Uncovering Injury Risks in Micromobility: A Comparative Machine Learning Study

Introduction

As cities promote sustainable and flexible transport options, **micromobility vehicles**—such as mopeds and pedal cycles—are becoming increasingly integrated into urban mobility systems. While these modes provide environmental and congestion-reducing benefits, they also introduce new safety challenges due to users' physical vulnerability and limited protection in the event of a crash. Compared to occupants of enclosed motor vehicles, micromobility users are more exposed to environmental hazards and direct impact forces. Understanding how **injury severity manifests across different transport modes** is critical for developing tailored safety policies. This study investigates and compares the factors influencing serious injury outcomes for micromobility users versus other road users using a large-scale traffic collision dataset and interpretable machine learning models.

Methodology

Leveraging a comprehensive dataset spanning from 2013 to 2023, the analysis focuses on the 2023 subset of the data, which includes key variables such as transport mode, user demographics (age and gender), accident characteristics (weather, light conditions, number of vehicles involved), and geographical context (urban versus rural areas). The dataset was segmented into two groups: (1) incidents involving micromobility users (mopeds and pedal cycles), and (2) those involving all other vehicle types. A binary target variable representing the presence of serious injury was constructed by aggregating fatal and severe injury outcomes.

Descriptive statistics in table 1 revealed that micromobility users, while less likely to suffer fatal injuries (3.7% vs. 6.4%), were more frequently involved in collisions that resulted in serious injuries (mean = 0.50 vs. 0.31) and slight injuries (mean = 2.46 vs. 2.06). Additionally, the variability of injury counts (STD = 7.74 for slight injuries) suggests more complex or multi-injury scenarios among micromobility incidents. These patterns motivated the development of mode-specific models to better understand distinct injury risk mechanisms.

Table 1: *Descriptive statistics*

	Fatally Injured (at 30 days)		
	count	**mean**	**std**
Other Vehicles	264893	0.064	0.28
Micromobility	63859	0.037	0.20
	Seriously Injured (as reported)		
	count	**mean**	**std**
Other Vehicles	264893	0.31	0.87
Micromobility	63859	0.50	1.41
	Slightly Injured		
	count	**mean**	**std**
Other Vehicles	264893	2.06	4.88
Micromobility	63859	2.46	7.74

For each data subset, a **Random Forest classifier** was trained within a machine learning pipeline that integrated **Synthetic Minority Over-sampling Technique (SMOTE)** to address class imbalance. Hyperparameter optimization was performed using **GridSearchCV** with stratified cross-validation to enhance model generalizability. Model performance was rigorously evaluated using standard classification metrics, including **accuracy, F1-score, and detailed classification reports.**

To enhance interpretability and extract meaningful insights from the models, we applied **SHapley Additive exPlanations (SHAP)**. SHAP assigns feature importance values based on cooperative game theory principles, allowing us to quantify each predictor's contribution to the model's output for individual predictions. This approach enabled a detailed understanding of which factors most influenced the likelihood of serious injury in each user group.

Results

The micromobility model achieved **63.3% accuracy**, **0.67 precision**, **0.63 recall**, and an **F1-score of 0.64**. The model for other vehicle types yielded marginally better results with **64.7% accuracy**, **0.72 precision**, **0.65 recall**, and an **F1-score of 0.67**. These results indicate moderate but useful predictive performance, supporting downstream interpretability analysis.

Beyond prediction, the models served as analytical tools to assess **feature importance** in determining injury severity. To understand the key drivers behind serious injury outcomes, we applied **SHAP (SHapley Additive exPlanations)**. For micromobility users, **age** emerged as the most important feature, with a clear **U-shaped risk distribution**: young adults (18–24) and seniors (65+) were more likely to sustain severe outcomes. **Weather** (notably rain or unknown conditions) and **low light conditions** (e.g., darkness) were also strongly associated with higher injury severity. In contrast, for other transport modes, **urban setting** was the most influential predictor, followed by **weather conditions** and **male gender**, both contributing moderately to serious injury outcomes.

