A PREDICTIVE MODEL OF ROAD TRAFFIC ACCIDENTS USING SURROGATE SAFETY MEASURES

By **Prithviraj – M22AI603**

Under the supervision of **Dr. Ranju Mohan**

in partial fulfilment of the requirements for the award of the degree of

MTech in Data and Computational Science

॥ त्वं ज्ञानमयो विज्ञानमयोऽसि ॥

Indian Institute of Technology Jodhpur

Data and Computational Science

November, 2023.

Declaration

I hereby declare that the work presented in this Project Report titled **A Predictive Model of Road Traffic Accident Using Surrogate Safety Measures** submitted to the **Indian Institute of Technology Jodhpur** in partial fulfilment of the requirements for the award of the degree of **MTech in Data and Computational Science**, is a bonafide record of the research work carried out under the supervision of **Dr. Ranju Mohan**. The contents of this Project Report in full or in parts, have not been submitted to, and will not be submitted by me to, any other Institute or University in India or abroad for the award of any degree or diploma.

Prithviraj

M22AI603

Certificate

This is to certify that the Project Report titled **A Predictive Model of Road Traffic Accident Using Surrogate Safety Measures**, submitted by **Prithviraj (M22AI603)** to the **Indian Institute of Technology Jodhpur** for the award of the degree of **MTech in Data and Computational Science**, is a bonafide record of the research work done by him under my supervision. To the best of my knowledge, the contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Ranju Mohan

Abstract

This project presents a comprehensive approach to develop a predictive model for road traffic accidents using Surrogate Safety Measures (SSM). The primary objective is to estimate the likelihood of accidents at both micro and macro levels, encompassing intersections and mid-blocks, as well as identifying accident-prone areas within city networks. SSM, derived from various traffic-related variables including road geometry, traffic flow characteristics, and driver behaviour, serve as indirect indicators for potential crash occurrences. A predictive model is developed using machine learning algorithms, adapting to the complexity of the data and desired accuracy. The performance of the model is evaluated through metrics like accuracy, precision, recall, and F1 score, which gauges its reliability in estimating accident risks. Furthermore, the model is deployed into practical applications, potentially integrated into traffic management systems, providing real-time accident risk predictions to drivers and relevant stakeholders. This research offers a holistic framework for predicting road traffic accidents, contributing to improved road safety and accident prevention.

Contents

1. INTRODUCTION

Addressing road traffic accidents on a global scale is of paramount importance due to the substantial loss of life, injuries, and economic damages they cause. In pursuit of enhanced road safety, various predictive models have been developed to reduce the occurrence of accidents. This research project introduces an innovative approach to road safety, focusing around the creation of a predictive model that utilizes Surrogate Safety Measures (SSM) to gauge the likelihood of accidents.

Unlike conventional accident prediction models, which heavily depend on historical accident data, this project incorporates SSM, encompassing traffic flow characteristics, road geometry, and driver behaviour. The aim is to proactively identify accident-prone areas, both at micro-levels such as intersections and mid-blocks, and at a macro-level within city networks. This approach has the potential to facilitate timely interventions and pre-emptive measures, thereby reducing the frequency of accidents.

The significance of Surrogate Safety Measures (SSM) in traffic safety evaluation arises from the lack of reliable statistical safety models in many scenarios. This is particularly relevant for transportation facilities with complex site characteristics and/or nontraditional traffic safety treatments, where historical crash data may be limited or unavailable for developing safety predictive models [6]. Research on SSM dates back to the early 1970s [3], and significant progress has been achieved in this field since then.

The ultimate objective of this research is to contribute to road safety by providing a predictive model that offers real-time risk estimations for accidents. Such a transportation model can seamlessly integrate into traffic management systems, providing valuable insights to drivers, traffic authorities, and other stakeholders. This integration aims to facilitate safer road usage and prevent accidents. This research aspires to be a pioneering step towards establishing a safer and more efficient road system.

2. BACKGROUND

Road accidents, on a scale are a concern when it comes to public health and safety. According to the World Health Organization (WHO) [4] these accidents lead to millions of deaths and injuries every year. Not do they cause harm but they also have significant social and economic consequences for societies around the world. Fortunately, many road accidents can be. Identifying risks early on is crucial in minimizing their impact.

To understand the likelihood of an accident occurring researchers have turned to Surrogate Safety Measures (SSM). These measures act as indicators of crashes since collecting and analysing actual crash data can be challenging and time consuming. SSM consider factors such as traffic flow characteristics (like speed and volume) road layout (including curves and slopes) weather conditions (such as rain or fog) driver behaviour (including driving or distractions) and other variables that influence the likelihood of road accidents.

