Analysis of Road Accident Data Using Data Science

Asst. Prof. Ravi H Gedam
Asst. Prof Dept. of Computer Science and Engineering
G.H. Raisoni University, Saikheda.

Akshay Nerkar, Ayushi Patil, Prachi Latkar, Samiksha Shrikhande
Department of CSE
GHRU
Saikheda

Abstract:
Road accidents result in a great deal of death and damage to the economy, making them a major global public safety concern. When analysing data from traffic accidents, data science can be beneficial in determining contributing elements, predicting models, and creating successful interventions. Data on traffic events, meteorological circumstances, and local government records are some sources of information about road accidents. However, because of problems like under-reporting and over-dispersion, gathering and analysing data on traffic accidents can be difficult. Road accident data analytics has demonstrated encouraging results in locating hotspots and lowering accidents despite these obstacles. In this situation, analysing the temporal and spatial patterns of traffic accidents can be of special use for spatial data science.

Keywords:
Juypter, Python.

Introduction:
Data science has emerged as a powerful tool in addressing some of society's most pressing challenges, including road safety. Every year, millions of lives are affected by road accidents worldwide, making it a critical issue for governments, transportation authorities, and the public. Leveraging data science methodologies and techniques to analyse road accident data has become instrumental in understanding the underlying causes, improving safety measures, and ultimately reducing the toll of these tragic incidents.

When applying data science to road accident data, a methodical approach is taken to the collection, cleaning, analysis, and extraction of valuable insights from sizable datasets about traffic incidents. In addition to helping to comprehend the trends and causes of accidents, this process directs stakeholders and policymakers in the development of evidence-based road safety initiatives.

The different steps of the data science process applied to road accident data will be examined in this investigation, emphasizing the value of predictive modelling, data collection, and preprocessing, and the potential applications of data-driven insights in preventing injuries, saving lives, and improving transportation systems. We can create a more secure and effective driving environment in the future by utilizing data science, which will ultimately reduce accidents and enhance people's quality of life in general.

Objective:
The primary objective of this data science project on road accident data is to leverage data-driven insights and predictive analytics to enhance road safety and reduce the frequency and severity of accidents. Specifically, the project aims to achieve the following key objectives:

Accident Understanding: - Analyse past accident data to get a thorough picture of traffic incidents. Determine the trends, patterns, and common elements—such as the type of road, the type of vehicle, the weather, and driving habits—that contribute to accidents.

Risk Prediction: - Create predictive models that can estimate the chance of mishaps happening in particular areas or with particular circumstances. By acting as early warning systems, these models will

help authorities spend resources wisely and put preventative measures in place.

Injury Severity Analysis: - Examine the variables that are associated with the seriousness of injuries received in collisions. The study intends to assist emergency responders and medical professionals in better planning for and handling accidents by recognizing these characteristics.

Optimizing Traffic Management: - Make the most of data insights to improve traffic control tactics. To lessen traffic jams and accident-prone conditions, this involves enhancing traffic signal timings, road signs, and speed limit enforcement.

Public Awareness: - Create data-driven strategies and campaigns to raise public awareness of traffic safety. Utilize the data's insights to customize interventions and messaging for particular populations and geographic areas.

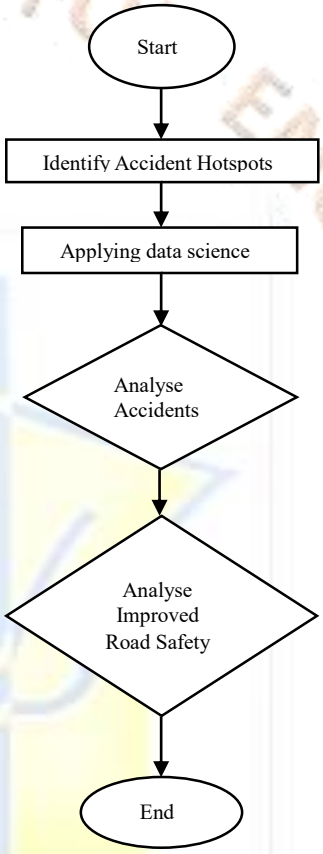
Resource Allocation: Real-time accident data and accident hotspot predictions help government agencies and first responders allocate resources more effectively. This may result in emergency services that are more efficient and have quicker reaction times.

Policy Recommendations: - Provide evidence-based suggestions for modifying policies and enhancing infrastructure that can improve road safety based on data analysis. Governmental organizations can use these suggestions to pu specific policies into action.

Evaluation of Interventions: - Analyse the impact of prior road safety policies and interventions on accident rates to determine their efficacy. Make use of this knowledge to suggest new tactics or improve current ones.

METHODOLOGY:

1)Problem Statement – The problem is the increasing rate of accidents day by day.
2)Gather the data – This may involve collecting data from internal sources like GitHub and Kaggle.
3)Clean & prepare the data – This step involves ensuring that the data is accurate, complete & consistent. If there is any missing value found then it will fill with by default value or may apply some parameters like mean, median, and mode & also transform the data into a format that is more suitable for analysis.
4)Apply Data Science – Then we will apply the data science process.
5)Explore the Data – This step involves using statistical & graphical techniques to gain insights into the data.
After applying all the processes, we will apply an algorithm to analyse future outcomes



Flowchart proposed for analysing road accident data using data science

ANALYSIS:

1) The goal of our project is to identify accident hotspots and why accidents are a major problem there.

2)At first, we will have to identify some hotspots by analysing accident data. It is possible to identify areas where accidents are more likely to occur. This information can be used to target safety improvements in these areas.

3)Data science can be used to develop models that analyse the risk of accidents based on factors such as weather conditions, traffic volume, and driver behaviour. This information can be used to warn drivers of potential hazards and to take steps to mitigate risk.

4)After analysis of risk improve road safety according to the reason for road accidents.

5)Then apply some developing new safety technologies to get results.

6)The result will provide what are the hotspots according to available l data and why there are accident problems.

CONCLUSIONS:

Understanding the underlying patterns, causes, and contributing factors to accidents has been made possible by the application of data science to the analysis of traffic accidents. It has made it possible to identify high-risk regions, accident-prone demographics, and particular environmental factors. Using this information will help in the development of focused road safety interventions.

REFERENCES:

1)Rabbani, M. et al. (2021) 'Road traffic accident data analysis and its visualization', Civil Engineering and Architecture, 9(5), pp. 1603–1614. doi:10.13189/cea.2021.090530.

2) D. Teja Santosh (2022) 'Road Accident Data Collection Systems in Developing and Developed Countries: A Review', Transportation Research Interdisciplinary Perspectives, 19, p. 100814. doi:10.1016/j.trip.2022.100814.

3) Mesquitela, J. et al. (2022) 'Data Analytics process over road accidents data—a case study of Lisbon City', ISPRS International Journal of Geo-Information, 11(2), p. 143. doi:10.3390/ijgi11020143.

4) Jones, A.P. and Jørgensen, S.H. (2003) 'The use of multilevel models for the prediction of road accident outcomes', Accident Analysis & Prevention, 35(1), pp. 59–69. doi:10.1016/s00014575(01)00086-0.

5) Naeem, M.M. (2020) 'Black Spots identification and accident analysis of Indus Highway (N-55)', JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES, 15(11). doi:10.26782/jmcms.2020.11.00015.