An Roinn Iompair
Department of Transport

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

Unfolding the dynamics of driving behavior: A machine learning analysis from Germany and Belgium

Stella Roussou

Transportation Engineer, PhD Candidate

Together with:
E. Michelaraki, C. Katrakazas, A. P. Afghari, C. Al Haddad, M. R. Alam, C. Antoniou, E. Papadimitriou, T. Brijs, G. Yannis

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

- **Driver behavior** is a contributory factor in over 90% of crashes

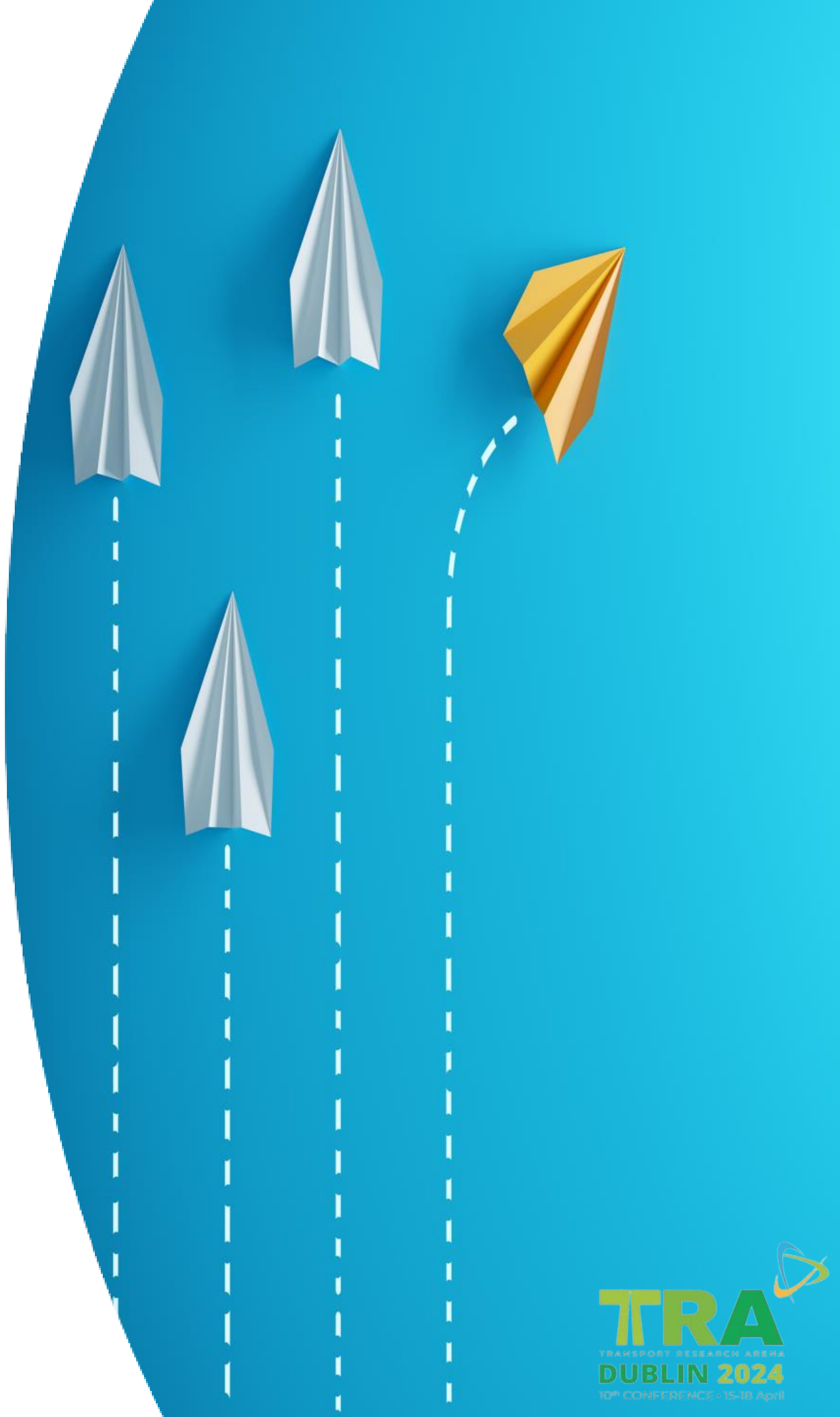
- Factors such as **driver's state,** environmental conditions, and traffic circumstances remain significant contributors to traffic collisions

- **Intelligent driving behavior monitoring systems** enable real-time interventions and demonstrate remarkable efficacy in improving road safety

