UTM
UNIVERSITI TEKNOLOGI MALAYSIA

PROJECT PROPOSAL PRESENTATION

OPTIMIZATION OF URBAN TRAFFIC IN MALAYSIA USING MACHINE LEARNING ALGORITHM

PRESENTED BY : YANG YUEFEI
MATRIC NO : MCS241011
VENUE : DISCUSSION ROOM 4, LEVEL 2, BLOCK N28A
DATE : 17TH JANUARY 2025
LECTURER :

INTRODUCTION

BACKGROUND

- With the rapid development of Malaysia's economy and the acceleration of urbanization, the number of motor vehicles shows an increasing trend year by year, and road traffic accidents have become an important issue of social concern.
- With the development of cities, more motor vehicles are registered, so there are more traffic safety risks. Traffic accidents will bring many challenges to Malaysia's economy, medical treatment, social environment and other aspects, so it is necessary to promote the modernization of traffic management through intelligent technology and data analysis to reduce the negative impact of traffic accidents on the country and its people.
- According to historical data, motorcycle accidents occupy the main part of traffic accidents, and the night and rush hour are the key time period of the accident. Traffic accidents not only pose a threat to people's life safety, but also have a significant impact on medical resources, traffic efficiency and social economy.

PROBLEM STATEMENT

- Traffic problem(including traffic congestion and traffic accident and so on) is a widespread issue influenced by various factorssuch as traffic volume, road conditions, weather, and public events.Analyzing and addressing this problem is complex. Traditionalmethods of managing traffic have often failed to effectively control traffic volume and provide a smoother experience for commuters..

RESEARCH QUESTIONS & OBJECTIVES

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Research Question

- **RQ 1**: What are the major contributorsthat cause traffic accident across different parts of Malaysia?
- **RQ2**: How can we model these factors using predictive models that can forecast and prevent traffic accident?
- **RQ3** : How do the traffic patterns andcongestion factors that contribute to traffic accident vary for differentregions within Malaysia?

Research Objectives

- To identify the major contributors that cause traffic accident across different parts of Malaysia.
- To develop and apply predictivemodels to forecast traffic congestion levels based on identified factors.
- To analyze the variation in traffic patterns and congestion factors across different regions within Malaysia.

SCOPE OF THE STUDY

Geographical Coverage:

Analysis of traffic accident across urban centers, smaller cities, and rural areas in different states.

Data Sources:

Traffic reports, , road work, and public events data from sources like Waze, Google Traffic, Third-party data platform, and Department of Public Works available in Malaysia open data portals.

Alalytical Technologies:

Use machine learning to analyze the spatiotemporal distribution characteristics of traffic accidents and identify high-risk periods and areas. Identify key factors that affect the occurrence of accidents, such as road user type, number of vehicles, and time factors

Focus on Outcome:

Construct an accident prediction framework and evaluate model performance. Propose targeted policy recommendations based on the analysis results, such as improving transportation infrastructure, strengthening traffic safety education, and adjusting regulations.

LITERATURE REVIEW

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

www.utm.my

INTRODUCTION

Traffic congestion is a prevalent problem in urban areas worldwide, leading to economic losses, environmental pollution, and a decline in quality of life. It is impossible to avoid it completely, but we can use appropriate strategies to reduce the risk of traffic accidents. Malaysia, a country undergoing rapid urban development and escalating private vehicle ownership, is subject to grim traffic accident, especially in major cities. The aim of this chapter is to study the relevant literature regarding traffic accident: discuss about causes, effects, and prior research that have developed machine learning algorithms to aid both the prediction and amelioration of the issue.

LITERATURE REVIEW

Key Components

Cases

- High Vehicle denisty
- Road capacity
- Bad weather
- Public transportation is inefficien

Impacts

- Economic losses
- Social issues
- Pressure on the medical system
- Environmental pollution

