TRA

TRANSPORT RESEARCH ARENA BUDAPEST

18-21/05/26

Factors Influencing Speed Limit Violations on Athens Road Network

Daphne Kyprouli,
Dimitrios Nikolaou,
George Yannis

19/05/2026

Introduction and Background

Study Objective

Analyse the factors influencing **speeding behavior** in the Athens road network

- Develop a tool to identify **critical parameters** and support **targeted road safety interventions**
- Investigate the relationship between **speeding and crash risk** using statistical models
- Propose measures to improve **compliance with speed limits** and reduce road accidents

Introduction 1/2

Speed Limit - Speeding

- Speeding is defined as driving **above the posted speed limit**
- It is a key factor in road crashes, affecting both **frequency and severity**
- It is considered one of the **‘fatal five’ risk factors** in road accidents
- Reducing the average speed leads to a decrease in crash risk and injury severity

Reasons for Speeding

- Driver-related factors
- Vehicle-related factors
- Road conditions
- Social and psychological factors

Introduction 2/2

Factors Affecting Road Safety

- Speeding
- Driver distraction
- Non-use of protective equipment (e.g., seat belts, helmets)
- Dangerous and aggressive driving (e.g., harsh acceleration, harsh braking)
- Driving under the influence of alcohol or drugs
- Adverse weather conditions

Data Collection

Data Collection 1/2

Sources: OSeven Telematics & OpenStreetMap

- Naturalistic real-time driving data via smartphone sensors:
  - GPS (speed, location)
  - Accelerometer, magnetometer & gyroscope
- Automatic trip detection (start/stop based on driving activity)

Data Processing & Analysis

- Central processing using Machine Learning algorithms

  Retains only information relevant to road safety and eco-driving
- Multi-level analysis transforms raw data into high-precision insights, including:

  Risk exposure metrics (peak-hour driving, total trip duration)

  Driving behavior indicators (harsh acceleration, harsh braking, turning intensity)

Data Collection 2/2

35,282 road segments in the Athens road network including:

- **Key variables** (ID, Lanes, Length, Efficiency)
- **Morphological characteristics indicators** (Slope, Trip count, Angle rate)
- **Safety and protective equipment usage indicators** (Seatbelt use, Helmet use)
- **Driving behavior indicators** (Mobile usage count, Speeding count)
- **Accident and severity indicators for 2016–2020** (Total crashes, Fatalities, Serious injuries)
- **Environmental footprint indicators** (CO_2, CO emissions)

Methodology

Methodology 1/2

Data Analysis Methods:

- **Binary Logistic Regression** → predicts probability of binary outcomes
- **Naïve Bayes** → probabilistic, assumes independent predictors
- **Decision Tree Regression** → recursive partitioning of observations
- **Random Forest** → ensemble of trees for better stability
- **Support Vector Machine (SVM)** → finds optimal separating hyperplane (supports nonlinear kernels)

Methodology 2/2

Data Processing:

Descriptive statistics → basic metrics

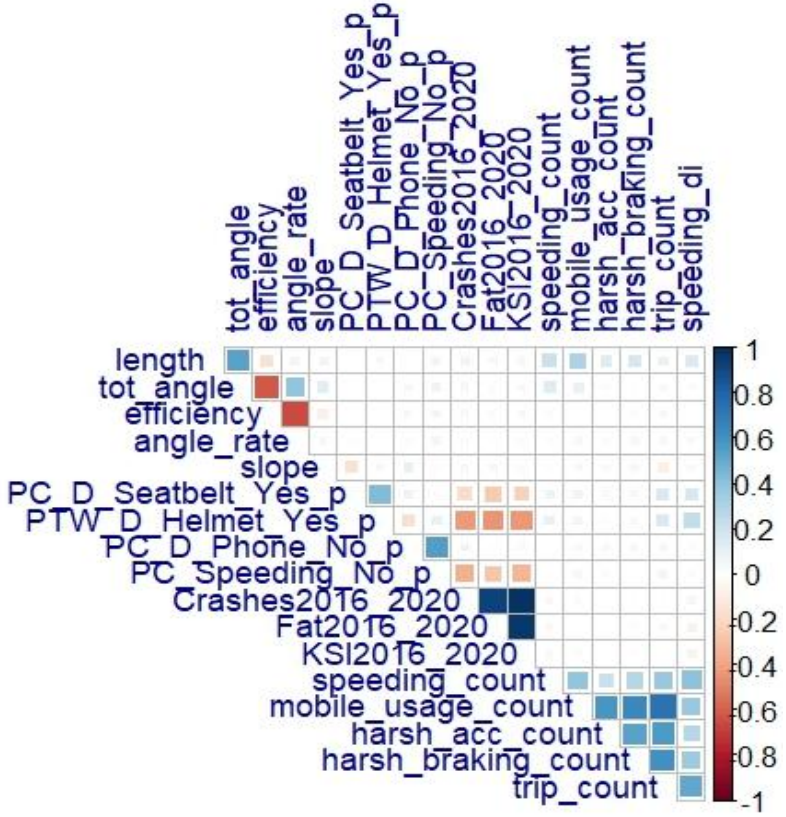
Data visualization → boxplots

Model Evaluation:

