

Machine Learning Tools for Predicting the Outcome of Pedestrian Crashes: Preliminary Findings in the Metropolitan City of Rome

Mauro D'Apuzzo[1(✉)], Giuseppe Cappelli[1,2], Sofia Nardoianni[1], Michele De Guidi[1], and Vittorio Nicolosi[3]

[1] University of Cassino and Southern Lazio, Via G. Di Biasio 43, 03043 Cassino, Italy
{dapuzzo,giuseppe.cappelli1,sofia.nardoianni}@unicas.it,
michele.deguidi@studentmail.unicas.it
[2] University School for Advanced Studies, IUSS, Piazza della Vittoria n. 15, 27100 Pavia, Italy
giuseppe.cappelli@iusspavia.it
[3] University of Rome "Tor Vergata", Via del Politecnico, 1, 00133 Rome, Italy
nicolosi@uniroma2.it

Abstract. Creating cities and roads at a human scale is the main challenge of a not-distant future. The spread of new modes of transport reduces the distances from this future at a high speed and changes in vulnerable user urban mobility pose new questions in terms of sustainability a safety issues. Actions to promote these two vital conditions constitute the main prerequisites to be ensured by experts in the context of policies, economics, planning, and design.

Road accidents are among the top ten causes of death worldwide, ranking seventh overall. On the other hand, they occupy the first place as regards deaths between the ages of 15 and 29. More than 1.2 million people in a single year die and most of them suffer from non-fatal injuries that may often impair quality of life.

Although the damages to the human body are unfortunately well known, thanks also to the current scientific literature, accidents may cause a large economic cost, and knowing and understanding this macroeconomic burden could give a strong direction in an effective policy-making decision process.

In this paper, following a general review of pedestrian accidents that occurred in the city of Rome, several machine learning models are proposed to predict the outcome of an accident involving a pedestrian.

Keywords: Sustainable Mobility · Road Safety · Pedestrian accidents · pedestrian safety · Random Forest · SVM · Machine Learning

1 Introduction

Road accidents unfortunately represent one of the leading causes of premature death and permanent disability in many parts of the world [1]. This is particularly true for pedestrians, cyclists, and motorcyclists, who constitute the most vulnerable road users

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
O. Gervasi et al. (Eds.): ICCSA 2024 Workshops, LNCS 14824, pp. 116–132, 2024.
https://doi.org/10.1007/978-3-031-65332-2_8

[2, 3]. Their vulnerability stems mainly from the lack of external protections, exposing them to significantly higher risks than other road users when interacting with the urban environment [3–5].

Pedestrians are exposed to risk by various factors including speeding vehicles, jaywalking, distractions, psychophysical impairments (such as alcohol, drugs, and disease), and poor infrastructure conditions [6]. These factors contribute to road accidents involving pedestrians, which often result in fatalities [2, 7–11]. Reducing risk factors and promoting responsible behaviors is a challenge for every national government, as fewer deaths and injuries on the roads would lead to fewer economic and logistical damages [12, 13]. According to some authors, it has been demonstrated that the promotion of pedestrian and cycling mobility requires changes to road infrastructure [14–25] aimed at increasing safety and encouraging this type of mobility [26–28].

Sustainability is not only about the environment but also about maintaining the balance between society and economy. To achieve this balance, it is essential to reduce the number of victims of road accidents [29]. In the pursuit of sustainable development and improvement of road safety, various initiatives have been adopted both nationally and internationally. The 2030 Sustainable Development Goals, established by the United Nations in 2015 [29], serve as a guiding framework for global efforts towards sustainable development. Complementing these objectives, Directive 2019/1936 in Europe highlights the importance of road safety rules and standards to mitigate risks on the roads [30]. Furthermore, the Next Generation EU initiative aims to address the socio-economic challenges posed by the COVID-19 pandemic and promote a sustainable recovery across Europe [31]. In Italy, the National Recovery and Resilience Plan (PNRR) is the plan for accessing the funds requested by the European Commission from each of the Member States to obtain the funds from the Next Generation EU plan [31, 32]. The PNRR adopts a proactive approach by allocating significant investments towards the strengthening and innovation of infrastructure to be implemented by 2026, following the ecological transition, for €61.5 billion. Italy has also established and updated the "National Plan for Road Safety 2030" to reduce road victims by 50% by 2030 with a series of investments and reforms [33]. These coordinated efforts underline the commitment to sustainable development while ensuring safer and more resilient transport systems.

Despite this great movement by international and European policies, as can be seen from the Global Status Report on Road Safety 2018 [1], the number of deaths and injuries worldwide has certainly increased, undermining the objectives of the Sustainable Development Goals (SDGs). As noted by WHO, it is possible to see a double trend that refers to countries with a high average income and those with a medium-income: in the former, mainly European Countries, it is associated with a decrease in the number of accidents [34], while in the latter (especially in African and South-East Asian countries in strong development) there is a general and dizzying increase in accidents leading to injury and death [2].

In a study on 166 countries, Chen et al. [35] evaluate the economic burden of all types of road accidents using a macro-economic model that explains the effects of fatal and non-fatal injuries on labor supply, differences in the education and expertise in involved people and fewer investments due to the diversion of treatment costs. The evidence suggests that road accidents cost the world economy US $1.8 trillion in 2015–2030,

which equates to an annual tax of 0.12% on global GDP: low-medium income countries have the greatest health burden but a lower economic burden than high-income countries.

Economic considerations, however, find time and place in the analysis of sustainability assessment. In the study of Karkalakos & Tselekounis [36], the authors study the impact of road safety (expressed as a number of deaths per million inhabitants) on sustainability performance in the Organisation for Economic Co-operation and Development area (OECD). The two authors suggest that about 61.88% of the variation in the sustainability score can be explained by a variation in the number of serious road accidents.