- The combination of autonomous vehicles and intelligent monitoring systems mitigate the impact of **human error** and create a **safer road environment** for all users

Objectives

- Development of a Neural Network Model and a Long-Short Term Memory Model
- Comparison and contrast of the two machine learning classifiers and comparison between the trips from two different countries of Germany and Belgium
- Identification of the level of risky driving behavior based the two machine learning techniques
- Investigation of the speeding and headway as the two key risk factors for road safety



Data Description

- The vehicle data collected from the naturalistic driving experiment consisted in total of two datasets:
  - 1. German drivers
    - 30 drivers
    - 5.344 trips and
    - 84.434 minutes
  - 2. Belgian drivers
    - 39 drivers (the ones remaining consistent)
    - 7.163 trips and
    - 147.337 minutes
- For both datasets, similar driving behaviour characteristics (such as speed, headway, TTC, etc.) were collected

Phase 1 Monitoring	Phase 2 Real time Interventions	Phase 3 Real time & post-trip interventions	Phase 4 Real time & post-trip interventions & gamification
German Dataset			
30 drivers	30 drivers	30 drivers	30 drivers
1.397 trips	1.322 trips	1.129 trips	1.496 trips
23.617 minutes	19.469 minutes	17.704 minutes	23.644 minutes
Belgian Dataset			
39 drivers	43 drivers	51 drivers	49 drivers
1.173 trips	1.549 trips	1.973 trips	2.468 trips
23.725 minutes	31.414 minutes	40.121 minutes	52.077 minutes

Experiment Phases

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

- A **Neural Network (NN)** was carried out involving 30 car drivers from Germany and 39 car drivers from Belgium and a large database consisting of 5.344 trips and 7.163 trips was collected and analyzed

- **Long-Short Term Memory Network (LSTM)** was also developed with the same dataset in order to compare the two machine learning techniques and the two different countries

- The **classification algorithms** are **evaluated** using the:
  - Accuracy
  - Precision
  - Recall
  - f1-score
  - False Alarm Rate

Neural Networks (NNs) Results

- NNs were employed to investigate if real-time prediction of the STZ is feasible
- Phase was considered as an independent variable and the analysis was performed for the whole dataset
- **German's** NN has an accuracy of more than 94% with a false alarm rate of only 6% whereas the **Belgian** results showcase a lower accuracy and false alarm rate
- Belgium's NN models, while strong, present difficulties in achieving high precision, especially for speeding incidents.
- The NNs classification algorithms act as a preparatory step towards the LSTM classification

Table 1: Confusion data matrix for headway and speeding

GERMANY

Variable	TP	FP	FN	TN	SUM
Headway	33378	0	1400	82	34860
Speeding	2178	1987	63	30632	34860

Table 2: Assessment of classification model for headway and speeding

GERMANY

Variables	Accuracy	Precision	Recall	f1-score	FA Rate
Headway	95.98%	100.00%	95.97%	97.95%	0.00%
Speeding	94.12%	52.29%	97.19%	68.00%	6.11%

Table 3: Confusion data matrix for headway and speeding

BELGIUM

Variable	TP	FP	FN	TN	SUM
Headway	37517	0	80	7918	45512
Speeding	30069	0	0	6193	36462

Table 4: Assessment of classification model for headway and speeding

BELGIUM

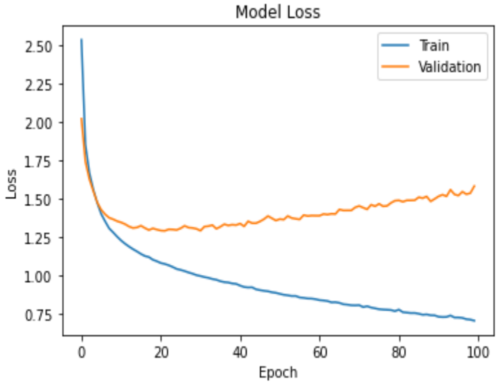
Variables	Accuracy	Precision	Recall	f1-score	FA Rate
Headway	77.19%	77.64%	77.19%	76.90%	22.81%
Speeding	83.51%	80.71%	83.51%	79.78%	16.49%

Long Short-Term Memory (LSTM) Results

- LSTMs were trained to predict «dangerous» speeding and headway level. The sequence is implicit in the way that the data was collected or organized, even if it's not immediately apparent from the predictors themselves
- The accuracy of less than 60% may not be sufficient, however, the required level of accuracy depends on the specific use case and the risks involved
- The output of prediction models can be used for ongoing analysis and monitoring of road safety performance, in order to identify trends and patterns that can inform future interventions and improvements
- The LSTM models in both countries show potential for capturing temporal patterns, but they currently lag behind the NN models in terms of overall accuracy and precision-recall balance
- A descending trend in both training and validation loss is achieved in the LSTM of German car drivers for headway and speeding