Dataset split: 80% training / 20% testing

Metrics:

- Correlation
- Statistical Significance
- **Accuracy → overall correct predictions**
- **Sensitivity → true positives detected**
- **Specificity → true negatives detected**



Model Selection
Balance between accuracy, sensitivity, specificity

Results

Classification models 1/5

The classification models were evaluated under two scenarios:

- **Scenario 1**:

Speeding indicator, where 0 = no speeding was recorded on the road segment,

1 = speeding was recorded for at least one second.

- **Scenario 2:**

Speeding indicator, where 0 = speeding duration on the road segment was less than 15.55 seconds,

1 = speeding lasted at least 15.55 seconds.

The value of 15.55 seconds corresponds to the mean speeding duration across all examined road segments.

Classification models 2/5

Independent Variables

length: road length

efficiency: road alignment index

slope_class: road slope category

highway: road type

mobile_usage_count: mobile phone usage (sec)

harsh_acc_count: number of harsh accelerations

harsh_braking_count: number of harsh brakings

trip_count: number of trips

Dependent Variable

speeding_di: speeding indicator

Classification models 3/5

Name	Estimate	Pr(>|z|)
(Intercept)	-37.359.162	< 2e-16 ***
length	0.0023779	< 2e-16 ***
efficiency	0.6267127	0.039356 *
slope_class0-3	10.487.670	1.96e-07 ***
slope_class3-5	0.9506093	3.30e-06 ***
slope_class5-8	0.7791038	0.000223 ***
slope_class8-10	0.7992437	0.001297 **
highwaymotorway	30.019.741	< 2e-16 ***
highwaymotorway_link	20.316.959	< 2e-16 ***
highwayprimary	10.810.562	1.15e-14 ***
highwayprimary_link	18.766.039	< 2e-16 ***
highwayresidential	0.0699595	0.554660
highwaysecondary	0.4190342	0.001174 **
highwaysecondary_link	15.609.873	< 2e-16 ***
highwaytertiary	0.7164621	6.10e-09 ***
highwaytertiary_link	0.7323863	0.007577 **
highwaytrunk	16.009.920	1.72e-05 ***
highwaytrunk_link	14.009.141	0.020153 *
mobile_usage_count	-0.0039140	4.19e-12 ***
harsh_acc_count	0.0283877	0.036268 *
harsh_braking_count	0.0781338	2.24e-12 ***
trip_count	0.0031773	< 2e-16 ***

Name	Estimate	Pr(>|z|)
(Intercept)	-62.446.160	2.05e-13 ***
length	0.0028611	< 2e-16 ***
efficiency	10.822.690	0.10508
slope_class0-3	0.1233906	0.75311
slope_class3-5	0.3118467	0.43216
slope_class5-8	-0.0704221	0.86465
slope_class8-10	0.2702110	0.58874
highwaymotorway	41.106.006	< 2e-16 ***
highwaymotorway_link	31.497.470	1.73e-13 ***
highwayprimary	20.847.150	8.02e-07 ***
highwayprimary_link	22.768.847	7.75e-07 ***
highwayresidential	-0.1735704	0.67717
highwaysecondary	20.414.313	9.63e-07 ***
highwaysecondary_link	25.713.696	1.91e-08 ***
highwaytertiary	18.642.344	7.03e-06 ***
highwaytertiary_link	21.618.771	0.00027 ***
highwaytrunk	38.601.927	4.75e-14 ***
highwaytrunk_link	-95.131.345	0.95428
mobile_usage_count	-0.0037988	2.81e-12 ***
harsh_acc_count	-0.0168832	0.20498
harsh_braking_count	0.0610185	7.96e-10 ***
trip_count	0.0031843	< 2e-16 ***