Several works have analyzed the occurrence patterns and risk factors of traffic accidents involving vulnerable road users in different contexts and some of the most important variables that influence vulnerable users' accidents are:

- The modal share of different road users, such as pedestrians, cyclists, and motorcyclists. Cities with higher rates of walking and cycling are safer for vulnerable road users than cities with higher rates of car occupancy [37], and a smaller amount of fatal and serious injuries [37, 38].
- The road infrastructure characteristics, such as segment length, driver sight clearance, land use, crash time, crash season, number of median openings, curves, and pedestrian crossings. These factors affect the visibility, maneuverability, and collision severity of vulnerable road users [39].
- The intelligent transport systems (ITS) that provide information and guidance to vulnerable road users through various devices and applications. ITS can improve the awareness, decision-making, and behavior of vulnerable road users in different scenarios [40].

However, it is important to emphasize that road safety responsibility is shared among all road users, including pedestrians [41]. Pedestrians must adhere to traffic rules and adopt cautious behaviors during urban movements to reduce the risk of accidents and protect their safety, according to the Italian Highway Code [41].

In this study conducted in the urban area of Rome, the understanding of pedestrian-involved accidents was deepened using data provided by ISTAT [42], referring to accidents involving at least one fatality or injury. Initially, an empirical approach to the data was adopted, employing descriptive analyses and categorizing accidents involving pedestrians based on their dynamics. Subsequently, the characteristics and information describing them are analyzed (road type, environmental conditions, circumstances, time, etc.).

Finally, machine-learning models [43–45] are used to predict the outcomes of such incidents, allowing designers to more accurately assess the risk of serious injury or death for pedestrians involved.

2 Crash Data

2.1 Descriptive Statistical Analysis

To analyze the accidents within the urban area of Rome, the database drawn up by ISTAT (Italian Institute of Statistics) [42] was used, which records accidents in which at least a road user involved is injured or dead, with a maximum observation time interval of

thirty days, further classifying deaths as having died within twenty-four hours or thirty days [46].

The event is recorded as a sample with various information and characteristics (independent variables or features). The 121 features are divided into information on the road users involved (sex and age, for instance), on the vehicles (type of vehicle, registration, etc.), dynamics, and environmental conditions. The analysis refers to the period from 2014 to 2020 for a total of 84,243 accidents that occurred in the city of Rome. Accidents with at least one pedestrian involved were extracted from this database, thus analyzing a total of 12,359 accidents.

As it is possible to see in Fig. 1 (and with particular detail in Table 1), the main circumstances that lead to a crash occur when the driver's car proceeds regularly, maneuvering or proceeding at high speed. A great part of accidents is also caused by non-direct accidents. Although there are other circumstances, for the descriptive statistical analysis, only the first six circumstances are described to give a preliminary overview.

Figure 1 and Table 1 are linked: the former shows a mean value of the frequency in the years analyzed, and in the latter, a focus on a single year.

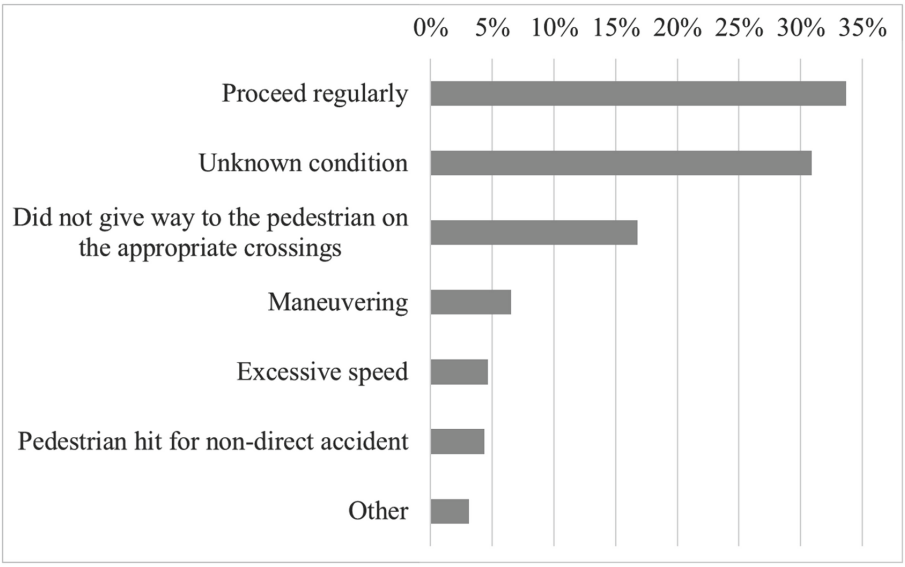

Fig. 1. Main circumstances that lead to a pedestrian accident.

The analysis also shows that the probability of being involved in a pedestrian accident, according to the circumstance analyzed, decreased in the period 2014–2020 and that shows a higher interest in last year's pedestrian safety than in the previous years, due to the awareness of road users and cities managers on the topic (see Fig. 2).

Although the number of accidents is decreasing, the problem of injured and dead road users needs to be analyzed in detail. As it is possible to see in Table 2 and Fig. 3 (in which the same results are proposed, in more detail in the table, and with a graphical output in the figure), according to each circumstance, the vast majority of crashes occur

on working days and this may be due to critical traffic conditions on city roads, while on the weekend, there is a large decreased of the number that involve a pedestrian. The hourly pattern in Fig. 4 shows that the number of accidents increases between 10 a.m. and 2 p.m. and between 6 p.m. and 10 p.m.