Table 5: Assessment of classification model for headway and speeding

GERMANY					
Variables	Accuracy	Precision	Recall	f1-score	FA Rate
Headway	45.57%	42.13%	45.57%	41.11%	54.43%
Speeding	53.14%	49.54%	53.14%	50.81%	46.86%



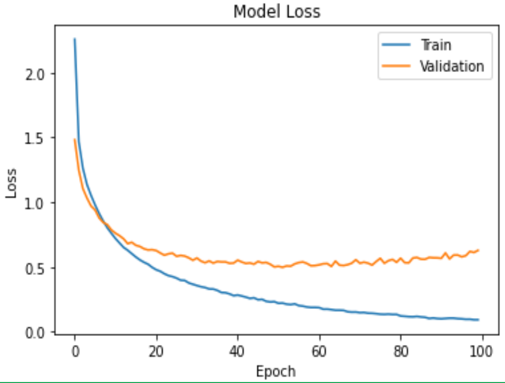
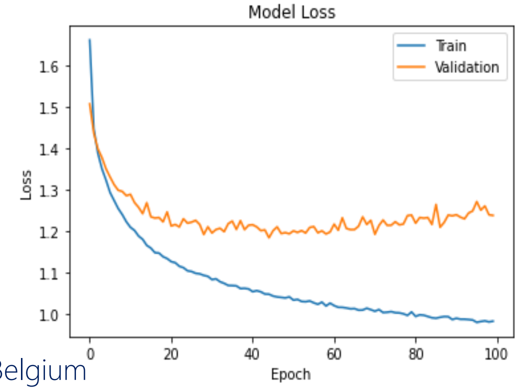
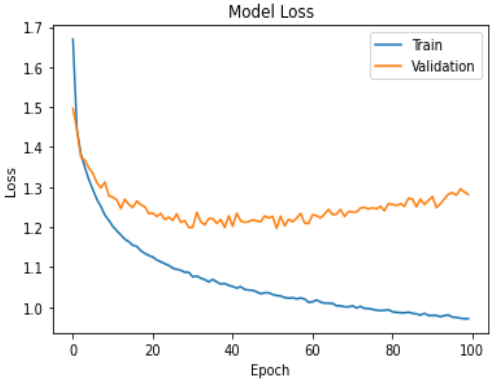


Table 6: Assessment of classification model for headway and speeding

BELGIUM					
Variables	Accuracy	Precision	Recall	f1-score	FA Rate
Headway	58.12%	35.65%	58.12%	37.33%	41.88%
Speeding	48.27%	25.75%	48.27%	32.59%	51.73%





Discussion

- Germany's NN models demonstrate superior performance in both accuracy and precision-recall balance compared to Belgium's models and their respective LSTM counterparts
- Predictive real-time analyses demonstrated that it is possible to predict the level of STZ with an accuracy of up to 95%
- Post-trip explanatory studies showcased the capacity of state-of-the-art econometric models to shed light on the complex relationship of risk with the interdependence of task complexity and coping capacity
- Machine learning algorithms can recognize specific driving patterns associated with safe driving. NNs proved to be the best approach to capture complex relationships between various driving parameters and predict the likelihood of potential risks or crashes
- The contrasting factor between the two methods lies in the ability to capture temporal dependencies in the data

Conclusions

- By leveraging machine learning algorithms and data-driven insights, it is possible to identify safe driving behavior, provide immediate feedback to drivers, and ultimately contribute to creating a safer driving environment
- The development and deployment of real-time applications based on these techniques can provide drivers with immediate feedback and guidance to help them make informed decisions, improve their driving habits, and reduce crash risk
- Future research could consider incorporating contextual information into the models
- Factors such as weather conditions, road infrastructure, and traffic patterns could be used to enhance the accuracy and applicability of the models
- Individual driver characteristics like gender, age, and driving style could be further examined

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

Unfolding the dynamics of driving behavior: A machine learning analysis from Germany and Belgium

Stella Roussou

Transportation Engineer, PhD Candidate

Together with:
E. Michelaraki, C. Katrakazas, A. P. Afghari, C. Al Haddad, M. R. Alam, C. Antoniou, E. Papadimitriou, T. Brijs, G. Yannis