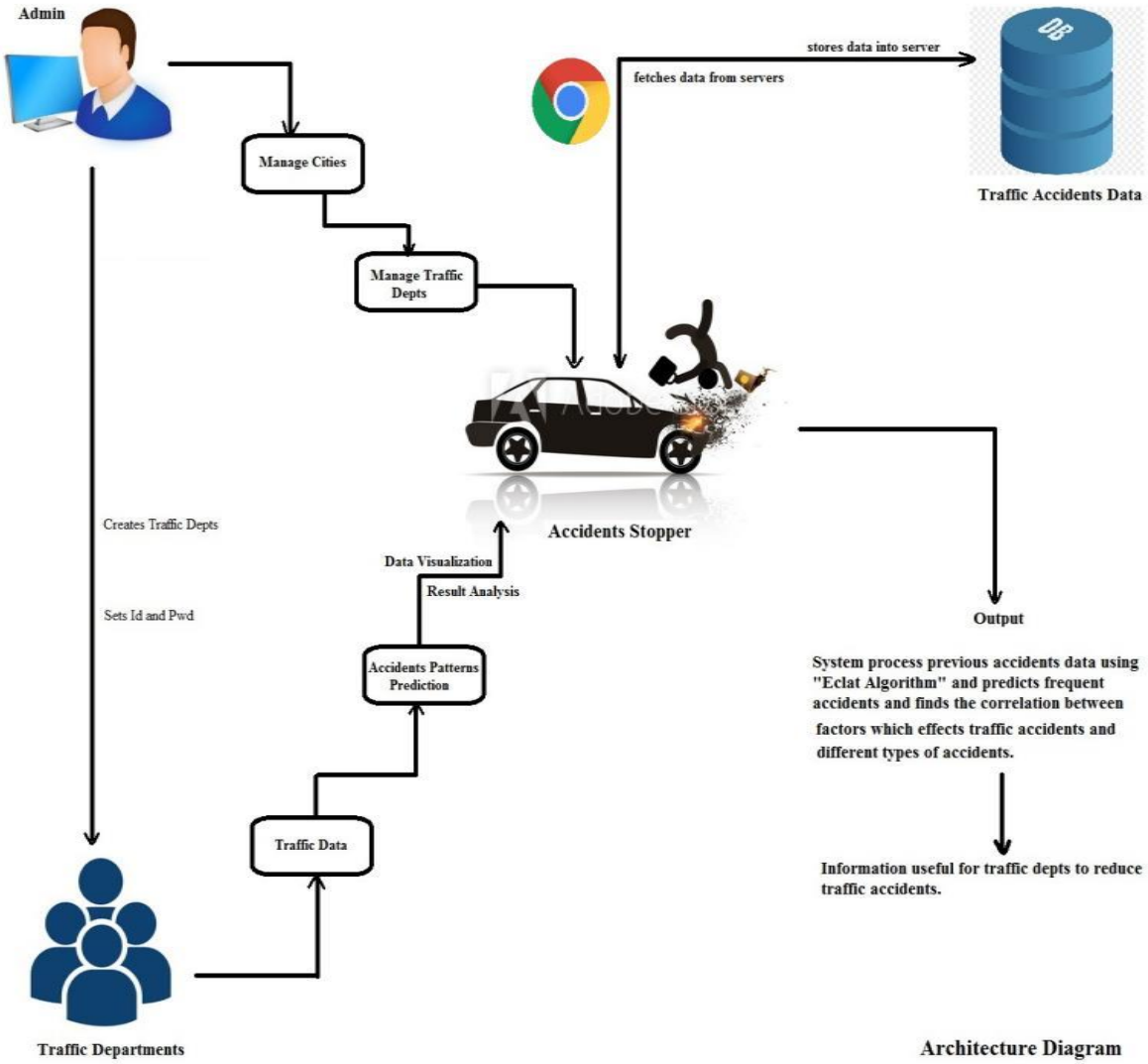
IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106119



Architecture Diagram

IV. METHODOLOGY

ML concerns with construction and study of system which will learn from data. For instance, ML are often utilized in E-mail message to find out the way to distinguish between spam and inbox messages.

There are three sorts of Machine learning(ML), they are

i. Supervised Machine Learning

Here we've labels and therefore the input is past examples.

Ex: 1-4

ii. Unsupervised Machine Learning

Extraction of patterns without labels.

Ex. 5 and 6

iii. Semi-Supervised Machine Learning

Unsupervised Learning

A Descriptive model is employed for tasks that might enjoy the insight gained from summarizing data in new and interesting ways. There are not any predefined labels in unsupervised learning technique. The goal is to explore the info and find some structure with in. Unsupervised learning works well on transactional data.