Half of all accidents happen on two-way carriageways for all circumstances identified: the second place is it possible to find two carriageways and this shows that the number of conflicts between motorized and pedestrian flows likely induce an accident (see Fig. 5 and Table 3, that show the same numbers differently). There is also a high probability of being involved in a pedestrian accident in a straight road and intersection for all circumstance (see Fig. 6). In Table 4, the majority of pedestrian crashes happen when vertical and horizontal road signs are present.

In Table 5 and Table 6, it is possible to see the probability that a pedestrian accident can happen according to weather and road surface conditions. There is a high probability that an accident will occur in sunny weather and on dry road surfaces: this is mainly because accident data are very unbalanced data, so there are more injured users than dead road users, so the frequency an accidents happening in normal conditions is high.

Table 1. Probabilities of accidents under specific circumstances: historical series [%].

Circumstances	2014	2015	2016	2017	2018	2019	2020
Proceed regularly	8.94	19.55	27.72	34.81	40.50	45.00	45.29
Unknown Condition	45.00	42.00	36.65	30.98	29.08	24.89	23.37
Did not give way to the pedestrians on the crossings	6.22	5.93	4.63	4.53	3.51	3.71	4.19
Maneuvering	0.53	0.16	0.22	0.16	0.17	0.05	0.16
Excessive speed	0.43	0.59	0.32	0.38	0.28	0.38	0.32
Pedestrian hit for non-direct accident	0.16	0.05	0.11	0.54	0.00	0.11	0.24
Maneuvered without complying with traffic signals	9.52	7.05	6.62	6.03	5.79	5.00	5.32
Running off-road and hitting the pedestrian	0.85	0.86	0.91	0.86	0.22	0.11	0.73
Hitting the pawn with the load	0.48	0.59	0.65	0.05	0.39	0.11	0.32
Proceed the wrong way	1.44	0.81	0.27	0.22	0.28	0.54	0.40
Backing without precaution from a driveway	20.32	16.86	16.47	16.97	14.43	15.43	16.76
Proceed without respecting the speed limits	0.16	0.27	0.22	0.32	0.33	0.05	0.08
Overtaking a vehicle stopped to allow crossing	0.85	0.54	0.59	0.48	0.28	0.38	0.32
Overtaking a vehicle on the move	0.00	0.16	0.00	0.00	0.00	0.00	0.00
Irregularly passed a tram at the stop	5.11	5.12	4.63	3.66	4.74	4.25	2.50

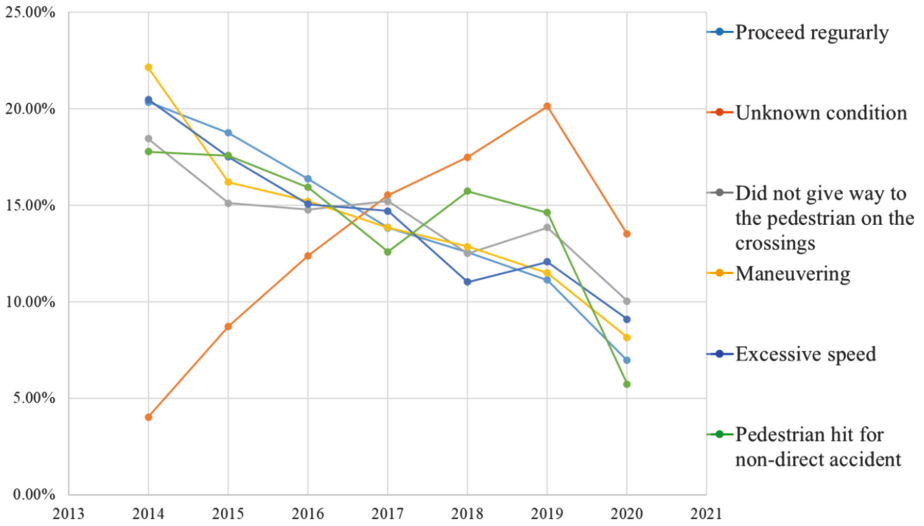

Fig. 2. Reduction of pedestrian accidents in the years 2014–2020.

Table 2. Number of accidents according to the circumstances and according to the day of the week (2014–2020).

Circumstances	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Proceed regularly	342	386	369	325	362	195	134
Unknown Condition	67	680	734	697	647	461	268
Did not give way to the pedestrians on the crossings	614	661	597	606	617	445	278
Maneuvering	134	155	140	119	120	89	51
Excessive speed	95	93	98	79	100	62	44
Pedestrian hit for non-direct accident	86	86	82	83	96	51	56

Evidence of that is shown in Table 7. The results show that the percentage of injured pedestrians is higher compared to the other types of outcomes.

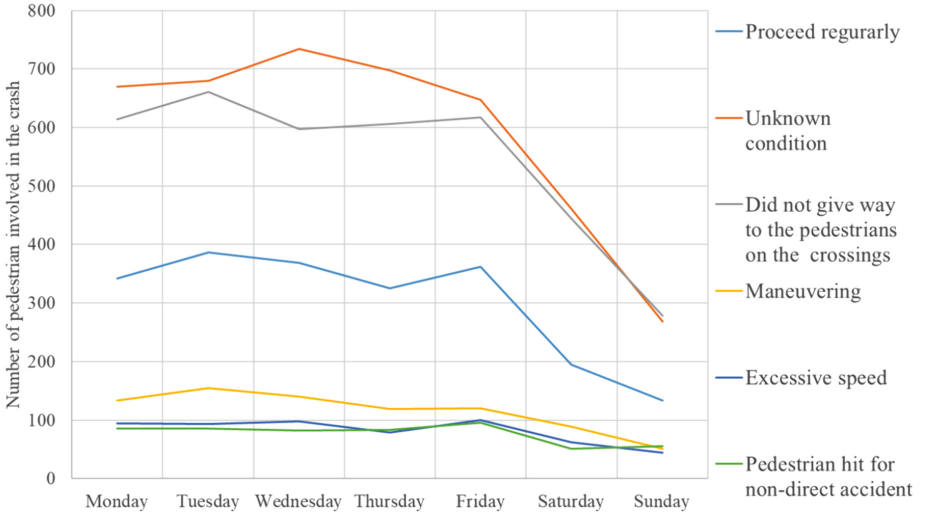

Fig. 3. Weekly pattern of pedestrian accidents in the years 2014–2020, according to the day of the week and the circumstances.