Developing models for road accident prediction holds promise. It provides insights for traffic authorities, transportation planners and stakeholders in implementing targeted safety measures allocating resources efficiently and ultimately reducing the occurrence of road accidents. By embracing data analysis techniques predicting road accidents using safety measures has gained attention as an approach, to enhancing road safety and saving lives.

3. PROBLEM STATEMENT

Addressing road accidents is a significant concern for public health and safety, and predicting accident risks early on is essential for minimizing their impact. However, the conventional reliance on actual crash data, such as First Inspection Reports (FIR), for accident prediction is not only time-consuming but also inefficient. There is a growing necessity to create predictive models capable of estimating the likelihood of road accidents using surrogate safety measures, indirect indicators that signal potential crash occurrences.

Within the realm of mobility, road safety emerges as a critical aspect with the primary goal of preventing accidents and mitigating their consequences. The problem statement encompasses key elements related to road safety, including accident prediction, accident prevention, and data analysis for safety improvement. In essence, the problem statement aligns with the broader objective of enhancing road safety by developing predictive models that leverage surrogate safety measures. This proactive approach aims to prevent road accidents, optimize resource allocation, and facilitate evidence-based decision-making to improve road safety.

4. SURROGATE SAFETY MEASURES (SSM)

The assessment of safety performance in designs, countermeasures, or systems often relies on crash frequency and severity, crucial indicators of their effectiveness. However, crashes are infrequent events, making it impractical and, to some extent, ethically questionable to rely solely on historical crash data for evaluating the performance of new safety strategies, such as the introduction of a new traffic sign. In response to this challenge, Surrogate Safety Measures (SSM) derived from traffic conflicts have gained popularity as an alternative.

Traffic conflicts represent observable non-crash events where interactions among multiple road users create a risk of collision unless their courses of movement are altered [1]. A conflict is deemed etiologically connected to a crash when a failure (e.g., human operator failure, road failure, vehicle failure) leading to the conflict cannot be adequately corrected [2] [7]. Due to this causal relationship, measures employed to identify traffic conflicts and assess their severities can be considered as SSM. Notably, traffic conflicts are more frequent compared to actual crashes.

It is important to highlight that numerous safety-related measures have been developed over time, but not all qualify as SSM. According to [6], two criteria for SSM qualification are:

- It must be derived from traffic conflicts directly linked to crashes.
- The relationship between traffic conflicts and the potential crash frequency and/or severity can be quantified using practical methods.

From this standpoint, traffic exposure/flow measurements such as Annual Average Daily Traffic (AADT), speed variation, and average operating speed do not meet the criteria for SSM, despite their proven associations with crash risk and occasional adoption as crash "surrogates." This paper specifically reviews and discusses safety measures that satisfy the two qualifying criteria mentioned above.

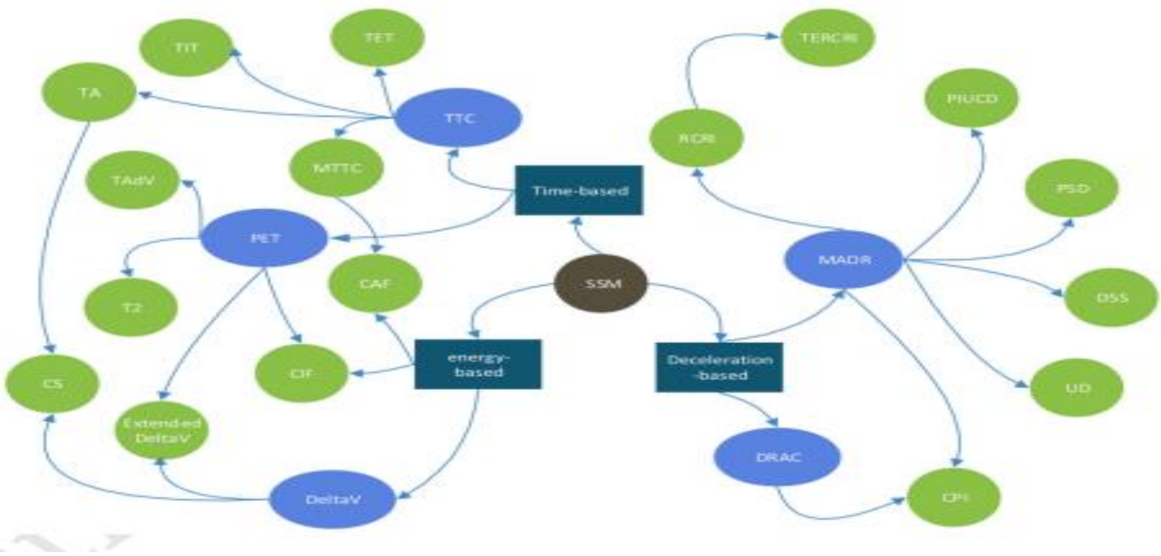

Figure 1: SSM Knowledge Map

As shown in Figure 2, there are three main sub-categories under SSM, which are time-based SSM, deceleration-based SSM, and energy-based SSM. These three sub-categories of SSM are described in detail below.

