Machine Learning-Based Categorisation of Central Roads in Athens

Paraskevi Koliou
Dr., Transport Engineer

Together with:
Stelios Peithis, George Yannis

Department of Transportation Planning and Engineering
National Technical University of Athens

12th European Conference on
Injury, Prevention and Safety Promotion
EU SAFETY 2025
1-2 October 2025 | Heraklion, Crete, Greece

EU | Safety 2025

Introduction

Why does it matter?

- Road crashes are a significant public health issue, with over 1.19 million annual fatalities worldwide
- Crash risk is shaped by driver behaviour (speeding, braking, acceleration), environmental conditions, and road infrastructure deficiencies
- Traditional crash analysis relies heavily on police reports and retrospective data, limiting proactive safety planning
- This study introduces a data-driven approach, integrating telematics, machine learning, and spatial analysis, to classify urban road segments into safe vs unsafe categories and guide proactive interventions

Objectives

1. Classify urban road segments by safety level
   - Develop a machine learning model to distinguish Safe vs. Unsafe segments for proactive risk assessment.
2. Integrate diverse data sources
   - Combine telematics (driver behaviours), crash records, and road network data for a holistic view of urban road safety.
3. Assess behavioural risk factors
   - Quantify the impact of speeding, harsh braking, and sudden acceleration on crash likelihood.
4. Address data imbalance in crash prediction
   - Apply resampling techniques (SMOTE) to ensure reliable detection of unsafe road segments.
5. Support evidence-based safety interventions
   - Translate predictive outputs into practical recommendations such as speed enforcement, better signage, and safer junction design.

Methodology

1. Study Area & Data Sources

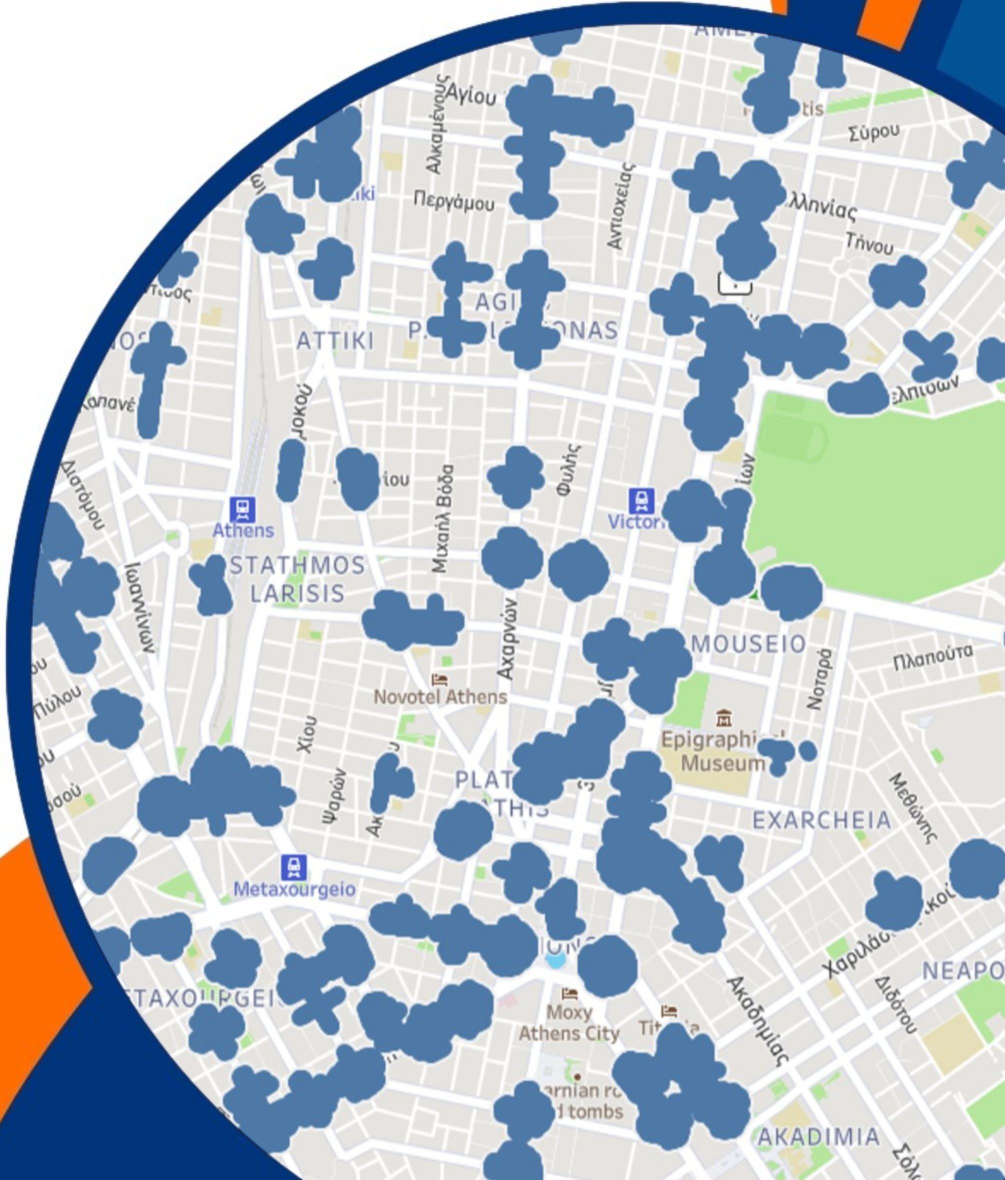
- Location: **Athens, Greece**.
- **Data inputs**:
  - **Crash data** from Greek authorities (geocoded via OSMnx).
  - **Telematics data**: 2,614 trips, 257 drivers, 8 municipalities.
  - **Road networks**: OpenStreetMap (OSM).

