Comparative Analysis of Machine Learning Models for Predicting Risky Driving Behaviors from a Naturalistic Experiment

Stella Roussou

PhD Candidate, Research Associate

Together with:
Virginia Petraki and George Yannis

Department of Transportation Planning and Engineering
National Technical University of Athens

Road Safety on Five Continents
20th international RS5C conference
Leeds, UK from 3–5 September 2025

Background

- Factors such as driver's state, environmental conditions, and traffic circumstances remain significant contributors to traffic collisions.
- Road crashes mainly result from driver behavior, but vehicle features also play a role.
- Previous research has demonstrated that the design and performance attributes of a vehicle can significantly impact how drivers respond to different driving conditions.
- Intelligent driving behavior monitoring systems enable real-time interventions and demonstrate remarkable efficacy in improving road safety
- Horizon 2020 i-DREAMS project: framework to define a Safety Tolerance Zone (STZ) and test interventions in real driving conditions.
- This study used naturalistic driving data from 48 drivers across 4,922 trips.
- This study focuses on predicting risky driving behaviors (harsh acceleration/deceleration) using machine learning models (CatBoost, Random Forest, XGBoost).

Objectives

- Compare predictive performance of ML models (CatBoost, XGBoost, Random Forest).
- Evaluate the influence of vehicle characteristics on risky driving behavior.
- Identify the most influential vehicle attributes (age, engine capacity, horsepower, fuel type, gearbox, brand).
- Assess the effect of phased interventions (baseline, post-trip feedback, gamification) on risky events.
- Provide evidence for data-driven safety measures and support personalized interventions to improve road safety.

Data Collection

Data sources:

- In-vehicle monitoring systems: **harsh acceleration, harsh deceleration**, speeding, distraction.
- **Driver questionnaires**: vehicle brand, model, age, fuel type, engine capacity, horsepower, gearbox.
- Dataset enriched with: safety-promoting goals (SPG), performance objectives (PO), intervention phase.
- Sample: **48 drivers in Greece; 4,922 trips recorded**.

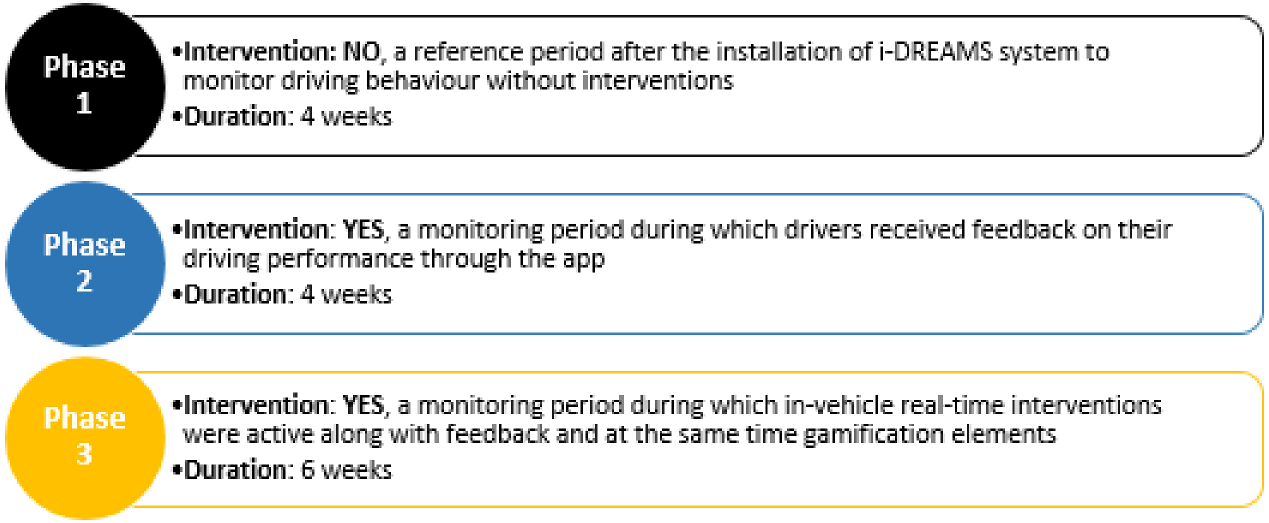


Figure 1: The i-DREAMS on-road experiment interventions

Safety Promoting Goals	Performance Objectives	Severity Level	Events per 100 km					
			Baseline Phase		Post-trip Phase		Post-trip & Gamification Phase	
			Mean	STD	Mean	STD	Mean	STD
Speed Management	Speeding	Medium	7	15	6	14	6	13
		High	25	33	24	32	21	29
Vehicle Control	Acceleration	Medium	3	10	4	13	2	9
		High	2	11	2	9	2	9
	Deceleration	Medium	6	15	6	14	5	13
		High	3	11	3	11	3	9
Driver Fitness	Distraction	n/a	23	60	23	57	17	48
Distance (km) per trip		n/a	9	15	8	15	10	18
Duration (min) per trip		n/a	17	15	17	14	18	15

Table 1: Descriptives of events per 100 km for SPG and PO

Machine Learning Models

CatBoost:

- Gradient Boosted Decision Trees
- Native handling of categorical features
- Robust, less preprocessing required.

