Leveraging Machine Learning Algorithms to Predict and Analyze Single-Vehicle and Multi-Vehicle Crash Occurrences on Motorways

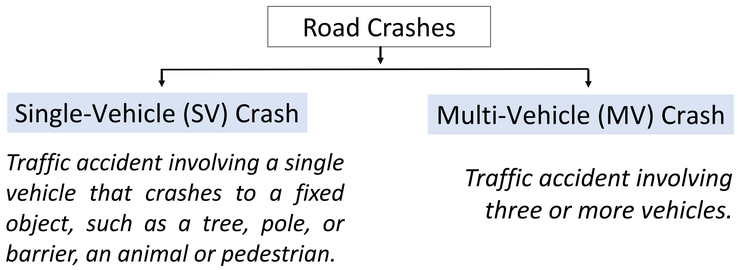
Saumik Sakib Bin Masud, Kirti Mahajan, Alexandra Kondyli, Katerina Deliali, George Yannis

KU School of Engineering, The University of Kansas

National Technical University of Athens

INTRODUCTION & MOTIVATION

Road crashes occur in all parts of the world causing significant loss of life, injury, and economic damage.



- ❖ Only a few previous studies have individually examined crash determinants for SV and MV crashes using ML, and fewer have compared prediction abilities across different modeling approaches.
- ❖ Moreover, the majority of previous research focused on predicting accident severity rather than crash occurrences.
- ❖ Additionally, none of the previous studies solely focused on roadway design parameters as crash determinants for SV and MV crashes.

OBJECTIVES

1. Developing separate prediction models for SV and MV crashes and identify prominent factors contributing to each crash type.
2. Comparing the classification accuracy of different models using ML and statistical analysis techniques.

DATA OVERVIEW

Crash records from the year **2015 to 2020**

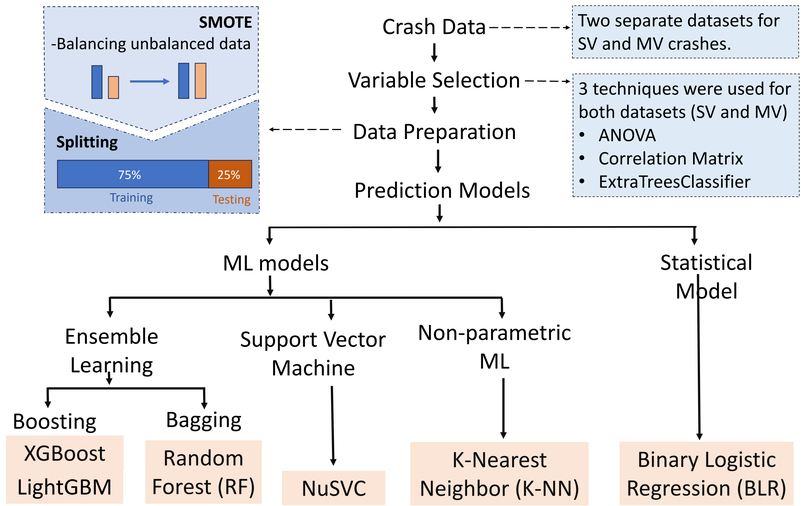
Olympia Odos motorway, a rural motorway located in southern Greece

- The dataset contained **1,306** road segments where **946 SV** and **492 MV** crashes occurred in six years.
- Contained information related to-

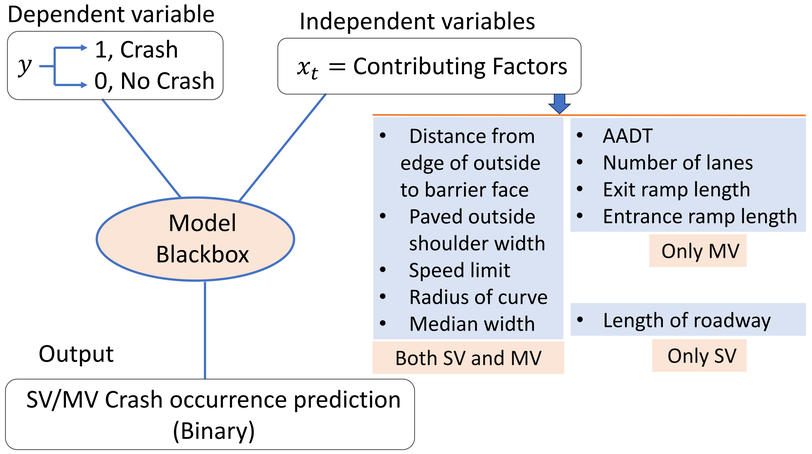
Traffic flow measurement (AADT) | Roadway design parameters

Three degrees of Accident Severity (PDO, Injury, Fatal)

METHODOLOGY



❖ Model Summary:



