Yuan Ze University International Bachelor Program In Engineering Graduation Project

Analyzing and Predicting Scooter Accident Severity — A Multi-Dimensional Approach

Student : Man-Ying Yen Bo-Cyuan Chen Chamma Avrelus Advisor : Yun-Jia Liang Professor

Research Purposes

The purpose of our research is to conduct an in-depth analysis of the complex nature and severity of road traffic scooter incidents. By exploring various facets of these accidents, our objective is to gain a deeper understanding of the underlying dynamics and contributing factors.

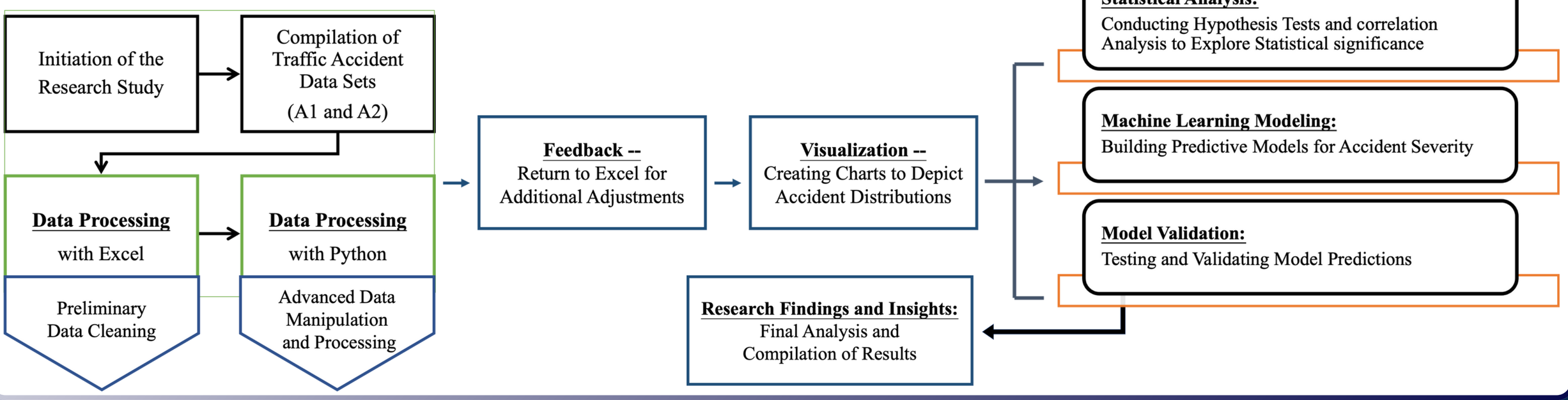
Our study strives to uncover patterns and correlations that can inform more targeted and efficient approaches to reducing traffic-related injuries and fatalities.

Research Methods

Our study employs a robust methodological framework, combining hypothesis testing, correlation analysis, and machine learning to dissect traffic accident dynamics.

We use **chi-square tests** for understanding relationships among categorical variables and conduct thorough **correlation analysis** to explore how driver age, gender, and accident characteristics interact across different conditions. In predictive analysis, we leverage advanced algorithms like **Logistic Regression**, **Decision Trees**, and **Neural Networks**, enabling us to effectively model and predict traffic scooter accident patterns and severity, thus providing a comprehensive view of traffic safety.

Research Process



Research Result

Hypothesis Testing Highlights:

- Gender significantly influences accident types ($\chi^2 = 267.54$, $p < 0.001$).
- Younger drivers (23-35 years) are predominantly involved in accidents.
- Strong correlation between road types and specific accident categories ($\chi^2 = 1141.80$, $p < 0.001$).

Demographic Trends:

- Younger drivers (23-35) most involved in accidents. Fig. 1

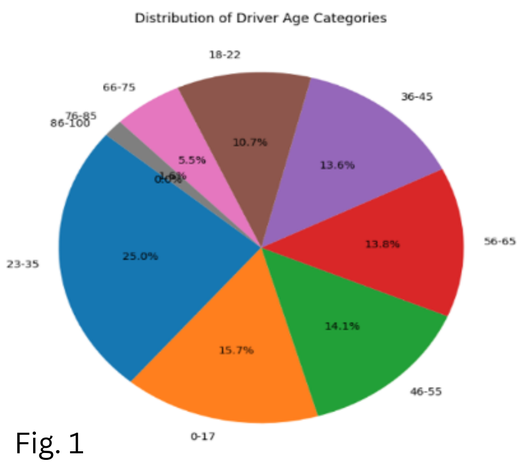


Fig. 1

Demographic Trends:

