Acta Polytechnica **66**(1):64–74, 2026
https://doi.org/10.14311/AP.2026.66.0064

Published by the Czech Technical University in Prague

IDENTIFICATION OF POTENTIAL SUICIDE ATTEMPTS BY TRAFFIC ACCIDENTS IN THE CZECH REPUBLIC

Vilém Kovač*, Ali Ibraheem, Tomáš Řezníček

Institute of Technology and Business in České Budějovice, Okružní 10, 370 01 České Budějovice, Czech Republic
* corresponding author: 33241@mail.vstecb.cz

Abstract. This study focuses on identifying hidden potential suicide attempts by traffic accidents (PSA-TA) in the Czech Republic by combining exploratory data analysis (EDA), classification algorithms (KNN, XGBoost), and interpretability (SHAP) to design a predictive model capable of distinguishing PSA-TA from ordinary fatal road traffic accidents (FRTA). The results of the analysis show that cases of suicidal behaviour exhibit specific characteristics. Typically, these involve collisions with a fixed obstacle outside residential areas, at times of low traffic, without the use of safety features, involving a single vehicle, and often involving male drivers. Based on these characteristics, the model identified 13 cases from 2024 that are likely to bear the hallmarks of intentional behaviour, even though they were officially recorded as ordinary FRTA. The results of the study confirm that advanced analytical tools can be used to detect hidden suicidal behaviour ex post, thereby contributing to more accurate statistics, forensic assessment of accidents, and the development of targeted preventive policies in the areas of traffic safety and mental health. We also outline ethical aspects and key limitations of administrative coding, including potential misclassification.

Keywords: Traffic, behavior, public safety, data analysis, classification, machine learning.

1. Introduction

This study focuses on identifying cases of potential suicide attempts by traffic accidents (PSA-TA) within fatal road traffic accidents (FRTA) in the Czech Republic. Given that official databases contain only a limited number of cases in which the driver survived and admitted suicidal intent, the aim of this study was to uncover other probable cases that were formally recorded as ordinary fatal accidents but, based on an analysis, show significant similarities to confirmed suicide attempts. To achieve this goal, an exploratory data analysis (EDA) was first performed, which revealed characteristic patterns and differences to ordinary FRTAs. The data were then transformed using principal component analysis (PCA), and the k-nearest neighbours (KNN) algorithm was applied to the transformed data, which enabled the cases to be classified based on their similarity to a reference set of suicide attempts. To further confirm the results and interpret the key variables, the XGBoost model was used, whose outputs were explained using the SHAP (SHapley Additive Explanations) method.

The results show that traffic accidents involving suicidal behaviour have different characteristics to those of ordinary FRTAs, in particular, involving single-vehicle incidents, collisions with fixed obstacles, occurrence outside residential areas, in the early morning or night hours, often without injury to other persons, and with a low number of safety features used. The model identified 13 cases that are highly likely to show signs of intentional behaviour, even though they were not formally classified as suicides.

The available literature shows that this topic remains underrepresented in research. Suicides in traffic are either completely overlooked or hidden under the broader category of traffic accidents, while suicidological research generally focuses on other, more common forms of suicidal behaviour. This study, therefore, contributes to bridging this gap and highlights the potential of data analysis as a tool for the early identification of risk cases.

The aim of this work is to contribute to a deeper understanding of suicidal behaviour involving traffic accidents in the Czech Republic. The analysis is based on both the confirmed suicide attempts and the FRTAs in which hidden intent is assumed based on a data analysis. The work focuses on identifying recurring patterns, characteristics, and factors that can help distinguish intentional behaviour from ordinary accidents. In order to achieve the objective of this study, the following research questions were formulated:

1.1. Research questions

- **RQ1:** What patterns of behaviour, time, and situational circumstances are typical of confirmed PSA-TA?

  Answering this question will help identify any recurring characteristics of these events, such as the time of the accident, day of the week, the type of the vehicle, or the gender of the driver. This information can be useful for the prevention and early detection of similar cases in the future

- **RQ2:** Which variables in the predictive model have the greatest influence on the classification of a case as a probable PSA-TA?

This question aims to identify the key factors involved in predicting suicidal intent. The goal is to determine whether there are patterns in the data (e.g. number of participants, use of safety features, location of the accident) that significantly distinguish these cases from normal accidents and can serve as input for the classification algorithm.

- **RQ3:** Based on the available data patterns and predictive model, is it possible to identify traffic accidents that are likely to represent a PSA-TA?

  Resolving this issue will determine whether it is possible to identify traffic accidents that are highly likely to be suicide attempts based on the existing data and the predictive model. If the model proves to be sufficiently accurate, it can serve as a tool for retroactively identifying accidents with a possible suicidal motive, which would make it possible to refine statistical reports, support preventive strategies, and expand forensic analysis of these events.

2. Literature review

Suicide is a complicated phenomenon caused by many biological, psychological, social, and environmental factors [1, 2]. According to the authors [3], approximately 700 000 people die by their own hand each year. Although the suicide rate has been declining significantly, up to 50 % for women since 1990, it is still a serious social problem. In 2021, the total number of cases was 746 000, 70 % of which were men [4]. There are significant differences in the incidence and methods of suicide between countries, which are related to cultural norms and the availability of lethal means [5].

According to the authors [6], who analysed 37 studies, the presence of any mental disorder increases the risk of completed suicide by more than tenfold and a previous suicide attempt by eightfold. Strongly associated with suicide is severe depression, borderline personality disorder, or schizophrenia. Among nonclinical factors, social isolation, unemployment, and low socio-economic status appear to be the most significant. This is corroborated by the authors, who specifically equated unemployment and poverty as the strongest predictors of suicidal behaviour. At the same time, according to the authors [7], low income, debt, or losing a job can lead to despair. Sociological research has repeatedly confirmed the significant influence of economic factors on the suicide curve. Psychosocial stressors, such as divorce and relationship strife, examined by [6], increase the likelihood of suicide severalfold.

According to the authors [3], globally, there was no statistically significant increase in completed suicides during the first year of the pandemic compared to previous trends. According to the authors [8], who conducted large studies in 33 countries, suicide rates in the first 9–15 months of the pandemic did not differ significantly from predictions. However, there is a considerable variation between countries. For example, in Japan, there was an increase in suicides among women and younger population in 2020, probably due to a combination of socioeconomic impacts and worsening mental health.