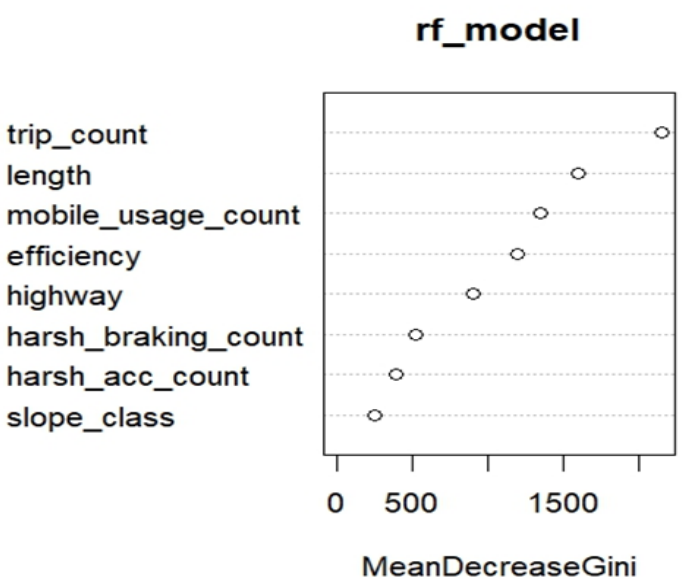
The tables show the results of the binary logistic regression.

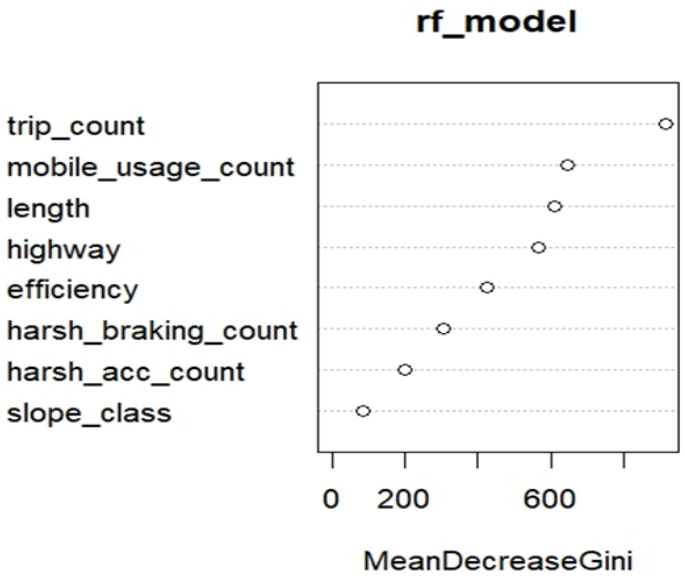
Estimate (+/-): Indicates whether the variable increases (+) or decreases (-) the probability of the event.

Pr(>|z|): The p-value shows the statistical significance.

Symbols (*): The more symbols, the more significant the effect of the variable

Classification models 4/5





Main Findings:

Trip count: most influential variable in both scenarios.
Harsh events & slope class: consistently low importance

Scenario comparison:

Scenario 1 → **road length & efficiency** slightly less important.

Scenario 2 → slight increase in importance of **mobile phone use**.

Classification models 5/5

Scenario 1

	Binary Logistic	Random Forest	Decision Tree	SVM	Naive Bayes
Accuracy	0.778	**0.784**	0.771	0.781	0.767
Sensitivity	0.355	**0.409**	0.330	0.319	0.323
Specificity	0.942	0.928	0.941	**0.958**	0.938

Scenario 2

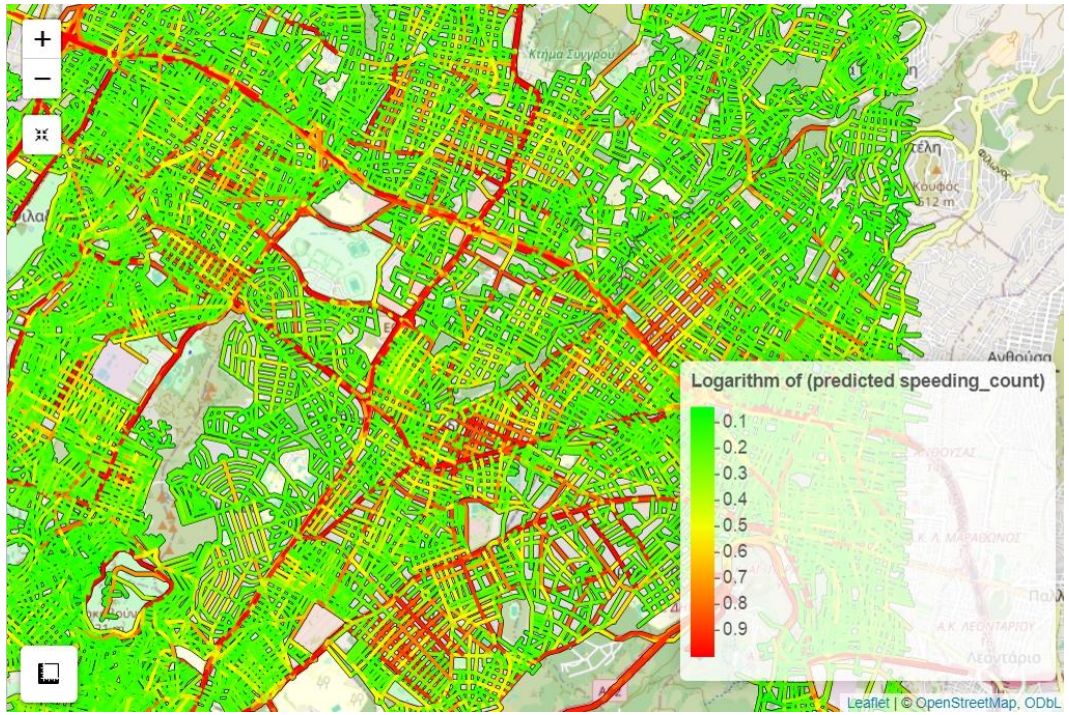
	Binary Logistic	Random Forest	Decision Tree	SVM	Naive Bayes
Accuracy	0.941	**0.945**	0.938	0.939	0.916
Sensitivity	0.355	0.392	0.265	0.275	**0.486**
Specificity	0.989	0.989	0.993	**0.994**	0.951