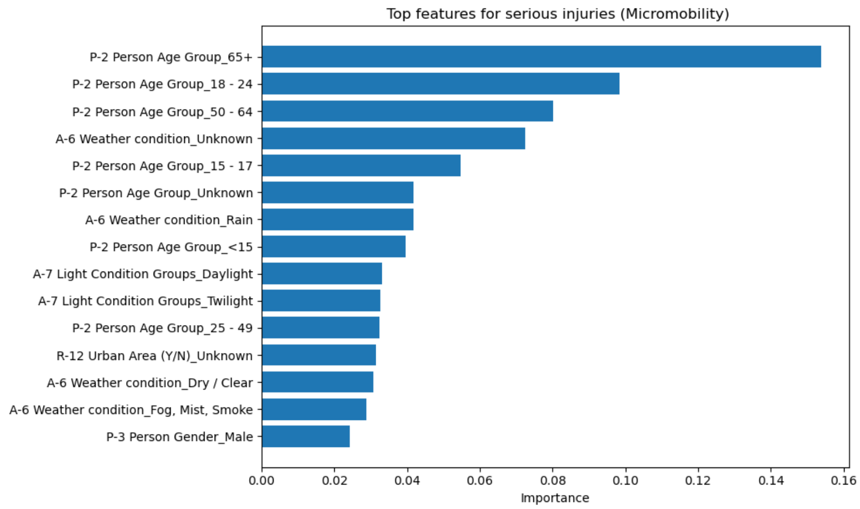

Figure 1: *Feature Importance Chart for micromobility*

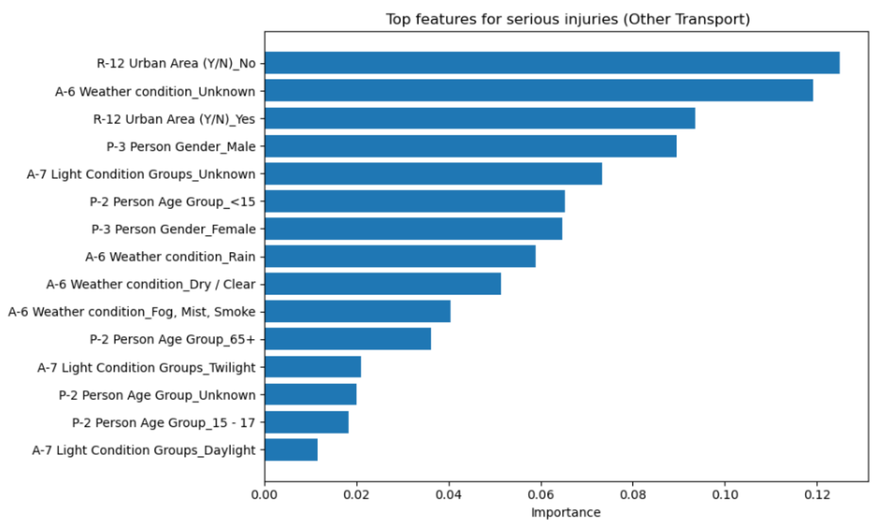

Figure 2: *Feature Importance Chart for other vehicles*

Figures 1 and 2 display the **top-ranked features** contributing to serious injury outcomes, as determined by feature importance scores from the Random Forest models, separately for **micromobility** and **other transport modes**.