Clustering and association learning techniques were used to create a descriptive model. Many efficient algorithms are available, including the "eclat algorithm," "AIT algorithm," "SFIT algorithm," "STEM Algorithm," "FP Growth algorithm," "K Means algorithm," "Fuzzy C Means algorithm," and so on.

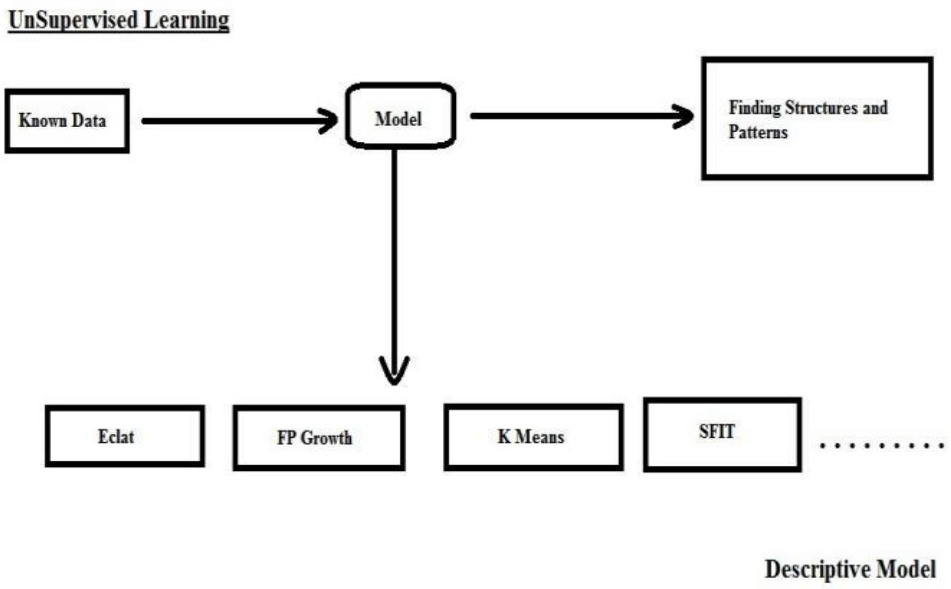
IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106119



In the project we use "Apriori algorithm" to find the relationship between traffic accidents and injuries. Apriori algorithm is one of the efficient algorithm and takes less time for data processing. This algorithm works fine for small data-sets as well as large data-sets.

Accidents Pattern Prediction Process

Step 1: Data Collection

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

Step 2: Data Preparation

Here data from servers extracted and analyzed. Complete data extracted and analyzed where we remove irrelevant data and retain data required for processing. According to the project only accidents and parameters are required to generate outputs.

Step 3: Specify Constraints

SUPPORT COUNT

The relationship between the entire number of transaction containing that item (A) with the entire number of transaction in data set.

CONFIDENCE

Confidence of item set defined as total number of transaction containing the item set to the total number of transaction containing LHS.

Step 4: Association Rules Mining (AprioriAlgorithm)

Association (or relation) is perhaps higher known and most familiar and easy data processing technique. Here, we make an easy correlation between two or more items, often of an equivalent type to spot patterns.

For example, Market-basket analysis, where we track people's buying habbits ,we'd identify that a customer always buys cream when they buy strawberries, and thus suggest that subsequent time that they buy strawberries they could also want to shop for cream.

We use apriori algorithm to process e commerce data and to find the patterns. Here we generate patterns related to traffic accidents.

Algorithm Is chosen due to subsequent reasons.

1. Quicker Results
2. Works fine for little data set also as Huge data set.
3. One scan of Database is Enough.
4. Works fine for multiple constraints.

Step 5: Patterns Prediction

Here system predicts the relationship between frequent traffic accidents.