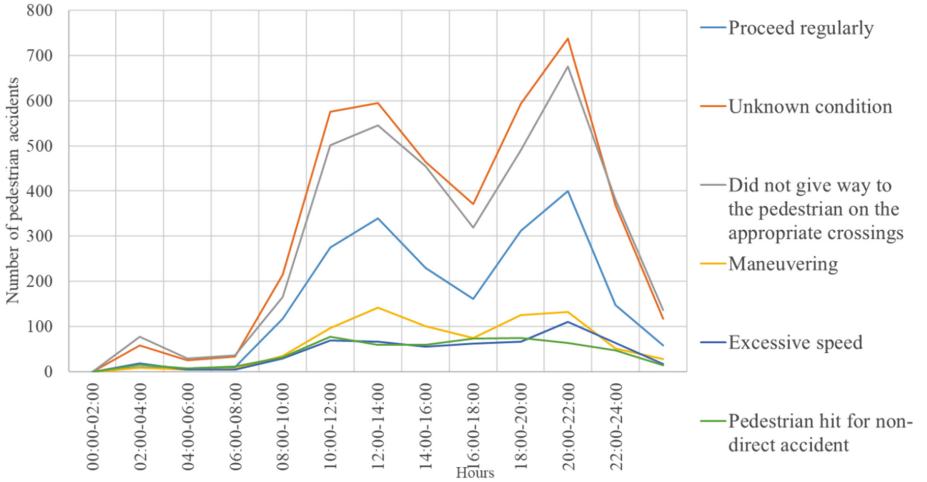

Fig. 4. Number of accidents according to the circumstances and the hour of the day (2014–2020): graphical output.

Table 3. Probability of accident [%] according to the circumstances and the road types (2014–2020).

Circumstances	One-way street	Two-way carriageway	Two carriageways	More than two carriageways
Proceed regularly	20.92	52.27	21.69	5.12
Unknown Condition	22.97	51.05	19.03	6.95
Did not give way to the pedestrians on the crossings	23.31	51.36	17.42	7.91
Maneuvering	26.98	51.11	16.96	1.05
Excessive speed	22.42	51.14	21.54	4.90
Pedestrian hit for non-direct accident	17.41	52.96	25.19	4.44

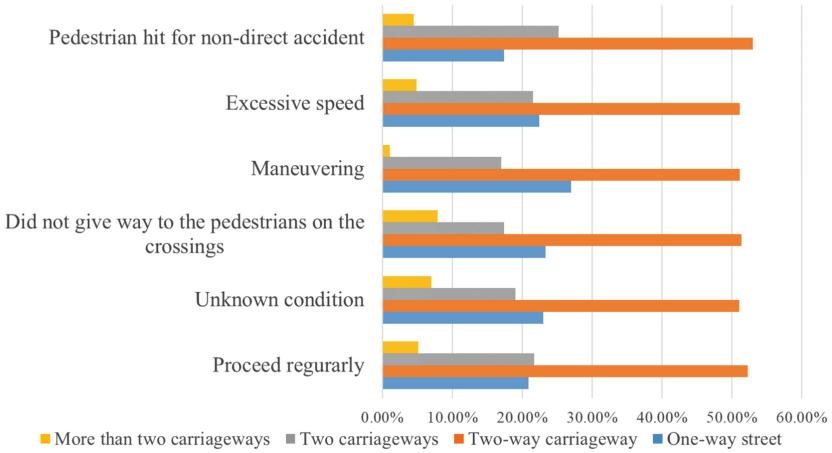

Fig. 5. Distribution of pedestrian accidents [%] according to the circumstances and the road types (2014–2020)

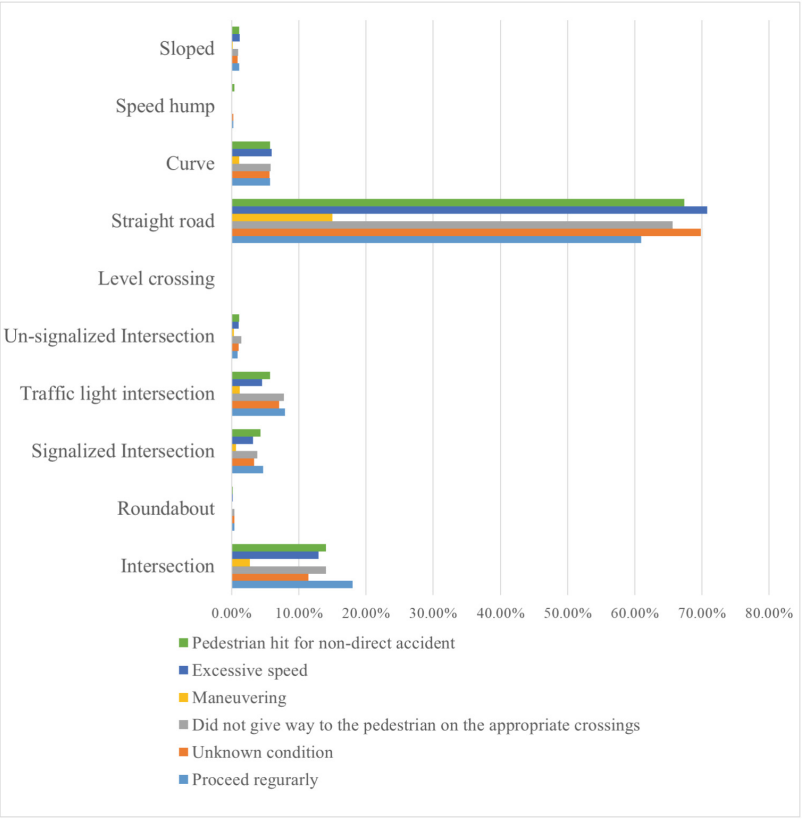

Fig. 6. Distribution of pedestrian accidents [%] according to road characteristics (2014–2020).

