Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University of Munich

Crash risk analysis of driver behaviour: a driving simulation study

A thesis presented in part fulfilment of the requirements of the Degree of Master of Science in Transportation Systems at TUM School of Engineering and Design, Technical University of Munich.

Supervisor Dr. Kui Yang
MSc. Christelle Al Haddad
Chair of Transportation Systems Engineering

Submitted by Isidora Gkena
Arcisstrasse 21
80333 München

Submitted on München, 10.06.2022

Abstract

Rear-end crashes have increased during the last years and according to researchers Forward Collision Warning (FCW) systems help significantly to mitigate the occurrence of these crashes. This thesis is based on a driving simulator experiment with 60 drivers that participated in three driving scenarios, monitoring, intervention and distraction. Participants during intervention scenario received warnings triggered based on the time headway with the leading vehicle. During distraction scenario drivers along with warnings were receiving text messages.

The first part of the research was a statistical analysis conducted with paired samples t-test between the scenarios. A significant difference in population means was observed for most of the variables between monitoring and intervention scenario and therefore was concluded that real time interventions by informing drivers timely about critical situations, had a significant impact. Further investigation was implemented between monitoring and intervention scenario focusing on the different road sections (urban, rural, highway). The results presented that interventions had a higher impact while driving in rural and highway road environment. At the second part of the thesis, some commonly used Machine Learning models (Logistic Regression, Support Vector Machine, Artificial Neural Network, Random Forest) implemented for the prediction of dangerous driving events. It was found that Random Forest overperformed with respect to the other models, by reaching 99% performance in Recall and F1-score for the minority class.

Keywords: driving simulation experiment, interventions, distractions, statistical analysis, Machine Learning models

Acknowledgments

I would like to express my sincere gratitude to my supervisors Dr. Kui Yang and Christelle Al Haddad. Their guidance and support during my research were crucial factors for the completion of my master thesis.

I would also like to thank my mother and the rest of my family as well as my beloved friends for their love and support.

Contents

List of Figures

List of Tables

Abbreviations

ADAS	Advanced Driver Assistance Systems
ANCOVA	Analysis of Covariance
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
DHW	Distance Headway
FCW	Forward Collision Warning
FN	False Negative
FP	False Positive
i-DREAMS	Driver and Road Environment Assessment and Monitoring System
SVM	Support Vector Machine
THW	Time Headway
TN	True Negative
TP	True Positive
TTC	Time to Collision
WHO	World Health Organization

1. Introduction

1.1. Background

According to World Health Organization (WHO), a road traffic crash is defined as a crash entailing at least one moving vehicle, that has as a result the injury or death of a person (Peden , 2004). As stated by WHO around 1.3 million deaths occur every year due to road traffic crashes and these deaths involve mainly children, teenagers and young people in the age group of 5 to 29 years (World Health Organization , 2021).

Some of the main risk factors leading to road traffic crashes as they are mentioned by WHO, are speeding, alcohol, distracted driving, the absence of seat belts and helmets as well as unsafe vehicles and road infrastructure. Furthermore, it is of great importance to add that 93% of the deaths mentioned above take place in low as well as middle income countries and they possess approximately 60% percent of the vehicles worldwide (World Health Organization , 2021).

According to (Wegman, 2016), countries such as Spain and France have already implemented measures regarding road safety and reached a decrease in number of road fatalities. Several interventions have been applied in order to improve road safety. For instance, legislation and campaigns have been used to improve human behaviour. Furthermore, new planning and designing strategies focus on improving the infrastructure, while active vehicle safety interventions contribute to make vehicles safer (Wegman, 2016).

More specifically, concerning the vehicle improvement, Advanced Driver Assistance Systems (ADAS) contribute significantly to road safety. As stated by the European Commission, ADAS support the driver during the driving task through assistance in situations that cannot be always easily handled by the driver. ADAS support the driver to pay attention to the leading vehicle and the traffic in front as well as to take into consideration the important information. Moreover, they assist in detecting other road users in the blind spot and last but not least they inform about the coming traffic situation (European Commission, 2022).

A significant advanced driver assistance system which is also implemented in this thesis is Forward Collision Warning (FCW). This system provides warnings to the drivers when they approach very close to the leading vehicle. The main goal of Forward Collision Warning is to mitigate the rear end collisions and give the opportunity to the driver to act timely in order to avoid a critical situation (Car ADAS, 2021).

Many studies have shown that Forward Collision Warning reduces the rear end crashes. Specifically, a study case based on large trucks estimated a 44% decrease in the rate of rear end crashes as well as a reduce of approximately 20% in the rate of police reported crashes (Teoh, 2021). Moreover, a driving simulator study found that a FCW system decreases the likelihood of a rear end crash significantly and presented the preference of participants to the adaptive system, due to the fact that they found it less stressful (Jamson, Lai, & Carsten, 2007).

Nowadays the usage of mobile phone while driving causes distraction of the driver which can lead in critical situations or collision. A meta-analysis examined the effects of mobile phone conversations during driving, taking into consideration twenty-three studies. The results have shown that there were significant impacts on drivers related with hazardous events while they were engaged in mobile phone conversations. The study also found slight differences regarding the impact of mobile phone usage while comparing simulator and field studies (Horrey & Wickens, 2006). At this case it is useful to use interventions during the driving task in order to inform the driver whether it is critical or not to use the mobile phone.

It is therefore important to collect data of driver simulator or field experiments in order to recognise different driving styles and predict the dangerous driving events with the implementation of machine learning algorithms. Different machine learning models have been applied for this purpose such as Support Vector Machine, Random Forest, K-Nearest Neighbor and Multi-Layer Perceptron (Xue, Wang, Lu, & Liu, 2019) as well as Recurrent Neural Network (Alvarez-Coello, et al., 2019).

1.2. Motivation

The implementation of a driving simulator in order to conduct experiments with different participants and investigate several topics in transportation field is a privilege nowadays. Several driving simulator experiments have been conducted, considering different topics, parameters, data and factors but all focus on improving road safety. Many driving scenarios cannot be tested in a vehicle during a real driving situation, because it can lead to critical events such as injuries or fatalities. Therefore, it is useful to conduct driving simulator experiments which provide a controllable, reproducible and standardized environment, where different scenarios, weather conditions, road sections and virtual traffic can be designed and applied (Winter, Leeuwen, & Happee, 2012).

It has been proved that driving simulator data can be easily collected and were more accurate than the ones from real vehicles (Santos, Merat, Mouta, Brookhuis, & de Waard, 2005). In addition, in a driver simulator it is easier to implement dangerous driving events without causing physical harm to the participants (Underwood, Crundall, & Chapman, 2011). Last but not least, through a driving simulator experiment feedback and suggestions for improvement can be received from the participants at the end of their drive (Winter, Leeuwen, & Happee, 2012).

Furthermore, the improvement and development of artificial intelligence technologies, machine learning models as well as big data provide many opportunities and different ways to utilize the transportation data obtained either from a driving simulator, surveillance cameras or sensors, such as radar and LIDAR data (He, Hu, Park, & Levin, 2019). Researchers implement machine learning technologies in order to gain insights on driving behavior, travel. Behavior as well as to overcome transportation challenges (Urban Mobility Lab at MIT, 2022).

Several studies have used driving simulator data in order to conduct a statistical analysis between different driving conditions and compare them, e.g. different driving scenarios (Babić, Babić, Cajner, Sruk, & Fiolić, 2020), (Guo, et al., 2019). Others have utilized these data in order to understand the driving behavior and predict dangerous driving events, defined by different states, implementing several machine learning algorithms (Alvarez-Coello, et al., 2019), (Ahangari, Dehzangi, & Jeihani, 2019). Moreover, by obtaining transportation data research can be done in several fields, such as time series analysis implementation for forecasting (Moorthy & Ratcliffe, 2007), (Ghosh, Basu, & O'Mahony, 2005).

1.3. Research Questions and Objectives

1) Do the interventions have an impact on driving behavior?

2) How to predict the dangerous driving event?

The objectives of this study will be presented below:

The main objective regarding the first research question is to compare scenarios with and without real-time interventions. Driving simulator data corresponding to a high time headway were excluded from the analysis. Moreover, a binary classification was implemented in order to separate the data to class 1 (dangerous driving events) and class 0 (safe condition). The aggregation of data followed based on the critical driving event which is defined by class 1. Last not least, a statistical analysis took place by implementing paired samples t-test between monitoring, intervention and distraction scenario.

According to the second research question, the main goal is to implement some important machine learning algorithms in order to predict the dangerous event (y) and investigate the importance of the independent variables for each model. At this part, data representing a high time headway were also excluded and were classified to class 1 and class 0 likewise to the first research question. Moreover, data of distraction scenario were implemented but they were not aggregated. In order to use the machine learning models a train test split was conducted, separating the data to training and testing data. Furthermore, Logistic Regression, Support Vector Machine, Artificial Neural Network and Random Forest were applied and evaluated based on their performance.

1.4. Contributions

This thesis focuses on contributing in the following parts:

- ◊ Investigate the impact of real time interventions on driving behavior by conducting a comparison between monitoring, intervention and distraction scenario.
- ◊ Compare drivers' behavior in different driving environments between monitoring and intervention scenario based on user, vehicle and road perspective.
- ◊ Employ some commonly used machine learning models (Logistic Regression, Support Vector Machine (SVM), Artificial Neural Network (ANN) and Random Forest) to predict dangerous events, which can be used in real time driving behavior management.

1.5. Thesis Outline

This thesis consists of six chapters, which will be explained in this paragraph. Chapter 1 refers to the background and motivation related to this work. Furthermore, it states the research questions and objectives as well as the contributions of the thesis. Chapter 2 introduces the literature review, which focuses on driving simulation studies related to interventions and distractions as well as studies that implemented machine learning models. Moreover, some important safety indicators and their thresholds were described. Chapter 3 presents the methodology that was implemented in order to answer the research questions. In chapter 4, experiment design is explained as well as the driving simulator data that were used. Chapter 5 consists of methodology results and is divided in two parts. The first part presents the results of statistical analysis regarding the evaluation effectiveness on driving behavior. The second part contains the performance results of the machine learning models that were used for the dangerous event prediction. Chapter 6 refers to the conclusions of the thesis, the limitations and some of the future work that would be interesting to be implemented.

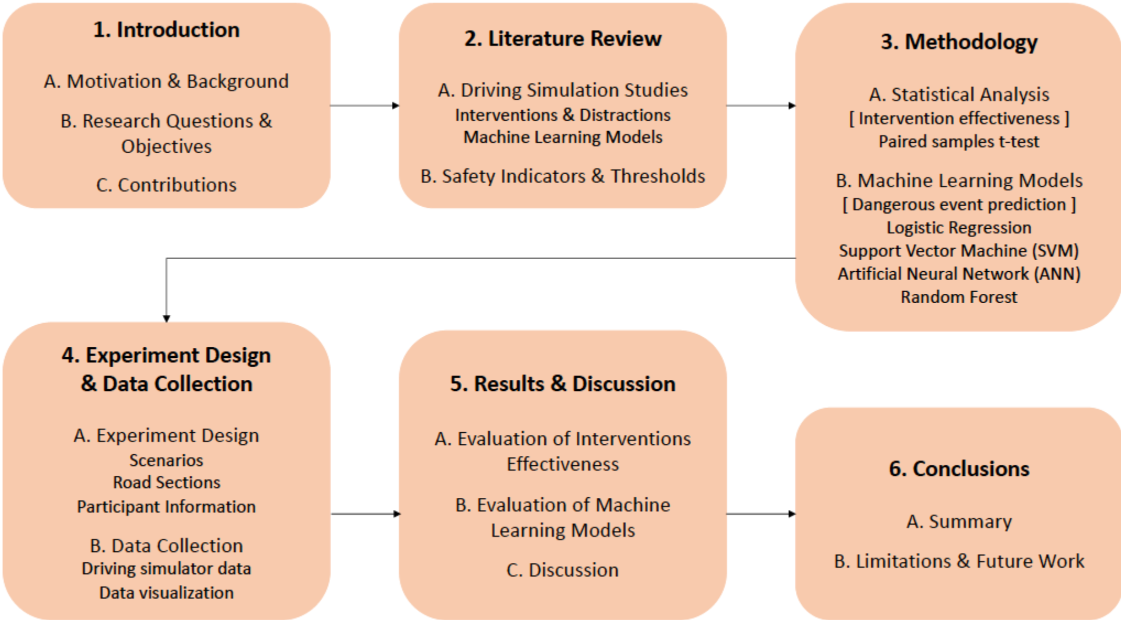

Figure 1 Thesis overview

2. Literature Review

2.1. Driving Simulation Studies

Various research studies were conducted either using driving simulation data or naturalistic driving data (Blana, 1996). Nevertheless, both of these experimental methods have advantages. More specifically, a driving simulation experiment is conducted under controllable conditions that is safe for the participants, with low costs and mostly an easy data collection. On the other hand, in experiments with real vehicles researchers obtain data which are real, reliable and practically applicable (Li, Guo , & Li , 2021).

A study investigated how young drivers' behaviour was affected by traffic signs and road markings during night-time driving. Drivers participated in two drives, one with traffic signalling and one without, conducted in a driving simulator with the implementation of eye tracking glasses as well as an electrocardiograph. A statistical analysis took place using a paired sample t-test comparing speed, acceleration, deceleration, lateral position and participants' eye movement between the two drives. The results presented for most of the variables statistically significant differences, which indicates that traffic signalling had a high impact on drivers' behaviour. Participants were able to modify their driving timely based on the traffic signs indications, something that contributed to a better driving performance and increased the traffic safety (Babić, Babić, Cajner, Sruk, & Fiolić, 2020).

In their research Guo et al. analyse the impact of anxiety on eye movement characteristics of female drivers. They conducted driving simulations as well as experiments with a real vehicle in order to compare participants' eye movement during calm and anxious situations. For this comparison a paired sample t-test was implemented with a 95% confidence interval. A significant statistical difference on eye movement between situations of calmness and anxiety was observed, especially when drivers were dealing with curved parts of the road and traffic accident scenes (Guo, et al., 2019).

2.1.1. Intervention

According to Calvi, D'Amico, Ferrante and Ciampoli a driving simulation study was conducted in order to test the effectiveness of Augmented Reality technology implementation on increasing the safety of drivers that approach a zebra crossing area. Participants during their drive were receiving warnings while approaching a zebra crossing. Their driving behaviour was compared based on their speed, deceleration and distance as well as using time to collision and time to zebra indicators. The results of their study showed that as the warning appeared, drivers started to reduce their speed, decelerate and reached high values of time to collision and time to zebra (Calvi, D'Amico, Ferrante, & Ciampoli, 2020).

Jamson, Lai and Carsten in a driving simulator study investigated the advantages of an adaptive Forward Collision Warning (FCW) system. This study tested on 45 participants a non-adaptive FCW system and an adaptive one. Drivers observed benefits in both FCW systems regarding their safety in terms of avoiding a rear end collision. More specifically, participants that were driving less aggressively did not observe a significant difference between the two systems. On the other hand, the aggressive drivers showed a higher preference on the adaptive FCW system and stated that it was less irritating and stressful (Jamson, Lai, & Carsten, 2007).

A study case of Abe and Richardson tested on a driving simulator the response of participants to a FCW system. Specifically, they investigated the impact of the alarm timing system on the driver behaviour for three different driving speeds (40, 60 and 70 m/h) as well as for different time headways (1.7 s and 2.2 s). Their results found that alarm promptness had a significant impact on the trust of the participants to the system and improved their braking performance. Participants had the opposite opinion when alarms were appearing after the braking. Therefore, they considered these alarms as late alarms and their trust was significantly reduced (Abe & Richardson, 2006).

2.1.2. Distraction

Distraction occurs when an external triggering shifts driver's attention from driving task to another object, e.g. mobile phone (World Health Organization, 2011). As stated by K. Young and M. Regan of Monash University, distraction of the driver is considered as a prior issue regarding road safety in Europe and Japan as well as in North America. Through their literature review found that the interaction with devices while driving impairs driving performance. Specifically, it degrades the ability of maintaining the speed, lateral position as well as vehicle control. Moreover, they stated that younger drivers with less driving experience and older drivers are highly susceptible while they are engaged with mobile phone usage in comparison with middle-aged people with experience in driving (Young & Regan, 2007).