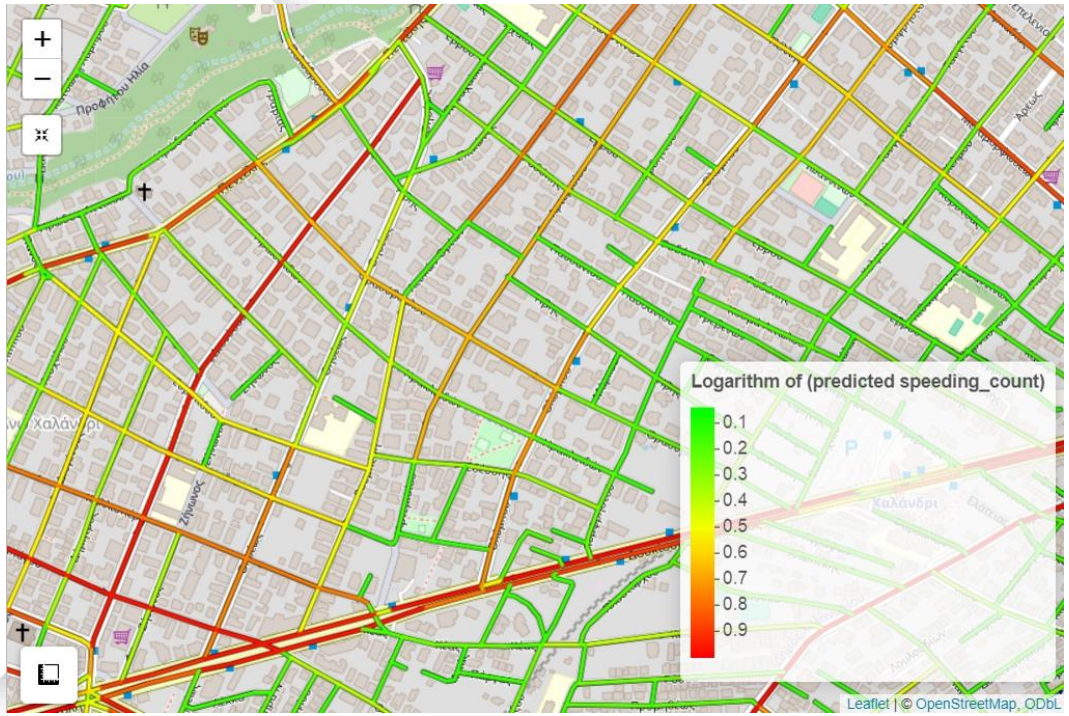
Map Visualization

Map Visualization

Model Used: **Binary Logistic Regression,** estimates probabilities based on linear combinations of variables

Scenario Used: **Scenario 1,** offers a balance between interpretability, practical value, and coverage of speeding events





Conclusion

Key Findings

- **Speeding** is a critical determinant of road safety, reduces reaction time, increases crash severity
- **Positively associated** with trip count, road length
- **Negatively associated** with mobile phone use, may reflect compensatory cautiousness, although distraction remains risky
- **Harsh acceleration / braking** linked to speeding due to higher reaction demands at high speeds
- **Random Forest outperformed** other models and identified the most important predictors
- **Low sensitivity:** difficulty detecting speeding events due to class imbalance

Policy & Practical Implications

Targeted interventions to reduce speeding:

- Adaptive speed management systems
- Telematics-based driver feedback
- Awareness campaigns on speeding & distracted driving

Interactive speeding map:

- Identifies high-risk road segments
- Supports evidence-based, location-specific safety measures

THANK YOU!

Kyprouli Daphne

TRA TRANSPORT RESEARCH ARENA BUDAPEST 18-21/05/26 Re-Generation in transport