Table 4. Probability of pedestrian accident [%] according to road signs (2014–2020).

Circumstances	Missing	Vertical	Horizontal	Vertical and Horizontal	Other
Proceed regularly	3.09	2.17	16.09	78.26	0.39
Unknown Condition	8.47	4.50	16.60	70.12	0.31
Did not give way to the pedestrians on the crossings	8.51	3.93	14.51	72.81	0.24
Maneuvering	15.35	6.56	17.45	60.40	0.05
Excessive speed	8.58	4.38	14.01	72.68	0.35
Pedestrian hit for non-direct accident	4.63	4.26	16.48	73.89	0.74

Table 5. Probability of accident [%] according to the weather conditions (2014–2020).

Circumstances	Sunny	Fog	Rain	Hail	Snow	Other
Proceed regularly	87.10	0.34	11.11	0.00	0.05	1.40
Unknown Condition	87.76	0.34	10.37	0.00	0.05	1.49
Did not give way to the pedestrians on the crossings	88.48	0.50	9.17	0.00	0.00	1.86
Maneuvering	90.72	0.50	7.55	0.00	0.00	1.24
Excessive speed	86.16	0.35	12.26	0.18	0.00	0.74
Pedestrian hit for non-direct accident	90.93	0.56	7.22	0.19	0.19	0.93

Table 6. Probability of accident [%] according to the road surface conditions (2014–2020).

Circumstances	Dry	Wet	Slippery	Frozen	Snow-covered
Proceed regularly	81.93	17.97	0.10	0.00	0.00
Unknown Condition	82.68	17.30	0.00	0.00	0.02
Did not give way to the pedestrians on the crossings	82.79	17.10	0.08	0.03	0.00
Maneuvering	84.78	15.10	0.12	0.00	0.00
Excessive speed	80.21	19.61	0.18	0.00	0.00
Pedestrian hit for non-direct accident	83.70	15.56	0.19	0.37	0.19

Table 7. Probability of accident [%] according to the outcome of the crash (2014–2020).

Circumstances	Dead driver	Injured driver	Dead pedestrian	Injured pedestrian	Dead pedestrian in 24 h	Dead pedestrians in 30 days
Proceed regularly	0.04	5.48	2.04	90.37	1.24	0.84
Unknown Condition	0.06	8.32	2.64	86.28	1.91	0.79
Did not give way to the pedestrians on the crossings	0.02	7.00	1.42	90.11	1.11	0.34
Maneuvering	0.11	4.43	1.33	92.80	1.11	0.22
Excessive speed	0.11	4.43	1.33	92.80	1.11	0.22
Pedestrian hit for non-direct accident	0.11	4.43	1.33	92.80	1.11	0.22

3 Method

3.1 Machine Learning Process

Machine Learning (ML) is a sector of artificial intelligence (AI) that allows machines and computers to develop and improve models (or even algorithms). These models are instructed through input data to perform specific tasks and make predictions without being explicitly programmed, so the machine is predisposed to learn something independently. Machine Learning means the ability to perform with some precision and accuracy new tasks, or different circumstances not previously addressed, based on the experience gained from a set of training data [47].

Two main types of approaches to Machine Learning exist [48]:

- Supervised Learning: examples are provided to the model in the form of possible inputs and their desired outputs and the goal is to extract a general rule that associates the input to the correct output.
- Unsupervised Learning: the model aims to find a structure in the input provided, without the input being labeled in any way.

Machine Learning process can be summarized in the following points:

- Data collection: collection of data needed to train the model. Data can come from various sources and must be relevant and representative of the problem.
- Data Preparation: clean, pre-process, and transform data so that it is suitable for model training.
- Model selection: choose the model type best suited to the problem and the available data. The most famous models are, for instance:

  1. Linear Discriminant Analysis (LDA) [49];
  2. Quadratic Discriminant Analysis (QDA) [50];
  3. K-Nearest Neighbour (KNN) [51];
  4. Decision Tree (DT) [52];
  5. Random Forest (RF) [53];
  6. Support Vector Machines (SVM) [54].

- Model training: using training data to teach the model to make predictions or perform the desired task.
- Model evaluation: model performance evaluation using separate data (test and validation datasets) to assess its generalization capability and accuracy on new data.

4 Results

Before moving on to the discussion of the results, it is necessary to introduce the metrics to evaluate the model's assumptions and then make comparisons: all will start from the introduction of the concept of the confusion matrix (see Fig. 7) to clarify in detail terms such as specificity, recall, F1-score, etc.

A confusion matrix is a table that summarizes the performance of a classification model on a set of test data. It shows the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) for each class. The output of the

	Real Values	
Predicted Values	True Positive (TP)	False Positive (FP)
	False Negative (FN)	True Negative (TN)

Fig. 7. An Example of Confusion Matrix.

model is the Confusion Matrix: the four elements that compose it are the relationship between the model's predicted and the actual values:

- True Positive (TP): are the cases where the model has correctly predicted a positive class;
- True False (TF): are the cases in which the model has correctly predicted a negative class;
- False Positive (FP): these are the cases where the model predicted a positive class that was negative (Type I error);
- False Negative (FN): these are the cases where the model predicted a negative class that was positive (Type II error).