A meta-analysis presented how reading text messages and typing affected the driving behaviour of participants. More specifically, it showed that during their interaction with the phone drivers were distracted, affected on responding to critical traffic events, maintaining headway and speed as well as presented a high deviation on their lateral position (Caird, Johnston, Willness, Asbridge, & Steel, 2014).

A driving simulator study investigated the impact of smartphones on driving behaviour, taking into consideration the age and gender of participants and implemented a statistical analysis using ANOVA and ANCOVA. Their results showed that age is a factor that affects significantly the driving performance and therefore older people noted to decrease their speed while they were using the mobile phone. Furthermore, it was also observed that the risk of collision increased for all the participants apart from their age during their interaction with the smartphone. Last but not least, the estimations showed that while using the smartphone participants were increasing their speed (Fancello, Adamu, Serra, & Fadda, 2020).

Another study conducted in a driving simulator focused on the impact of hands-free and hand-held smartphones on driver behaviour. The drivers participated in four drives in which they were distracted with a conversation on the smartphone. The results of this study found that during the secondary task of talking on the phone, the average speed of participants as well as the standard deviation of acceleration were significantly reduced (Haigney, Taylor, & Westerman, 2000).

Choudhary and Velaga used reaction time as an indicator to evaluate the impact of phone usage during the driving task. The participants in the experiment received four kinds of distractions related to message texting and talking conversations and faced two critical events in a pedestrian crossing and a road crossing next to a parking slot, in order to evaluate their time of reaction. The study results showed that in both critical events happening while drivers were engaged to secondary tasks their reaction time increased significantly. Therefore, they conclude that mobile phone usage while driving can lead to a reduced awareness of the driver and cause delay to the response in critical situations that may cause accidents (Choudhary & Velaga, 2017).

2.1.3. Machine Learning Models

On their research Alvarez-Coello et al. classified dangerous driving behavior by taking into consideration aggressive maneuvers of drivers. Furthermore, they used in-vehicle data and implemented Random Forest and Recurrent Neural Network classifiers to model dangerous driving events. Their outcomes showed that sensor data provide a low frequency as well as that a dataset can be defined as limited with respect to driving events transitions. Last but not least, they pointed out that the classifiers of dangerous driving event can be implemented for discriminating the data and provide integration (Alvarez-Coello, et al., 2019).

Another study focused on creating rules for drivers' cognitive distractions implementing eye-tracking data as well as driving simulation data. Specifically, eighteen drivers participated in two drives with approximate duration of 15 minutes, one with no load and one with distractions, in which the load was a cognitive task. In order to create the rules for driving under distraction, Support Vector Machine (SVM) was implemented based on a constant time interval data transformation and reached as a result qualitative data for the model (Yoshizawa, Nishiyama, Iwasaki, & Mizoguch, 2016).

Xue, Wang, Lu and Liu on their experiment focused on the recognition of different driving styles by implementing trajectory data from a surveillance video. The indicators that were selected for the evaluation of crash risk of vehicle trajectory were inversed time to collision, modified margin to collision and time headway. Moreover, data were labelled based on the risk of rear end collision by implementing K-mean algorithm. Last but not least, several models were applied in order to recognize the different driving styles such as Support Vector Machine, Random Forest, K-Nearest Neighbour and Multi-Layer Perceptron. It was observed that Support Vector Machine performed with the highest accuracy (approximately 90%) with respect to the other models that mentioned above (Xue, Wang, Lu, & Liu, 2019).

In their study case Ahangari, Jeihani and Dehzangi implemented driving simulator data in order to detect participants' distraction. Specifically, they conducted an experiment with 92 drivers which participated in six driving scenarios dealing with different forms of distraction such as conversations with a hand-held and hands-free mobile phone, message texting etc. and drove on four types of road sections. In order to predict driving performance of participants while they were engaged with secondary tasks, they applied a Bayesian Network. They used various variables to evaluate the model some of them are velocity, acceleration, lane deviation, collision and brake. Implementing the Bayesian Network for predicting distraction of drivers they reached approximately 68% accuracy in the performance of the model (Ahangari, Dehzangi, & Jeihani, 2019).

Tango and Botta conducted research in order to detect real time distraction of drivers that participated in a driving simulator experiment. Driving simulator data from the moments that participants were distracted with a visual task were used to train different machine learning algorithms. The models were compared regarding their characteristics, feature importance as well as their performance. It was observed that Support Vector Machine (SVM) reached the highest performance with respect to the other machine learning models for detecting the visual distraction (Tango & Botta, 2013).

2.2. Safety Indicators and Thresholds

2.2.1. Safety Indicators

Time headway (THW), time to collision (TTC) and distance headway (DHW) have been used in different study cases as safety indicators (Khansari, Nejad, & Moogehi, 2020), (Liu & Fu, 2018). THW is defined as the time span between two vehicles passing an index point, estimated from the front of the leading vehicle to the front of following vehicle (Rossi & Gastaldi, 2012), (Khansari, Nejad, & Moogehi, 2020).

$$THW = t_i - t_{i-1}$$

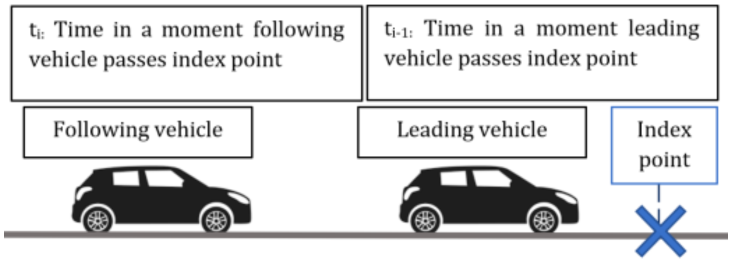

Figure 2 Time headway (Khansari, Nejad, & Moogehi, 2020)

Another important indicator is TTC, which is defined as the remaining time before a rear end collision if the speed difference of the vehicles is maintained (Saffarzadeh, Nadimi, Naseralavi, & Mamdoohi). It can be estimated with the following formula (Khansari, Nejad, & Moogehi, 2020):

$$TTC_i = \frac{X_{i-1}(t) - X_i(t) - l_i}{\dot{X}_i(t) - \dot{X}_{i-1}(t)}$$

$$\forall\, \dot{X}_i(t) > \dot{X}_{i-1}(t)$$

Where $\dot{X}_i$ denotes the speed, X_i denotes the position and l_i the length of the following vehicle and $\dot{X}_{i-1}$, X_{i-1} denote the speed and position of the leading vehicle respectively (Khansari, Nejad, & Moogehi, 2020).

Another indicator that should also be considered is DHW, which is equal to the gap distance adding the length of the leading vehicle. Gap distance indicates the space from the back bumper of the leading vehicle to the front bumper of the vehicle that follows (Liu & Fu, 2018). All the indicators mentioned above are effective measures for distinguishing a normal driving behaviour from a critical one in car-following events (Saffarzadeh, Nadimi, Naseralavi, & Mamdoohi).

Car-following behaviour composes of the acceptable distance that drivers maintain from the leading vehicle and the acceleration based on the driving behaviour of the front vehicle (Sato & Akamatsu, 2012). It has been observed that critical car-following events can be caused due to tailgating behaviour. More specifically, tailgating behaviour is driving closely to the leading vehicle with significantly short time headway (Hassan, Sarhan, Garib, & Harthei, 2017). (Rämä & Kulmala, 2000) mention in their research that tailgating is an aggressive and significantly dangerous driving behaviour and a major cause leading to rear-end collision.

2.2.2. Thresholds

Taking into consideration various study cases, different thresholds have been used to define indicators such as THW, TTC and DHW as critical. Crashes that are caused due to tailgating behaviour can be minimized if drivers keep at least 2 seconds THW or maintain a distance equal to the length of one car for each 16 kilometres per hour from the leading vehicle (Monteiro, Balogun, Kote, & Tlhabano, 2014). Another study stated that a THW less than 2 seconds is neither safe nor sufficient (Wang & Song, 2011).

A study in order to distinguish the car-following event from the free-flow situation used gap distance as an indicator. Taking into consideration that the leading vehicle is in the same lane should have a gap distance of maximum 120 meters (Mai, Wang, & Prokop, 2017). Another study mentions a DHW of 150 meters as critical threshold for discriminating a car-following situation with free traffic flow (Transportation Research Board , 2015). Participants that took part in an experiment were asked to keep a safe distance without trying to pass behind the leading car with a speed range between 50 and 100 kilometres per hour. Based on the results it was noted that participants adjusted the DHW with respect to speed from 9.5 meters to 19 meters at 50 and 100 kilometres per hour respectively (Loulizi, Bichiou, & Rakha, 2019).

TTC is an important safety indicator with respect to rear-end collisions, which has been used profitably in safety analysis (Khansari, Nejad, & Moogehi, 2020). On their research Hirst and Graham stated that a TTC equal to 4 seconds can differentiate the condition where drivers feel safe and have control of driving with the one that drivers are involved in an unsafe and dangerous situation (Hirst & Graham, 1997). Another study reported that due to the driving behaviour variation during different driving situations, no specific threshold can be defined for TTC in order to differentiate the safe from a dangerous car-following event. Therefore, it was stated that a range of thresholds was selected from 0.5 to 10 seconds (Saffarzadeh, Nadimi, Naseralavi, & Mamdoohi).

Khansari, Nejad and Moogehi conducted research using a driving simulator in order to compare THW and TTC indicators. They used two different types of THW, which was braking THW (the moment of breaking) and following THW (while following the leading car). Their results showed that braking THW is the most important indicator for discriminating critical car-following situations as well as that most of the drivers were trying to maintain a braking THW of 1.1 seconds during the whole drive and not proceed closer to the lead vehicle (Khansari, Nejad, & Moogehi, 2020). In the research study of Vogel, it was concluded that THW should be used for enforcement purposes, since low values of THW can cause dangerous driving situation and TTC, since it indicates the occurrence of dangerous events, should be implied for the safety evaluation of a specific traffic environment (Vogel, 2003).

The following table provides a summary of the thresholds that mentioned above which considered safe for each indicator:

Indicator	**Threshold**	**Source**
Time Headway (THW)	THW ≥ 2 s	Monteiro et al., 2014, Wang & Song, 2011
Distance Headway (DHW)	9.5 m ≥ DHW ≥ 19 m	Loulizi, Bichiou & Rakha, 2019
Time to Collision (TTC)	TTC ≥ 4 s	Hirst & Graham, 1997

Table 1 Thresholds of THW, DHW, TTC

3. Methodology

3.1. Paired Samples t-test

According to Kent State University libraries a Paired Samples t-test compares the means of two measurements collected from the same object or individual under two different conditions. The main goal of the test is to ascertain whether the mean difference between paired observations appears to have significant difference from zero. The test is also called dependent t-test or repeated measures t-test (Kent State University, 2021).

The t-test is based on the null hypothesis (H_0) and the alternative hypothesis (H_1). The null hypothesis ($H_0: \mu_1 = \mu_2$) indicates that the means of paired population are equal and the alternative hypothesis ($H_1: \mu_1 \neq \mu_2$) indicates that the means of paired population are not equal, where μ_1 represents the mean of variable 1 and μ_2 represents the mean of variable 2 (Kent State University, 2021).

The Paired Samples t-test can be calculated as the one sample t-test and accordingly the test statistic can be estimated with the following formula:

$$t = \frac{\bar{x}_{diff} - 0}{S_{\bar{x}}}$$

Where:

$$S_{\bar{x}} = \frac{S_{diff}}{\sqrt{n}}$$

where n indicates the sample size, $S_{\bar{x}}$ the standard error, $\bar{x}_{diff}$ the sample average of the differences and S_{diff} the sample standard deviation of the differences (Kent State University, 2021).

In order to determine if the null hypothesis will be rejected or not, the test statistic value should be compared for a selected confidence interval (95%) with the critical t value, which can be extracted from the t distribution table. Specifically, the null hypothesis can be rejected if the estimated t value is higher than the critical value from the t distribution table. Therefore, this indicates that the paired population means are significantly different (Kent State University, 2021). Another approach for this decision is to use the significance level and the p-value. Taking into consideration a significance level (α) of 0.05, the null hypothesis can be rejected when the p-value is less or equal to the significance level, $p - value \leq 0.05 = \alpha$ (Eberly College of Science, n.d.). It is important to be noted that a significance level of 0.05 represents a risk of 5% which leads to the conclusion that there is a difference when in reality no difference exists (Frost, 2022).

3.2. Machine Learning Models

3.2.1. Logistic Regression

Logistic regression is a machine learning algorithm that is mainly used in binary classification problems. Data are categorized into specific number of classes and label can be assigned to each class (Ezukwoke & Zareian, 2019). According to the literature, linear regression model is not an appropriate model to be implemented in a binary group and therefore will lead to undesirable results. To overcome this issue logistic model is used, which for all values of X gives results between 0 and 1. The logistic function that is used in logistic regression is the following (James, Witten, Hastie, & Tibshirani, 2021):

$$logit(p) = log(\frac{p}{1-p}) = b_0 + x_1 b_1 + \; x_2 b_2 + \; . \; . \; . \; + x_n b_n$$

In the formula above Y denotes the dependent variable, X the independent variables b_0 the intercept and b_i the coefficients. The logistic regression function produces an S-shaped curve with the following curve (James, Witten, Hastie, & Tibshirani, 2021), (Varsheni, 2021):

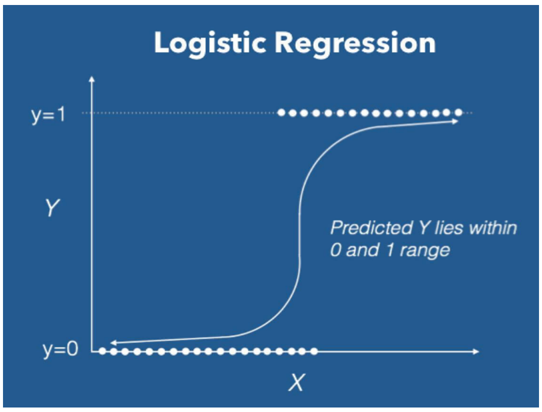

Figure 3 Logistic regression (Varsheni, 2021)

- **Interpretation of Coefficients**

In order to interpret the coefficients of logistic regression, it is important to understand the impact (increase or decrease) of a coefficient with respect to $logit(p)$ or $log\left(\frac{p}{1-p}\right)$. For instance, increasing variable x_1 by 1 unit will lead to a b_1 increase in $log\left(\frac{p}{1-p}\right)$. Hence, if $log\left(\frac{p}{1-p}\right)$ has an increase of b_1 this has consequence an increase of $(\frac{p}{1-p})$ by $\exp{(b_1)}$. Last but not least the increase (percentage) in the odds that an event will occur, taking into consideration that the rest variables of the function will remain fixed, can be estimated (Jankovic, 2021).

3.2.2. Support Vector Machines (SVM)

On their research Kunapuli, Bennett, Hu and Pang mention that SVM nowadays are a popular machine learning algorithm due to the fact that they can easily capture nonlinear relationships as well as they can be implemented to high-dimensional datasets that consist of thousand points. SVM can be applied to several problems, such as classification, ranking, regression and novelty detection. Although they are successful and reach high performance, some problematic points can be detected usually in model selection (Kunapuli, Bennett, Hu, & Pang, 2008).

According to Gandhi the main goal of support vector machine algorithm is to classify the data points finding an optimal hyperplane (or decision surface) in a space of n-dimensions. In the figure below possible hyperplanes and an optimal hyperplane can be observed (Gandhi, 2018):

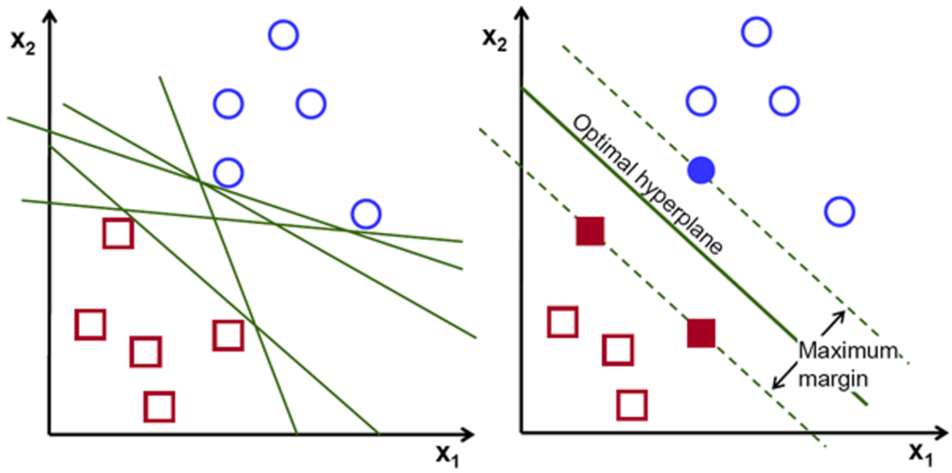

Figure 4 Possible hyperplanes (left) and optimal hyperplane (right), (Gandhi, 2018)

There are several possible hyperplanes that can be selected in order to separate two classes. Althought it is of great importance to find the optimal hyperplane, which is the one with the maximum distance between the data of two classes (maximum margin).