4.1 Time-based SSM

Surrogate Safety Measures (SSM) that are time-based assess the risk of an interaction based on its temporal proximity to a potential collision. One prevalent time-based SSM is Time-to-Collision (TTC), first introduced by Hayward [3]. TTC is defined as "the time remaining until a crash between two vehicles would occur if the current course and speed difference are maintained." The fundamental assumption underlying TTC is the consistent maintenance of both the speeds and directions of the vehicles involved. Mathematically, it is expressed as:

$$TTC_i(\mathrm{t}) = \frac{(X_{i-1}(t) - X_i) - l_i}{v_i(t) - v_{i-1}(t)}$$

Where,

X = vehicle position,

i = following vehicle,

i-1 = leading vehicle,

l = vehicle length,

v = vehicle velocity.

4.2 Deceleration-based SSM

Deceleration-based Surrogate Safety Measures (SSM) shift their focus from measuring time proximity to assessing how a vehicle's deceleration can avert a potential crash. In 1976, Cooper and Ferguson introduced DRAC (Deceleration Rate to Avoid the Crash) as one of the pioneering deceleration-based SSM. DRAC determines the severity of an interaction by calculating the minimum braking rate necessary for a vehicle to avoid colliding with another. The calculation is predicated on the assumption that one vehicle takes evasive actions while the other maintains its speed and direction. To evaluate the risk of collision, specific thresholds for DRAC are also required. Mathematically, it is expressed as:

$$DRAC_{i,t+1} = \frac{(V_{i,t} - V_{i-1,t})^2}{2\left[(P_{i-1,t} - P_{i,t}) - L_{i-1}\right]}$$

Where,

t = time interval,

P = position of a vehicle,

i = following vehicle,

i-1 = leading vehicle,

L = vehicle length,

V = velocity

4.3 Energy-based SSM

In contrast to time- and deceleration-based Surrogate Safety Measures (SSM), which gauge the proximity of a conflict to a potential crash, energy-based SSM were introduced to assess the severity of an interaction. Notably, DeltaV is a crucial measure within this category. It quantifies the change in velocity experienced by road users due to a collision, taking into account the speed, mass of each involved road user, and the approach angle between them [5]. Utilizing the momentum conservation assumption, specifically in an inelastic crash scenario, DeltaV can be calculated as:

$$\Delta v_1 = \frac{m_2}{m_1 + m_2}(v_2 - v_1)$$

$$\Delta v_2 = \frac{m_1}{m_1 + m_2}(v_1 - v_2)$$

Where, "v1" and "v2" represent the velocities of the two vehicles involved in a potential collision before the crash, while "m1" and "m2" denote the masses of these respective vehicles.

5. METHODOLOGY

This outlines the systematic process of developing a predictive model for road traffic accidents using Surrogate Safety Measures, starting from data collection and pre-processing, feature selection, model development, and evaluation, and finally, deployment and validation. It ensures that the model is accurate, reliable, and effective in preventing road accidents.

5.1 Data Collection

Gathering relevant data related to road accidents, surrogate safety measures, and other relevant variables such as road characteristics, weather conditions, and traffic patterns. This data may come from various sources, including traffic cameras, accident databases, and road surveys.

5.2 Data Preprocessing

Cleaning and preparing the collected data for analysis, which may involve data validation, outlier detection, and data normalization. This step is crucial to ensure that the data used for modelling is accurate and reliable.

5.3 Feature Selection

Identifying the most relevant surrogate safety measures that have a significant impact on road accidents. This may involve statistical analysis, feature ranking, and feature engineering techniques to select the most informative variables for the predictive model. Identify the most relevant surrogate safety measures (SSM) for accident prediction.

Potential SSM may include:

- Traffic density: Measured as vehicles per unit time or distance.
- Speed differentials: Variations in vehicle speeds on the same road.
- Road curvature: Degree of road bends or curves.
- Lane width: Width of road lanes.

Using statistical analysis and feature ranking techniques to prioritize SSM and Conduct feature engineering to create new features if necessary.

5.4 Model Development

Building a predictive model that can estimate the likelihood of road accidents based on the selected surrogate safety measures. This may involve machine learning techniques such as:

- Regression models: Linear regression, logistic regression.
- Decision tree models: Random Forest, Gradient Boosting.
- Neural network models: Feedforward neural networks.

Then the data can be split into training and testing sets for model training and evaluation. Train the model using the training dataset and optimize hyperparameters. Finally, Validate the model's accuracy and performance through cross-validation techniques.