Machine Learning

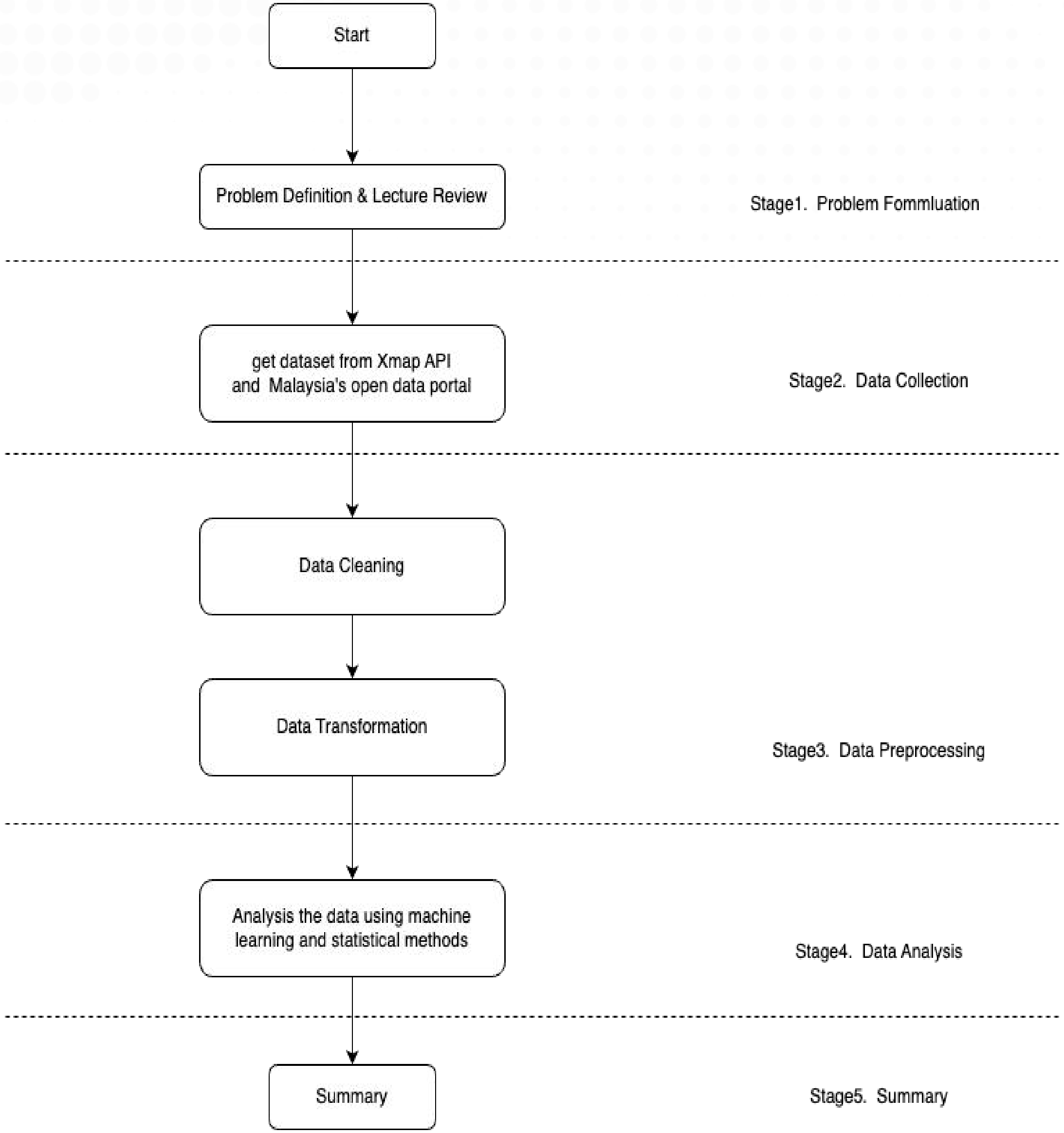
- Linear Regression
- Logical Regression
- Decision Tree
- Random Forests