According to the authors [9], suicidal behaviour in traffic can manifest in various ways. Examples include deliberately directing one's vehicle into an accident, driving into oncoming traffic or a fixed obstacle, deliberately stepping in front of a moving vehicle, or falling off a bridge onto the roadway. The authors [10] point out that the actual proportion of these cases is probably higher than the official statistics indicate. Road traffic suicides tend to be reported and disguised as accidents. An international literature review by the authors [11] provides estimates that globally, intentional self-harm by driving could account for about 1–8 % of all road traffic collisions.

According to the authors' qualitative research [12], which included interviews with people contemplating suicide by road traffic, this method is appealing to many because it is quick, highly lethal, easily accessible, and the potential to be seen as a random accident. This finding is supported by a study by the authors [13] from Australia, which showed that adolescents with a history of self-harm had a 23 % higher risk of road traffic accidents in subsequent years, even after accounting for for common factors (age, alcohol, experience). The authors explain that for some young drivers, latent self-injurious tendencies may lead to risky driving styles and accidents without a conscious suicide attempt.

According to the authors [14], it was the threat that someone else might be killed or traumatised during their attempt that represented a strong deterrent for many respondents from intending to commit suicide in this way. However, according to the authors [15], there are known cases of a desperate driver deliberately driving into an oncoming car full of people. Motivation may also be vindictive, in which case the person chooses to die and takes other victims with them, or hopes that their family will receive a life insurance payout. The authors [10] found that milder methods were sometimes chosen out of compassion for the survivors, but this leads to an underestimation of traffic suicide statistics.

According to the authors [16], road traffic accidents involve many factors, but 80–95 % of traffic accidents involve human factor. However, analyses of the causes of fatal collisions show that a single cause cannot always be identified. Many circumstances contribute to a traffic accident, from the driver's momentary distraction, their health and mental state, to experience, risk-taking tendencies, or external stressors. According to the authors [17], "risky drivers" have certain personality traits, such as higher impulsivity, aggressiveness, or, on the contrary, tendencies towards depression and self-neglect, which can increase the likelihood of fatal accidents. Once a person has decided to take their own life, vehicle safety can be the difference between life and death. This was investigated by [18], whose results show that the newly available vehicles are safer than older ones. Furthermore, the authors of [19] concluded that the price of a vehicle does not play a major role in its safety. These two factors can influence the overall numbers of successful suicides by traffic accidents.

The authors [20] further highlight the role of vehicle design and safety standards, by highlighting that demographic parameters such as age and gender significantly influence the likelihood and pattern of traffic accidents across regions. Their findings emphasise the importance of incorporating demographic analyses into traffic safety planning. Similarly, the authors [19] investigated the relationship between a vehicle's market value and its safety ratings. Although one might expect that more expensive cars would be safer, the data suggest only a weak correlation between price and safety, indicating that perceived value does not necessarily equate to protection in critical situations. Furthermore, the study by the authors [21] presents a practical approach to analysing railway traffic accidents by proposing structured intervention plans for crisis scenarios. This serves as an example of how preventive planning and response exercises can be used to reduce the impacts of both accidental and intentional transport incidents.

A systematic review by the authors [22] identified younger age, low level of driving experience and education, alcohol use, speeding, talking on the phone while driving, not wearing protective equipment (helmets, seat belts), and overall risky driving as the most common risk factors for crashes involving young drivers. According to the authors [23], driving while fatigued or distracted is also dangerous. A modern problem is the use of mobile phones, where texting while driving impairs attention so much that the risk of a crash increases by up to 6 times. Another factor affecting crash rates is the weather and outdoor conditions. The authors [24] investigated traffic accidents where light was considered an influencing factor and it was found to be. However, the accident rate itself showed interesting parallels.

Furthermore, in their review, the authors [25] found that most published studies confirmed an increased risk of traffic accidents among drivers with psychiatric disorders. This is complemented by the authors [26], who suggest that untreated psychosis, severe depression, or active addiction can lead to inattention and slowed reactions, or to risky behaviour behind the wheel, which stems from disturbances in the perception of reality.

3. Methodology

3.1. Data

The data file used for this analysis of PSA-TA contains 75 confirmed cases recorded between January 1, 2019, and April 30, 2025, which the police assessed as suicide attempts (i.e. confirmed PSA-TA based on survivor admission). The data on individual accidents and their factors were obtained via a filtered export from an online database [27] that collects data from the Czech Police accident database. All data are anonymised.

For the training cohort (A) we applied the database filter corresponding to the administrative label "suicide attempt", which in practice appears only when a survivor explicitly admits intent; these cases are therefore confirmed non-fatal suicide attempts through traffic accidents (PSA-TAs). After the export, we selected/engineered analytical variables for (A) from the fields available in the registry (e.g. time, location context, vehicle type, lane departure, and the use of safety features). For the screening cohort (B), we exported the complete census of fatal road traffic accidents (FRTAs) for the year 2024, without any intent labels. Some variables (e.g. alcohol/drugs for deceased drivers) are not systematically available for fatalities, and are therefore missing in (B); such fields were not used in the comparative or SHAP analyses.

All values for FRTAs and confirmed PSA-TAs reported in the text and figures are raw annual totals (yearly counts) from the same export; they are not annual averages or rates. Flagged fatalities are screening signals produced by the model (trained on dataset A) and do not constitute ground truth.

One of the main problems in analysing traffic accidents with a possible suicidal motive is the lack of systematic recording of cases where the driver dies. The police can only classify an accident as a suicide attempt if the driver survives and admits their intention. In the case of FRTAs, this scenario is impossible. Even if indicators of suicidal behaviour are discovered later, this information is not recorded in the traffic accident registry. Statistically, only non-fatal accidents are recorded, while the motive remains part of the criminal investigation without affecting the aggregated data [28].

Data on FRTAs in which we sought to identify potential PSA-TAs were also obtained from the database [27], where the data collection was carried out in a similar manner, with the exception of variables such as alcohol or drug influence, as these data are not monitored in traffic accident victims.

