An Roinn Iompair
Department of Transport

European Commission

Bonneagar Iompair Éireann
Transport Infrastructure Ireland

UCD DUBLIN

Enterprise Ireland

IDA Ireland

CONFERENCE PARTNERS INTERNATIONAL

TRA
TRANSPORT RESEARCH ARENA

10th Transport Research Arena Conference
Advancing Sustainable and Inclusive Mobility
Dublin, Ireland, April 15-18, 2024

Predicting risky driving behavior with classification algorithms: Results from a large-scale field-trial and simulator experiment

Thodoris Garefalakis
Transportation Engineer, Research Associate

Together with:
Eva Michelaraki, Stella Roussou, Christos Katrakazas, Tom Brijs, George Yannis

UHASSELT

The i-DREAMS project

- 13 Project partners:
  - National Technical University of Athens

    Universiteit Hasselt, Loughborough University, Technische Universität München, Kuratorium für Verkehrssicherheit, Delft University of Technology, University of Maribor, OSeven Telematics, DriveSimSolutions, CardioID Technologies, European Transport Safety Council, POLIS Network, Barraqueiro Transportes S.A.

- Duration of the project:
  - 48 months (May 2019 – April 2023)

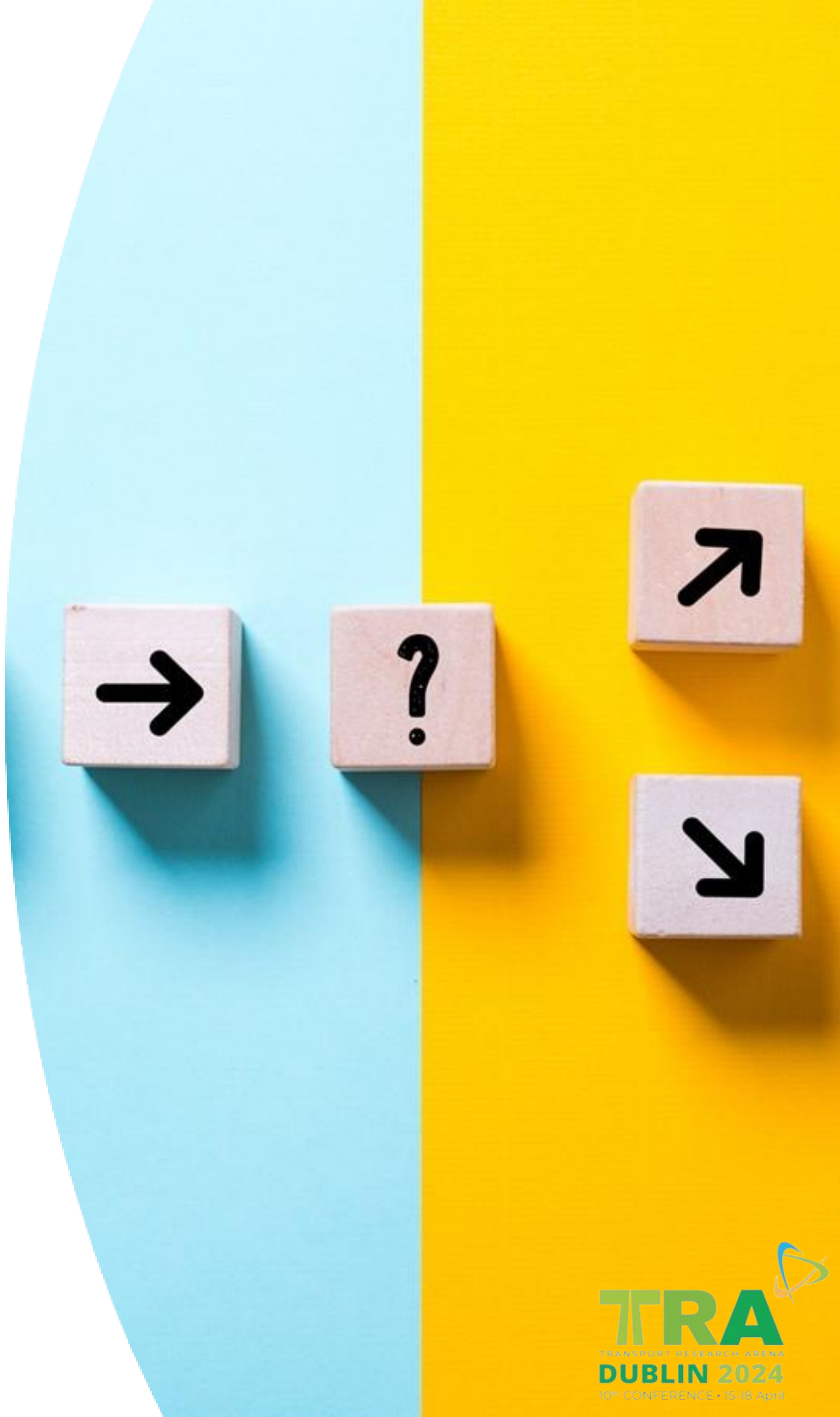
- Framework Program:
  - Horizon 2020 - The EU Union Framework Programme for Research and Innovation - Mobility for Growth