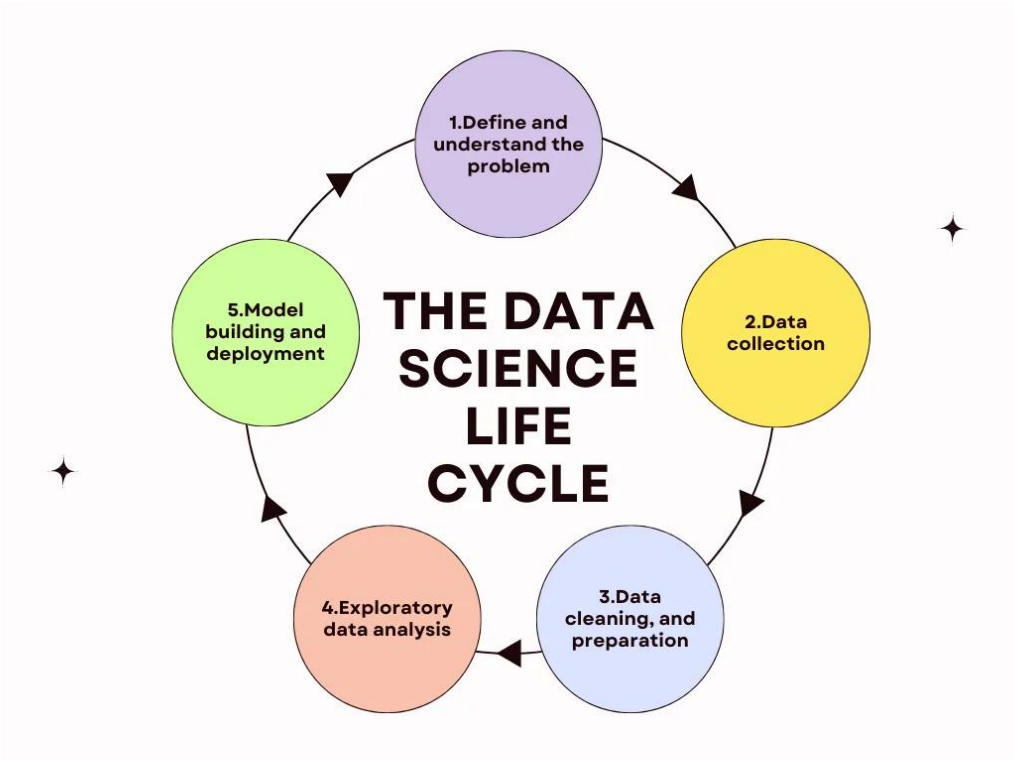
Future Work

- Expanding data scope to include environmental features
- Explore deep learning methods
- Optimizing feature engineering

METHODOLOGY

A research framework for using machine learning to optimize Malaysia’s traffic





KEY STEPS

Data Collection:
Obtain motor vehicle registration data, traffic accident loss data, traffic accident and age relationship data from Malaysia Open Data Portal, traffic throughput data and toll road length data from Third Party Data Platform (CEIC)

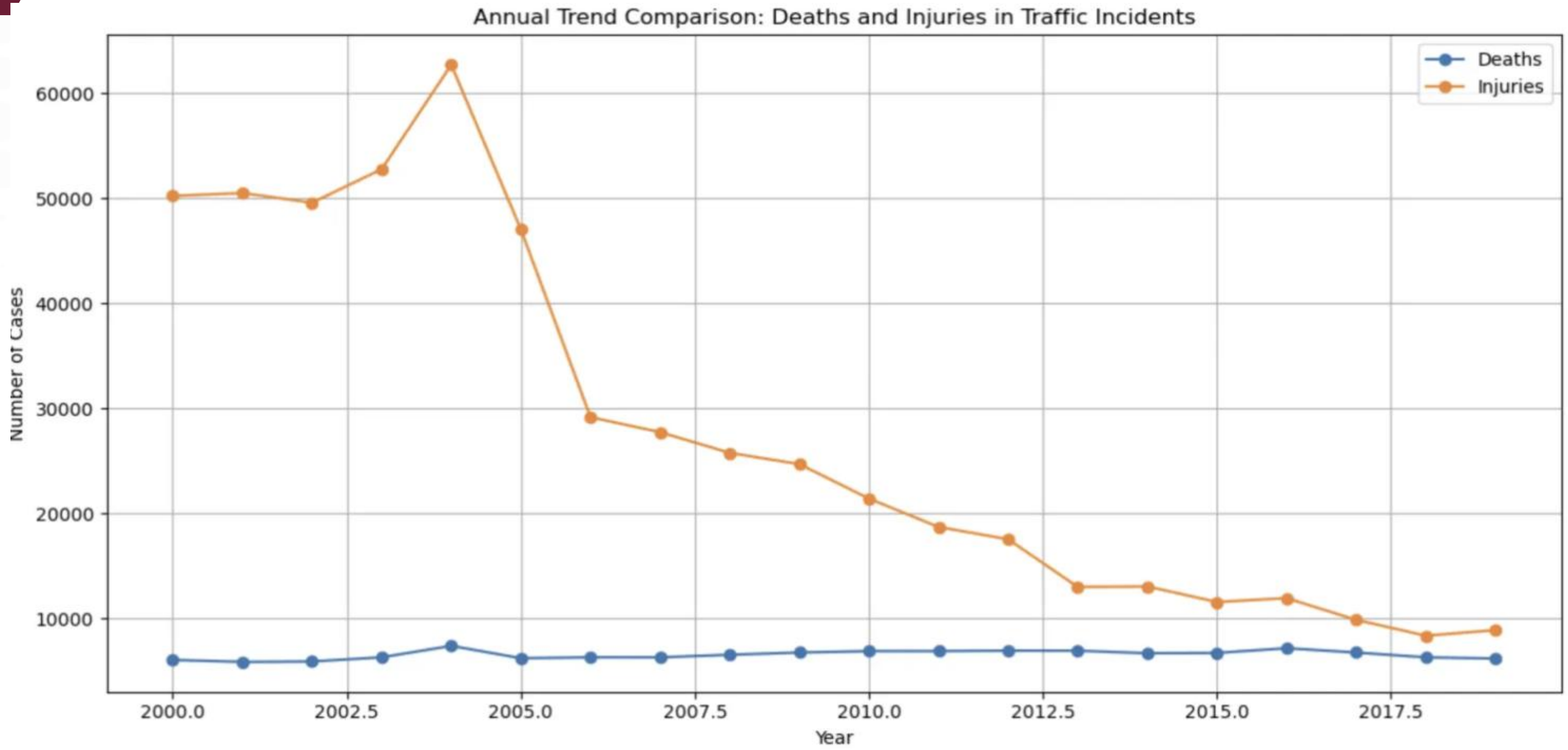
Data Cleaning:
- Remove the duplicate data. It is essential to check and eliminate duplicate data to avoid redundancy and ensure that the data is unique.
- Missing value processing: For numerical features, the mean or median is used to fill; For classification features, use mode padding.

