



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

ROAD ACCIDENT ANALYSIS USING LINEAR REGRESSION

T.L.DEEPIKA ROY

*Assistant Professor
Department of CSE*

CMR TECHNICAL CAMPUS, HYDERABAD

Dr.SUWARNA GOTHANE

*Associate Professor
Department of CSE*

CMR TECHNICALCAMPUS, HYDERABAD

P.DEEPA

*B.Tech student
Department of CSE*

CMR TECHNICAL CAMPUS, HYDERABAD

K.MADHAVILATHA

*B.Tech student
Department of CSE*

CMR TECHNICALCAMPUS, HYDERABAD

CH.SHIRISHA

*B.Tech student
Department of CSE*

CMR TECHNICAL CAMPUS, HYDERABAD

SUMAIR ALI

*B.Tech student
Department of CSE*

CMR TECHNICALCAMPUS, HYDERABAD

ABSTRACT: *In automobiles industries there are many inventories for building safety measures to automobiles but traffic accidents are said to be unavoidable . The probability of occurrence of accidents depends on numerous factors like weather, age of the driver, road structure etc. India is one among the developing nation which experience the highest rate of such accidents and in recent years the road accidents has been considered as one of the major issues and remarked as ninth prominent cause in the world.it is entirely not admissible and saddening to allow its citizens to be killed by road accidents. consequently, to handle this situation, a precise analysis is required. **KEYWORD:** Nature of Accidents, Cause of Accidents, Road Feature, Classification of Accident, Road Safety.*

high income counties indicates a downward trend in recent decades. There is a way to improve the safety. Although the number of lives lost in road accidents in high income countries indicates a downward trend in recent decades, for most of the worlds population, the burden of road traffic injury in terms of societal and economic costs is rising substantially. Injuries and death due to road traffic accidents(RTA) are a major public problem in developing countries where more than 85% of all deaths and 90% of disability adjusted life years were lost from road traffic injuries. There is a huge impact on society due to the traffic accidents where there is a great cost of fatalities and injuries. The effects of injuries due to road traffic accidents have a tremendous impact on socio-economic development of a country. In recent years, there is an increase in the researcher's attention to determine the significantly affect the severity of the driver's injuries which is caused due to the road accidents. The effective use of accident records depends on some factors, like the accuracy of the data and data analysis. Machine learning which is a sub-branch of artificial intelligence supplies machine learning of computers taking advantage of data warehouses.

I.INTRODUCTION

This paper focuses on the accidents that are occurring in recent times. The road accidents are occurring lot continuously and has marked as one of the major problem. In addition, according to the reports of traffic accidents nearly 90% of accidents caused by human factors, and the humans become one of the most unstable factors in causing traffic safety problems. Although the number of lives lost in road accidents in

II. LITERATURE SURVEY

Krishnavi and Hemalatha [5] worked with some classification models to predict the severity of injuries that occurred during traffic accidents. The research carried out actually shows that the effect of seat belt use at front seats of around 40. The numerous studies were carried out to identify the various parameters influenced for the causes of road accidents. India already accounts for about 9.5% of the total 1.2 million fatal accidents in the world. Road accidents are a complex interaction of different parameters like road, vehicle etc. The various measures are used to estimate the severity of a traffic conflict which are limited to estimate the road accidents utilization of machine learning of useful since it is able to extract the data and can use the statistical methods that are required. It was also identified that the accident rates are highest in intersections than the other portion of the road [4].

III. PROPOSED SYSTEM

The proposed system is a middleware which uses techniques that involve data slicing, data analysis and data pre-processing for secured and optimized results. Data pre-processing is an important and mandatory step for any machine learning model because it involves steps like feature scaling to get exact values. As machine learning models deal with values with close proximity, splitting the dataset in training is necessary as the dataset contains a huge number of features within which various unwanted features are also present that are not required. After pre-processing the important features are extracted. The data pre-processing is the raw data and making it suitable for machine learning model. It is the first and crucial step while creating a machine learning model. When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So for this, we use data pre-processing task.

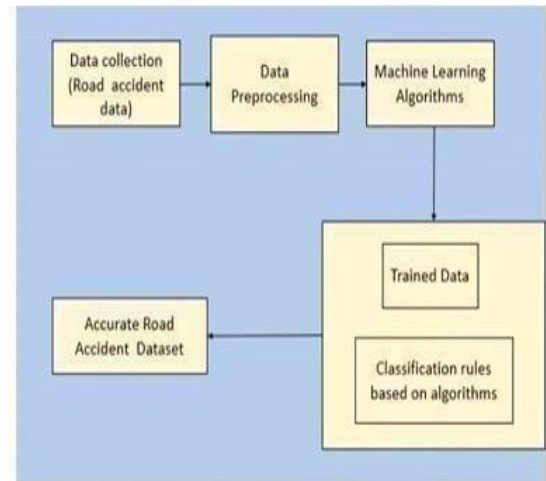


Fig 1: Proposed System Architecture

Data Collection: Real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

Data Pre Processing: Data pre-processing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So for this, we use data pre-processing task. The steps involved in data pre-processing are as follows

- ✦ Getting the dataset
- ✦ Importing libraries
- ✦ Importing datasets
- ✦ Finding Missing Data
- ✦ Encoding Categorical Data
- ✦ Feature scaling

IV. IMPLEMENTATION

We see the output page as some buttons, when we press the buttons it analyses several data from given sheets and predicts the output as minimum number of accidents, maximum number and area in which maximum and minimum number of accidents occur.