Finding the maximum margin distance can lead to a better performance in the classification of the future data (Gandhi, 2018).

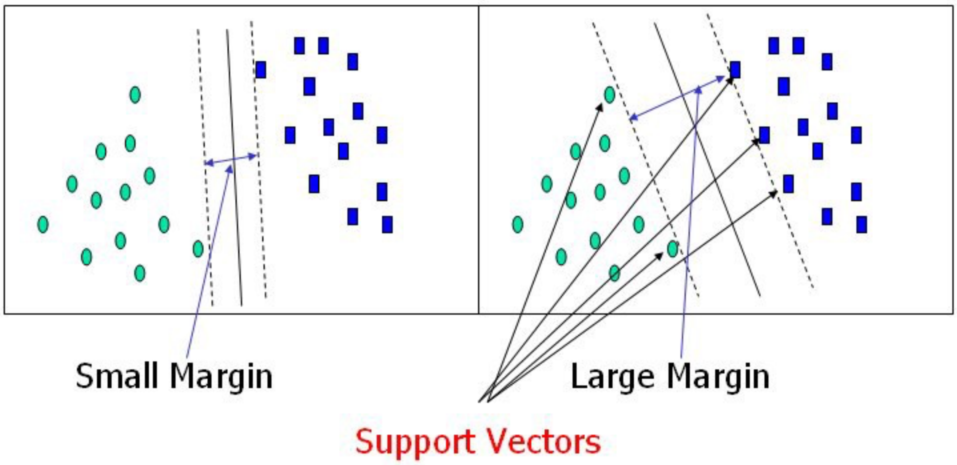

Figure 5 Support vectors (Gandhi, 2018)

Furthermore, Gandhi mentions that support vectors are important because their position can influence hyperplane's orientation, since they are the data points located closer to it. Therefore, support vectors contribute to maximizing the margin distance of the classifier and to the creation of the SVM model (Gandhi, 2018).

According to Misra, the main perspective of soft-margin SVM is to keep a high margin by allowing some mistakes, thus other data points can be classified accurately (Misra, 2019). The formula of soft-margin SVM is presented below (Kunapuli, Bennett, Hu, & Pang, 2008):

$$\min \ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i$$

$$s.t. \ \ y_i(\acute{w}x_i - b) \geq 1 - \xi_i \quad \forall i = 1 \ldots n, \xi_i \geq 0$$

Where C denotes a hyperparameter that minimizes the mistakes and maximizes the margin between the points. The value x represents the data points and ξ indicates the distance of a data point from the related margin of the class (Misra, 2019).

3.2.3. Artificial Neural Network (ANN)

The structure of an Artificial Neural Network (ANN) is similar to a human brain and is composed of neurons as well as synapses, which are ordered in layers (Gavrilova, 2020). In ANN, the perceptron model represents a biological neuron and was first created in 1957 at Cornell Aeronautical Laboratory, U.S., for image recognition based on binary classification and used data which were mainly linearly divided (Tyagi, 2020).

As mentioned by Tyagi there are two categories of perceptron models, the single-layered and the multi-layered. The single-layered perceptron is the simplest ANN which can be implemented in binary classification problems (class 0 and 1). At this model the inputs should be first weighted and if the outcome is the same with the expected result, this indicates that the model performed well and the weights should not be changed (Tyagi, 2020). The following figure depicts a single-layered perceptron model (Sayad, 2022):

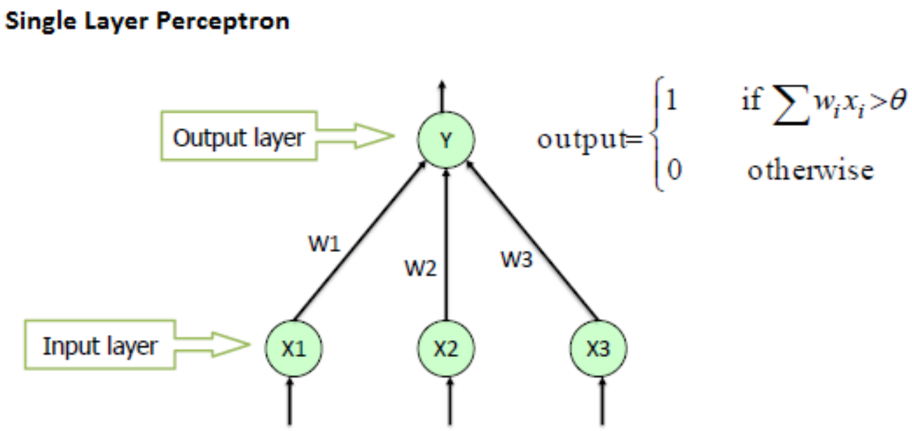

Figure 6 Single-layered perceptron model (Sayad, 2022)

The structure of a multi-layered perceptron model is similar to a single-layered, but consists of two or more hidden layers. This model carries out the forward stage as well as the backward stages. More specifically, in forward stage the activation functions are implied from the input until the output, while in the backward stage, the output begins backward in order to alter values of weights and bias. At this case, due to the multiple layers the activation function is not linear and it can be functions such as sigmoid, relu a.o. (Tyagi, 2020). Below a representation of the multi-layered perceptron model is depicted (Mohamed, Negm, Zahran, & Saavedra, 2015).

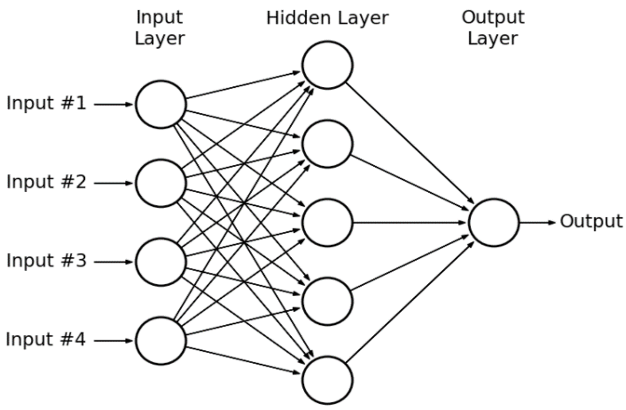

Figure 7 Multi-layered perceptron model (Mohamed, Negm, Zahran, & Saavedra, 2015)

3.2.4. Random Forest

Another popular and widely used machine learning algorithm is Random Forest, which can be implemented either in regression or classification problems. The algorithm creates decision trees based on different samples and chooses their average for regression problems and their majority vote in case of classification. Although Random Forest can handle data that contain continuous variables in case of regression problems, it has been observed that the algorithm performs better with datasets containing categorical variables and therefore in classification problems (Sruthi, 2021).

The following figure represents the steps that take place in Random Forest algorithm (Sruthi, 2021):

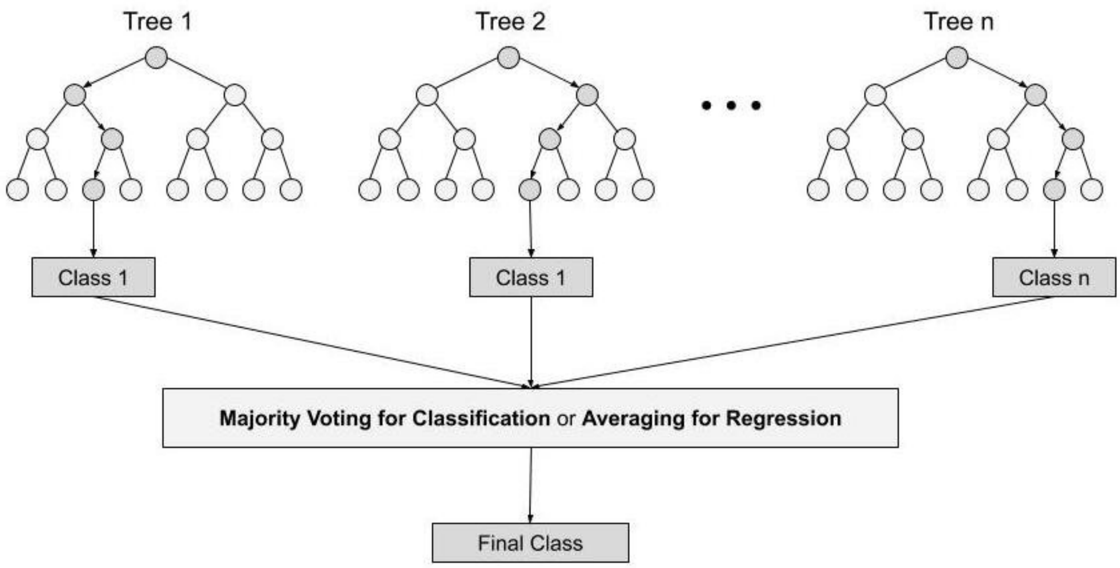

Figure 8 Steps of Random Forest algorithm (Sruthi, 2021)

As it can be observed in the figure above, firstly n-number of random records are chosen from the data, then different decision trees are created for each sample. At the third step each decision will have an output. Last but not least, the final result is estimated based on the average for regression problems and the majority vote for classification problems accordingly (Sruthi, 2021).

3.3. Model Evaluation Metrics

Evaluation of a machine learning model is of great importance, because through this it can be concluded if a model achieved a high or low performance (Mishra, 2018). There are several metrics that can be taken into consideration for this purpose such as Confusion matrix, Accuracy, Precision, Recall, F1 Score and Precision-Recall Curve (Yoshizawa, Nishiyama, Iwasaki, & Mizoguch, 2016). The explanation of these metrics is presented in the following part:

3.3.1. Confusion Matrix

An important evaluation metric for machine learning models is the confusion matrix. Specifically, it is a matrix that contains four combinations of actual and predicted values. These combinations are the following (Mohajon, 2020):

True Positive (TP) indicates that the predicted value is positive and the actual value is also positive.

True Negative (TN) represents that the predicted value is negative and the actual value is also negative.

False Positive (FP) or Type 1 Error denotes that the predicted value is positive and the actual value is negative.

False Negative (FN) or Type 2 Error designates that the predicted value is negative and the actual value is positive.

A confusion matrix for binary classification can be presented as follows (Mohajon, 2020):

Predicted Class \ True Class	Positive	Negative
Positive	TP	FP
Negative	FN	TN

Table 2 Confusion matrix (Mohajon, 2020)

To achieve a high performance of a machine learning model it is significantly important to minimize both False Negative (FN) and False Positive (FP) cases. Therefore, the higher number in the results should be observed in True Positive (TP) as well as True negative (TN) cases (Singh, 2021).

3.3.2. Accuracy

Represents the ratio of correct predictions to the total number of predictions (Mishra, 2018).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

3.3.3. Precision

Refers to the ratio of correct positive results to the number of positives predicted by the classifiers (Mishra, 2018).

$$Precision = \frac{TP}{TP + FP}$$

3.3.4. Recall

Indicates the ratio of correct positive results to all the samples which should have been predicted as positive (Mishra, 2018).

$$Recall = \frac{TP}{TP + FN}$$

3.3.5. F1 Score

Represents the harmonic mean between Precision and Recall (Mishra, 2018).

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Preecision + Recall}$$

The range of the metrics presented above is [0,1]. The higher the result of a metric the better the performance of the model. It is important to mention that in most of the cases a high accuracy is not a real indication of high model performance (Mishra, 2018).

3.4. Cross Validation

K-Fold Cross validation is not a metric but a procedure for evaluating a machine learning model on a specific data sample. It is a well-known and intuitive method, mainly used to assess the expected performance of the model to predict on unseen data, referring to data that have not been considered for the training of the model (Brownlee, Machine Learning Mastery, 2020). This method provides an assurance that the chosen model is low on variance and bias, thus most of data patterns were interpreted correctly and the noise of the data was excluded (Gupta, 2017).

Conforming to Baheti, cross validation technique allows the hyperparameters of a machine learning model to be tuned as well as contributes on preventing overfitting. Specifically, at k-fold cross validation data are distributed into k same sized smaller sets, which are known as folds. One fold is used for testing the model and the rest of the folds for training. In every iteration one fold is used for evaluating the performance of the model. Lastly, the average of scores of all iterations (total folds) is estimated, in order to result the overall performance of the selected model (Baheti, 2022).

Figure 9 K-fold cross validation (Baheti, 2022)

4. Experiment Design and Data Implementation

4.1. i-DREAMS Project

A driving simulator study was conducted due to a collaboration between i-DREAMS project and the chair of Transportation System Engineering (TSE) of Technical University of Munich. The i-DREAMS project is a founded research programme from European Union, which focuses on defining, developing, testing as well as validating a Safety Tolerance Zone in order to prevent unsafe and dangerous situations by implementing real time interventions and distractions during the diving task (i-DREAMS, 2022).

Figure 10 i-DREAMS driving simulator, Chair of Transportation System Engineering (TSE, 2021)

4.2. Experiment Design

4.2.1. Scenarios

A driving simulator study was conducted at the chair of Transportation System Engineering with 60 participants. Specifically, the drivers participated in 3 driving scenarios with approximate duration of 15 minutes. First driving scenario was monitoring, in which each participant was driving without any intervention. Intervention scenario was the second drive in which the participant was receiving warnings from the i-Dreams warning system about speed and time headway. The third scenario was distraction, in which the participant along with the warnings was receiving text messages, that had to be read and sometimes answered.

- **Monitoring**

At monitoring scenario participants were driving without receiving any intervention or distraction. The duration of the drive was approximately 15 to 17 minutes. Furthermore, during monitoring scenario participants drove through all the road sections, which were highway, rural and urban road section. The same road sections appeared in the other two scenarios (intervention, distraction) as well, but with a different sequence.

- **Interventions**

During intervention scenario there were warnings appearing from the from the i-Dreams warning system based on the time headway of the driver with the leading vehicle. The warnings were triggered for different time headway values depending on the speed of the vehicle. The thresholds of the speed that were taken into consideration were less or equal to 50 km/h, less or equal to 90 km/h and higher than 90 km/h.

For instance, the moment that the participant drives with a speed less or equal to 50 kilometres per hour and has a time headway between 2.5 s and higher 1.4 s the time headway will be displayed in the screen with green color. The first stage of headway warning appears in red color when the driver reaches a time headway between 1.4 s and higher than 0.6 s. Lastly the second stage of headway warning appears for headway values lower or equal to 0.6 s and comes along with a red blinking warning. The following tables depict the time headway thresholds based on the different driving speed:

Speed $\leq$ 50 km/h		
THW thresholds	Status	Warning displayed
1.4 s $<$ THW $\leq$ 2.5 s	Vehicle detected	THW in green
0.6 s $<$ THW $\leq$ 1.4 s	First warning stage	Vehicle symbol and THW in red
THW $\leq$ 0.6 s	Second warning stage	Vehicle symbol red and blinking, THW in red

Table 3 Warnings corresponding to speed ≤ 50 km/h (i-Dreams project)

Speed ≤ 90 km/h		
THW thresholds	Status	Warning displayed
1.2 s < THW ≤ 2.5 s	Vehicle detected	THW in green
0.6 s < THW ≤ 1.2 s	First warning stage	Vehicle symbol and THW in red
THW ≤ 0.6 s	Second warning stage	Vehicle symbol red and blinking, THW in red

Table 4 Warnings corresponding to speed ≤ 90 km/h (i-Dreams project)

Speed > 90 km/h		
THW thresholds	Status	Warning displayed
1.0 s < THW ≤ 2.5 s	Vehicle detected	THW in green
0.6 s < THW ≤ 1.0 s	First warning stage	Vehicle symbol and THW in red
THW ≤ 0.6 s	Second warning stage	Vehicle symbol red and blinking, THW in red

Table 5 Warnings corresponding to speed >90 km/h (i-Dreams project)

The following picture presents a second stage headway warning with a red blinking vehicle appearing in the i-Dreams warning system.