Data Transforming:
- Column name normalization: Converts all column names to lowercase, removes Spaces, and replaces special characters with underscores.
- Time feature extraction: Extract the year, month, and time period (e.g., peak hour, night) from the date.

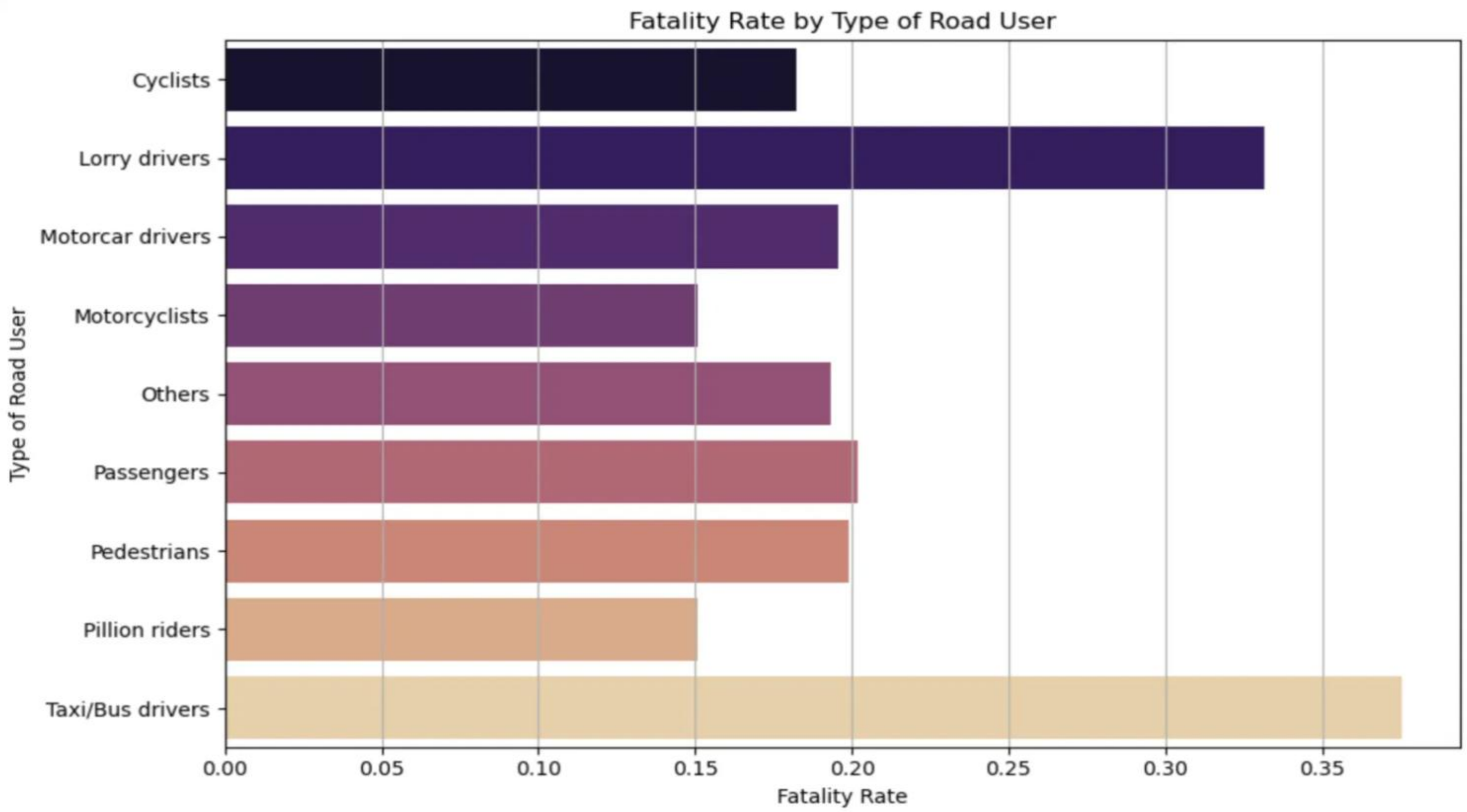
EDA:
Comprehensive visualization and analysis of the data will be conducted using graphical libraries. This step will help identify important patterns, trends, and insights within the data. Graphical representations will make it easier tounderstand the data’s behavior and identify any anomalies or significant trends.