The Police of the Czech Republic classify incidents as fatal accidents when a death occurs, and as suicide attempts when the driver intentionally caused the accident but survived and confirmed the intent during investigation. The full list and definition of the variables used in the analysis are presented in Table 1.

3.2. Exploratory data analysis (EDA)

In the initial phase, exploratory data analysis (EDA) was performed to understand the basic characteristics of the records and identify differences between fatal accidents and suicide attempts. Temporal, demographic, and contextual variables were analyzed, such as the time of the accident, type of collision, driver age, seat belt use, and location. Qualitative data were

Variable	Description
Date	Date of traffic accident
Day_of_the_week	Day of the week
Time	Time of accident (HH:MM:SS)
Region	Region of occurrence
Type_of_traffic_accident	Accident type
Type of vehicle	Vehicle type involved
Sex	Gender of vehicle occupants
In/Out city	Localization (In, Out)
Cause of accident	Primary cause
Number of cars involved	Vehicles involved
Age of vehicle	Vehicle age category
Age of person in car	Driver age group
Location	Vehicle location after crash
Seatbelt/Helmet	Use of safety gear
Drift / Lane departure	(Yes/No)

TABLE 1. Overview of variables used for modeling.

coded and numerical variables were standardized. Visualization and frequency analyses revealed recurring patterns of behavior in suicide attempts, which were later used to classify suspicious cases.

The exploratory analysis was followed by data processing using dimensionality reduction and classification. The aim of this phase was to identify patterns of behavior corresponding to suicide attempts and to reveal potentially hidden cases in the group of fatal accidents.

3.2.1. Principal component analysis (PCA)

To ensure efficient data processing and eliminate redundant correlations between variables, the principal component analysis (PCA) method was used. This transformation converted multidimensional standardized data into a new space of smaller dimensions while maintaining the maximum possible information variance:

$$Z = X \cdot W, \tag{1}$$

where:

X is the matrix of input (standardized) data,

W is the matrix of eigenvectors derived from the covariance matrix,

Z is the projection of the original data into the principal component space.

PCA facilitated the visualization of the data structure and improved the conditions for subsequent classification.

3.3. K-nearest neighbors (KNN)

The k-nearest neighbors (KNN) algorithm was applied to the transformed data, which allows new records to be assigned to known categories based on similarity to the training set (in this case, suicide attempts). Each record was classified according to the majority vote of its nearest neighbors:

$$\text{label}(x) = \text{majority\_vote}\big\{\text{label}(x_i) \mid x_i \in \text{nearest}_k(x)\big\}. \tag{2}$$

The degree of similarity (distance) between points was calculated using Euclidean metrics:

$$d(x, x_i) = \sqrt{\sum_{j=1}^{n} (x_j - x_{i,j})^2}, \tag{3}$$

where:

x is an unclassified vector (accident with unknown outcome),

x_i is a vector from the training set (suicide attempt),

n is the number of variables.

The output of the classification was a similarity score and an indicator of whether the accident in question exhibited characteristics consistent with cases classified as intentional.

3.3.1. XGBoost objective

In this study, we employ the XGBoost algorithm, a gradient boosting framework that optimizes an objective function combining a differentiable loss function and a regularization term. The objective function at iteration t is formulated as:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \ell\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t), \tag{4}$$

where:

n denotes the number of training instances,

y_i represents the true label for the i^{th} instance,

$\hat{y}_i^{(t-1)}$ is the prediction from the ensemble model after $t-1$ iterations,

$f_t(x_i)$ is the newly added regression tree at iteration t,

ℓ is a differentiable convex loss function measuring the discrepancy between predicted and true values,

$\Omega(f_t)$ is a regularization term penalizing model complexity to prevent overfitting.

The objective function formalizes the balance between fitting the training data and maintaining model simplicity. By minimizing this objective iteratively, XGBoost incrementally improves predictions while controlling overfitting through the regularization term, which is critical for achieving robust generalization on unseen data.

The regularization term is defined as:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \tag{5}$$

where:

T is the number of leaves in the tree,

w_j denotes the weight of leaf j,

γ and λ are hyperparameters controlling the regularization strength.

Regularization discourages the model from becoming excessively complex by penalizing the number of leaves and the magnitude of leaf weights. This helps prevent overfitting, improves model interpretability, and enhances generalization performance, especially when dealing with noisy or limited data.

To efficiently optimize the objective, a second-order Taylor expansion of the loss function around the previous predictions is employed:

$$\begin{aligned}\ell\Big(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\Big) \\ \approx \ell\Big(y_i, \hat{y}_i^{(t-1)}\Big) + g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2,\end{aligned} \tag{6}$$

where:

$g_i = \frac{\partial \ell(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}$ is the first derivative (gradient),

$h_i = \frac{\partial^2 \ell(y_i, \hat{y}_i^{(t-1)})}{\partial (\hat{y}_i^{(t-1)})^2}$ is the second derivative (Hessian).

Using the second-order Taylor expansion enables XGBoost to leverage both gradient and curvature information of the loss function. This results in more accurate and efficient optimization compared to first-order methods, allowing faster convergence and better fitting of complex patterns in the data.

By minimizing this quadratic approximation combined with the regularization term, XGBoost constructs each new tree to optimally reduce the overall objective, thus progressively enhancing model performance with each iteration.

3.3.2. Model interpretability with SHAP

To interpret the feature contributions to each prediction made by XGBoost, we used SHapley Additive exPlanations (SHAP). SHAP assigns each feature a value representing its contribution to the final model output. The explanation model is additive and defined as:

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i, \tag{7}$$

where:

$f(x)$ is the model's prediction for input x,

ϕ_0 is the base value (mean prediction),

ϕ_i is the SHAP value for feature i,

M is the total number of features.

SHAP values allow for both global and local interpretability by explaining how much each feature contributes to a particular prediction. This helped identify key variables linked with intentional behavior.

4. Results

To understand the spatial distribution of fatal car accidents and suicide attempts across the Czech Republic, we analyzed incident counts per region, differentiating between general fatal accidents and those classified as suicide attempts. This distinction is crucial for our model's objective to detect potentially hidden suicide attempts within the broader category of fatal car accidents. The bar graph (Figure 1) presents a comparative visual of these two incident types across 14 administrative regions.