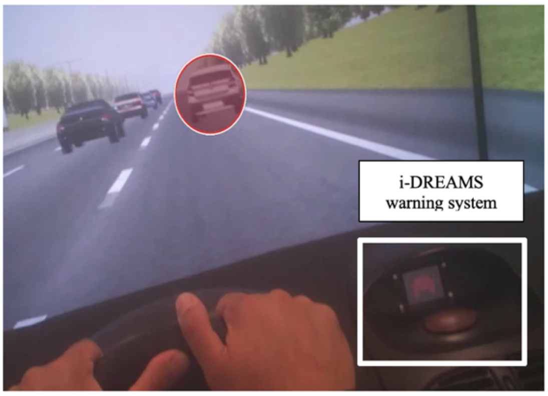

Figure 11 i-Dreams warning system and eye movement (red dot)

- **Distraction**

During the distraction scenario participants were receiving text messages, which they had to read and/or reply. There were eight text messages in total, from which the six were sent before a dangerous event and two of them during normal driving conditions (no event). Before the beginning of the drive participants received the instructions to answer only in messages that included a question and read the rest of messages (Ezzati Amini, et al., 2022). The following table presents the distraction task (reading, replying to text messages), the complexity of the task as well as the script of the text messages (Ezzati Amini, et al., 2021):

Distraction task	Complexity	Script of text message
Reading	simple	“Thank you for participating in the experiment”
Reading & replying	complex	“Can you name two cities you want to visit?”
Reading	simple	Your dentist appointment is scheduled for 30/04/2021 at 14:15”
Reading & replying	simple	“Where is your hometown?”
Reading	simple	“Nice to see you at the café yesterday”
Reading	simple	“50% off on online orders! Today only!”
Reading & replying	complex	“What are two things that you enjoy doing the most?”
Reading & replying	complex	“27+30=?”

Table 6 Distraction tasks (Ezzati Amini, et al., 2021)

4.2.2. Road Sections

According to i-Dreams project here were three road sections appearing while driving, highway, urban and rural road section. These road sections were appearing with a different sequence in each driving scenario. The three possible ways that the road sections were appearing in the simulator as well as the total distance of each road section is presented in the tables that follow:

Case A	Distance (m)
1. Rural	0-4400
2. Urban	4400-8000
3. Highway	8000-13500

Table 7 Road sections sequence of case A (i-Dreams project)

Case B	Distance (m)
1. Urban	0-5440
2. Highway	5440-9040
3. Rural	9040-13440

Table 8 Road sections sequence of case B (i-Dreams project)

Case C	Distance (m)
1. Highway	0-3600
2. Rural	3600-8000
3. Urban	8000-13500

Table 9 Road sections sequence of case C (i-Dreams project)

For each participant these cases (A, B, C) were allocated before the beginning of the driving task. For instance, if one participant was selected to have case A for monitoring scenario, this means that in the simulator screen it was appearing first the rural road environment, then the urban and last the highway. The figures below depict the different road environments:

Figure 12 Highway road section (i-Dreams project)

Figure 13 Rural road section (i-Dreams project)

Figure 14 Urban road section (i-Dreams project)

It is important to be noted that each road section had a different speed limit as well as a different number of lanes per direction. More specifically, the highway section had no speed limit and three lanes per direction. In the rural area there was a speed limit of 70 km/h and two lanes per direction and last but not least the urban area had a speed limit of 50 km/h and one lane per direction. The following table presents the different speed limits and number of lanes for each road section separately:

Road section	**Speed limit**	**Number of lanes**
Highway	No speed limit	3 per direction
Rural	70 km/h	2 per direction
Urban	50 km/h	1 per direction

Table 10 Speed limit and number of lanes per road section (i-Dreams project)

4.2.3. Participant Information

Driving simulator data were collected from 60 drivers that participated in all driving scenarios and completed successfully the driving task. Furthermore, participants did not show signs of dizziness nor felt uncomfortable during driving. The experiments took place between June and September 2021 at the chair of Transportation System Engineering (TSE) of Technical University of Munich.

From the 60 drivers that participated in the experiment, 35 were males and 25 females. Regarding the age participants were divided in three groups. Specifically, the first group consisted of participants between 18 and 25 years old, the second was between 26 and 45 years and the last group was between 46 and 64 years old. These characteristics are presented in the following figures.

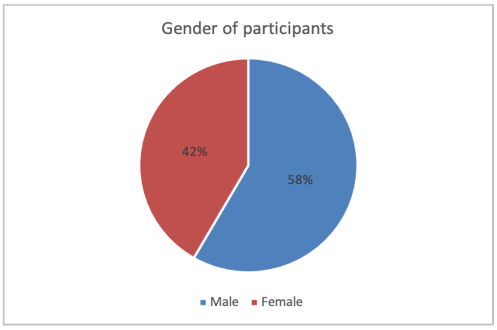

Figure 15 Gender of participants (i-Dreams project)

In the bar chart below, the higher number of female participants is observed in the first group, on the other hand a higher number of males is noted both in the second and third group.

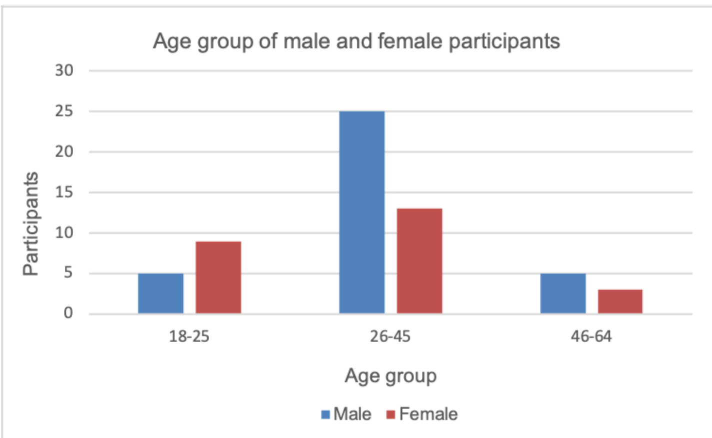

Figure 16 Age group of male and female participants (i-Dreams project)

Moreover, it should be mentioned that participants come from several different countries. As it can be observed in the following figure, the majority of drivers comes from Germany and the second higher number of participants has Egyptian and Greek nationality.

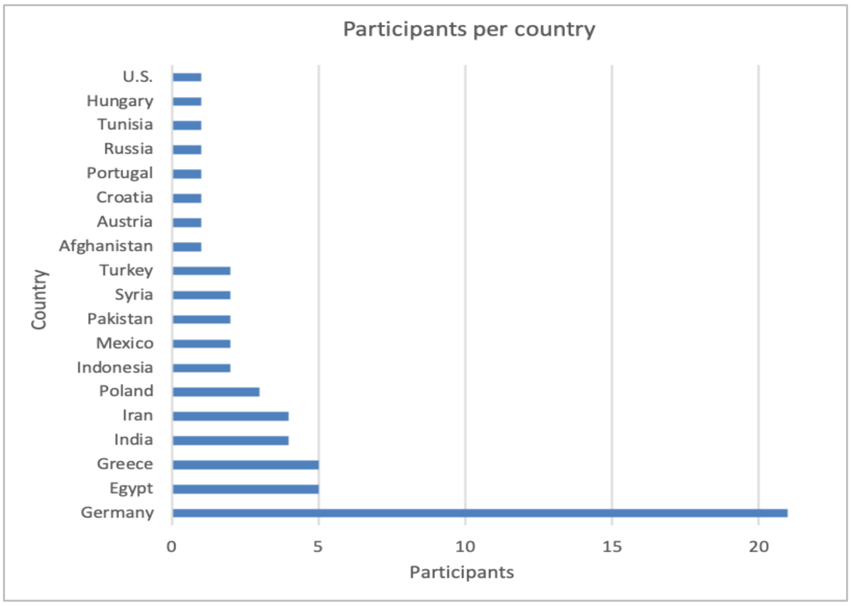

Figure 17 Participants per country (i-Dreams project)

4.3. Data Implementation

4.3.1. Driving Simulator Data

The variables that were extracted from the driving simulator are depicted on the following table:

Variable	**Description**	**Unit**
ElapsedTime	Time since start	s
LongAcc	Longitudinal acceleration	m/s^2
LatAcc	Lateral acceleration	m/s^2
LongVelocity	Longitudinal velocity	m/s
LatVelocity	Lateral Velocity	m/s
TotalLongDistTravelled	Total distance driving	m
LatPos	Lateral position	m
SteeringWheelAngle	Steering wheel input	$degrees$
SpeedLimitsMs	Current speed limit	m/s

Variable	Description	Unit
SpeedLimitsKph	Current speed limit	km/h
Headway	Time headway to vehicle ahead	s
TTC	Time to collision with vehicle ahead	s
GasPedalPercentageDisplayed	Percentage of max gas pedal	0-1
BrakePedalPercentageDisplayed	Input count of brake pedal	0-1

Table 11 Driving simulator variables (i-Dreams project)

4.3.2. Data Visualization

The following plots present some of the variables by distance travelled, using as an example the data of one participant for the three scenarios (monitoring, intervention, distraction):

- **Longitudinal velocity by distance**

The following plot depicts the longitudinal velocity by distance for one participant during the three scenarios. It can be observed that the participant during monitoring scenario maintains a lower velocity at the first part of the drive, which increases at the second part. On the other hand, during intervention and distraction scenario the participant reaches higher velocity values in the beginning of the drive, which decrease to approximately 10 m/s until the end. In order to understand better both plot and driving behavior, it is meaningful to examine the plot "Headway by Distance travelled" that follows.

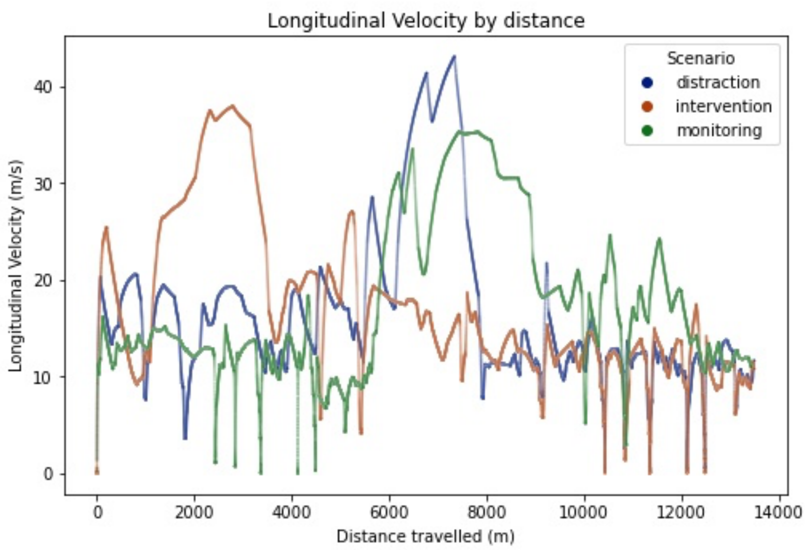

Figure 18 Longitudinal velocity by distance (own plot)

- **Headway by distance**

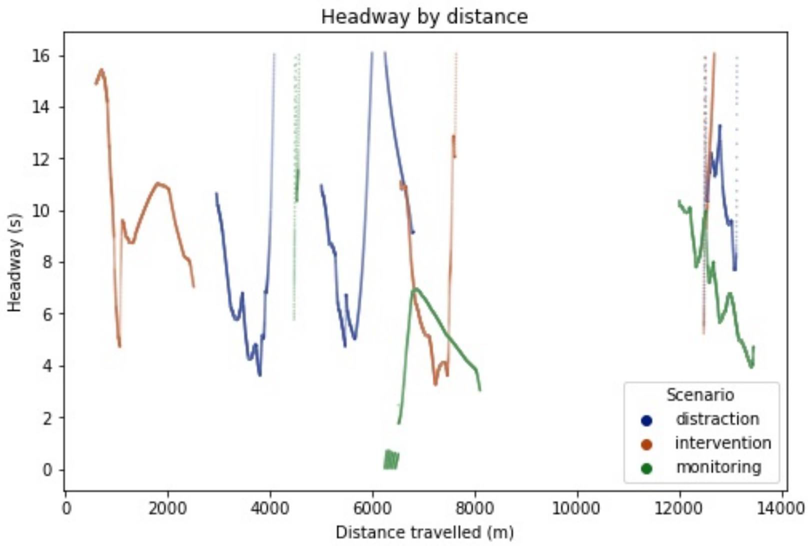

Figure 19 Headway by distance (own plot)

The plot above presents the headway per distance travelled for the same participant during the three driving scenarios. Different headway values are observed during the drive with the lowest value to be reached at monitoring scenario in the middle of the drive.

The observation of both plots simultaneously can lead to the following conclusion. At the road sections that low values of time headway are noted the participant drives with high velocity values. For instance, the lowest headway values for monitoring scenario are observed between 6000 and 8000 m. At the same part of the road are also noted the highest velocity values, between 30 and 35 m/s. This fact indicates a dangerous driving behavior, where the participant drives significantly close to the leading vehicle.

- **Longitudinal velocity by distance**

The following plot presents the longitudinal velocity by distance for another participant during the monitoring, intervention and distraction scenario. In this plot a different driving behavior is observed. More specifically, the participant drives with high velocity at the first part of the drive during distraction scenario. At the middle part (5000-9000 m) higher velocity values are noticed during intervention scenario. At the end of the drive there are similar values for intervention and distraction scenario, while during monitoring are slightly higher. The plot longitudinal velocity by distance for this participant is depicted below:

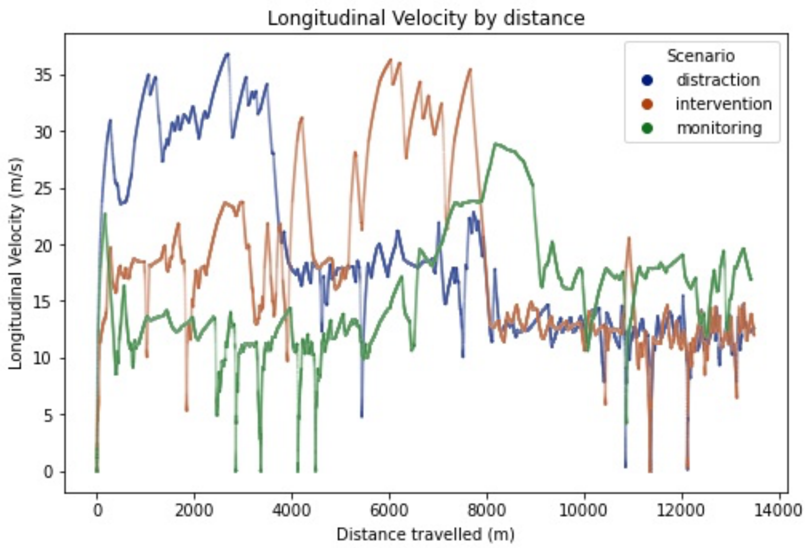

Figure 20 Longitudinal velocity by distance (own plot)

- **Headway by distance**

The plot below presents the headway per distance travelled for participant Id=47 during the driving scenarios. Significantly low headway values are observed mainly during distraction and intervention scenario, while for monitoring scenario lower headway values can be seen at the end of the drive. This plot describes probably a behavior, in which the driver took into consideration the warnings displayed during intervention and distraction and therefore reached lower headway values with the leading vehicle and drove safer.