Introduction

- Road crashes are the leading cause of death for children and young people aged 5 to 29 (WHO)
- Driving behavior is a key factor in over 90% of these crashes
- Cognitive processes (i.e., attention, perception, decision-making are crucial for drivers to adapt to changing road conditions and make split-second decisions
- The integration of emerging technologies presents an opportunity to mitigate human errors and enhance road safety by reducing cognitive overload

Objectives

- Implementation of a dual-source methodology (e.g., simulator driving experiment and naturalistic driving experiment) to enhance the diversity of the dataset and the robustness of the model
- Investigation of key risk factors for road safety such as average speed, harsh acceleration, harsh braking, headway, overtaking, mobile phone use, and fatigue
- Development and evaluation of various classification models to predict risky driving behaviour



Data Description (1/2)

- **Simulator Driving Experiment dataset**
  - The simulator experiment involved 36 car drivers from Belgium and conducted from December 2020 to January 2021
  - The experiment comprised three scenarios, each with distinct road sections, number of lanes, and speed limits
  - Each participant completed three drives:
    - Drive 1 without interventions,
    - Drive 2 with interventions,
    - Drive 3 with interventions and modifying conditions
  - Key driving behaviour characteristics were collected

Scenario	Road Section	Number of lanes	Speed Limits
A	0-6300 m	1x1	70 km/h
	6300-11300 m	2x2	90 km/h
	11300-16500 m	2x2	120 km/h
B	0-6100 m	2x2	90 km/h
	6100-12000 m	2x2	120 km/h
	12000-18200 m	1x1	70 km/h
C	0-6000 m	2x2	120 km/h
	6000-11000 m	2x2	90 km/h
	11000-17200 m	1x1	70 km/h

Data Description (2/2)

- **Naturalistic Driving Experiment dataset**
  - The naturalistic driving experiment consisted of for four phases, involving 51 car drivers from Belgium conducted between February to September 2022
  - A large database consisting 7,163 trips (147,337 minutes) was collected and analyzed
- For both datasets, similar driving behaviour characteristics (such as Speed, Headway, TTC, etc.) were collected

Phase 1 (Baseline)
- **Intervention:** No
- **Description:** a reference period after the installation of the i-DREAMS system in order to monitor driving behavior without interventions
- **Duration:** 4 weeks

Phase 2
- **Intervention:** Real-time
- **Description:** a monitoring period during which only in vehicle real-time warnings provided using adaptive ADAS
- **Duration:** 4 weeks

Phase 3
- **Intervention:** Real-time + Post-trip
- **Description:** a monitoring period during which in addition to real-time in vehicle warnings, drivers received feedback on their driving performance through the app
- **Duration:** 4 weeks

Phase 4
- **Intervention:** Real-time + Post-trip + Gamification
- **Description:** a monitoring period during which in vehicle real-time interventions were active along with feedback but at the same time gamification elements were also active
- **Duration:** 6 weeks