The Central Bohemia Region recorded the highest number of total incidents, with 61 fatal accidents and a notable count of suicide attempts, suggesting a region of particular interest for further analysis. South Moravian and Plzeň regions also exhibited elevated fatal accident rates (47 and 36 respectively), though suicide attempts formed a smaller portion of those totals. Regions like Liberec, Moravian-Silesian, and Ústí nad Labem showed a more balanced distribution between the two categories, indicating possible overlaps in features that may aid the classification model. In contrast, Prague had relatively low numbers for both categories. These regional patterns provide foundational insight for evaluating the performance of the model in distinguishing suicide-related cases among fatal accidents.

To explore behavioral and contextual distinctions between regular fatal accidents and suicide attempts via traffic incidents, we analyzed accident occurrences by vehicle type. This categorization is particularly relevant for identifying patterns associated with intentional crashes, as vehicle accessibility and perceived lethality may influence individual behavior in suicide attempts. Figure 2 presents the frequently both labels normal accidents and suicides attempts are distributed across various vehicle types.

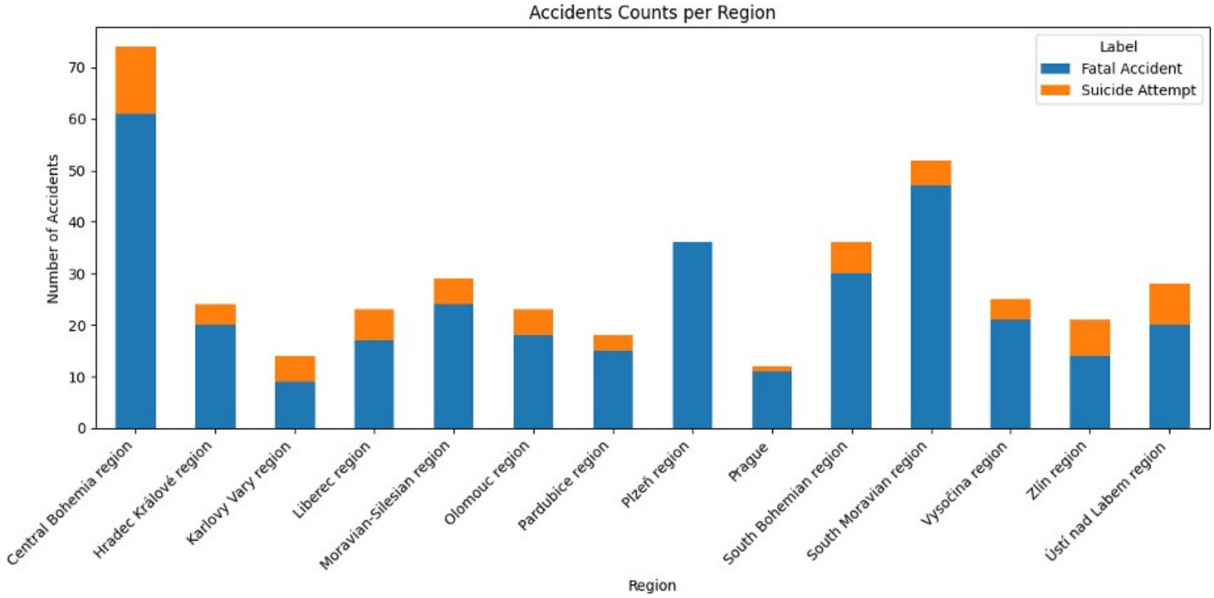

Figure 1. Distribution of fatal accidents and suicide attempts in individual regions of the Czech Republic.

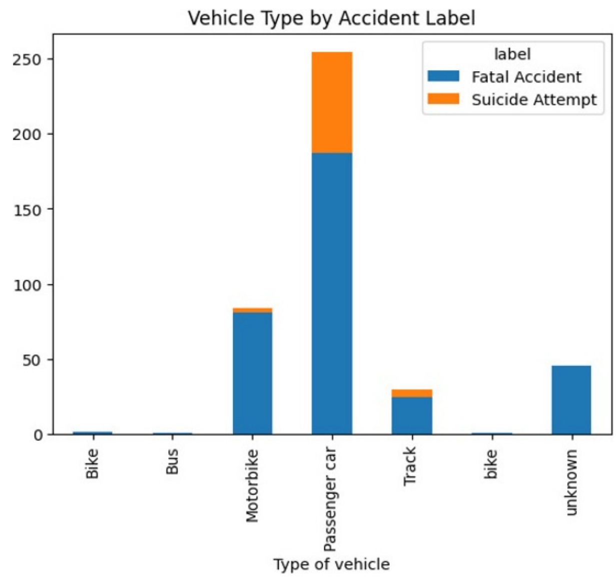

Figure 2. Classification of accidents by vehicle type and classification into fatal accidents and suicide attempts.

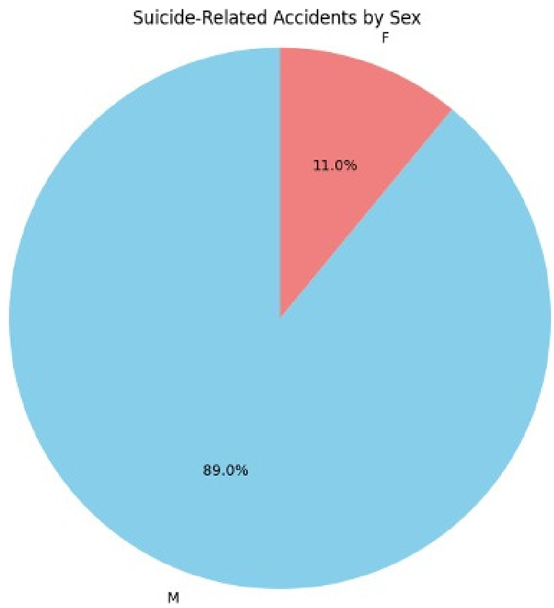

Figure 3. Distribution of suicide attempts by traffic accidents according to sex.