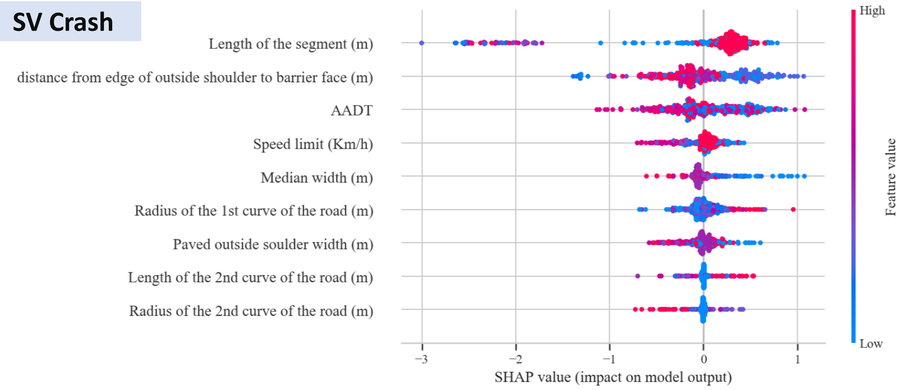
RESULTS

Model Type	Model names	MV Crash Prediction Performance metrics			SV Crash Prediction Performance metrics		
		Accuracy	ROC-AUC	F1 score	Accuracy	ROC-AUC	F1 score
Boosting-based ensemble ML	LightGBM	0.74	0.82	0.74	**0.76**	**0.82**	**0.76**
Boosting-based ensemble ML	XGBoost	**0.75**	**0.83**	**0.76**	0.75	0.80	0.76
Bagging ensemble ML	Random Forest (RF) Classifier	0.71	0.74	0.71	0.73	0.81	0.73
Support Vector Machine	NuSVC	0.69	0.69	0.69	0.65	0.66	0.65
Non-parametric ML	K-Nearest Neighbor (K-NN)	0.67	0.66	0.66	0.62	0.63	0.63
Statistical model	Binary Logistic Regression (BLR)	0.64	0.71	0.64	0.59	0.61	0.59

FINDINGS

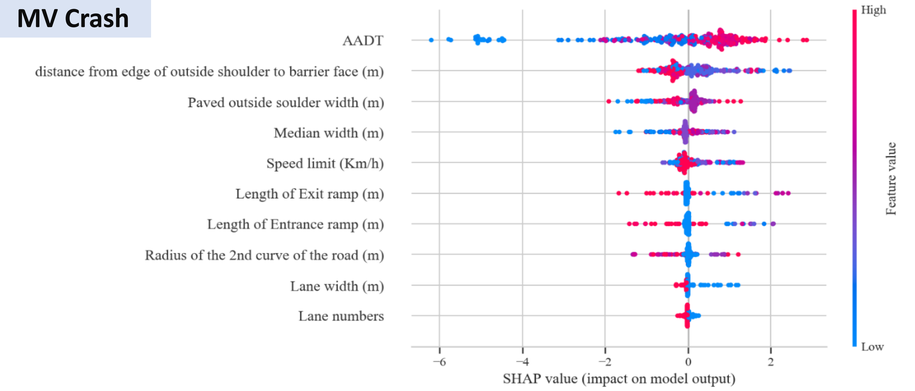
❖ Prominent Factors of SV and MV Crash Occurrence (SHAP Analysis):

SV Crash



- Longer length of road segment, speed limit, radius of 1st curve of the road, and the length of the 2nd curve of the road, increase the probability of SV crash occurrences on motorways (positively correlated).
- Higher AADT, distance from the edge of outer shoulder to barrier face, median width, paved outside shoulder width, and radius of 2nd curve of the road decrease the SV crash occurrences (negatively correlated).

MV Crash



- Higher AADT, speed limit, and wider median width increase the probability of the MV crash occurrences on motorways.
- Higher number of lanes and wider lane width decrease the MV crash occurrences.
- Higher length of exit and entrance ramp, and the curvature of the road decrease the MV crash occurrences on rural motorways.

KEY FINDINGS

- ➢ All the ML models outperformed the BLR model, and among the ML models, the boosting-based algorithms (LightGBM and XGBoost) showed the best performances.
- ➢ AADT, median width, and shoulder width showed the opposite contributions to the SV and MV crashes.
- ➢ The contribution of some variables was not found to be straightforward positive and negative, rather the combinations of other variables were necessary to extract the exact relationship. For example, the effect of shoulder width on MV crashes also depended on the speed limit and the number of lanes of that facility.

Saumik Sakib Bin Masud, M.Sc, EIT
Department of Civil, Environmental and Architectural Engineering, University of Kansas
Email- saumiksakib.masud@ku.edu

Kirti Mahajan, PhD
Department of Civil Environmental and Architectural Engineering, University of Kansas
Email: kirtimahajan@ku.edu

Alexandra Kondyli, PhD
Department of Civil Environmental and Architectural Engineering, University of Kansas
Email: akondyli@ku.edu

Katerina Deliali, PhD
Department of Civil Engineering National Technical University of Athens, Athens, Greece.
Email: kdeliali@mail.ntua.gr

George Yannis, PhD
Department of Civil Engineering National Technical University of Athens, Athens, Greece.
Email: geyannis@central.ntua.gr