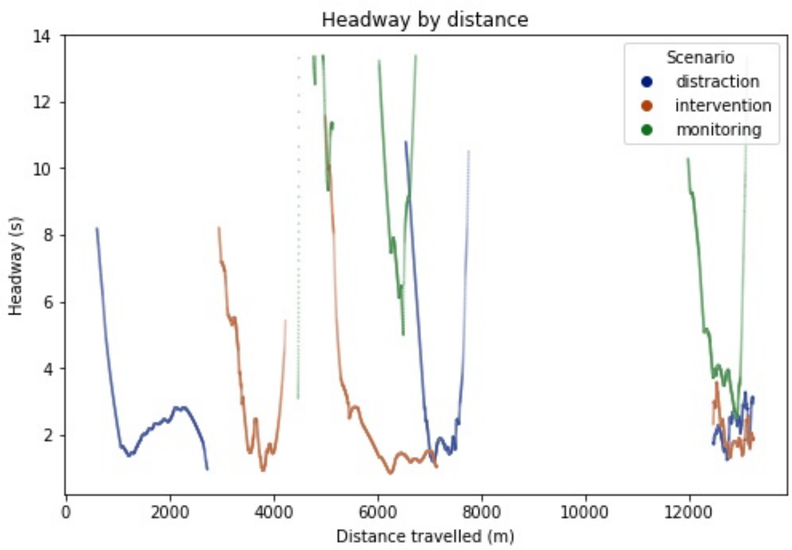

Figure 21 Headway by distance (own plot)

5. Results and Discussion

5.1. Effectiveness Evaluation of Interventions

5.1.1. Data Preparation

Data of 60 participants from monitoring, intervention and distraction scenario were used at this part. Furthermore, data corresponding to a time headway greater than 2.5 s were excluded. This threshold was selected based on the i-DREAMS warning system, in which the leading vehicle was detected in a time headway less or equal to 2.5 s. Moreover, taking into consideration the literature, a time headway higher or equal to 2 s is considered safe (Monteiro, Balogun, Kote, & Tlhabano, 2014), (Wang & Song, 2011). Therefore, a threshold of higher than 2.5 s was considered as safe and data referring to this were removed.

Moreover, data were labelled based on the time headway, where a safe condition (class 0) was referring to data corresponding to a time headway between 1.4 s and 2.5 s, while an unsafe condition (class 1) was defined by a time headway less or equal to 1.4 s. These thresholds were also considered based on the i-DREAMS warning system design. As already mentioned, the i-DREAMS warning system has three thresholds (1.4 s, 1.2 s, 1.0 s) depending on the speed, in which the first warning is triggered. Therefore, a time headway of 1.4 s was selected as the threshold for classifying the data, since it also includes the data that correspond to the lower values of time headway (1.2 s, 1.0s).

Variables from the driving simulator data were selected, such as elapsed time, distance travelled, longitudinal and lateral acceleration, longitudinal and lateral velocity, lateral position, headway and TTC. Data were aggregated based on the dangerous event. A dangerous event starts, when the time headway with the leading vehicle is THW $\leq$ 1.4s and lasts until this value becomes higher than 1.4 s, when the event stops. Due to the events new values were created like duration and distance of the event. Last but not least, before the implementation of statistical analysis, one critical event was chosen for each participant, that included the minimum variables of all the events for one scenario.

In order to examine the impact of interventions in driving task a statistical analysis was implemented. More specifically, a paired sample t test was conducted between monitoring, intervention and distraction scenario. The variables that are statistically significant correspond to a 95% confidence interval, which is denoted with a p-value less or equal to 0.05 (noted in bold). The results of the statistical analysis are presented in the following tables.

5.1.2. Statistical Analysis Results

⇒ **T-test between Monitoring and Intervention scenario**

Variable		T-value	P-value
Longitudinal Acceleration	Minimum	5.147	**0.000**
	Maximum	0.792	0.434
	Mean	3.546	**0.001**
	Standard deviation	-3.961	**0.000**
Lateral Acceleration	Minimum	1.364	0.182
	Maximum	0.947	0.350
	Mean	-0.405	0.688
	Standard deviation	1.161	0.254
Longitudinal Velocity	Minimum	-3.667	**0.001**
	Maximum	-3.957	**0.000**
	Mean	-3.928	**0.000**
	Standard deviation	-3.166	**0.003**
Lateral Velocity	Minimum	0.316	0.754
	Maximum	0.914	0.367
	Mean	0.656	0.517
	Standard deviation	1.161	0.254
Lateral Position	Standard deviation	0.355	0.725
Headway	Minimum	-4.679	**0.000**
	Standard deviation	-1.042	0.305
TTC	Minimum	-2.853	**0.007**
	Standard deviation	-2.419	**0.021**
Duration	Minimum	-2.346	**0.025**
Distance	Minimum	-2.143	**0.039**

Table 12 T-test results between Monitoring and Intervention scenario

In the table above, it can be observed that the population means are not significantly different for Lateral Acceleration, Lateral Velocity and Lateral Position, since the p-value is higher than 0.05. On the other hand, population means appear to be significantly different for Longitudinal Acceleration, Longitudinal Velocity, Headway, TTC, Duration and Distance, where p-value is less or equal to 0.05.

Since a significant difference in population means is observed for most of the variables, it can be concluded that there is a significant impact of interventions on driving behaviour. For instance, the negative T-values of Longitudinal Velocity and Headway indicate that the sample means of both variables increased from monitoring to intervention scenario. More specifically, at intervention scenario during critical events, although participants were driving with a higher longitudinal velocity, they managed to maintain a higher time headway from the leading vehicle. It is therefore observed that real time interventions from the i-DREAMS warning system by informing drivers timely about critical situation contributed to their safety.

The changes between monitoring and intervention scenario are depicted in the box plots below for the variables longitudinal and lateral velocity, longitudinal and lateral acceleration and headway.

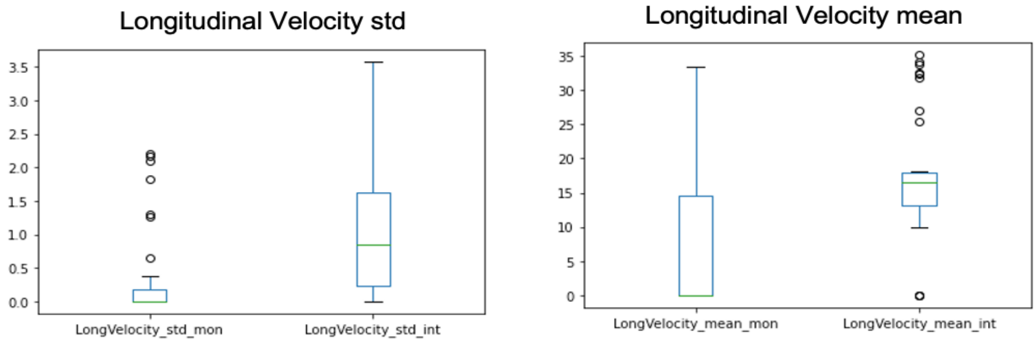

Figure 22 Box plots of longitudinal velocity std (left) and longitudinal velocity mean (right)

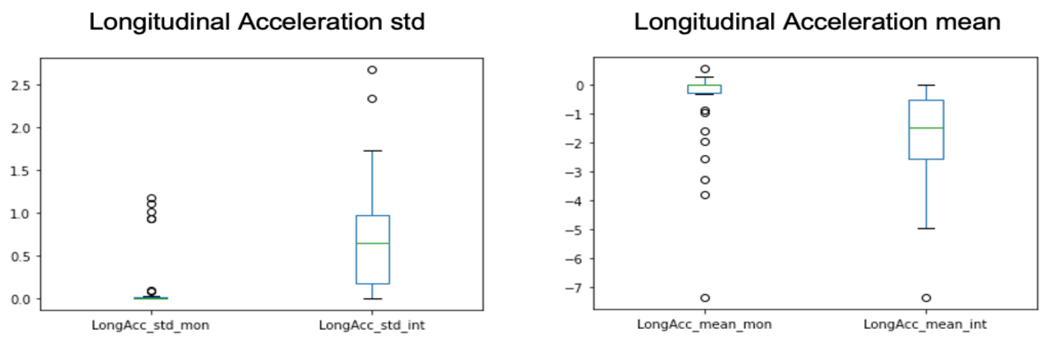

Figure 23 Box plots of longitudinal velocity std (left) and longitudinal velocity mean (right)

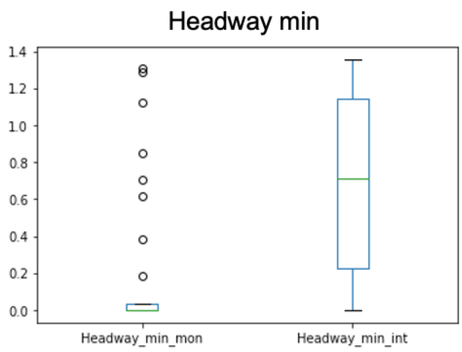

Figure 24 Box plot of headway min

⇒ **T-test between Monitoring and Distraction scenario**

Variable		T-value	P-value
Longitudinal Acceleration	Minimum	2.669	**0.012**
	Maximum	-0.296	0.690
	Mean	1.280	0.210
	Standard deviation	-2.003	0.054
Lateral Acceleration	Minimum	2.091	**0.045**
	Maximum	1.291	0.206
	Mean	-0.152	0.880
	Standard deviation	0.877	0.387
Longitudinal Velocity	Minimum	-0.970	0.340
	Maximum	-1.503	0.143
	Mean	-1.317	0.198
	Standard deviation	-2.031	0.051
Lateral Velocity	Minimum	0.092	0.927
	Maximum	0.130	0.900
	Mean	-0.097	0.924
	Standard deviation	1.133	0.266
Lateral Position	Standard deviation	0.166	0.869
Headway	Minimum	-2.014	0.053
	Standard deviation	-1.347	0.188
TTC	Minimum	-2.463	**0.019**
	Standard deviation	-2.600	**0.014**
Duration	Minimum	-2.241	**0.032**
Distance	Minimum	-2.022	0.052

Table 13 T-test results between Monitoring and Distraction scenario

Comparing monitoring and distraction scenario a significant difference in population means can be noted for time to collision, duration of critical event, minimum longitudinal acceleration and minimum lateral acceleration. Regarding the T-value can be observed that minimum longitudinal acceleration as well as lateral acceleration were decreased between monitoring and distraction scenario. Furthermore, the minimum duration of critical event during distraction increased, similarly to standard deviation of TTC as well as minimum TTC. Although, participants during distraction were receiving interventions and text messages, they had a similar driving behaviour with monitoring scenario (no interventions, no text messages), but their Tme to Collision increased.

Box plots of minimum longitudinal acceleration and minimum duration of the event are depicted below:

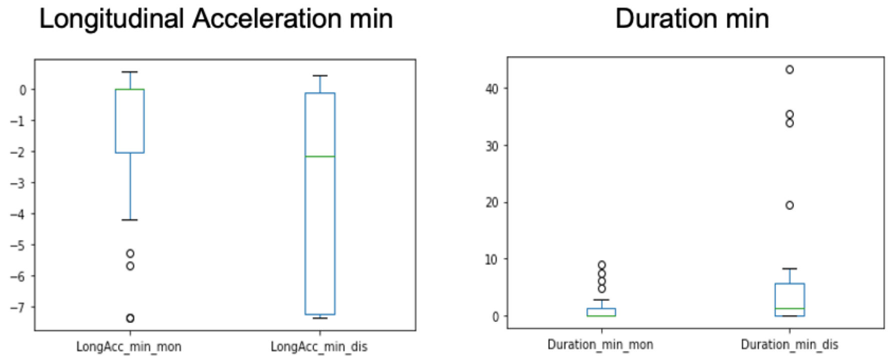

Figure 25 Box plots of longitudinal acceleration min (left) and duration min (right)

⇒ **T-test between Intervention and Distraction scenario**

Variable		**T-value**	**P-value**
Longitudinal Acceleration	Minimum	-1.557	0.128
	Maximum	-0.967	0.340
	Mean	-1.383	0.175
	Standard deviation	2.287	**0.028**
Lateral Acceleration	Minimum	0.478	0.636
	Maximum	0.954	0.346
	Mean	0.418	0.678
	Standard deviation	-0.275	0.785
Longitudinal Velocity	Minimum	2.607	**0.013**
	Maximum	2.319	**0.026**
	Mean	2.536	**0.016**
	Standard deviation	0.933	0.357
Lateral Velocity	Minimum	-0.194	0.847
	Maximum	-0.485	0.631
	Mean	-0.555	0.582
	Standard deviation	0.331	0.743
Lateral Position	Standard deviation	-0.176	0.861
Headway	Minimum	1.990	0.054
	Standard deviation	-0.598	0.553
TTC	Minimum	1.492	0.144
	Standard deviation	0.762	0.451
Duration	Minimum	-0.778	0.442
Distance	Minimum	0.144	0.887

Table 14 T-test results between Intervention and Distraction scenario

A comparison between intervention and distraction scenario appears to have less variables with a significant difference in population means. Specifically, this is observed in longitudinal velocity and standard deviation of longitudinal acceleration. Furthermore, this indicates that in comparison with intervention drivers during distraction were driving with lower longitudinal velocity while dealing with critical situations. At this case the interventions helped the participants to minimize their speed and avoid a collision.

Box plots of standard deviation of longitudinal acceleration and average longitudinal velocity follow:

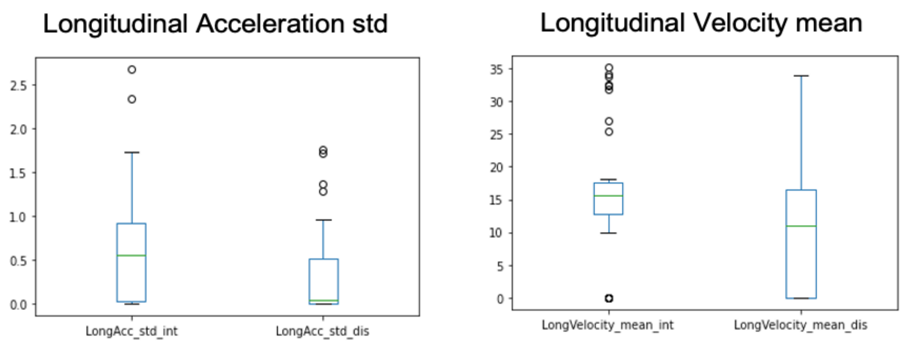

Figure 26 Box plots of longitudinal acceleration std (left) and longitudinal velocity mean (right)

Comparing the three scenarios showed that the most significant impact of interventions appears in the results of the statistical test between monitoring and intervention scenario. Thus, the driving behaviour during intervention while dealing with critical events changed. It is meaningful to observe the results of statistical analysis between monitoring and intervention while driving in different road environments (urban, rural, highway) during critical driving situations.

⇒ **T-test between Monitoring and Intervention scenario in Urban Road section**

The following table depicts the statistical test results between monitoring and intervention scenario for critical events that happened in urban road section.

It is observed that only minimum longitudinal velocity and minimum headway have a p-value lower than 0.05. This indicates that there is no significant change between the samples of monitoring and intervention for driving at the urban section. Moreover, T-values of both longitudinal velocity und headway denote an increase in intervention scenario. Thus, the participants during urban road section were driving with a higher speed and maintaining at the same time higher headway. It can be concluded that at this road section the drivers took into consideration the warnings from the i-DREAMS warning system and maintained higher and safer time headway from the leading vehicle.

Variable		T-value	P-value
Longitudinal Acceleration	Minimum	1.002	0.343
	Maximum	0.391	0.705
	Mean	0.667	0.521
	Standard deviation	-1.539	0.158
Lateral Acceleration	Minimum	-0.943	0.370
	Maximum	0.967	0.359
	Mean	-0.777	0.457
	Standard deviation	0.905	0.389
Longitudinal Velocity	Minimum	-2.313	**0.046**
	Maximum	-2.184	0.057
	Mean	-2.203	0.055
	Standard deviation	-1.028	0.331
Lateral Velocity	Minimum	-0.757	0.467
	Maximum	-0.620	0.551
	Mean	-0.566	0.585
	Standard deviation	0.854	0.415
Lateral Position	Standard deviation	0.565	0.586
Headway	Minimum	-2.400	**0.040**
	Standard deviation	0.966	0.359
TTC	Minimum	-2.078	0.067
	Standard deviation	1.416	0.190
Duration	Minimum	-1.603	0.143
Distance	Minimum	-1.571	0.151

Table 15 T-test results between Monitoring and Intervention scenario in Urban Road section

Box plots of minimum longitudinal velocity and minimum headway are presented below:

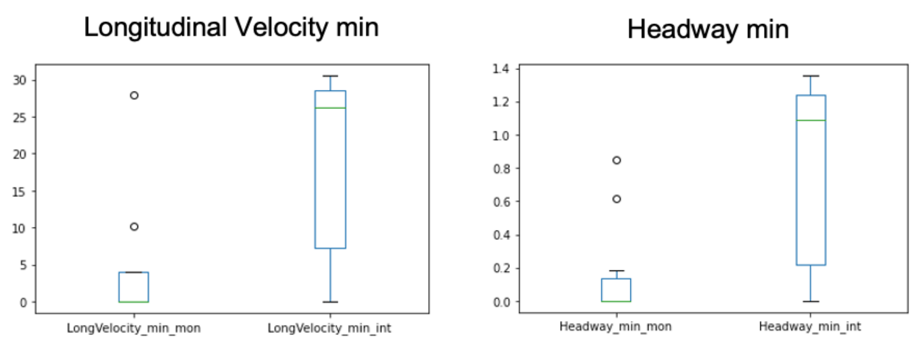

Figure 27 Box plots of longitudinal velocity min (left) and headway min (right)

⇒ **T-test between Monitoring and Intervention scenario in Rural Road section**

The following table presents the statistical test results between monitoring and intervention scenario for critical events that take place in rural road section.