Passenger cars were overwhelmingly the most common vehicle type involved in both fatal accidents and suicide attempts, with approximately 185 normal accidents and a significant proportion (around 70 cases) attributed to suicide attempts. Followed by Motorbikes with over 80 incidents, though only a minimal fraction were linked to suicide attempts. Notably, trucks and unknown vehicle types appeared in both categories, but in substantially lower numbers. Public transport and non motorized vehicles (bikes, buses) had negligible or no recorded suicide attempts. These findings indicate that passenger cars represent the primary mode through which individuals engage in both accidental and intentional high risk driving behavior, a trend that warrants focused attention in subsequent model evaluation and discussion.

An analysis of the gender distribution among individuals involved in attempted suicide by traffic accident reveals significant gender inequality. The pie chart shows that men account for the majority of these incidents, while women represent only 11 %. The significant predominance of men in cases of suicide attempts using a vehicle points to the need for targeted interventions focused on their specific behavioral patterns. Psychological screening, crisis intervention, and education focused on stress management may be key to reducing the number of such cases among men. The gender distribution of individuals involved in attempted suicide by traffic accident is shown in Figure 3.

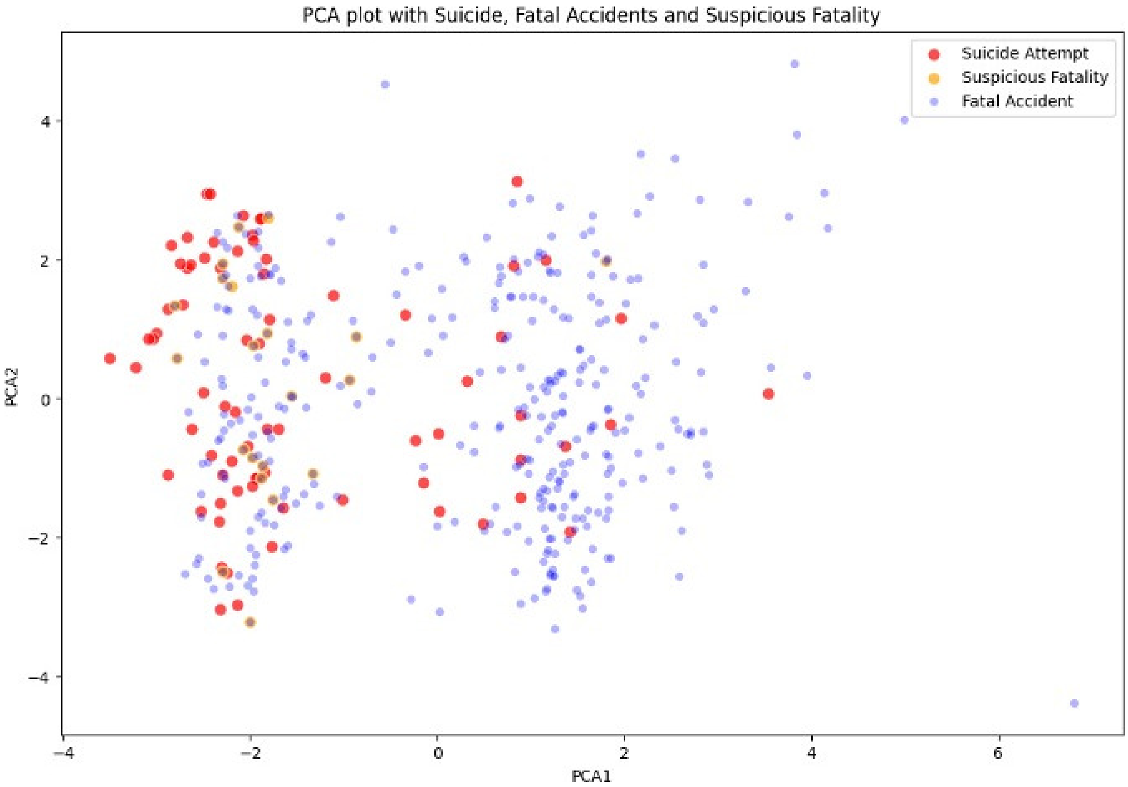

Figure 4. PCA visualization of the distribution of suicide cases, suspicious accidents, and fatal traffic accidents.

To evaluate the representation of individual factors in confirmed suicide attempts by traffic accident, a table was created showing the percentage occurrence of all monitored variables. Due to its size, it is included separately in the appendices to this work.

In addition to the aforementioned gender, regional distribution, and vehicle type, the table shows that 79 % of accidents involved a collision with a fixed obstacle, most events took place outside of towns (86 %), and 75 % of cases involved a single vehicle. Nearly half of the individuals did not use a seat belt. Older vehicles (over 15 years old) were involved in more than half of the cases. In terms of time, accidents dominated in the afternoon and evening hours, with more frequent occurrences in the spring months. These findings complement the profile of a typical suicide attempt and can be used in the prediction and prevention of similar events. Other variables, including their percentage representation, are listed in a separate table in the appendix for reasons of scope and clarity.

For a deeper understanding of the causes of fatal traffic accidents, a pattern analysis was performed using dimensionality reduction and classification, which revealed that some cases officially labeled as Normal (Fatal Traffic Accidents) or Suspicious Normal (fatal accidents that, according to our analysis, show similarities to suicidal behavior) share characteristics with confirmed Suicide (Confirmed Suicide Attempts). Figure 4 shows that suspicious cases are spatially clustered near confirmed suicidal incidents, suggesting the possibility of misclassification. This approach underscores the importance of more detailed investigation of the circumstances of fatal accidents for more accurate identification and prevention of suicides.

After the grouping and classification phase, the model identified a subset of fatal traffic accidents that closely resemble known suicide attempts in terms of their profile. Although these cases were not officially categorized as suicides, they were flagged based on their similarity to patterns commonly observed in confirmed intentional incidents. This result supports the hypothesis that certain deaths, originally considered standard or inconclusive, may require review due to common behavioral and situational features with suicide-related events.

The model identified 13 cases formally classified as ordinary fatal accidents but showing a high probability of suicidal behavior. All cases involved single-vehicle accidents, often outside urban areas, at night or in the early morning hours, and involving a collision with a fixed obstacle. As a rule, no injuries were reported and the drivers did not use seat belts. Some records lacked important data, which complicates the evaluation. Recurring patterns suggest that these cases may have been intentional, highlighting the value of predictive models in uncovering hidden risks in traffic data.