5.5 Model Evaluation

Assessing the performance of the developed model using appropriate evaluation metrics, such as:

- Accuracy: The proportion of correct predictions.
- Precision: The ability to make accurate positive predictions.
- Recall: The ability to identify all relevant instances.
- F1 Score: A combined metric of precision and recall.

5.6 Model Deployment

Implementing the developed model into a practical application, such as a traffic management system, which can provide real-time predictions of road accident risks to drivers, traffic authorities, or other relevant stakeholders.

This involves, Integration with Traffic Management System, Realtime Predictions, Alert Generation, Visualization and User Interface, and Response Mechanism.

5.7 Model Validation

Validating the predictive model using real-world data to ensure its accuracy and reliability in real-world scenarios. This may involve comparing the model's predictions with actual road accidents data to assess its performance and make necessary refinements.

6. RESULT

The result of this project includes the development of a heat-map of accident-prone areas and a user interface for traffic authorities. These outcomes are crucial for providing actionable insights to enhance road safety.

6.1 Heat-Map of Accident-Prone Areas

The heat-map is a visual representation of accident-prone areas within the city network. It provides a graphical and intuitive way to identify high-risk locations for road traffic accidents.

Figure 2: Head Map

6.2 User Interface for Traffic Authorities

The user interface is a practical tool designed for traffic authorities and operators. It serves as a dashboard where users can interact with the predictive model, view real-time data, and receive actionable information. Key components of the user interface include:

- Real-Time Data Display: The interface presents real-time traffic data, including current traffic flow, weather conditions, and alerts generated by the predictive model.
- Heat-Map Integration: The heat-map of accident-prone areas is integrated into the user interface. Traffic authorities can access this map to identify high-risk locations instantly.
- Alert Management: The interface includes an alert management system that prioritizes alerts and provides relevant information about accident risks. Users can acknowledge alerts and trigger appropriate responses.
- Customization: The interface can be customized to display information specific to the city's needs. This can include setting alert thresholds, defining intervention actions, and configuring the display layout.
- Historical Data Analysis: Traffic authorities can access historical accident data through the interface to analyse trends and patterns. This can assist in long-term road safety planning.
- Communication Tools: The interface may include communication tools to coordinate responses with other stakeholders, such as law enforcement, emergency services, and maintenance crews.

The user interface empowers traffic authorities with valuable tools to make data-driven decisions in real-time. It enhances their ability to respond promptly to potential accidents, allocate resources effectively, and implement safety measures where they are needed most Overall, this project result significantly contributes to the proactive management of road safety and accident prevention within the city network.

7. CONCLUSION

In the face of escalating road traffic accidents and their devastating consequences, this project aimed to pioneer a proactive approach to road safety through the development of a predictive model utilizing Surrogate Safety Measures (SSM).

In conclusion, this project serves as a pioneering step toward safer and more efficient road transportation systems. By equipping traffic authorities and stakeholders with advanced tools and insights, it reinforces the mission to reduce road traffic accidents and create safer roadways for all. The project underscores the significance of data-driven innovation in addressing one of the most pressing challenges of our time — Road Safety.

REFERENCES

[1] Finn H Amundsen and Guro Ranes. 2000. Studies on traffic accidents in Norwegian road tunnels. Tunnelling and underground space technology 15, 1 (2000), 3–11.

[2] Gary A. Davis, John Hourdos, Hui Xiong, and Indrajit Chatterjee. 2011. Outline for a causal model of traffic conflicts and crashes. Accident Analysis Prevention 43, 6 (2011), 1907–1919. https://doi.org/10.1016/j.aap.2011.05.001

[3] John Hayward. 1971. Near misses as a measure of safety at urban intersections . Pennsylvania Transportation and Traffic Safety Center.

[4] Chinmoy Pal, Shigeru Hirayama, Sangolla Narahari, Manoharan Jeyabharath, Gopinath Prakash, and Vimalathithan Kulothungan. 2018. An insight of World Health Organization (WHO) accident database by cluster analysis with selforganizing map (SOM). Traffic injury prevention 19, sup1 (2018), S15–S20.

[5] Steven G Shelby et al . 2011. Delta-V as a measure of traffic conflict severity. In 3rd International Conference on Road Safety and Simulati. September. 14–16.

[6] Andrew P Tarko. 2018. Surrogate measures of safety. In Safe mobility: challenges, methodology and solutions. Vol. 11. Emerald Publishing Limited, 383–405.

[7] Andrew P. Tarko. 2020. Chapter 3 - Traffic conflicts as crash surrogates. In Measuring Road Safety Using Surrogate Events, Andrew P. Tarko (Ed.). Elsevier, 31–45. https://doi.org/10.1016/B978-0-12-810504-7.00003-3