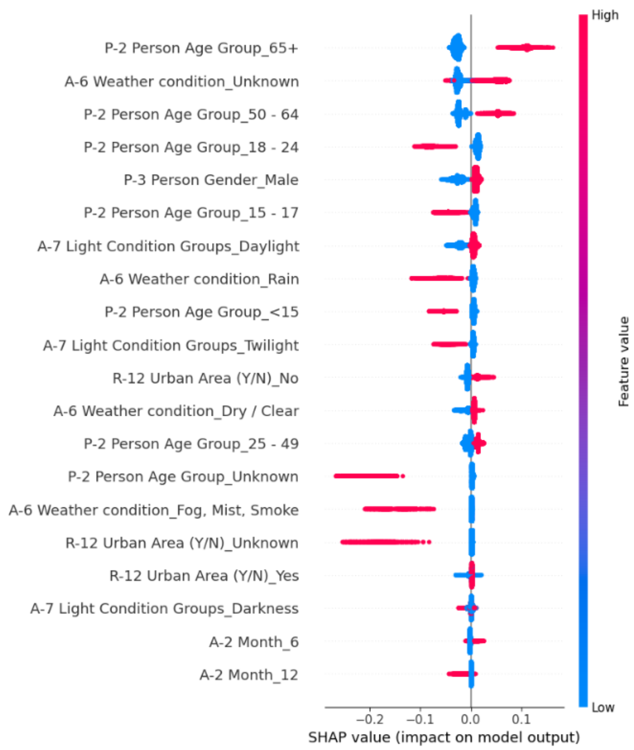

Figure 3: *SHAP Summary Plot for micromobility*

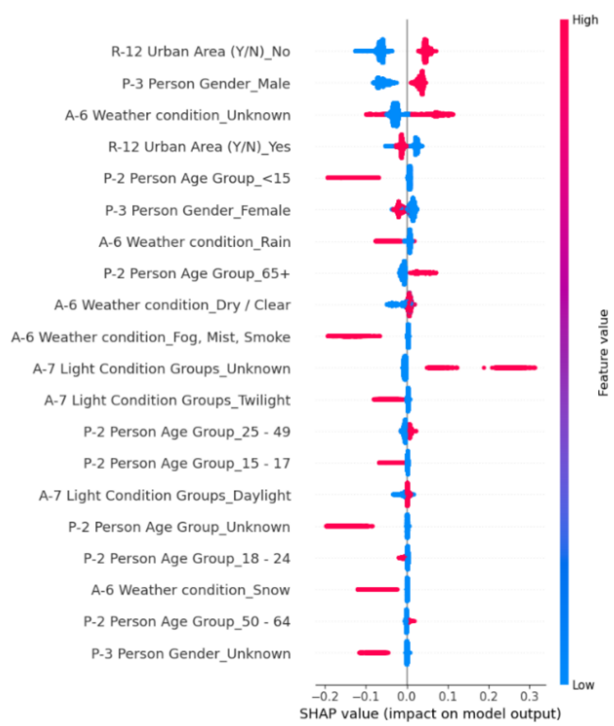

Figure 4: *SHAP Summary Plot for other vehicles*

To complement this, **Figures 3 and 4** present the corresponding **SHAP summary plots**, which visualize the magnitude and direction of each feature's contribution to individual predictions, offering deeper insight into how specific variables influence the model's classification decisions.

Discussion

These results reveal **distinct injury risk profiles** between micromobility users and traditional vehicle occupants. The elevated serious and slight injury rates among micromobility users—despite lower fatality rates—highlight **the increased vulnerability associated with limited physical protection**. The strong influence of age for micromobility users suggests a need for targeted interventions aimed at age-specific risks, such as training programs for young riders and adaptive infrastructure for older adults. Meanwhile, the dominant role of urban context and gender among other vehicle users underscores the importance of **location-sensitive planning** and **gender-informed policy design**.

Moreover, the SHAP-based interpretability confirms that injury severity is shaped not only by collision dynamics but also by demographic and environmental contexts, emphasizing the limitations of generic, one-size-fits-all road safety strategies.

Conclusion

This study provides evidence for the heterogeneous nature of injury severity determinants across transport modes. While micromobility users face greater risks from environmental exposure and demographic vulnerability, other vehicle users are more impacted by contextual and spatial factors. These findings support the development of mode-specific safety strategies, including infrastructure adjustments, educational campaigns, and policy reforms tailored to the unique risk mechanisms of each group. As micromobility adoption grows, integrating such evidence into safety planning is essential for ensuring equitable and effective protection across all user categories. Future work should explore behavioral variables and longitudinal injury patterns to further refine these insights.