To interpret the influence of individual variables on the classification model output, we used the XGBoost algorithm in conjunction with the SHAP method. The model was trained on data containing both confirmed suicide attempts and common fatal traffic accidents, with SHAP values allowing us to quantify the im-

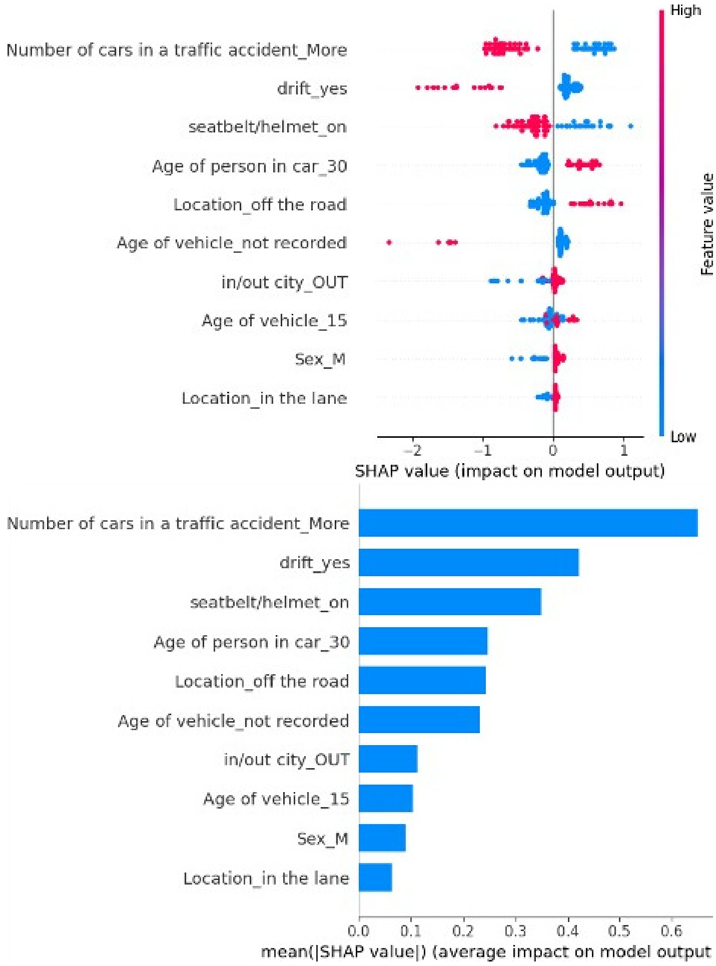

FIGURE 5. SHAP analysis of key factors of suicidal behavior in accidents.

pact of each variable on the model's decision-making. In order to increase interpretability and emphasize driver behavior and accident context, we deliberately excluded variables such as region, day of the week, and severity of injury from the analysis.

Both graphs in Figure 5 show the results of SHAP analysis, which evaluates the influence of individual variables on the classification of traffic accidents in relation to suspected suicidal behavior. The left part of the graph shows the dispersion of SHAP values for each variable separately. Each point represents one case from the test set, with the color of the point indicating the value of the variable (red = high, blue = low) and its position on the X-axis indicating the impact on the model output. It is clear that, for example, high values of the variable "multi-vehicle accident" (red) strongly contribute to a lower probability of suicidal behavior, while low values (blue) support this output. The same is true for variables such as "skid", "seat belt/helmet", or "off-road location", which have clearly defined directional effects.

The right side of the graph summarizes the average absolute SHAP values for each variable, expressing its overall importance in the model. The most influential factor here is "number of vehicles in the accident", followed by the occurrence of skidding, the use of safety features, and the location of the accident. Together, these variables carry the most weight in the model's decision-making process and reveal the dominant patterns that distinguish ordinary fatal accidents from incidents involving intentional behavior.

5. DISCUSSION OF RESULTS

5.1. RQ1: WHAT PATTERNS OF BEHAVIOR, TEMPORAL AND SITUATIONAL CIRCUMSTANCES ARE TYPICAL FOR CONFIRMED SUICIDE ATTEMPTS BY TRAFFIC ACCIDENT?

An analysis of confirmed cases of suicide attempts by traffic accident revealed distinct patterns in both behavior and spatiotemporal context. The most common day of occurrence was Sunday (20.8 %), followed by Saturday (19.4 %) and Thursday (18.1 %), which may reflect weekend psychological stress or a change in routine. Most events took place outside the municipality (86.1 %), which may be related to a lower risk of detection or a preference for open space. The dominant type of accident was a collision with a fixed object (79.2 %), most often a passenger car (88.9 %). In 75 % of cases, it was a single-vehicle accident, which further confirms intentional behavior without the presence of other participants. Driver error was cited as the cause in 100 % of cases. The presence of older vehicles was also significant, with 54.2 % of vehicles being over 15 years old and a lower average driver age (38.9 % were under 30 years old). The location of the accident, mainly in the lane (48.6 %) or off the road (40.3 %), confirms that the events occur without any significant attempt at correction. Approximately half of the drivers did not use a seat belt or helmet (48.6 %), which again indicates an intention to minimize the chance of survival.

This finding is consistent with the conclusions of the authors [9], who identified typical forms of suicidal behavior in traffic as crashing into a solid obstacle, driving off the road, or driving into oncoming traffic, with these incidents occurring more frequently in areas with lower traffic density. Similarly, [10] point out that fatal accidents of an intentional nature often show patterns of single-vehicle involvement and the absence of sudden evasive maneuvers. Furthermore, [12] confirm in their qualitative study that many individuals consider traffic suicide to be a “quick and concealable” form of ending their lives, which explains the preference for isolated, difficult-to-detect accidents.

5.2. RQ2: Which variables in the predictive model have the greatest influence on the classification of a case as a probable suicide attempt?

The predictive model using the SHAP method identified key variables that have the most significant influence on the classification of cases as potential suicide attempts. The greatest positive impact on the model’s output was the involvement of multiple vehicles in the accident (i.e. its absence increased the likelihood of a suicidal nature). This was followed by factors such as “drift”, the use of seat belts or helmets, the age of the person in the vehicle (especially under 30), the location of the accident off the road, and the older age of the vehicle. These variables show consistent agreement with the results of statistical analysis of cases and confirm the presence of behavioral and technical indicators that can be associated with intentional behavior with a higher degree of probability.