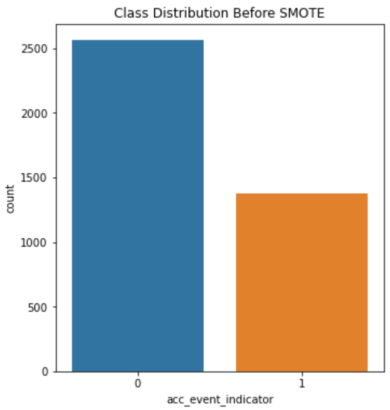
XGBoost:

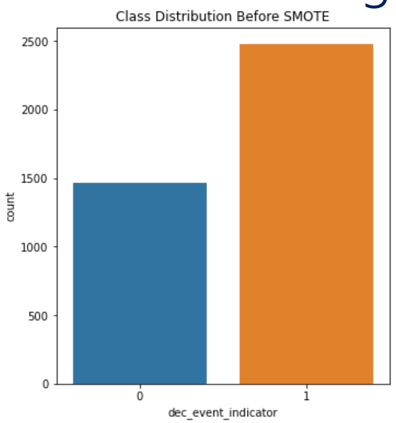
- Optimized Gradient Boosting
- Strong recall & speed
- Regularization reduces overfitting.

Random Forest:

- Ensemble of decision trees (bagging)
- Stable, easy to interpret
- Biased toward majority class in imbalance.

- Dataset was split by 80% training and 20% testing, preserving class distribution.
- Class imbalance addressed with SMOTE during model training.





Model Results

	0 (No acc. events)	1 (acc. events)	
precision	0.76	0.61	0.73
recall	0.83	0.50	0.72
f1- score	0.79	0.55	0.72
Accuracy			0.70
Macro avg	0.68	0.68	0.68
Weighted avg	0.71	0.71	0.71

Random Forest:

- Lower AUC (68% accel, 58% decel).
- Easy to interpret, but poor with rare risky events.
- Biased towards majority (non-risky) class.

XGBoost:

- Recall 72% (accel), 63% (decel).
- Sensitive to outliers but superior minority-class detection.
- Best at detecting risky behaviors.

CatBoost:

- Precision 76% (accel), 65% (decel).
- Stable across conditions (confirmed by cross-validation)
- Balanced performance, strong with categorical data.

	Predicted No acc. events	Predicted acc. events
Actual Non Acc. Events	538	109
Actual Acc. Events	170	168

Feature Importance

- Top predictors: vehicle age, vehicle brand, engine capacity, intervention phase.
- Most influential predictors for acceleration events:
  - Vehicle Age (older vehicles linked with risky accel. events).
  - Engine Capacity & Horsepower (higher values → more aggressive driving).
- Most influential predictors for deceleration events:
  - Intervention Phase (baseline vs. feedback vs. gamification), followed by
  - Vehicle Brand and Engine CC.
- Other relevant factors: gearbox type, fuel type.