A confusion matrix can help in evaluating the accuracy, precision, recall, and F1-score of models, as well as identify the classes that are most often confused by the model itself. Some common performance metrics for classification problems are:

- Accuracy: the proportion of correct predictions among all predictions. It is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

- Precision: the proportion of correct positive predictions among all positive predictions. It is calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

- Recall: the proportion of correct positive predictions among all actual positive predictions. It is calculated as:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

- Specificity: the specificity of a confusion matrix quantifies the true negative rate or the number of observations the model defined as negative that were truly negative:

$$Specificity = \frac{TN}{TP + FN} \tag{4}$$

- F1-score: the harmonic means of precision and recall. It is calculated as:

$$F1 - score = \frac{2 * Precision * Recll}{Precision + Recall} \tag{5}$$

There are also two other important metrics. The F2 and F3 scores are variations of the F1 score, which is a measure of a model's accuracy on a dataset. The F1 score is the harmonic mean of the model's precision and recall, and it ranges from 0 to 1. A perfect model has an F1 score of 1.

The F2 and F3 scores are adjusted F-scores that give more importance to recall than precision. Recall is the fraction of correct positive predictions among all actual positive instances. The F2 score considers recall to be twice as important as precision, while the F3 score considers recall to be three times as important as precision. The formulas for the F2 and F3 scores are:

$$F2 - score = \frac{5 * Precision * Recll}{4 * Precision + Recall} \tag{6}$$

$$F3 - score = \frac{10 * Precision * Recll}{9 * Precision + Recall} \tag{7}$$

In Table 8 below the main metrics analyzed for each model used to predict the outcome of a pedestrian accident.

Table 8. Main performance metrics used to assess the prediction accuracy of the model for the classification task.

Performance metrics	LDA	QDA	KNN	DT	RF	SVM-L	SVM-R	SVM-P
Accuracy	0.6787	0.7377	0.8824	0.9306	0.9378	0.4763	0.7818	0.6302
Recall	0.6901	0.7068	0.8939	0.9478	0.9542	0.9330	0.9845	0.9903
Specificity	0.6680	0.7815	0.3625	0.1605	0.1875	0.0325	0.0379	0.0414
Precision	0.6588	0.8205	0.9845	0.9806	0.9817	0.4685	0.7897	0.6281
F1-score	0.6741	0.7594	0.9370	0.9639	0.9678	0.6367	0.8764	0.7687
F2-score	0.6836	0.7270	0.9107	0.9542	0.9596	0.8114	0.9382	0.8879
F3-score	0.6869	0.7167	0.9022	0.9510	0.9569	0.8930	0.9608	0.9363

According to the F1-score, which is a harmonic mean value of precision and recall, the best model seems to be the Random Forest (RF) model. Another important metric to observe is the F3-score which ranges from 0 to 1, where 1 is the best possible score and 0 is the worst. The F3 score is useful for evaluating machine learning models that deal with unbalanced data or rare events, where recall is more important than precision. For example, in medical diagnosis, it might be preferable to have a high recall to avoid missing any potential cases of a disease, even if that means some false alarms. And that is the case of pedestrian road accidents, where the dataset is very unbalanced and the number of fatal injured pedestrians is very low compared to the only injured ones. The

best model according to the F3-score is the Support Vector Machine with a radial kernel model (SVM-R).

5 Conclusions

Although the number of road accidents involving pedestrians in Europe is decreasing, globally, especially in emerging or developing countries, there is a trend opposite to that in Europe. This may be due to increased interest on the part of governments to reduce the number of accidents by promoting projects, information campaigns, and establishing funds for road safety. Furthermore, the objective of reducing the number of accidents fits perfectly into the sustainable development objectives of Agenda 30, to make European and world cities safer, and reduce the number of pollutants released into the atmosphere.

Taking a look at the current scientific technical literature, there is a need to understand what are the main factors and the main characteristics that lead to the cause of an accident and what actions need to be undertaken to reduce the severity and degree of injury of a vulnerable road user. Being the problem of not-so-simple resolution, models based on artificial intelligence and supervised learning could be useful. For this purpose, a first step in this paper was made to predict the outcome of an accident. Taking as reference the ISTAT data of the metropolitan city of Rome, first a descriptive analysis was conducted, and then several models were trained and then tested. The performance of the model was evaluated on the test set and not on the training set. Furthermore, since the death of a pedestrian is very rare in comparison to the number of accidents leading to injury, the resulting dataset is very unbalanced.

In this regard, oversampling techniques have been applied to the training set to balance the minority class. Without this, the model would only have predicted injured pedestrians and dead pedestrians. Applying models based on their degree of complexity, without any prejudice to one or the other, the most performing model in terms of F1-score was the Random Forest, while according to the F3-score, which gives more importance to recall than precision, was found to Support Vector Machine with a radial kernel the most relevant model to the data collected. Evaluating also the F3-score metric is necessary when the dataset is very unbalanced towards a certain class. In this case, it is reiterated, that the outcome of the death of the pedestrian belonged to the minority class, so models that predict many more false positives than models that predict few events in the minority class are preferred. The novelty of this research consists of applying several Machine Learning models and making an exhaustive comparison among them, for a single Italian city.

Further analysis is necessary and other metrics of evaluation of the performance of the model could be used. The main limitation of these models is that they work as black boxes and therefore there is a lack in the interpretability side compared to simple models like Logistic Regression that have their own more explicit mathematical structure, so within this research, it is also difficult to understand which variables are strongly related to pedestrian crashes. Obviously, as future developments, it is necessary to apply techniques, tools, and methods that allow not only to predict the outcome of an accident but also to understand what variables are the most influencing the outcome of an accident.