In contrast, other variables such as gender, the location of the accident in the lane, or whether the accident occurred in or outside a built-up area had lower predictive power and their influence on the model was limited. This suggests that for prediction purposes, it is more appropriate to focus on the specific circumstances of the accident rather than on more general demographic characteristics. The findings support the usefulness of predictive models in identifying atypical and potentially suicidal traffic accidents, especially when combined with detailed descriptive variables related to driver behavior and the nature of the event.

These findings correspond with the results of a systematic which described young driver age, low driving skill, and non-compliance with safety measures as key factors in increased risk. Furthermore, [2, 25] report that the presence of psychiatric disorders, untreated depression, or impulsivity can be strong predictors of suicidal behavior, which may be secondarily reflected in driving behavior. Furthermore [13] empirically demonstrates that young people with a history of self-harm are 23 % more likely to be involved in a serious traffic accident, which supports the results of this model.

5.3. RQ3: Can potential suicide attempts be reliably identified among ordinary traffic accidents based on the identified patterns and predictive model?

Based on the predictive model created, which was trained on confirmed cases of suicide attempts, its use in classifying ordinary fatal traffic accidents was subsequently tested. The results show that the model was able to identify 13 cases out of all fatal accidents in 2024 that bore strong similarities to previously analyzed suicide attempts. In these cases, the degree of suspicion was quantified, expressed as a percentage probability, with the highest suspicion reaching over 80 %.

The model’s ability to effectively recognize patterns in new data confirms its practical usefulness in the expost analysis of fatal traffic accidents and highlights the potential of this approach in the field of forensic traffic analysis. Although it is not possible to confirm the driver’s intent with absolute certainty in this way, it is an essential tool for selecting cases that deserve deeper psychological or criminological review.

The model’s ability to recognize patterns in new data is significant in terms of forensic traffic analysis. According to [11, 29], a significant proportion of traffic suicides remain hidden in official statistics, precisely because of the absence of a direct admission of intent or insufficient investigation of the circumstances of the accident. Study [15] also points to the role of systemic hesitation to classify an incident as a suicide, even when there are clear indications. In this context, the predictive model represents a tool that can support the review of cases that would otherwise remain undiagnosed.

6. Conclusion

The aim of this work was to identify potential cases of suicidal behavior through traffic accidents in the Czech Republic, based on data analysis combining exploratory data analysis (EDA), classification algorithms (KNN, XGBoost), and interpretation tools (SHAP). The study set out to distinguish patterns of behavior typical of confirmed suicide attempts from common fatal accidents and, on this basis, to propose a model for recognizing hidden cases.

The results of the analysis showed that suicidal accidents have specific characteristics, most often involving collisions with a fixed obstacle outside a built-up area in the morning or at night, where safety features are not used and the accident involves a single vehicle. Based on these characteristics, the model identified 13 cases that were highly likely to show signs of intentional behavior, even though they were listed as ordinary accidents in official statistics. The predictive model showed stable performance, and the key variables corresponded to the results of previous research in the fields of suicidology and traffic safety.

The contribution of this work is not only to map a previously neglected phenomenon, but also to use a tool that can serve as a basis for ex-post evaluation of fatal accidents with the aim of revealing hidden suicidal intent. This approach has the potential to contribute to more accurate statistics, targeted prevention, and forensic assessment of controversial cases.

The main limitations of the work include the limited number of confirmed suicide attempts on which the model was trained, as well as missing data in some records. Furthermore, misclassification by law enforcement agencies, which was reflected in the official data, cannot be ruled out.

In the future, it would be appropriate to expand the dataset with additional variables, including psychological and social factors, and to validate the model in other geographical areas or periods. Cooperation with the police and emergency services also appears promising, as it would make it possible to refine the recording of accident motives and identify at-risk individuals in a timely manner. Furthermore, an international comparison with similar cases in Australia, Finland, or Japan, where partial studies focusing on suicidal behavior in road traffic already exist, is recommended. Such comparative approaches could reveal culturally conditioned patterns and contribute to the creation of an internationally applicable system for the early detection and prevention of these specific cases.

7. Data availability statement

The detailed data and analyses are available in the attachments [30].

Attachment's first sheet summarizes the frequency of factors involved in suicide attempts by traffic accidents, such as day of week, region, accident type, vehicle type, and driver characteristics. Most cases involve collisions with fixed obstacles, male drivers, passenger cars, single vehicles, and occur outside city limits due to driver error.

Attachment's second sheet presents traffic accidents with a high predicted probability of suicidal behavior ($>= 0.4$), detailing factors such as date, time, region, accident type, vehicle, and driver characteristics. Most cases involve collisions with fixed obstacles, predominantly male drivers, and occur outside city limits due to driver error.

References

[1] S. Stack. Contributing factors to suicide: Political, social, cultural and economic. *Preventive Medicine* **152**:106498, 2021. https://doi.org/10.1016/j.ypmed.2021.106498

[2] A. Pigoni, G. Delvecchio, N. Turtulici, et al. Machine learning and the prediction of suicide in psychiatric populations: A systematic review. *Translational Psychiatry* **14**(1):140, 2024. https://doi.org/10.1038/s41398-024-02852-9

[3] Š. Selak, N. Crnkovič, A. Šorgo, et al. Resilience and social support as protective factors against suicidal ideation among tertiary students during COVID-19: A cross-sectional study. *BMC Public Health* **24**(1):1942, 2024. https://doi.org/10.1186/s12889-024-19470-1

[4] N. D. Weaver, G. J. Bertolacci, E. Rosenblad, et al. Global, regional, and national burden of suicide, 1990–2021: A systematic analysis for the global burden of disease study 2021. *The Lancet Public Health* **10**(3):e189–e202, 2025. https://doi.org/10.1016/S2468-2667(25)00006-4

[5] L. Schölin, D. Knipe, P. Bandara, et al. Banning highly hazardous pesticides saves the lives of young people, particularly females, in low- and middle-income countries. *BMC Public Health* **23**(1):2249, 2023. https://doi.org/10.1186/s12889-023-17071-y

[6] L. Favril, R. Yu, A. Uyar, et al. Risk factors for suicide in adults: Systematic review and meta-analysis of psychological autopsy studies. *Evidence Based Mental Health* **25**(4):148–155, 2022. https://doi.org/10.1136/ebmental-2022-300549