Methodological Overview

- 'Safety Tolerance Zone' (STZ) levels were defined based on thresholds of the time headway between two vehicles:
  - 'Normal' level: Headway > 2 sec
  - 'Dangerous' level: 1.4 sec < Headway < 2 sec
  - 'Avoidable Accident' level: Headway < 1.4 sec

- Four classification algorithms (e.g., Support Vector Machines, Random Forest, AdaBoost and Multilayer Perceptron) were employed to predict STZ level of driving behaviour

- The input variables (x) for the classification were selected through feature importance, with three variables emerging as significant: a) Speed, b) Distance Travelled, and c) Speed Limit, while the output variable (y) was set to be the STZ level

- The evaluation of the classification models was based on:

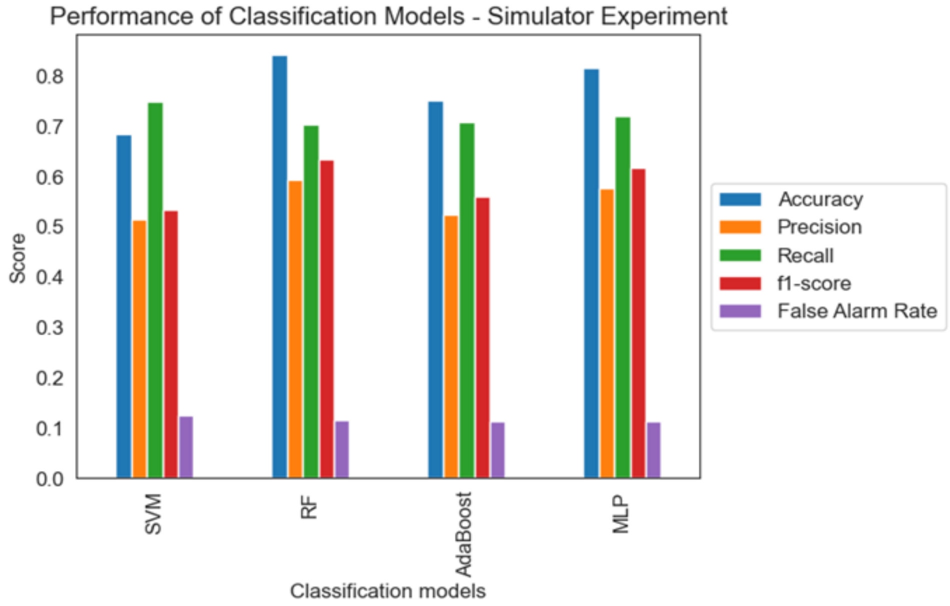
Accuracy | Precision | Recall | f1-score | False Alarm Rate

Results - Simulator Driving Experiment

- Random Forest (RF) excels with the highest accuracy of 84%, showcasing its proficiency in accurately classifying driving behaviors, while its well-balanced precision and recall make it a robust choice for driving behavior classification

- MLP's competitive accuracy and balanced metrics underscore its effectiveness in capturing diverse driving behaviors within the simulation framework

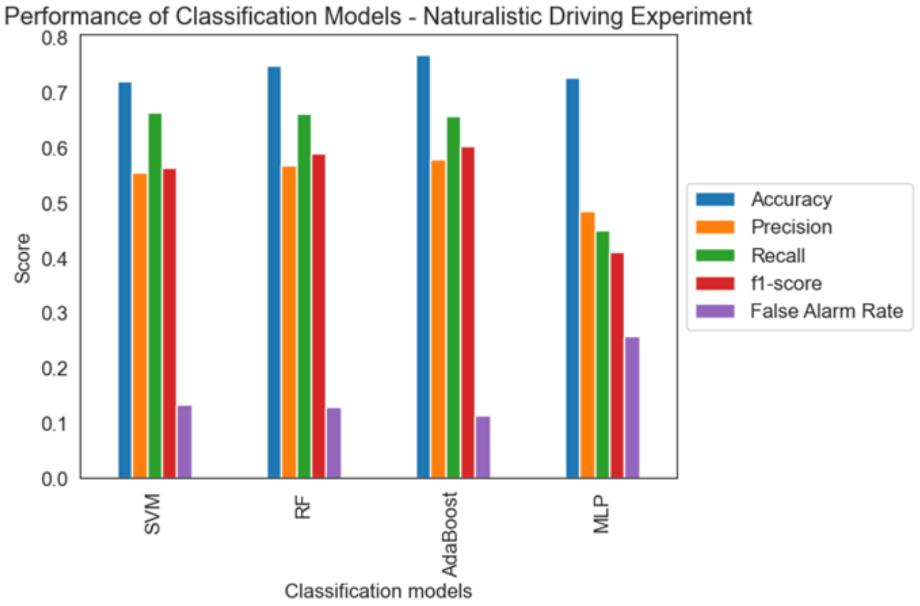
- While AdaBoost and Support Vector Machine (SVM) achieve reasonable accuracy, they show trade-offs between precision and recall, highlighting the importance of understanding algorithm nuances