Acknowledgments. The research leading to these results has received funding from Project "Ecosistema dell'innovazione Rome Technopole" financed by the EU in the NextGenerationEU plan through MUR Decree n. 1051 23.06.2022 - CUP H33C22000420001. This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

1. Global Status Report on Road Safety. WHO World Health Organisation, p. X (2015). https://www.who.int/publications/i/item/9789241565066. Accessed Jan 2024
2. Pedestrian Safety. WHO World Health Organization et al. Cap, pp. 1–2 (2013). https://iris.who.int/bitstream/handle/10665/79753/9789241505352_eng.pdf?sequence=1. Accessed Jan 2024
3. Yannis, G., et al.: Vulnerable road users: Cross-cultural perspectives on performance and attitudes **44**(3), 220–229 (2020)
4. Sun, Z., et al.: Exploring injury severity of vulnerable road user involved crashes across seasons: a hybrid method integrating random parameter logit model and Bayesian network. ScienceDirect **150**, 105682 (2022)
5. Wolfensohn, J., Cherpitel, D. (2002)
6. Pedestrian Safety. WHO World Health Organization et al., pp. 17–20 (2013). https://iris.who.int/bitstream/handle/10665/79753/9789241505352_eng.pdf?sequence=1. Accessed:Jan 2024
7. Odero, W.: Alcohol-related road traffic injuries in Eldoret, Kenya. East Afr. Med. J. **75**, 708–711 (1998)
8. Drinking and Driving: A Road Safety Manual for Decision-Makers. Global Road Safety Partnership, Geneva (2007)
9. Peden, M., et al. (eds.): World Report on Road Traffic Injury Prevention. World Health Organization, Geneva (2004)
10. World Health Organization. Mobile Phone Use: A Growing Problem of Driver Distraction. WHO, Geneva (2011)
11. McLean, A.J., et al.: Vehicle Travel Speeds and the Incidence of Fatal Pedestrian Collisions, vol. 1. Federal Office of Road Safety, Canberra (1994) (CR 146)
12. Equivalence of Distribution of Social and Economic Damage Caused by Road Accidents in the Territories of the Largest Countries of the World: A New Pattern? vol. 36, pp. 577–585 (2018)
13. The Global Macroeconomic Burden of Road Injuries: Estimates and Projections for 166 Countries, vol. 3, no. 9, pp. e390–e398 (2019)
14. Pedestrian Safety. WHO World Health Organization et al. cap. 4 Implementing Pedestrian Safety Interventions (2013). https://iris.who.int/bitstream/handle/10665/79753/9789241505352_eng.pdf?sequence=1. Accessed Jan 2024
15. Abdelfattah, L., et al.: The 15-minute city: interpreting the model to bring out urban resiliencies. Transp. Res. Procedia **60**, 330–337 (2022)
16. Santilli, D., D'Apuzzo, M., Evangelisti, A., Nicolosi, V.: Towards sustainability: new tools for planning urban pedestrian mobility. Sustainability **13**(16), 9371 (2021)
17. D'Apuzzo, M., Santilli, D., Evangelisti, A., Pelagalli, V., Montanaro, O., Nicolosi, V.: An exploratory step to evaluate the pedestrian flow in urban environment. In: Gervasi, O., et al. (eds.) ICCSA 2020. LNCS, vol. 12255, pp. 645–657. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58820-5_47

18. D'Apuzzo, M., Santilli, D., Evangelisti, A., Nicolosi, V., Cappelli, G.: Estimation of pedestrian flows in urban context: a comparison between the pre and post pandemic period. In: Gervasi, O., Murgante, B., Misra, S., Ana, M.A., Rocha, C., Garau, C. (eds.) Computational Science and Its Applications – ICCSA 2022 Workshops: Malaga, Spain, July 4–7, 2022, Proceedings, Part IV, pp. 484–495. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10542-5_33
19. D'Apuzzo, M., Santilli, D., Evangelisti, A., Nicolosi, V.: A conceptual framework for risk assessment in road safety of vulnerable users. In: Gervasi, O., et al. (eds.) ICCSA 2021. LNCS, vol. 12958, pp. 542–556. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87016-4_39
20. D'Apuzzo, M., Santilli, D., Evangelisti, A., Nicolosi, V.: Some remarks on soft mobility: a new engineered approach to the cycling infrastructure design. In: Gervasi, O., et al. (eds.) ICCSA 2021. LNCS, vol. 12958, pp. 441–456. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87016-4_33
21. D'Apuzzo, M., et al.: Toward the development of a hybrid approach to speed estimation in urban and rural areas. Traffic Inj. Prev. **22**(6), 449–454 (2021)
22. D'Apuzzo, M., Evangelisti, A., Santilli, D., Nardoianni, S., Cappelli, G., Nicolosi, V.: Towards a new design methodology for vertical traffic calming devices. Sustainability **15**(18), 13381 (2023)
23. D'Apuzzo, M., Evangelisti, A., Cappelli, G., Nardoianni, S., Nicolosi, V.: Some remarks on automatic braking system related to pedestrian safety. In: Gervasi, O., et al. (eds.) Computational Science and Its Applications – ICCSA 2023 Workshops: Athens, Greece, July 3–6, 2023, Proceedings, Part VIII, pp. 194–209. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37126-4_14
24. D'Apuzzo, M., Evangelisti, A., Cappelli, G., Nicolosi, V.: Safety oriented road asset management methodology for urban areas. In: Gervasi, O., Murgante, B., Misra, S., Ana, M.A., Rocha, C., Garau, C. (eds.) Computational Science and Its Applications – ICCSA 2022 Workshops: Malaga, Spain, July 4–7, 2022, Proceedings, Part VI, pp. 650–662. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10592-0_47
25. D'Apuzzo, M., Cappelli, G., Evangelisti, A., Santilli, D.: A design method for pedestrian areas by a simplified approach to predict the future traffic scenarios. In: Gervasi, O., Murgante, B., Misra, S., Ana, M.A., Rocha, C., Garau, C. (eds.) Computational Science and Its Applications – ICCSA 2022 Workshops: Malaga, Spain, July 4–7, 2022, Proceedings, Part IV, pp. 360–373. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10542-5_25
26. D'Apuzzo, M., Cappelli, G., Nardoianni, S., Nicolosi, V., Evangelisti, A.: A preliminary effort to develop a framework of distance decay functions for new urban active mobility. In: Gervasi, O., et al. (eds.) Computational Science and Its Applications – ICCSA 2023 Workshops: Athens, Greece, July 3–6, 2023, Proceedings, Part VII, pp. 194–208. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_16
27. D'Apuzzo, M., Evangelisti, A., Cappelli, G., Nicolosi, V.: An introductory step to develop Distance Decay Functions in the Italian context to assess the modal split to e-bike and e-scooter. In: 2022 Second International Conference on Sustainable Mobility Applications, Renewables and Technology (SMART), pp. 1–8. IEEE (2022)
28. Cappelli, G., D'Apuzzo, M., Nardoianni, S., Nicolosi, V.: Exploring the influences of safety and energy expenditure parameters on cycling. Sustainability **16**(7), 2739 (2024)
29. UN. Sustainable Development Goals 11 (11.5)
30. Directive 2019/1936 Directive - 2019/1936 - EN - EUR-Lex. europa.eu. Accessed Jan 2024
31. Next Generatio Eu 2020 NextGenerationEU - European Union. europa.eu. Accessed Jan 2024
32. PNRR. The National Recovery and Resilience Plan (NRRP) - Ministry of Economy and Finance. mef.gov.it. Accessed Jan 2024
33. National Plan for Road Safety 2030 Piano Nazionale Sicurezza Stradale 2030 | Ministero delle infrastrutture e dei trasporti. mit.gov.it. Accessed Jan 2024