IMPLEMENTATION STEPS-

- ✦ Gathering the dataset from database
- ✦ Pre-processing the dataset and analysis dataset
- ✦ Splitting the datasets into training and testing in the ration80% and applying the linear regression to analyse the data
- ✦ Obtain the accuracy in prediction

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 (y) and one or more independent (x) 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.

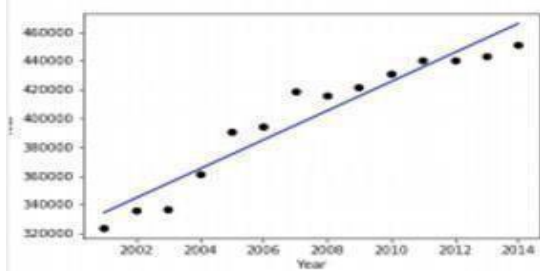


Fig 2: Regression Line

ANALYSIS AND PREDICTION

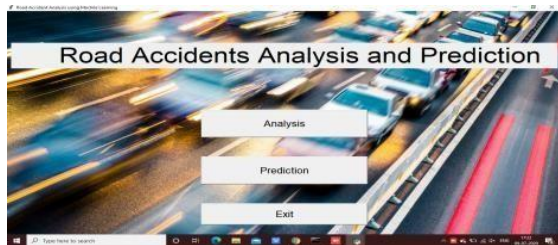


Fig 3: Analysis and Prediction

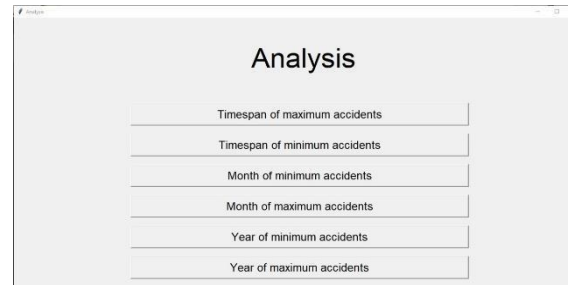


Fig 4: Analysis



Fig 5: Maximum Number of Accidents in a Year



Fig 6: Minimum Number of Accidents in a Year



Fig 7: Maximum Number of Accidents in a Month



Fig 8: Minimum Number of Accidents in a Month

V. CONCLUSION

Losses in road accidents are unbearable, to the society as well as a developing country like us. So, it has become an essential requirement to control and arrange traffic with an advanced system to decrease the number of road accidents in our country. Taking simple precautions, based on prediction or warnings of a sophisticated system may prevent traffic accidents. The road accidents are actually caused due to the factors such as age of driver, weather conditions etc. So it has become an essential requirement to control and arrange traffic. The police database and newspaper reports provide insights into the magnitude and nature of fatalities due to road traffic crashes. The limitations of the police database, which is the legal source of information on fatalities accuracy to predict traffic accident severity. Moreover, to make it more feasible, we will try to make a recommender system by using these approaches that can give a prediction to the traffic accident and can warn the road user. In the future, it will be our try to create a mobile application by implementing this methodology to provide an accurate prediction to the user and make it very useful and beneficial also. More effort is needed to have a comprehensive understanding of the various aspects of road traffic crashes, and the recommendations made for strengthening surveillance could serve as an initial step towards reducing fatalities and injuries due to road crashes in the long term.

VI. FUTURE WORK

In the future, it will be our try to create an application by implementing methodology to provide an accurate prediction to the user and make it very useful and beneficial also. Moreover, to make it more feasible, we will try to make a recommender system s by using the approaches that can give a prediction to the traffic accident and can warn the road user.

VII. ACKNOWLEDGEMENT

Apart from the efforts of us, the success of paper depends largely on encouragement and guidelines of many others. We take this opportunity to express my profound gratitude to CMR Technical Campus College Management for motivating us and for providing all the facilities required for this work. We are deeply indebted to Chairman Shri C.Gopal Reddy, Secretary Smt.C.Vasanth Latha, Director Dr.A.Raji Reddy, HOD CSE Dr.K.Srujan Raju, PRC Co-ordinator Dr.Suwarna Gothane, Project Guide T.L.Deepika Roy who always has been a constant source of inspiration.

VIII. REFERENCES

- [1]. Sachin Kumar, Durga Toshniwal, "A data mining approach to characterize road accident locations", *J. Mod. Transport.* (2016) 24(1):62–72.
- [2]. S. Shanthi and Dr. R. Geetha Ramani, "Gender Specific Classification of Road Accident Patterns through Data Mining Techniques", *IEEE-International Conference on Advances in Engineering, Science and*
- [3]. <https://towardsdatascience.com/5beginnerfriendly-steps-to-learn-machinelearninganddata-science-withpythonbf69e211ade5>
- [4]. K. Jayasudha and C Chandrashekar. An overview of Data Mining in Road Traffic and Accident Analysis.
- [5]. S. Krishnaveni and M. Hemalatha. A perspective of Analysis of Traffic Accidents using Data Mining techniques.
- [6]. 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*.
- [7]. 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.
- [8]. P. A. Nandurde 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), pp. 604-610, 2017.
- [9]. <https://www.irjet.net/archives/v7/i12/IRJEV7I12129>
- [10]. <https://www.ripublication.com/ijtam17/ijtamv12n1>
Management (ICAESM -2012) March 30, 31, 2012.
- [11]. <https://www.enggjournals.com/ijet/docs/IJET17-09-03-156>