- Males (Gender) show higher accident rates. Fig. 2

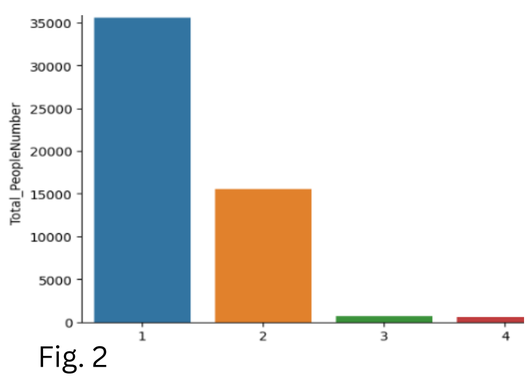


Fig. 2

District Trends:

- Shilin District records the highest number of fatal accidents. Fig. 3

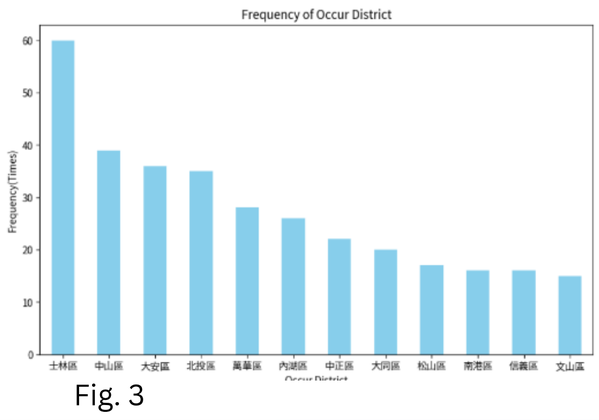


Fig. 3

Model Performance

Logistic Regression: 89.7% accuracy. Fig. 4

Neural Network: 90.07% validation accuracy.

Feature Significance:

Gender, vehicle type, and alcohol use are influential predictors.

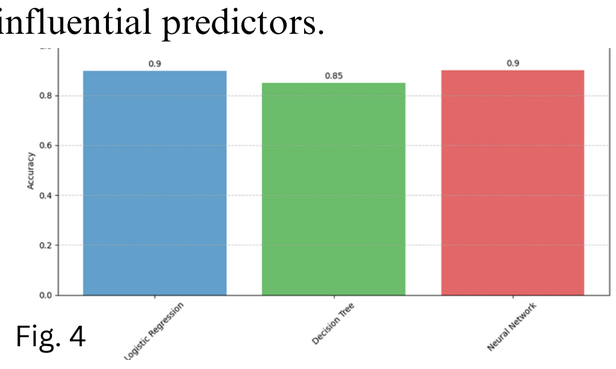


Fig. 4

The confusion matrices show that Logistic Regression excels at identifying 'Uninjured' cases, while the Decision Tree is effective but less distinct for 'Injured' cases. The Neural Network offers a balanced prediction across both categories, showcasing its pattern recognition strength. Fig. 5

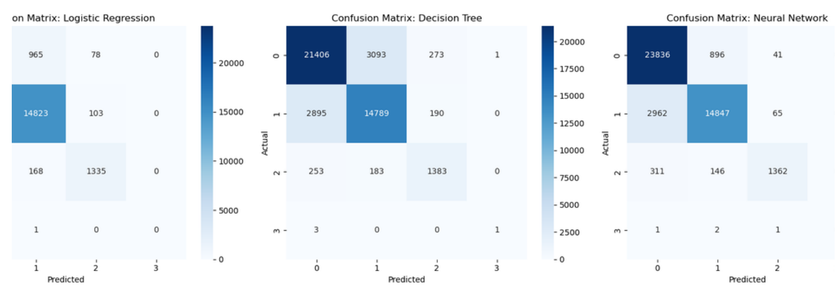


Fig. 5

Conclusion

- Key Takeaways: The study provides comprehensive insights into the factors influencing traffic accidents, emphasizing the need for targeted educational and infrastructural interventions.
- The predictive models demonstrate robust capabilities in classifying accident severity, with Logistic Regression showing high accuracy for non-fatal accidents and the Neural Network excelling in the fatal accident domain.
- Future Directions: Recommendations for ongoing research and adaptation of strategies to evolving traffic conditions, societal behaviors.