



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Road Accident Analysis Using Machine Learning

B. ANUJA BEATRICE,
ASSISTANT PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE
SRI KRISHNA ARTS AND SCIENCE COLLEGE

AFZAL AHAMED. A,
PG STUDENT
DEPARTMENT OF COMPUTER SCIENCE
SRI KRISHNA ARTS AND SCIENCE COLLEGE

JACINTH. P,
PG STUDENT
DEPARTMENT OF COMPUTER SCIENCE
SRI KRISHNA ARTS AND SCIENCE COLLEGE

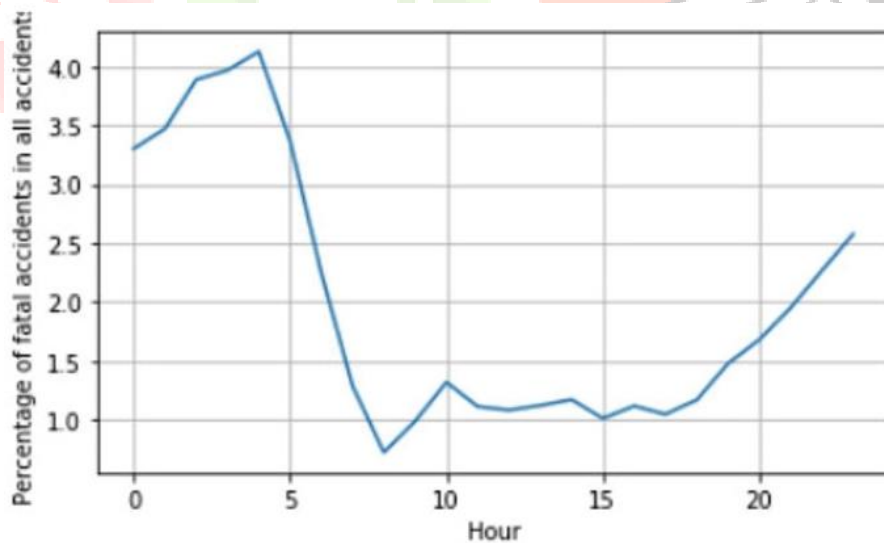
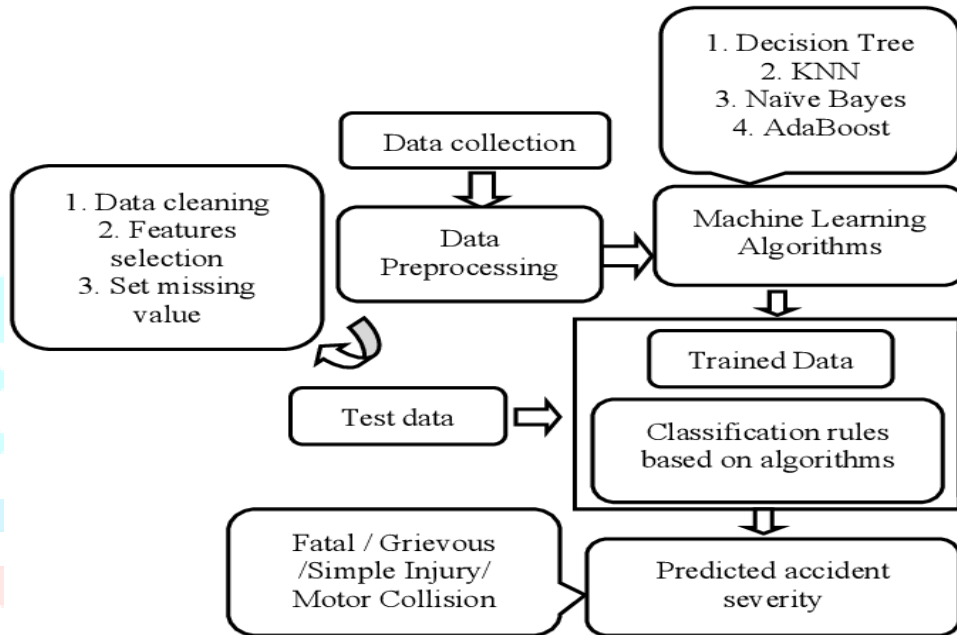
PRIYADHARSAN. S,
PG STUDENT
DEPARTMENT OF COMPUTER SCIENCE
SRI KRISHNA ARTS AND SCIENCE COLLEGE