2. Data Processing

- Crash records aligned with **road segments**.
- Driving behaviour features (speed variability, braking frequency, acceleration).
- **Normalisation** to account for traffic volume/exposure variations.

3. Modelling Approach

- **XGBoost classifier** for risk classification.
- **SMOTE** to balance Safe vs Unsafe classes.
- Performance measured with **precision, recall, F1-score, k-fold cross-validation**.
- **Comparative models**: Logistic Regression, SVM (for benchmarking).



Results

1. Classification Outcomes:
- Safe F1-score = 0.86, Unsafe F1-score = 0.80.
- **Strong performance** across both classes

2. Feature Importance:
- **Top predictors** → Speed variability, braking frequency, acceleration patterns

3. Behavioral Correlations:
- **High crash counts** in areas with frequent harsh braking and sudden accelerations.
- **Roads with high congestion** → fewer acceleration spikes, but
- **Low-density roads** → more aggressive driving

4. Spatial Insights:
- Scatter plots confirm a **strong correlation** between crash counts and risky driving metrics
- Spatial heatmaps identified arterial roads and intersections as the **riskiest segments**

Table 1: Summary of Dataset Characteristics

Category	*Counts*
Count of Drivers	257
Count of Trips	2614
Count of Attica Municipalities	8

Table 2: Classification Report for XGBoost Model Performance

	Precision	Recall	F1-Score	Support
Safe	0.88	0.83	0.86	18
Unsafe	0.77	0.83	0.8	12

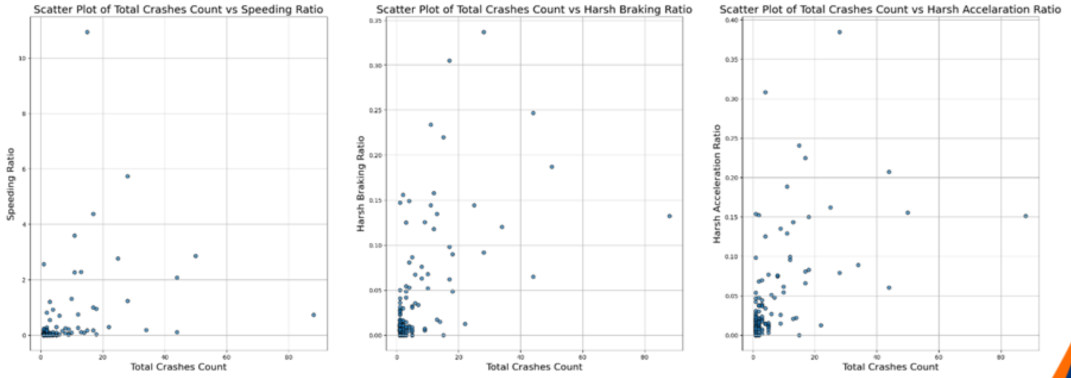


Figure 1: Scatter Plots of Total Crash Counts vs. Key Driving Behavior Metrics

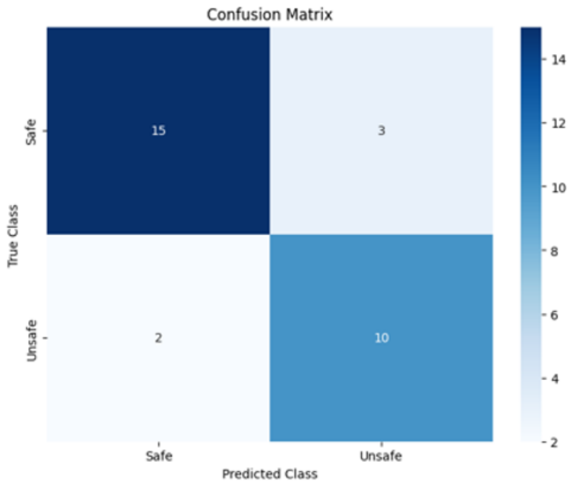


Figure 2: Confusion Matrix for XGBoost Model Performance

Koliou et al. , National Technical University of Athens

Conclusions

- **Telematics, ML, and Spatial Analysis** enable proactive crash risk classification.
- **Aggressive driving behaviours** are strongly linked to unsafe road segments.
- **Predictive modelling** supports targeted interventions: speed enforcement, signage, lane redesigns.
- **Practical applications:** real-time monitoring, driver feedback systems, smarter infrastructure investments.
- **Future work:** Incorporate environmental data (weather, lighting) and test across different cities for broader scalability.

Machine Learning-Based Categorisation of Central Roads in Athens

Paraskevi Koliou

Dr., Transport Engineer

Together with:

Stelios Peithis, George Yannis

Department of Transportation Planning and Engineering
National Technical University of Athens

Contact Information:
Dr. Paraskevi Koliou, Research Associate, NTUA
Department of Transportation Planning and Engineering
Email: evi_koliou@mail.ntua.gr
Website: https://www.nrso.ntua.gr/p/evikoliou/

EU | Safety 2025