[7] M. Sinyor, M. Silverman, J. Pirkis, K. Hawton. The effect of economic downturn, financial hardship, unemployment, and relevant government responses on suicide. *The Lancet Public Health* **9**(10):e802–e806, 2024. https://doi.org/10.1016/S2468-2667(24)00152-X

[8] K. Hawton, J. Pirkis. Preventing suicide: A call to action. *The Lancet Public Health* **9**(10):e825–e830, 2024. https://doi.org/10.1016/S2468-2667(24)00159-2

[9] P. C. F. Law, L. S. Too, M. J. Spittal, et al. Road vehicle collision suicide in Australia: Trends, collision types, and individual characteristics. *PLOS ONE* **19**(4):1–13, 2024. https://doi.org/10.1371/journal.pone.0299590

[10] I. Radun, J. Radun, M. Sutela, M. Tolvanen. Deliberate fatal crashes involving a motor vehicle and a cyclist or pedestrian. *Journal of Transport & Health* **30**:101619, 2023. https://doi.org/10.1016/j.jth.2023.101619

[11] D. Yelo, A. Parkes, J. Freeman, J. Davey. Complex and underreported? A study into the prevalence of suicide by motor vehicle in the state of Queensland. *Transportation Research Part F: Traffic Psychology and Behaviour* **81**:445–456, 2021. https://doi.org/10.1016/j.trf.2021.06.015

[12] H. Norman, L. Marzano, R. Winter, et al. Factors prompting and deterring suicides on the roads. *BJPsych Open* **9**(3):e81, 2023. https://doi.org/10.1192/bjo.2023.52

[13] P. Cullen, H. Mőller, R. Baffsky, et al. Self-harm in adolescence and risk of crash: A 13-year cohort study of novice drivers in New South Wales, Australia. *Injury Prevention* **29**(4):302–308, 2023. https://doi.org/10.1136/ip-2022-044807

[14] Assembly Committee on Transportation. SB 915 (Cortese) – As amended May 16, 2024, 2024.

[15] G. Tait, S. Jowett, B. Carpenter. Coronial decision-making and the management, classification and conceptualisation of the finding of 'suicide'. *Mortality* **25**(3):297–312, 2020. https://doi.org/10.1080/13576275.2019.1645104

[16] K. Bucsuházy, R. Zůvala, J. Ambros. Analysis of Covid restrictions' influence on road traffic crashes and related road users' behaviour in the Czech Republic. *Archives of Transport* **66**(2):109–121, 2023. https://doi.org/10.5604/01.3001.0053.6084

[17] S. S. Alavi, M. R. Mohammadi, H. Souri, et al. Personality, driving behavior and mental disorders factors as predictors of road traffic accidents based on logistic regression. *Iranian Journal of Medical Sciences* **42**(1):24–31, 2017.

[18] V. Kovač, N. Kováčiková, L. Turinská. Development of passenger car safety. *Ad alta: Journal of Interdisciplinary Research* **13**(2):361–367, 2023.

[19] V. Kovač, S. Alexandros, X. Panagiota. Relationship between vehicle price and its safety ratings. *LOGI – Scientific Journal on Transport and Logistics* **15**(1):37–48, 2024. https://doi.org/10.2478/logi-2024-0004

[20] T. Řezníček, V. Kovač. The impact of demographic factors on the traffic accident rate in Czech Republic. *Transport Problems* **20**(2):5–17, 2025. https://doi.org/10.20858/tp.2025.20.2.01

[21] M. Hlatká, G. Fedorko, N. Mikušová, et al. Example of an exercise plan for solving an accident on the railway track – Case study. *Transportation Research Procedia* **77**:151–156, 2024. https://doi.org/10.1016/j.trpro.2024.01.020

[22] G. Kiwango, D. Katopola, F. Francis, et al. A systematic review of risk factors associated with road traffic crashes and injuries among commercial motorcycle drivers. *International Journal of Injury Control and Safety Promotion* **31**(2):332–345, 2024. https://doi.org/10.1080/17457300.2024.2319628

[23] G.-D. Voinea, R. G. Boboc, I.-D. Buzdugan, et al. Texting while driving: A literature review on driving simulator studies. *International Journal of Environmental Research and Public Health* **20**(5):4354, 2023. https://doi.org/10.3390/ijerph20054354

[24] T. Řezníček, V. Sloup. Car accidents in the Czech Republic and light as a contributory factor. *Ad alta: Journal of Interdisciplinary Research* **14**(1):348–354, 2024. [2024-12-11]. https://www.webofscience.com/wos/woscc/full-record/WOS:001318682300001

[25] M. J. Rapoport, J. N. Chee, T. Prabha, et al. A systematic review of the risks of motor vehicle crashes associated with psychiatric disorders. *The Canadian Journal of Psychiatry* **68**(4):221–240, 2023. https://doi.org/10.1177/07067437221128468

[26] J. A. Staples, D. Daly-Grafstein, M. Khan, et al. Acute psychosis and the risk of motor vehicle crash. *medRxiv* pp. 1–27, 2024. https://doi.org/10.1101/2024.12.10.24318820

[27] Transport Research Centre. Dopravní nehody v ČR [In Czech; Road traffic accidents in the Czech Republic]. [2025-02-01]. https://nehody.cdv.cz/statistics.php

[28] M. Kostíková, K. Bucsuházy, P. Moravcová. Případová studie: Komplexní porovnání dvou téměř identických nehod [In Czech; Fall or collision? A comprehensive comparison of causes in traffic accidents]. *Soudní lékařství* **66**(1):6–10, 2021.

[29] M. Pompili, G. Serafini, M. Innamorati, et al. Car accidents as a method of suicide: A comprehensive overview. *Forensic Science International* **223**(1–3):1–9, 2012. https://doi.org/10.1016/j.forsciint.2012.04.012

[30] V. Kovač, T. Řezníček, A. Ibraheem. Supplementary materials for the article "Potential suicide attempts through traffic accidents (PSA-TA)", 2025. [2026-03-10]. https://drive.google.com/drive/folders/12eS2dclJSC5kD-lqzOKI5byIBSh7jUDm