Variable		T-value	P-value
Longitudinal Acceleration	Minimum	4.330	**0.000**
	Maximum	-1.053	0.304
	Mean	3.591	**0.002**
	Standard deviation	-4.481	**0.000**
Lateral Acceleration	Minimum	0.413	0.684
	Maximum	-1.665	0.110
	Mean	-1.057	0.302
	Standard deviation	-0.981	0.337
Longitudinal Velocity	Minimum	-3.917	**0.000**
	Maximum	-4.486	**0.000**
	Mean	-4.437	**0.000**
	Standard deviation	-2.831	**0.009**
Lateral Velocity	Minimum	0.536	0.597
	Maximum	0.806	0.429
	Mean	0.944	0.356
	Standard deviation	0.196	0.847
Lateral Position	Standard deviation	-0.551	0.587
Headway	Minimum	-5.220	**0.000**
	Standard deviation	-2.466	**0.022**
TTC	Minimum	-2.240	**0.036**
	Standard deviation	-2.746	**0.012**
Duration	Minimum	-2.122	**0.045**
Distance	Minimum	-2.098	**0.048**

Table 16 T-test between Monitoring and Intervention scenario in Rural Road section

The statistical analysis results showed that most of the variables reached a p-value lower than 0.05, which denotes a significant change between the samples of monitoring and intervention during driving in rural road environment. More specifically, an increase is observed for longitudinal velocity, longitudinal acceleration, time headway, time to collision, duration and distance of the event. This indicates that the participants use the warnings in order to keep a safe time headway and time to collision and this contributes to their road safety.

Box plots of average longitudinal velocity and minimum headway are presented below:

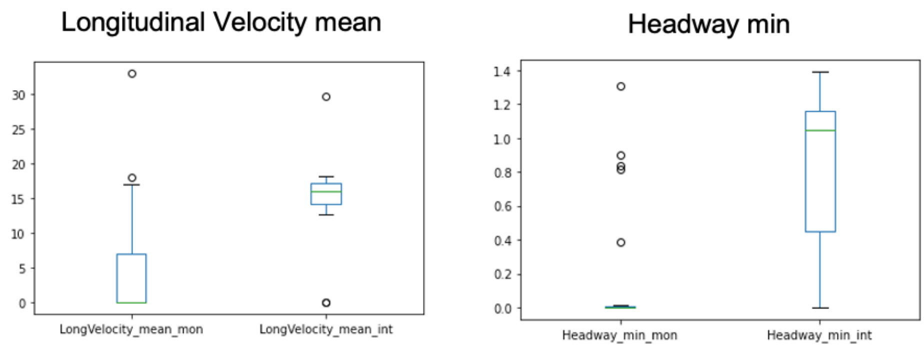

Figure 28 Box plots of longitudinal velocity min (left) and headway min (right)

⇒ T-test between Monitoring and Intervention scenario in Highway section

The following table presents the statistical analysis results between monitoring and intervention scenario for critical events that happened in the highway road section.

Variable		T-value	P-value
Longitudinal Acceleration	Minimum	3.736	**0.001**
	Maximum	-0.118	0.907
	Mean	2.605	**0.015**
	Standard deviation	-3.483	**0.002**
Lateral Acceleration	Minimum	2.507	**0.019**
	Maximum	-0.415	0.681
	Mean	-0.576	0.570
	Standard deviation	-0.028	0.978
Longitudinal Velocity	Minimum	-2.456	**0.021**
	Maximum	-2.805	**0.010**
	Mean	-2.746	**0.011**
	Standard deviation	-3.039	**0.005**
Lateral Velocity	Minimum	0.776	0.445
	Maximum	0.746	0.463
	Mean	1.361	0.186
	Standard deviation	0.011	0.992
Lateral Position	Standard deviation	-0.737	0.468
Headway	Minimum	-2.638	**0.014**
	Standard deviation	-2.177	**0.039**
TTC	Minimum	2.325	**0.028**
	Standard deviation	-1.985	0.058
Duration	Minimum	-2.846	**0.009**
Distance	Minimum	-2.715	**0.012**

Table 17 T-test results between Monitoring and Intervention scenario in Highway section

Similar results with the statistical analysis in the rural road section are observed in the highway section. Most of the variables have a low p-value (less than 0.05) and this indicates a high impact of interventions on driving in the highway. Specifically, longitudinal velocity, headway, duration and distance of the events increased from monitoring to intervention scenario. While minimum longitudinal acceleration and minimum TTC decreased from one scenario to the other.

Box plots of minimum headway and average longitudinal velocity are depicted below:

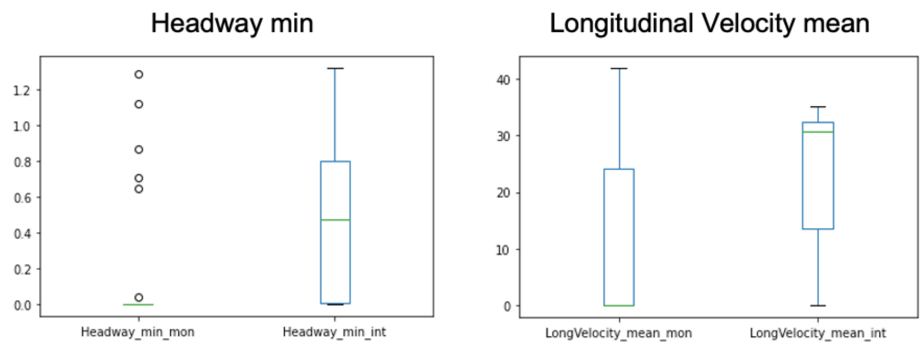

Figure 29 Box plots of longitudinal velocity min (left) and headway min (right)

5.2. Dangerous Event Prediction

5.2.1. Data Preparation

Data of 60 participants from the distraction scenario were implemented for the prediction of dangerous event. Furthermore, data corresponding to a time headway greater than 2.5 s were excluded. This threshold was selected based on the i-DREAMS warning system, in which the leading vehicle was detected in a time headway less or equal to 2.5 s. Moreover, taking into consideration the literature, a time headway higher or equal to 2 s is considered safe (Monteiro, Balogun, Kote, & Tlhabano, 2014), (Wang & Song, 2011). Therefore, a threshold of higher than 2.5 s was considered as safe and data referring to this were removed.

A binary classification of the data based on the time headway followed, where a safe condition (class 0) was referring to data corresponding to a time headway between 1.4 s and 2.5 s, while an unsafe condition (class 1) was defined by a time headway less or equal to 1.4 s. These thresholds were also considered based on the i-DREAMS warning system design. As already mentioned previously, the i-DREAMS warning system has three thresholds (1.4 s, 1.2 s, 1.0 s) depending on the speed, in which the first warning is triggered. Therefore, a time headway of 1.4 s was selected as the threshold for classifying the data, since it also includes the data that correspond to the lower values of time headway (1.2 s, 1.0s).

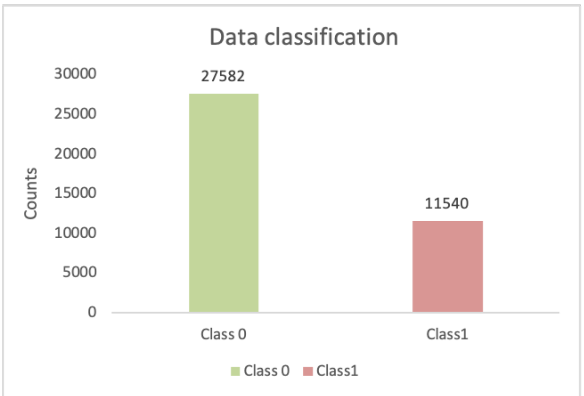

Figure 30 Data classification based on time headway

In order to implement the machine learning models, the data were first divided in training and testing data. More specifically, as training data were used the 80% of the total and for the testing the 20%. It is important these two groups of data to be different in order to test the model properly. Furthermore, the data were scaled between 0 and 1 with the MinMaxScaler in order to reach a higher performance of the models.

The following table presents the independent variables that were used for the implementation of the machine learning algorithms:

Variables	
1) Elapsed time	7) Lateral position
2) Longitudinal acceleration	8) Steering wheel angle
3) Lateral acceleration	9) Gas pedal percentage
4) Longitudinal velocity	10) Brake pedal percentage
5) Lateral velocity	11) Leading vehicle lateral position
6) Distance travelled	12) Leading vehicle longitudinal position

Table 18 Variables used for the dangerous event prediction

5.2.2. Logistic Regression

The first machine learning algorithm that was employed for the prediction of dangerous events was Logistic Regression. In order to observe the significance of the independent variables the logit model was implemented. The results showed that all the variables were statistically significant with respect to a confidence interval 95% (P-value $\leq$ 0.05). It should be noted that a variable with a P-value lower or equal to 0.05 is significantly important to the model. The following table depicts the results of the logit model, such as the Standard Error, z-value, which is the division of regression coefficient and Standard Error as well as the P-value for each variable.

- **Logit model results**

Variable	Coefficient	Standard Error	P > |z|
Elapsed time	-0.42	0.00	0.00
Longitudinal acceleration	-0.28	0.20	0.00
Lateral acceleration	0.24	0.71	0.00
Longitudinal velocity	0.12	0.10	0.00
Lateral velocity	-0.34	0.60	0.01
Distance travelled	0.38	0.06	0.00
Lateral position	-0.13	0.18	0.00
Steering wheel angle	-0.04	0.03	0.00
Gas pedal percentage	-0.29	0.65	0.00
Brake pedal percentage	0.31	1.46	0.00
Leading vehicle lateral position	0.36	0.17	0.00
Leading vehicle longitudinal position	-0.34	0.07	0.00

Table 19 Logit model results

- **Logistic regression model**

Let p ∈ $[0,1]$ be the probability of an event. The logistic regression model is defined as follows:

$$\begin{aligned} logit(p) = & -1.78 - 0.42 * ElapsedTime - 0.28 * LongAcc + 0.24 * LatAcc + 0.12 \\ & * LongVelocity - 0.34 * LatVelocity + 0.38 * DistanceTravelled - 0.13 \\ & * LatPos - 0.04 * SteeringWheelAngle - 0.29 * GasPedalPercentage + 0.31 \\ & * BrakePedalPercentage + 0.36 * LeadingVehicleLateralPosition - 0.34 \\ & * LeadingVehicleLongitudinalPosition \end{aligned}$$

where α = -1.78 denotes the intercept, β1 = -0.42, β2 = -0.28, β3 = 0.24, β4 = 0.12, β5 = -0.34, β6 = 0.38, β7 = -0.13, β8 = -0.04, β9 = -0.29, β10 = 0.31, β11 = 0.36, β12 = -0.34 the coefficients and p = P{Y = 1}, the probability of a dangerous event to occur. It can be noticed that variables such as lateral acceleration, longitudinal velocity, distance travelled, brake pedal percentage and leading vehicle lateral position are higher than zero (β > 0) and the rest of the variables are lower than zero (β < 0). The impact of each independent variable to the y variable (dangerous event) will be estimated in the following part.

- **Impact of independent variables on y**

Increasing the longitudinal acceleration by 1 unit ($1\, m/s^2$) will lead to an decrease by 0.28 in logit(p), which also can be written as log(p /1 - p). If an decrease of 0.28 occurs with respect to log(p /1 - p), this can be interpreted as a decrease in the odds ratio or (p /1 – p) by $e^{-0.28}$ = 0.76. This implies a decrease of 24% in the odds of a dangerous event appearance, if it is assumed that the rest variables will remain fixed. If the lateral acceleration be increased for 1 unit ($1\, m/s^2$), this means that the log(p /1 - p) will increase by 0.24 and based on the odds ratio, this indicates a 27% increase in the odds that a dangerous event will occur, if the rest of the variable remain fixed. The following table presents the estimated odds ratio (=exp(β)) based on the coefficient of each variable as well as the increase or decrease in odds.

Variable	**Coefficient**	**Odds ratio = exp(β)**	**Odds**
Elapsed time	-0.42	0.66	(-) 34%
Longitudinal acceleration	-0.28	0.76	(-) 24%
Lateral acceleration	0.24	1.27	(+) 27%
Longitudinal velocity	0.12	1.13	(+) 13%
Lateral velocity	-0.34	0.71	(-) 29%
Distance travelled	0.38	1.46	(+) 46%
Lateral position	-0.13	0.88	(-) 12%
Steering wheel angle	-0.04	0.96	(-) 4%

Variable	Coefficient	Odds ratio = exp(β)	Odds
Gas pedal percentage	-0.29	0.75	(-) 25%
Brake pedal percentage	0.31	1.36	(+) 36%
Leading vehicle lateral position	0.36	1.43	(+) 43%
Leading vehicle longitudinal position	-0.34	0.71	(-) 29%

Table 20 Odds that a dangerous event will occur (+): Increase, (-): Decrease

Taking into consideration the calculated odds ratios, the increase or decrease in odds that a dangerous event will happened can be easily estimated for the rest of the variables. For instance, an increase in longitudinal velocity by 1 unit (1 m/s) will result in an increase by 0.13 with respect to log(p /1 - p) and an increase by 13% in the odds that an event will happened. On the other hand, increasing lateral velocity by 0.34 will cause a logit(p) drop by 0.34 as well as a decrease of 29% in the odds. Moreover, increasing lateral position by 1 unit (1 m) will decrease the logit(p) by 0.13 and the odds will be decreased by 12%. Last but not least, an increase in steering wheel angle by 1 unit (1 degree) will have as a result a decrease by 0.04 in log odds ratio as well as a decrease by 4% in the odds that an event will occur. For every assumption above should be considered that the rest of the independent variables remain fixed. Likewise can be interpreted the increase or decrease in the odds that a dangerous event will occur for the rest of the independent variables.

- **Evaluation of Logistic Regression model**

  ◊ **Classification report**

At the end of the process, the performance of the model is estimated by implementing the evaluation metrics that were mentioned in Methodology section. The values of these metrics are included in the results of classification report, which is presented below:

Class	Precision	Recall	F1-Score
0	0.75	0.95	0.84
1	0.65	0.24	0.35

Table 21 Classification report results Logistic Regression

It can be observed that the model reached low performance for minority class (class 1), especially in recall (24%) and F1-score (35%).

◊ **Confusion matrix**

Another way to evaluate the model is the confusion matrix, which depicts the predictions of the classifier that were correct and the ones that were incorrect. The confusion matrix results of logistic regression model appear below:

TP (5224)	**FP** (296)
FN (1752)	**TN** (553)

Table 22 Confusion matrix results Logistic Regression

As it can be seen there is a high number of false negatives (1752), which indicates that logistic regression is not the highest performing model. The number of false negatives and false positives should be as low as possible, therefore improvement is still needed.

◊ **Cross validation**

Furthermore, cross validation was implemented by using recall as scoring metric and k folds number was 10. The estimated recall score through this method was equal to 0.245, which represents a low performance for the minority class.

5.2.3. Support Vector Machine (SVM)

The second model that was implemented, was Support Vector Machine (SVM), a more complex machine learning algorithm. The independent variables that were considered for this model, were the same with the ones that were taken into consideration for the logistic regression model.

While implementing the model different values of C and functions were selected. After several trials the highest performance for this model was reach for C=100, kernel='rbf' and gamma='scale'.