Classifier	Accuracy	Precision	Recall	False Alarm Rate	f1-score
SVM	68.67 %	51.35 %	74.72 %	12.47 %	53.22 %
RF	**84.00 %**	**59.41 %**	**70.27 %**	**11.47 %**	**63.42 %**
AdaBoost	75.08 %	52.31 %	70.71 %	11.30 %	55.87 %
MLP	81.28 %	57.51 %	72.04 %	11.37 %	61.79 %

Results - Naturalistic Driving Experiment

- AdaBoost emerged as the top performer with the highest accuracy (76%) among the models, maintaining competitive precision (58%) and recall (66%), consistent with the simulator data and achieving the highest f1-score of 60%
- Random Forest (RF) demonstrated consistent strong performance across both the simulation and naturalistic driving experiment datasets
- In comparison with the simulation environment, MLP demonstrated poor performance, suggesting potential difficulties in its ability to generalize effectively to real-world driving scenarios



Classifier	Accuracy	Precision	Recall	False Alarm Rate	f1-score
SVM	72.05 %	55.51 %	66.31 %	13.39 %	56.37 %
RF	75.00 %	56.77 %	66.28 %	12.97 %	59.03 %
AdaBoost	**76.76 %**	**57.91 %**	**65.81 %**	**11.47 %**	**60.19 %**
MLP	73.26 %	52.14 %	56.57 %	16.66 %	52.65 %

Discussion

- Notable performance variations observed among machine learning classifiers, with Random Forest (RF) demonstrating superior accuracy in simulation environment, and AdaBoost showing robust performance in real-world driving datasets

- The observed differences in model performance between the simulator and naturalistic driving experiments highlight the challenges posed by real-world driving scenarios, such as diverse traffic conditions and unexpected events, underscoring the need for adaptable models capable of addressing these complexities effectively

- The findings underscore the significant need for proactive measures in traffic safety management, emphasizing the potential of predictive models

Conclusions

- Evaluating model performance across diverse scenarios underscores the challenge of adaptability from simulation to real-world settings, emphasizing the need for robust and versatile modeling approaches

- Random Forest (RF) model emerges as a promising performer, exhibiting balanced precision and recall in both simulated and real-world driving scenarios

- Future research could explore integrating deep learning techniques to improve the accuracy of driver behavior classification models, and diversifying datasets by conducting driving experiments from diverse regions and transport modes to enhance the models' generalization

An Roinn Iompair
Department of Transport

European Commission

Bonneagar Iompair Éireann
Transport Infrastructure Ireland

UCD DUBLIN

Enterprise Ireland

IDA Ireland

CONFERENCE PARTNERS INTERNATIONAL

TRA
TRANSPORT RESEARCH ARENA

10th Transport Research Arena Conference
Advancing Sustainable and Inclusive Mobility
Dublin, Ireland, April 15-18, 2024

Predicting risky driving behavior with classification algorithms: Results from a large-scale field-trial and simulator experiment

Thodoris Garefalakis

Transportation Engineer, Research Associate

Together with:
Eva Michelaraki, Stella Roussou, Christos Katrakazas, Tom Brijs, George Yannis

UHASSELT