Figure 1: Feature Importance Plot for Acceleration Events

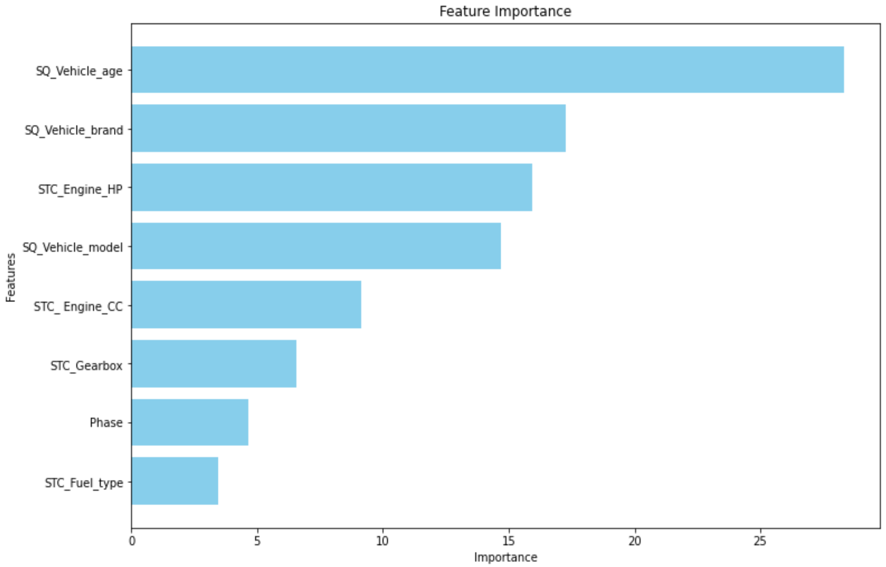
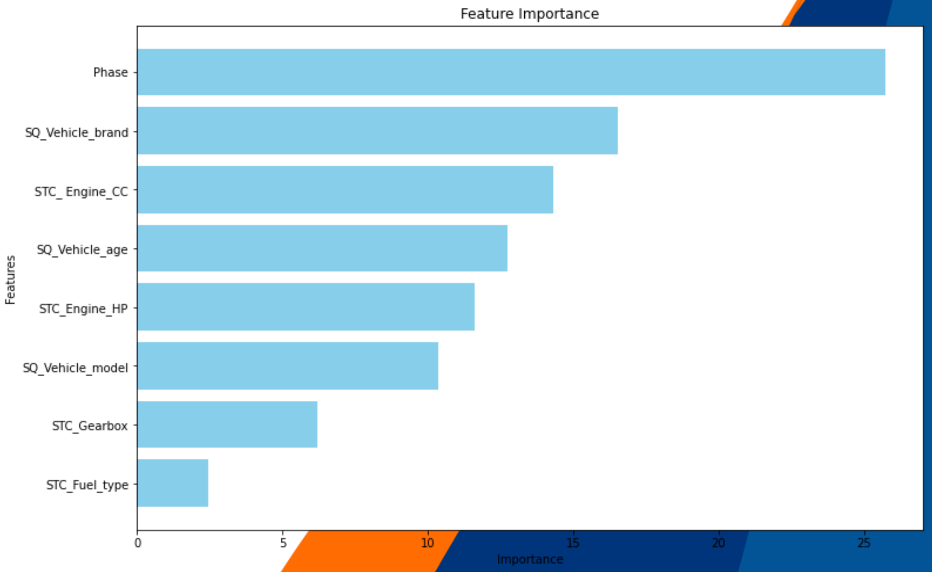


Figure 2: Feature Importance Plot for Deceleration Events



Hyperparameter	Examined range	Optimized Value
1	Vehicle Age	28.323
2	Vehicle Brand	17.258
3	Horsepower	15.928
4	Vehicle Model	14.700
5	Engine CC	9.158
6	Gearbox Type	6.543
7	Phase	4.646
8	Fuel type	3.440

Table 1: Feature Importance Plot for Acceleration Events

Hyperparameter	Examined range	Optimized Value
1	Phase	25.730
2	Vehicle Brand	16.547
3	Engine CC	14.330
4	Horsepower	12.727
5	Vehicle Age	11.615
6	Vehicle Model	10.374
7	Gearbox Type	6.221
8	Fuel type	2.452

Table 2: Feature Importance Plot for Deceleration Events

Discussion

- CatBoost: best balance of precision and accuracy.
- XGBoost: most effective at identifying risky behavior.
- Random Forest method is easy to interpret, but weaker on imbalanced data.
- Post-trip feedback and gamification reduced risky events.
- Vehicle characteristics significantly influence risky driving behaviors.
- ML models provide richer insights than traditional methods for road safety.
- Results confirm that combining vehicle data with interventions can guide effective safety policies

Conclusion & Future Work

- Machine Learning enables targeted vehicle safety interventions.
- XGBoost and CatBoost show promise in real-world risk prediction.
- **Vehicle age and performance attributes** (engine, horsepower) strongly linked to risky events.
- Phased interventions (especially gamification) improve safe driving.
- For Future Work, **expansion of the dataset** (more drivers, diverse vehicle types, varied road conditions) is needed.
- Integrate real-time traffic, **weather**, and driver state factors.
- Explore **deep learning** (LSTMs, transformers) for sequential driving behavior.
- Embed predictive models into ADAS & insurance risk profiling tools.

Comparative Analysis of Machine Learning Models for Predicting Risky Driving Behaviors from a Naturalistic Experiment

Stella Roussou
PhD Candidate, Research Associate

Together with:
Virginia Petraki and George Yannis

Department of Transportation Planning and Engineering
National Technical University of Athens

Road Safety on Five Continents
20th international RS5C conference
Leeds, UK from 3–5 September 2025