- **Feature importance**

To receive some insights for the SVM model, it is useful to observe the importance of the independent variables. The figure below presents the results of feature importance:

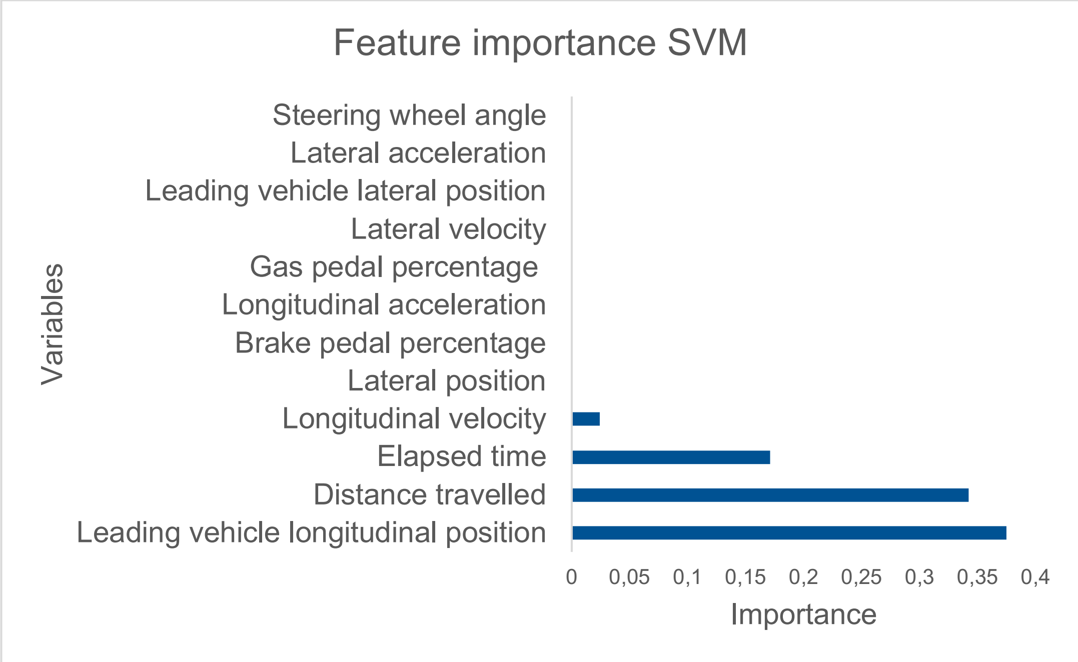

Figure 31 Feature importance SVM

It can be observed that the most important variable for the SVM model is leading vehicle longitudinal position and the second most important is total longitudinal distance travelled. Then follows the ellapsed time variable and less significant is longitudinal velocity.

- **Evaluation of Support Vector Machine model**
  - ◊ **Classification report**

The SVM model performed better with respect to Logistic Regression as it can be observed in the results of classification report. More specifically, both classes reached a good performance in precision, while in recall and F1-Score class 0 performed better. On the other hand, class 1 reached a performance of 72% and 80% respectively. In the table below follow the classification report results of both classes.

Class	Precision	Recall	F1-Score
0	0.89	0.97	0.93
1	0.90	0.72	0.80

Table 23 Classification report results SVM

  - ◊ **Confusion matrix**

In the confusion matrix results, it can be noticed a lower number of FN and FP in comparison with the results of Logistic Regression. More specifically, there were 648 false negatives and 186 false positives, as it can be observed in the following table:

TP (5334)	**FP** (186)
FN (648)	**TN** (1657)

Table 24 Confusion matrix results SVM

◊ **Cross validation**

Cross validation was implemented also at this model by using recall as scoring metric and k folds number was 10. The recall score was equal to 0.71, which indicates a highly improved performance for the minority class.

5.2.4. Artificial Neural Network (ANN)

The ANN model implemented next with the same independent variables as in the previous two models. It is another complex machine learning model and was chosen in order to be tested for a higher performance.

At this case, there were also many trials implemented related to the number of epochs as well as the hidden layers, in order to reach a high performance. The best combination was found for epoch=500, hidden layers: 12-20-30-40-50, activation function='relu', optimizer='adam'.

- **Feature importance**

For the ANN there were more variables significant with respect to the SVM model. As it can be seen in the following figure the most important variable for the ANN model is brake pedal percentage. Furthermore, lateral acceleration is in the second place and is followed by lateral velocity. Less important variables for this model are steering wheel angle, gas pedal percentage, leading vehicle lateral position and lateral position. The importance of the above-mentioned independent variables is depicted in the following figure:

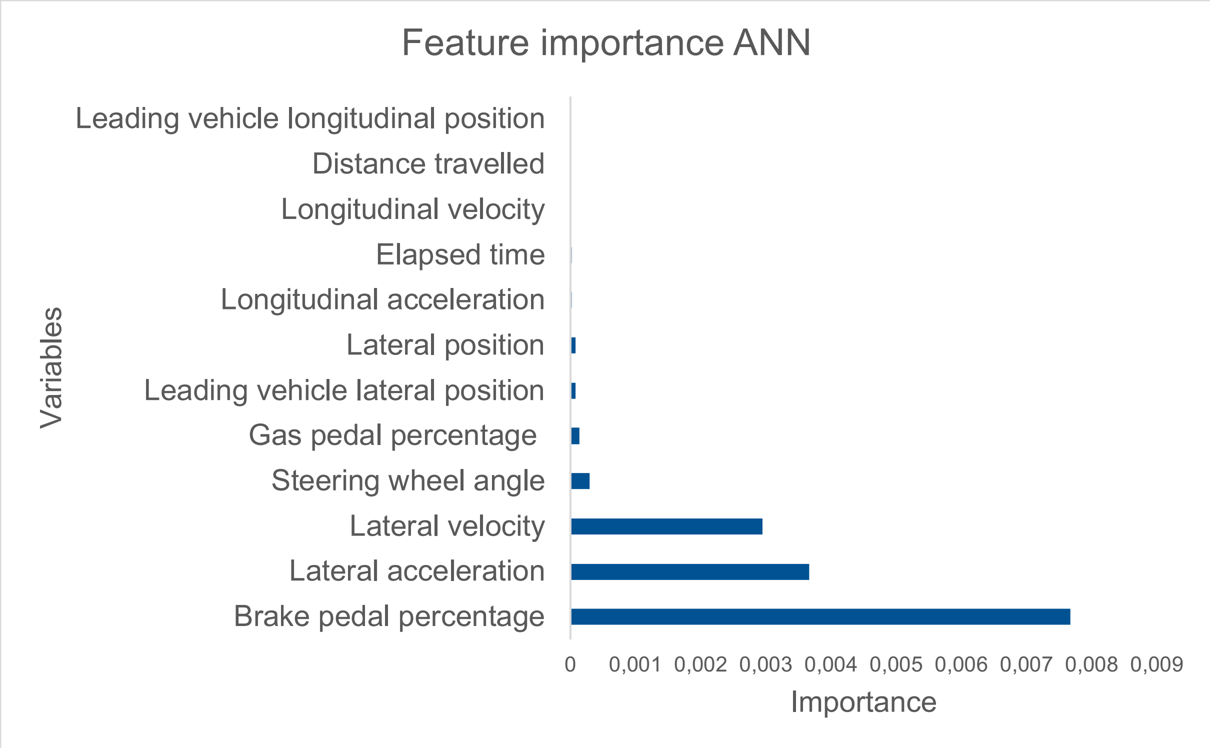

Figure 32 Feature importance ANN

- **Evaluation of the Artificial Neural Network (ANN)**

After reaching the best combination for the ANN model (hidden layers:12-20-30-40-50, epochs=500), the validation and training loss was calculated. As it is depicted in the following figure the curves of validation and training loss are close to each other, which indicates a high performance and that there is no overfitting of the data.

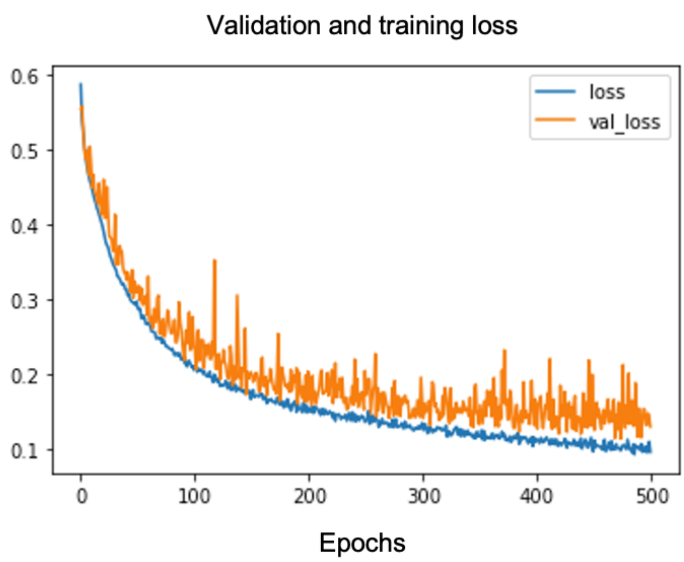

Figure 33 Validation and training loss of ANN model

◊ **Classification report**

The results of classification report showed a higher performance for both classes. Specifically, class 0 reached a performance of around 96%, while class 1 had a performance of around 91%. The outcome of the metrics is depicted in the classification report table below:

Class	Precision	Recall	F1-Score
0	0.96	0.97	0.96
1	0.92	0.91	0.91

Table 25 Classification report results ANN

◊ **Confusion matrix**

In the confusion matrix it was observed a significantly lower number of false negatives (216) and a slight drop in false positives (175) with respect to SVM results. The confusion matrix outcome is presented in the following table:

TP (5345)	**FP** (175)
FN (216)	**TN** (2089)

Table 26 Confusion matrix results ANN

◊ **Cross validation**

Cross validation was implemented also at this model by using recall as scoring metric and k folds number was 10. The recall score was equal to 0.91, which indicates a high performance.

5.2.5. Random Forest

The last machine learning algorithm implemented, was Random Forest. In order to compare the performance of Random Forest with the other models that mentioned above, the same independent variables were considered.

In order to reach a high performance during the implementation of random forest different numbers of estimators (200, 300, 400, 500) were implemented. The highest performance was reached with number of estimators= 500.

- **Feature importance**

At this case, it was also insightful to explore the importance of independent variables for Random Forest model. The following figure presents the results of feature importance. It can be noted that elapsed time is the most important variable for random forest model. Then in a similar level of importance is distance travelled, leading vehicle longitudinal position and lateral position. In a lower level of importance are variables such as longitudinal velocity, brake pedal percentage, longitudinal acceleration and gas pedal percentage. Last but not least, less important variables for the model are lateral velocity, leading vehicle lateral position, lateral acceleration and steering wheel angle.

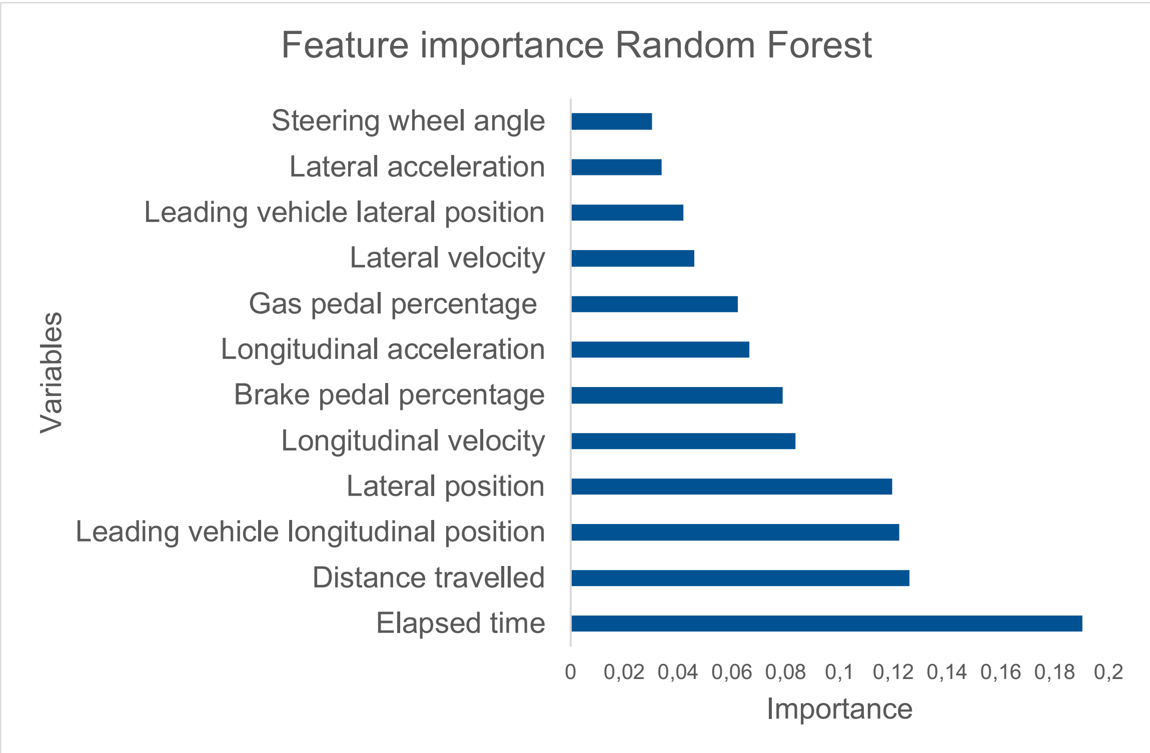

Figure 34 Feature importance of Random Forest

- **Evaluation of Random Forest model**
  - ◊ **Classification report**

Taking into consideration the results of classification report, a significantly high performance of both classes can be observed. More specifically, the highest performance for minority class (class 1) is noted in precision. A significantly improved performance is reported for recall (99%) as well as for f1-score (99%). The detailed results of classification report for both classes appear in the following table:

Class	Precision	Recall	F1-Score
0	1.00	1.00	1.00
1	1.00	0.99	0.99

Table 27 Classification report results Random Forest

◊ **Confusion matrix**

The significantly improved performance can also be noticed in the output of the confusion matrix. Specifically, the false negatives (21) as well as the false positives (8) presented a high drop, which denotes a significantly high performing model. The confusion matrix table is presented below:

TP (5512)	**FP** (8)
FN (21)	**TN** (2284)

Table 28 Confusion matrix results Random Forest

◊ **Cross validation**

Cross validation was implemented also at this model by using recall as scoring metric and k folds number was 10. The recall score was equal to 0.99 which shows a significantly high performance of minority class.

5.3. Discussion

The first part of the thesis focuses on the effectiveness of interventions in driving behavior. Driving simulation data were labelled based on time headway and defined as dangerous condition the moments that time headway was less or equal to 1.4 s, based on the design of i-DREAMS warning system. The results showed that interventions had a significant impact in the comparison between monitoring and intervention scenario. More specifically, between these two scenarios, participants considered the interventions mainly in rural and highway road sections.

Companies and researchers that work on Forward Collision Warning (FCW) systems or related projects could take into account these findings to improve their products/research. For instance, they could consider the threshold of 1.4 s in order to adjust their thresholds (increase or decrease) to make them safer for the users. Another point could be to improve these systems based on the urban road environment requirements, since the results of the research showed that interventions did not have a significant impact while driving in an urban area. In this area, there are many users that should be taken into consideration, such as pedestrians, bicyclists and drivers as well as the infrastructure and the vehicle.

The second part referred to dangerous event prediction by implementing machine learning models, such as Logistic Regression, Support Vector Machine (SVM), Artificial Neural Network (ANN) as well as Random Forest. The performance evaluation of the models presented that Random Forest overperformed in comparison with the other models and therefore reached the highest performance for minority class (99% Recall, 99% F1-score). At this case, researchers could use the highest performing model (Random forest) and probably the second highest one (ANN) in order to predict dangerous events in similar projects or conditions and gain a better insight concerning the driver behavior.

6. Conclusions

6.1. Summary

6.1.1. Effectiveness Evaluation of Interventions

The research at this part focuses on comparing monitoring, intervention and distraction scenario in order to investigate the effectiveness of interventions in driving behavior. Statistical analysis was implemented between monitoring and intervention scenario, monitoring and distraction as well as between intervention and distraction. More specifically, a paired samples T-test was employed to compare the driving behavior in the three different scenarios.

A significant difference in population means was found between monitoring and intervention scenario, this indicates a high impact of interventions on driving behavior. At this comparison longitudinal Velocity and headway variables increased during driving at intervention scenario. This can be interpreted that, participants while handling critical situation at intervention scenario, although they were driving with a higher longitudinal velocity, they managed to maintain a higher time headway from the leading vehicle. It is therefore observed that real time interventions from the i-DREAMS warning system by informing drivers timely about critical situations contributed to their safety.