Abstract

There are many inventories in automobile industries to design and build safety measures for automobiles, but traffic accidents are unavoidable. There is a huge number of accidents prevailing in all urban and rural areas. Patterns involved with different circumstances can be detected by developing an accurate prediction model which will be capable of automatic separation of various accidental scenarios. These clusters will be useful to prevent accidents and develop safety measures. We believe to acquire maximum possibilities of accident reduction using low budget resources by using some scientific measures. There is a huge impact on the society due to traffic accidents where there is a great cost of fatalities and injuries. In recent years, there is an increase in the researches' attention to determine the factors that significantly affect the severity of the drivers' injuries which is caused due to the road accidents. Accurate and comprehensive accident records are the basis of accident analysis. The effective use of accident records depends on some factors, like the accuracy of the data, record retention, and data analysis. There are many approaches applied to this scenario to study this problem. A recent study illustrated that the residential and shopping sites are more hazardous than village areas as might have been predicted, the frequencies of the casualties were higher near the zones of residence possibly because of the higher exposure. A study revealed that the casualty rates among the residential areas are classified as relatively deprived and significantly higher than those from relatively affluent areas.

1. Introduction

The traffic has been transformed into the difficult structure in points of designing and managing by the reason of increasing number of vehicles. This situation has discovered road accidents problem, influenced public health and country economy and done the studies on solution of the problem. Large calibrated data agglomerations have increased by the reasons of the technological improvements and data storage with low cost. Arising the need of accession to information from this large calibrated data obtained the corner stone of the data mining. In this study, assignment of the most compatible machine learning classification techniques for road accidents estimation by data mining.





2. Proposed method

The main objective of the road accident prediction system is Analyze the previously occurred accidents in the locality which will help us to determine the most accident-prone area and help us to set up the immediate required help for them, and to make predictions based on constraints like weather, pollution, road structure, etc. Models are created using accident data records which can help to understand the characteristics of many features like drivers' behavior, roadway conditions, light condition, weather conditions and so on. This can help the users to compute the safety measures which is useful to avoid accidents. It can be illustrated how statistical method based on directed graphs, by comparing two scenarios based on out-of-sample forecasts. the model is performed to identify statistically significant factors which can be able to predict the probabilities of crashes and injury that can be used to perform a risk factor and reduce it.

Here the road accident study is done by analyzing some data by giving some queries which is relevant to the study. The queries like what is the most dangerous time to drive, what fractions of accidents occur in rural, urban and other areas. What is the trend in the number of accidents that occur each year, do accidents in high-speed limit areas have more casualties and so on ...? These data can be accessed using Microsoft excel sheet and the required answer can be obtained. This analysis aims to highlight the data of the most importance in a road traffic accident and allow predictions to be made. The results from this methodology can be seen in the next section of the report.

3. Experimental results

All the tests should be traceable to customer requirements the focus of testing will shift progressively from programs Exhaustive testing is not possible to be more effective testing should be which has probability of finding errors

The following are the attributes of good test

1. A good test has a probability of finding an error
2. A good test should be "best of breeds"
3. A good test to neither simple nor too complex

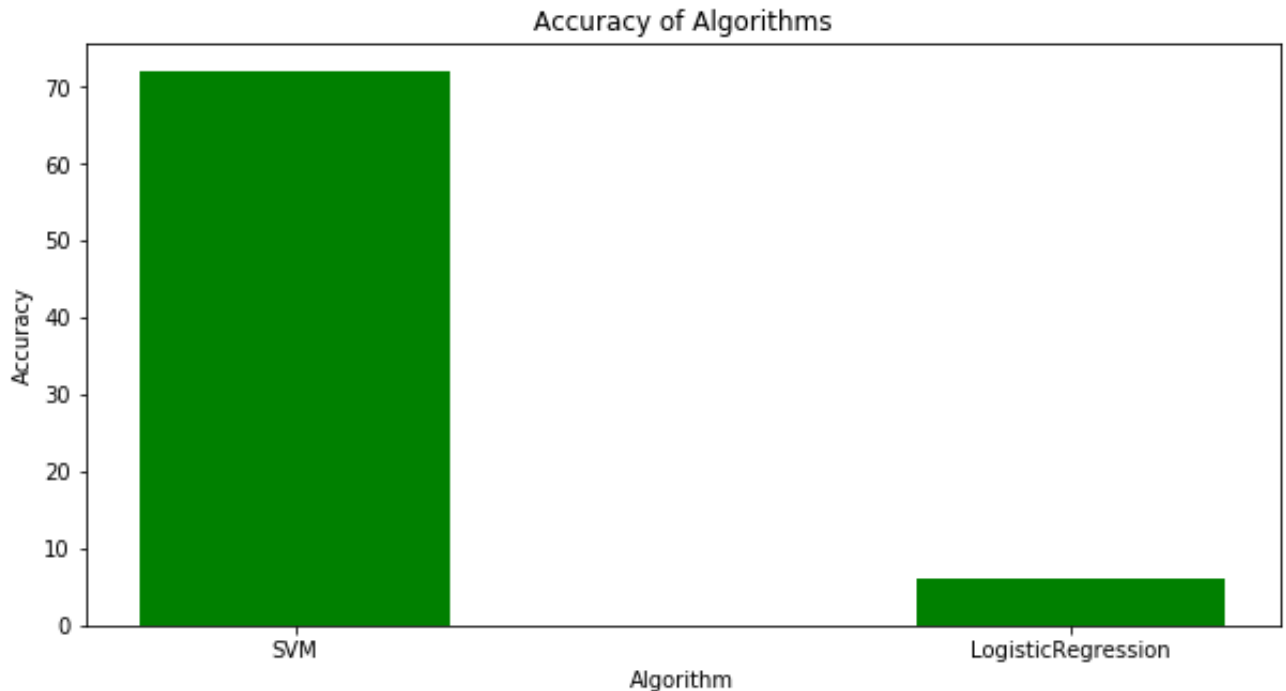
Algorithms accuracy score and prediction values

```
In [10]: print(model.predict([[2,3,1,1,1,2,9,1,2,80]]))  
[ 2.83301058]
```

```
In [11]: print(model.predict([[1,6,1,1,1,1,9,3,2,70]]))  
[ 2.63386]
```

```
In [12]: print(model.predict([[3,3,1,1,1,2,9,1,2,29]]))  
[ 2.96620897]
```