34. European Commission. EU Road Safety Policy Framework 2021–2030—Next Steps towards "Vision Zero" (2019)
35. Chen, S., Kuhn, M., Prettner, K., Bloom, D.E.: The global macroeconomic burden of road injuries: estimates and projections for 166 countries. The Lancet Planetary Health **3**(9), e390–e398 (2019)
36. Karkalakos, S., Tselekounis, M.: Road safety and sustainability performance: a cross-country analysis. E3S Web Conf. **436**, 11005 (2023). EDP Sciences
37. Klanjčić, M., Gauvin, L., Tizzoni, M., Szell, M.: Identifying urban features for vulnerable road user safety in Europe. EPJ Data Sci. **11**(1), 1–15 (2022)
38. Jacobsen, P.L.: Safety in numbers: more walkers and bicyclists, safer walking and bicycling. Inj. Prev. **21**(4), 271–275 (2015)
39. Gandupalli, S.R., Kokkeragadda, P., Dangeti, M.R.: Analysis and modelling of crash severity of vulnerable road users through discrete methods: a case study approach. Innov. Infrastruct. Solut. **8**(11), 298 (2023)
40. Scholliers, J.: Topical collection on the influence of intelligent transport systems on vulnerable road user accidents. Eur. Transp. Res. Rev. **10**, 60 (2018). https://doi.org/10.1186/s12544-018-0337-1
41. Titolo V - Codice della strada (D.lgs. 30 aprile 1992, n. 285)
42. ISTAT. Istat.it. Accessed Jan 2024
43. Montella, A., Aria, M., D'Ambrosio, A., Mauriello, F.: Data-mining techniques for exploratory analysis of pedestrian crashes. Transp. Res. Rec. **2237**(1), 107–116 (2011)
44. Rella Riccardi, M., Galante, F., Scarano, A., Montella, A.: Econometric and machine learning methods to identify pedestrian crash patterns. Sustainability **14**(22), 15471 (2022)
45. Rella Riccardi, M., Mauriello, F., Scarano, A., Montella, A.: Analysis of contributory factors of fatal pedestrian crashes by mixed logit model and association rules. Int. J. Inj. Contr. Saf. Promot. **30**(2), 195–209 (2023)
46. ISTAT, Methodological Note. https://www.istat.it/it/files/2016/11/nota-metodologica_incidenti-stradali_anno2015.pdf. Accessed Jan 2024
47. Bishop, C.M., Nasrabadi, N.M.: Pattern Recognition and Machine Learning, vol. 4, no. 4, p. 738. Springer, New York (2006)
48. Russell, S.J., Norvig, P.: Artificial Intelligence a Modern Approach. London (2010)
49. Tharwat, A.: Linear vs. quadratic discriminant analysis classifier: a tutorial. Int. J. Appl. Pattern Recogn. **3**(2), 145–180 (2016).
50. Mika, S., Ratsch, G., Weston, J., Scholkopf, B., Mullers, K.R.: Fisher discriminant analysis with kernels. In: Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (cat. no. 98th8468), pp. 41–48. IEEE (1999)
51. Fix, E., Hodges, J.: Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties. Technical Report 4, US AF School of Av iation Medicine, Randolph Field, Texas (1951)
52. Ho, T.K.: Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, pp. 278–282, vol. 1 (1995)
53. Elyassami, S., Hamid, Y., Habuza, T.: Road crashes analysis and prediction using gradient boosted and random forest trees. In 2020 6th IEEE Congress on Information Science and Technology (CiSt), pp. 520–525. IEEE (2021)
54. Cesarini, L., Figueiredo, R., Monteleone, B., Martina, M.L.: The potential of machine learning for weather index insurance. Nat. Hazard. **21**(8), 2379–2405 (2021)