Comparing monitoring to distraction scenario, a significant difference in population means was observed for less variables. Variables such as TTC and duration of the critical event increased during distraction scenario. The fact that less variables changed at this case indicates that participants although during distraction were receiving interventions and text messages, they had a similar driving behavior with monitoring scenario (no interventions, no text messages), with the exception that the Time to Collision increased.

The T-test results between intervention and distraction showed that even less variables appeared to have a significant difference in population means. A difference between the two scenarios was observed in variables longitudinal velocity and standard deviation of longitudinal acceleration. Furthermore, this indicates that in comparison with intervention drivers during distraction were driving with lower longitudinal velocity while dealing with critical situations. At this case the interventions helped the participants to minimize their speed and avoid a collision.

The fact that between monitoring and intervention scenario a significant difference in population means was observed in most of the variables, triggered the interest to further investigate these to scenarios. More specifically, comparison conducted between monitoring and intervention scenario for the three road sections (urban, rural and highway) separately.

Comparing monitoring to intervention during driving in urban road section it was noted that only two variables changed, minimum longitudinal velocity and minimum headway were both increased. Although, the overall driving behavior did not show a high difference between the two scenarios, participants during urban road section were driving with a higher speed and maintaining at the same time higher headway. It can be concluded that at this road section the drivers took into consideration the warnings from the i-DREAMS warning system and maintained higher and safer time headway from the leading vehicle.

In rural road section as well as in highway similar results were observed, thus, similar driving behavior between the two scenarios. Most of the variables changed and this indicates a high impact of interventions in rural and highway road environment. More specifically, longitudinal velocity, headway, duration and distance of the events increased from monitoring to intervention scenario, while minimum longitudinal acceleration and minimum TTC decreased from one scenario to the other.

6.1.2. Dangerous Event Prediction

At this section driving simulation data were implemented in order to predict dangerous driving events. For this purpose, some commonly used Machine Learning models, such as Logistic Regression, Support Vector Machine (SVM), Artificial Neural Network (ANN) as well as Random Forest, were employed.

The importance of the independent variables was different for each model and for this reason was investigated separately. For the model evaluation classification report including precision, recall and F1-score, confusion matrix and cross validation was used. The evaluation showed that the lowest performance was reached by Logistic Regression with 24% Recall and 35% F1-score for the minority class. Then Support Vector Machine (SVM) followed with 72% Recall and 80% F1-score (class 1). The second-best performing model was Artificial Neural Network (ANN) with 91% Recall and 91% F1-score for the minority class.

The best performance was reached by the implementation of Random Forest for both classes. More specifically, the highest performance for minority class (class 1) is noted in precision. A significantly improved performance is reported for Recall (99%) as well as for F1-score (99%).

It is also important to mention for Random Forest, which was the best performing model, the feature importance. It was observed that the most important variables for random forest model was elapsed time, distance travelled, leading vehicle longitudinal position and lateral position. While, less important variables were lateral velocity, leading vehicle lateral position, lateral acceleration and steering wheel angle.

The following table presents a summary of the performance evaluation for the machine learning models that implemented.

Classification Report				
Machine Learning Model	**Class**	**Precision**	**Recall**	**F1-score**
Logistic Regression	0	0.75	0.95	0.84
	1	0.65	0.24	0.35
Support Vector Machine	0	0.89	0.97	0.93
	1	0.90	0.72	0.80
Neural Network	0	0.96	0.97	0.96
	1	0.92	0.91	0.91
Random Forest	0	1.00	1.00	1.00
	1	1.00	0.99	0.99

Table 29 Classification report summary

6.2. Limitations and Future Work

During data preparation some inaccuracies regarding Time to Collision (TTC) variable were observed. Therefore, Time Headway was used as the main indicator for the data classification, into dangerous situation (class 1) and safe (class 0). In the future it would be interesting to investigate the implementation of more indicators, such as Time to Collision and Distance Headway, in the prediction of dangerous driving events.

Several variables were used from the extracted driving simulator data, such as longitudinal and lateral velocity, longitudinal and lateral acceleration, lateral position, steering wheel angle, headway etc. It should be mentioned that longitudinal and lateral acceleration was used with negative and positive values. It should be probably used in the future with the absolute value for accomplishing better results.

Regarding the statistical analysis, it would be meaningful for other methods to be also implemented. For instance, ANOVA, which is analysis of variance, is a method that could be employed. This method is suitable for a statistical analysis between three or more independent groups (Laerd statistics, 2018). Therefore, the differences between monitoring, intervention and distraction scenario can be compared all together.

Taking into consideration the dangerous event classification and prediction by implementing Machine Learning models, it would be an improvement to implement more complex models, such as Bayesian Logistic Regression, XGBoost, Bayesian Network, Recurrent Neural Network etc.

7. References

Abe, G., & Richardson, J. (2006). Alarm timing, trust and driver expectation for forward collision warning systems. *Applied Ergonomics*.

Ahangari, S., Dehzangi, I., & Jeihani, M. (2019). A Machine Learning Distracted Driving Prediction Model.

Alvarez-Coello, D., Klotz, B., Wilms, D., Fejji, S., Gomez, J., & Troncy, R. (2019). Modeling dangerous driving events based on in-vehicle data using Random Forest and Recurrent Neural Network.

Babić, D., Babić, D., Cajner, H., Sruk, A., & Fiolić, M. (2020). Effect of Road Markings and Traffic Signs Presence on Young Driver Stress Level, Eye Movement and Behaviour in Night-Time Conditions: A Driving Simulator Study.

Baheti, P. (2022). What is overfitting in Deep Learning and how to avoid It: https://www.v7labs.com/blog/overfitting#techniques-to-avoid-overfitting

Blana, E. (1996). Driving Simulator Validation Studies: A Literature Review.

Brooks, J., Goodenough, R., Crisler, M., Klein, N., R, A., Koon, B., . . . Wills, R. (2010). Simulator sickness during driving simulation studies. *Accident Analysis and Prevention 42*.

Brownlee , J. (2021). *Machine Learning Mastery*. How to Use ROC Curves and Precision-Recall Curves for Classification in Python: https://machinelearningmastery.com/roc-curves-and-precision-recall-curves-for-classification-in-python/

Brownlee, J. (2020). *Machine Learning Mastery*. A Gentle Introduction to k-fold Cross-Validation: https://machinelearningmastery.com/k-fold-cross-validation/

Caird, J., Johnston, K., Willness, C., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving.

Calvi, A., D'Amico, F., Ferrante, C., & Ciampoli, L. (2020). Effectiveness of augmented reality warnings on driving behaviour whilst approaching pedestrian crossings: A driving simulator study.

Car ADAS. (2021). *Understanding ADAS: Forward Collision Warning*. Car ADAS: https://caradas.com/understanding-adas-forward-collision-warning/

Choudhary, P., & Velaga, N. (2017). Modelling driver distraction effects due to mobile phone use on reaction time. *Transportation Research Part C: Emerging Technologies*.

Eberly College of Science. (kein Datum). *Introduction to Mathematical Statistics*. https://online.stat.psu.edu/stat415/lesson/10/10.3

European Commission. (2022). *ADAS*. European Commission: https://road-safety.transport.ec.europa.eu/statistics-and-analysis/statistics-and-analysis-archive/miscellaneous/adas_en

Ezukwoke, K., & Zareian, S. (2019). A comparative study of logistic regression and kernel logistic regression for binary classification.

Ezzati Amini, R., Al Haddad, C., Batabyal, D., Ryan, A., Gkena, I., De Vos, B., . . . Antoniou, C. (2022). Analyzing driver eye movements to investigate the impact of distraction on driving behavior. *Accepted to the 8th Road Safety & Simulation International Conference (RSS).*

Ezzati Amini, R., Michelaraki, E., Katrakazas, C., Al Haddad, C., De Vos, B., Cuenen, A., . . . Antoniou, C. (2021). Risk scenario designs for driving simulator experiments. *7th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems.*

Fancello, G., Adamu, M., Serra, P., & Fadda, P. (2020). Comparative analysis of the effects of mobilephone use on driving performance usingANOVA and ANCOVA. *IET Intelligent Transport Systems*.

Frost, J. (2022). *Significance level*. https://statisticsbyjim.com/glossary/significance-level/

Gandhi, R. (2018). *Towards Data Science*. Support Vector Machine — Introduction to Machine Learning Algorithms: https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47

Gavrilova, Y. (2020). A Guide to Deep Learning and Neural Networks: https://serokell.io/blog/deep-learning-and-neural-network-guide

Ghosh, B., Basu, B., & O'Mahony, M. (2005). Time series modelling for forecasting vehicular traffic flow in Dublin. *Transportation Research Board.*

Guo, Y., Wang, X., Xu, Q., Liu, F., Liu, Y., & Xia, Y. (2019). Change-point analysis of eye movement characteristics for female drivers in anxiety.

Gupta, P. (2017). *Towards Data Science*. Cross-Validation in Machine Learning: https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f

Haigney, D., Taylor, R., & Westerman, S. (2000). Concurrent mobile (cellular) phone use and driving performance: task demand characteristics and compensatory processes. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Hassan, H., Sarhan, M., Garib, A., & Harthei, H. (2017). Drivers' time headway characteristics and factors affecting tailgating crashes. *96st Annual Meeting of the Transportation Research Board.*

He, Z., Hu, J., Park, B., & Levin, M. (2019). Vehicle sensor data-based transportation research: Modeling, analysis, and management. *Journal of Intelligent Transportation Systems.*

Hirst, S., & Graham, R. (1997). The Format and Presentation of Collision Warnings. In *Ergonomics and Safety of Intelligent Driver Interfaces.*

Horrey, W., & Wickens, C. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.

i-DREAMS. (2022). i-DREAMS: https://idreamsproject.eu/wp/

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning.*

Jamson, H., Lai, F., & Carsten, O. (2007). Potential benefits of an adaptive forward collision warning system. *Transportation Research Part C: Emerging Technologies*.

Jankovic, D. (2021). *Towards Data Science*. A simple interpretation of Logistic Regression coefficients: https://towardsdatascience.com/a-simple-interpretation-of-logistic-regression-coefficients-e3a40a62e8cf

Kent State University. (2021). *University Libraries* . SPSS Tutorials: Paired samples t test: https://libguides.library.kent.edu/spss/pairedsamplesttest

Khansari, E., Nejad, F., & Moogehi, S. (2020). Comparing time to collision and time headway as safety criteria. *Pamukkale University Journal of Engineering Sciences*.

Kunapuli, G., Bennett, K., Hu, J., & Pang, J.-S. (2008). Bilevel model selection for Support Vector Machines.

Laerd statistics. (2018). *Laerd statistics*. One-way ANOVA: https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide.php

Li, X., Guo , Z., & Li , Y. (2021). Driver operational level identification of driving risk and graded time-based alarm under near-crash conditions: A driving simulator study.

Liu, T., & Fu, S. (2018). The relationship between different safety indicators in car-following situations. *IEEE Intelligent Vehicles Symposium (IV)*.

Loulizi, A., Bichiou, Y., & Rakha, H. (2019). Steady-state car-following time gaps: An empirical study using naturalistic driving data.

Mai, M., Wang, L., & Prokop, G. (2017). Advancement of the car following model of Wiedemann onlower velocity ranges for urban traffic simulation.

Massera Filho, C., Wolf, D., Grassi, V., & Osório, F. (2014). Longitudinal and lateral control for autonomous ground vehicles.

Mishra, A. (2018). *Towards Data Science*. Metrics to Evaluate your Machine Learning Algorithm: https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234

Misra, R. (2019). *Towards Data Science*. Support Vector Machines — Soft Margin Formulation and Kernel Trick: https://towardsdatascience.com/support-vector-machines-soft-margin-formulation-and-kernel-trick-4c9729dc8efe

Mohajon, J. (2020). *Towards Data Science*. Confusion Matrix for Your Multi-Class Machine Learning Model: https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826

Mohamed, H., Negm, A., Zahran, M., & Saavedra, O. (2015). Assessment of Artificial Neural Network for bathymetry estimation using High Resolution Satellite imagery in shallow lakes: Case Study El Burullus Lake. *EighteenthInternational Water Technology Conference.*

Monteiro, N., Balogun, S., Kote, M., & Tlhabano, K. (2014). Stationary tailgating in Gaborone, Botswana: the influence of gender, time of day, type of vehicle and presence of traffic officer.

Moorthy, C., & Ratcliffe, B. (2007). Short term traffic forecasting using time series methods. *Transportation Planning and Technology*.

Papantoniou, P., Papadimitriou, E., & Yannis, G. (2017). Review of driving performance parameters critical for distracted driving research.

Peden , M. (2004). *World Report on Road Traffic Injury Prevention.*

Rämä, P., & Kulmala, R. (Pirkko Rämä Risto Kulmala 2000). Effects of variable message signs for slippery road conditions 17 on driving speed and headways.

Rossi, R., & Gastaldi, M. (2012). An empirical analysis of vehicle time headways on rural two-lane two-way roads. *15th meeting of the EURO Working Group on Transportation*.

Saffarzadeh, M., Nadimi, N., Naseralavi, S., & Mamdoohi, A. (kein Datum). A general formulation for time-to-collision safety indicator.

Santos, J., Merat, N., Mouta, S., Brookhuis, K., & de Waard, D. (2005). The Interaction Between Driving and In-Vehicle Information Systems: Comparison of Results from Laboratory, Simulator and Real-World Studies. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Sato, T., & Akamatsu, M. (2012). *Understanding driver car-following behavior using a Fuzzy Logic car-following model.*

Sayad, S. (2022). Artificial Neural Network - Perceptron: https://www.saedsayad.com/artificial_neural_network_bkp.htm

Scikit learn. (2022). *Scikit learn*. Precision-Recall: https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html

Singh, S. (2021). An Emphasis on the Minimization of False Negatives/False Positives in Binary Classification: https://medium.com/@Sanskriti.Singh/an-emphasis-on-the-minimization-of-false-negatives-false-positives-in-binary-classification-9c22f3f9f73

Sruthi, E. (2021). *Analytics Vidhya*. Understanding Random Forest: https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/

Tango, F., & Botta, M. (2013). Real-Time Detection System of Driver Distraction Using Machine Learning. *IEEE Transactions on Intelligent Transportation Systems* .

Teoh, E. (2021). Effectiveness of front crash prevention systems in reducing large truck real-world crash rates. *Traffic Injury Prevention*.

Transportation Research Board . (2015). *The highway capacity manual.*

TSE. (2021). *Technical University of Munich*. Chair of Transportation System Engineering (TSE): https://www.mos.ed.tum.de/vvs/forschung/projekte/idreams/

Tyagi, N. (2020). Understanding the Perceptron Model in a Neural Network: https://medium.com/analytics-steps/understanding-the-perceptron-model-in-a-neural-network-2b3737ed70a2

Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Urban Mobility Lab at MIT. (2022). Machine Learning for Transportation : https://mobility.mit.edu/machine-learning

Varsheni, S. (2021). *Analytics Vidhya*. Perform Logistic Regression with PyTorch Seamlessly: https://www.analyticsvidhya.com/blog/2021/07/perform-logistic-regression-with-pytorch-seamlessly/

Vogel, K. (2003). A comparison of headway and time to collision as safety indicators.

Wang, J.-H., & Song, M. (2011). Assessing drivers' tailgating behavior and the effect of advisory signs in mitigating tailgating.

Wegman, F. (2016). The future of road safety: A worldwide perspective. *IATSS Research*.

Winter, J., Leeuwen, P., & Happee, R. (2012). Advantages and Disadvantages of Driving Simulators: A Discussion.

World Health Organization . (2021). *Road traffic injuries*. World Health Organization : https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

World Health Organization. (2011). Mobile phone use: A growing problem of driver distraction.

Xue, Q., Wang, K., Lu, J., & Liu, Y. (2019). Rapid driving style recognition in car-following using Machine Learning and vehicle trajectory data.

Yoshizawa, A., Nishiyama, H., Iwasaki, H., & Mizoguch, F. (2016). Machine-Learning approach to analysis of driving simulation data.

Young, K., & Regan, M. (2007). Driver Distraction: A Review of the Literature.

Declaration

I hereby confirm that the presented thesis work has been done independently and using only the sources and resources as are listed. This thesis has not previously been submitted elsewhere for purposes of assessment.

Munich, 10.06.2022, Isidora Gkena

Place, Date, Signature