Model Evaluation:
Machine learning models will be developed and trained to gain insights and learn complex patterns in the dataset.This involves selecting appropriate algorithms, tuning model parameters, and training the models on the preprocessed data to ensure they can accurately predict traffic accident and optimize it.

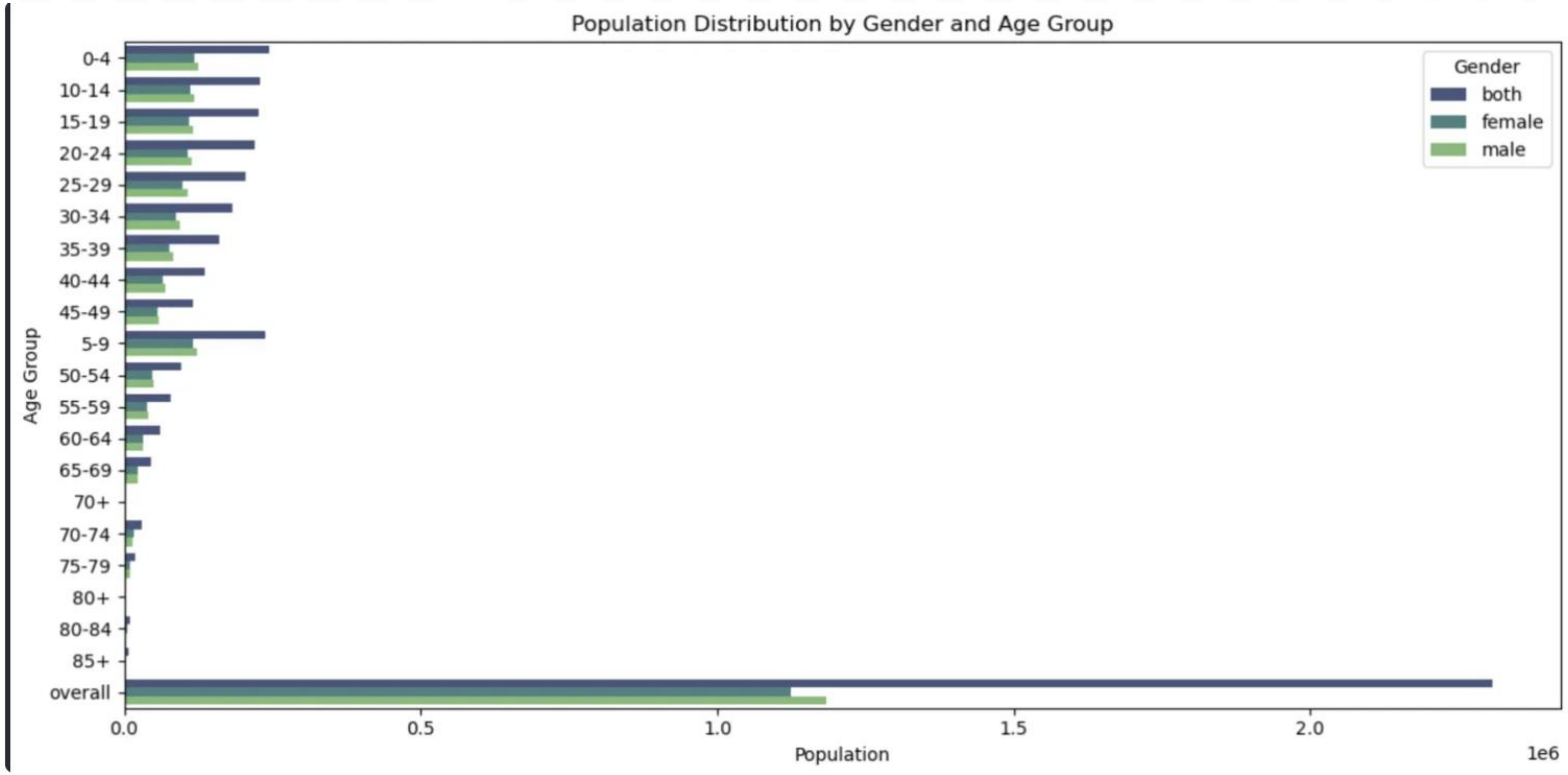
INITIAL RESULTS



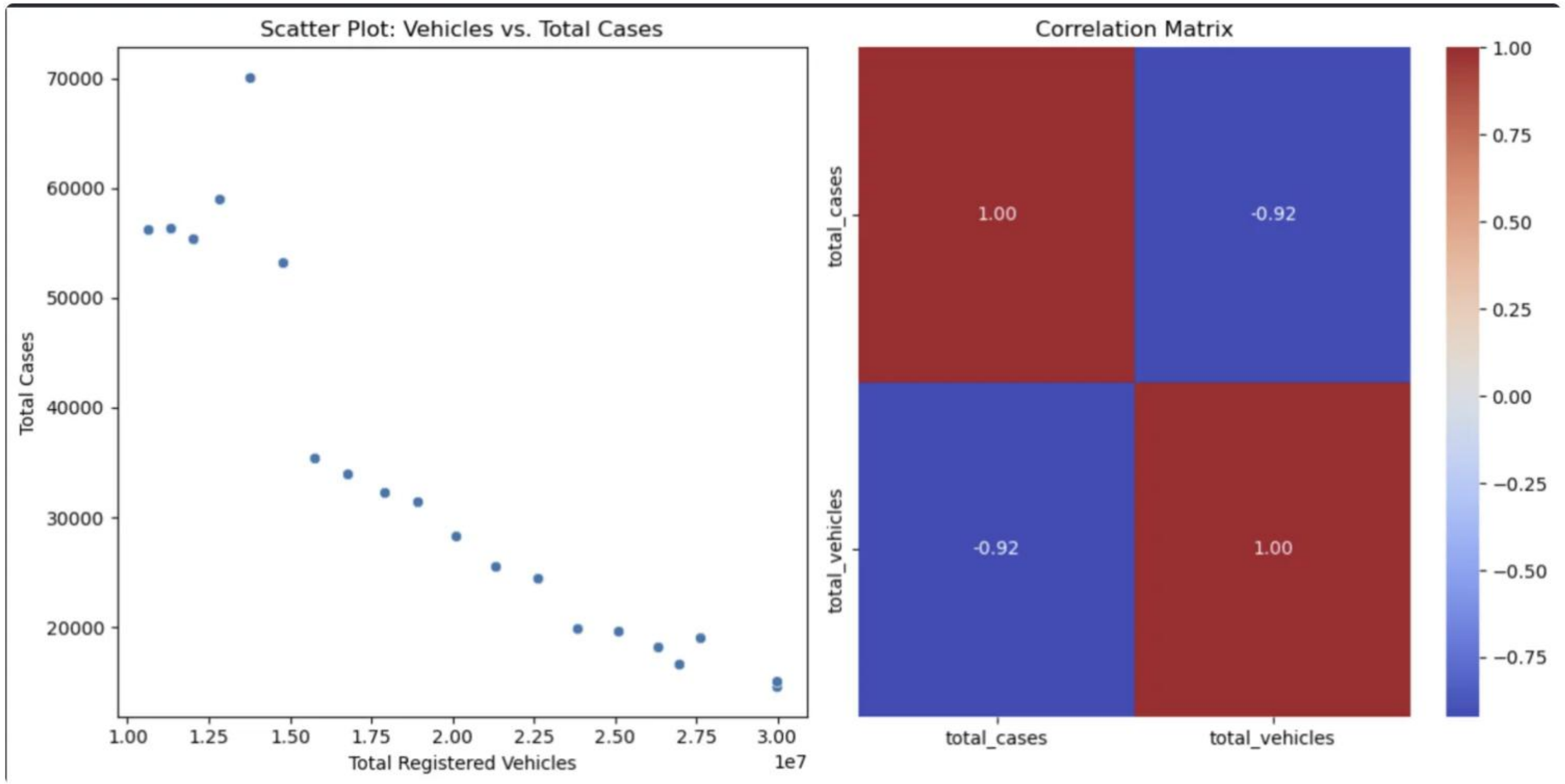
Comparison of traffic accident trends: From the data from 2000 to 2021, the number of deaths and injuries in traffic accidents shows an overall downward trend, especially the decline in fatalities, indicating that improvements in traffic management policies are having an effect. In some years, however, there have been short-term fluctuations in the number of deaths and injuries. For example, the number of deaths in some years is significantly higher than average, which can be related to bad weather, road construction, or increased traffic during holidays. These unusual fluctuations remind us that while the overall trend is positive, more stringent traffic control measures are still needed for specific hours.