Graph



4. Algorithms Used

Linear Regression: Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc. Linear regression algorithm shows a linear relationship between a dependent and one or more independent variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable. The linear regression model provides a sloped straight line representing the relationship between the variables.

Support vector machine (SVM): was founded by Vladimir Vapnik in 1992 [17]. It is an algorithm for prediction and classification of linear and non-linear data. The aim of SVM is to find the maximum margins of hyperplane. The maximum margin hyperplane gives the maximum distance between the separation decision classes. The training examples that are closest to the largest margin hyperplane are called support vectors [18][19]. A brief mathematical description of SVM algorithm is provided as follows. Assume a training set with input vector and target labels, according to Vapnic Formula, satisfies the following conditions: (1) Which is equivalent to: (2) Where the weight vector (maximum margin) and b is the bias.

5. Conclusion

Road Accidents are caused by various factors. By going through all the research papers, it can be concluded that Road Accident cases are hugely affected by the factors such as types of vehicles, age of the driver, age of the vehicle, weather condition, road structure and so on. Thus, we have built an application which gives efficient prediction of road accidents based on the above-mentioned factors.

References

1. T. Rahman, "Road Accidents in Bangladesh: An Alarming Issue", The World Bank, 2012.
2. K. M. Habibullah, A. Alam, S. Saha, A. Amin and A. K. Das, " A Driver-Centric Carpooling: Optimal Route-Finding Model using Heuristic Multi-Objective Search," 2019 4th International Conference on Computer and Communication Systems (ICCCS), Singapore, 2019.
3. M. M. L. Elahi, R. Yasir, M. A. Syrus, M. S. Q. Z. Nine, I. Hossain and N. Ahmed, "Computer vision based road traffic accident and anomaly detection in the context of Bangladesh," 2014 International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, 2014, pp. 1-6.
4. M. S. Satu, S. Ahamed, F. Hossain, T. Akter and D. M. Farid, "Mining traffic accident data of N5 national highway in Bangladesh employing decision trees," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017, pp. 722-725.
5. H. İ. Bülbül, T. Kaya and Y. Tulgar, "Analysis for Status of the Road Accident Occurrence and Determination of the Risk of Accident by Machine Learning in Istanbul," 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, 2016, pp. 426-430.
6. P. A. Nandurge and N. V. Dharwadkar, "Analyzing road accident data using machine learning paradigms," 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC), Palladam, 2017, pp. 604-610.
7. S. Kumar and D. Toshniwal, "A data mining approach to characterize road accident locations," Journal of Modern Transportation(2016), vol. 24, issue no. 1, pp. 62-72.
8. Beshah, Tibebe, and Shawndra Hill, "Mining Road Traffic Accident Data to Improve Safety: Role of Road-Related Factors on Accident Severity in Ethiopia," AAAI Spring Symposium: Artificial Intelligence for Development (2010).
9. M. A. A. Mamun, J. A. Puspo and A. K. Das, "An intelligent smartphone based approach using IoT for ensuring safe driving," 2017 International Conference on Electrical Engineering and Computer Science (ICECOS), Palembang, 2017, pp. 217-223