IJARCCE

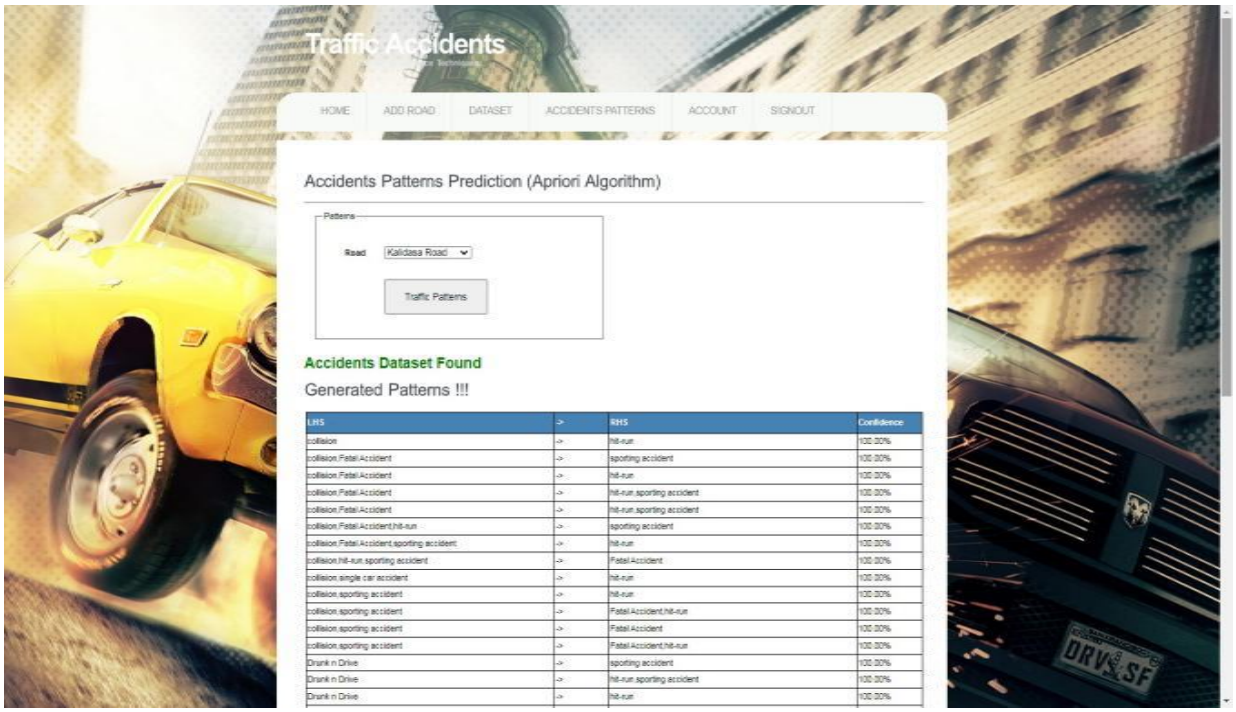
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June Month 2021

DOI 10.17148/IJARCCE.2021.106119

V. EXPERIMENT RESULTS



VI. CONCLUSION

Because road safety is such an important aspect of our lives, we must strive to enhance it as much as possible using all available resources and possibilities. Descriptive or predictive mining combined with other essential data such as weather or traffic conditions offers a fascinating alternative with potentially beneficial and helpful consequences for all parties involved. These reasons prompted the establishment of this study to examine existing data samples depicting road accidents in the United Kingdom, which represented a significant quantity of data and necessitated the adoption of a relatively new approach in this area, in-memory data processing.

Future Enhancements

- We can add public Notifications it helps to public.
- We can add query module for the interaction between administrator and member.
- We can predict reasons for accidents which helps traffic departments to take precautionary measures.

VII. REFERENCES

[1] Worrall R D, Bullen A G R. AN EMPIRICAL ANALYSIS OF LANE CHANGING ON MULTILANE HIGHWAYS[J]. Highway Research Record, 1970(303).
[2] Xu H, Cheng G, Pei Y. Study on effect of lane-changing behavioral characteristic to capacity[J]. Sciencepaper Online, 2010.
[3] Yulong P, Ji M, Research on countermeasures for road condition causes of traffic accidents [J]. China Journal of Highway and Transport, 2003, 16(4):77-82.
[4] Das S, Sun X. Investigating the Pattern of Traffic Crashes Under Rainy Weather by Association Rules in Data Mining[J]. Transportation Research Board Annual Meeting, 2014.
[5] He S B, Wang Y J, Sun Y K, et al. The Research of Multidimensional Association Rule in Traffic Accidents[C]// International Conference on Wireless Communications, NETWORKING and Mobile Computing. IEEE, 2008:1-4.
[6] Mennis J, Liu J W. Mining Association Rules in Spatio-Temporal Data[C]// 2003:699–700.
[7] Kumar S, Toshniwal D. A data mining framework to analyze road accident data[J]. Journal of Big Data, 2015, 2(1):26.
[8] López G, Abellán J, Montella A, et al. Patterns of Single-Vehicle Crashes on Two-Lane Rural Highways in Granada Province, Spain: In- Depth Analysis Through Decision Rules[J]. Transportation Research Record Journal of the Transportation Research Board, 2014, 2432(1):133-141.
[9] Weng J, Zhu J Z, Yan X, et al. Investigation of work zone crash casualty patterns using association rules[J]. Accident; analysis and prevention, 2016, 92:43.
[10] Sun J, Sun J. A dynamic Bayesian network model for real-time crash prediction using traffic speed conditions data[J]. Transportation Research Part C Emerging Technologies, 2015, 54:176-186.
[11] Apiletti D, Baralis E, Cerquitelli T, et al. Characterizing network traffic by means of the N et M ine, framework[J]. Computer Networks, 2009, 53(6):774-789.
[12] Pande A, Abdel-Aty M. Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool[J]. Safety Science, 2009, 47(1):145-154.