Death rate among different road users: Taxi drivers have the highest death rate among all types of road users (about 0.35). This may be related to the long hours taxi drivers drive in cities, the complex traffic environment they face, and the higher work pressure. Second, the significantly higher fatality rates for motorcycles, passengers, and pedestrians than for other groups may be related to their lack of physical protection (as compared to cars, for example), while the higher fatality rates for pedestrians may reflect the greater vulnerability of pedestrians in vehicle-pedestrian collisions.



Distribution of accidents by sex and age group: It can be seen that the proportion of males in traffic accidents is significantly higher than that of females, especially in the young adults aged 20-40. Male drivers are more likely to engage in high-risk driving behaviors, such as speeding, illegal lane changes and drunk driving. In addition, young drivers (especially those aged 20-30) and the elderly (aged 60 and above) are two groups that need attention. Young people are more likely to be involved in accidents due to inexperience driving and risky behavior, while the elderly are a high-risk group due to reduced physical responsiveness.



With the increase of the total number of registered vehicles, the total number of traffic accidents shows a significant downward trend and a negative correlation. This may be because the areas with a higher number of vehicle registrations are usually economically developed areas, and these areas may have better traffic management systems and higher quality road infrastructure, which can effectively reduce traffic accidents. This phenomenon highlights the importance and complexity of traffic safety: the presence of more vehicles does not necessarily lead to more accidents, but may lead to fewer accidents due to better traffic management, driver awareness and infrastructure.

INITIAL INSIGHTS GAINED FROM EDA

The Random Forest model showed moderate predictive power with a 9.7% lower accuracy. Key factors included motorcycle rider proportion, peak hours, nighttime, and vehicle registrations. Pedestrian accidents highlighted safety infrastructure gaps. The findings stress targeted policies for safety, better lighting, and improved traffic management during high-risk times, providing a basis for refined models and measures.

SIGNIFICANCES OF THE FINDINGS

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

The study's results have practical significance. Firstly, it identifies motorcycles and walkers as high-risk groups for safety education. Secondly, it links time with accident risk, suggesting traffic regulation and infrastructure optimization. Also, the increase in vehicle registrations calls for safety measures. The accident prediction framework offers data support despite its limitations.

FUTURE WORKS

Expand the data range	• include environmental characteristics (such as weather and road conditions) and driving behaviors (such as speed and drunk driving records);
Explore deep learning methods	• such as LSTM or graph neural networks to capture the spatiotemporal dependence and complex patterns of traffic accidents;
Optimize feature engineering	• further refine time and location features, and improve model prediction accuracy;
Carry out regionalized studies	•focusing on the detailed characteristics and treatment measures of accident black spots. By combining more comprehensive data with advanced modeling techniques, accident prediction capabilities can be further improved to provide stronger support for accurate implementation of traffic safety policies.
Chart diversification	•Use a variety of visualizations to increase viewing

THANK YOU

univteknologimalaysia utm.my utmofficial