

Article

An Advanced Machine Learning Approach to Predicting Pedestrian Fatality Caused by Road Crashes: A Step toward Sustainable Pedestrian Safety

Wenlong Tao ^{1,*}, Mahdi Aghaabbasi ^{2,*} , Mujahid Ali ^{3,*} , Abdulrazak H. Almaliki ⁴ , Rosilawati Zainol ² , Abdulrhman A. Almaliki ⁵  and Enas E. Hussein ^{6,*} 

¹ Department of Automotive Technology, Zhejiang Agricultural Business College, Shaoxing 312000, China

² Centre for Sustainable Urban Planning and Real Estate (SUPRE), Department of Urban and Regional Planning, Faculty of Built Environment, University of Malaya, Kuala Lumpur 50603, Wilayah Persekutuan Kuala Lumpur, Malaysia; rosilawatizai@um.edu.my

³ Department of Civil and Environmental Engineering, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Perak, Malaysia

⁴ Civil Engineering Department, College of Engineering, Taif University, Taif 21944, Saudi Arabia; a.almaliki@tu.edu.sa

⁵ Independent Researcher, Jeddah 12462, Saudi Arabia; a.almaliki1222@gmail.com

⁶ National Water Research Center, Shubra El-Kheima 13411, Egypt

* Correspondence: twl87910@163.com (W.T.); mahdi@um.edu.my (M.A.); mujahid_19001704@utp.edu.my (M.A.); enas_el-sayed@nwrc.gov.eg (E.E.H.)



Citation: Tao, W.; Aghaabbasi, M.; Ali, M.; Almaliki, A.H.; Zainol, R.; Almaliki, A.A.; Hussein, E.E. An Advanced Machine Learning Approach to Predicting Pedestrian Fatality Caused by Road Crashes: A Step toward Sustainable Pedestrian Safety. *Sustainability* **2022**, *14*, 2436. <https://doi.org/10.3390/su14042436>

Academic Editor: Marco Guerrieri

Received: 27 December 2021

Accepted: 16 February 2022

Published: 20 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: More than 8000 pedestrians were killed due to road crashes in Australia over the last 30 years. Pedestrians are assumed to be the most vulnerable users of roads. This susceptibility of pedestrians to road crashes conflicts with sustainable transportation objectives. It is critical to know the causes of pedestrian injuries in order to enhance the safety of these vulnerable road users. To achieve this, traditional statistical models are used frequently. However, they have been criticized for their inflexibility in handling outliers and missing or noisy data, and their strict pre-assumptions. This study applied an advanced machine learning algorithm, a Bayesian neural network, which has the characters of both Bayesian theory and neural networks. Several structures of this model were built, and the best structure was selected, which included three hidden neuron layers—sixteen hidden nodes in the first layer and eight hidden nodes in the second and third layers. The performance of this model was compared with the performances of some other machine learning techniques, including standard Bayesian networks, a standard neural network, and a random forest model. The Bayesian neural network model outperformed the other models. In addition, a study on the importance of the features showed that the individuals' characteristics, time, and circumstantial factors were essential. They greatly increased model performance if the model used them. This research lays the groundwork for using machine learning approaches to alleviate pedestrian deaths caused by road accidents.

Keywords: pedestrian fatality; road accident; Bayesian neural network; Bayesian theorem; sustainable road network development; machine learning

1. Introduction

Pedestrians are the most susceptible road users. Pedestrians also are an important component of the sustainable development of road networks. However, their vulnerability to road crashes conflicts with sustainable transportation objectives. Pedestrian deaths and injuries in road crashes have major socio-economic consequences. This is particularly important in view of developed countries' ongoing efforts to enhance road safety. Since practically anyone can be a pedestrian, pedestrians make up the biggest single road user category. People walk for a variety of reasons, including recreation; traveling to work,

study, or small retailers; and linking up with other means of transportation. In the National Road Safety Strategy, pedestrians are designated as a susceptible road user category. When compared to other road users, they have very limited defence in collisions [1]. Over 50,000 people have died on Australian roads in the last 30 years. Pedestrians accounted for 15.6% of all road accident deaths, even though pedestrians cover fewer miles than other road users. [2]. However, the pedestrian death toll has decreased by almost 57% over the past 30 years. Pedestrians account for a significant portion of fatalities in Australian collisions involving large vehicles and buses. Pedestrians, for example, account for around 30% of those killed in bus collisions [3]. Pedestrians, motorcyclists, and bicyclists make up around a quarter of all deaths in truck crashes [3].

Despite the decrease in pedestrian deaths due to road crashes in Australia, scholars have continued to look for opportunities to acquire a deeper understanding of the factors that impact crash probability in the hopes of effectively estimating the probability of pedestrian-involved crashes and guiding policy initiatives and prevention methods to decrease the incidence of pedestrian-involving crashes [4–7].

There have been several significant data flaws in the literature on pedestrian-related crashes. These problems could lead to erroneous pedestrian crash forecasts and inaccurate conclusions about the factors that cause crashes if analytical models are poorly specified. Imprecision in crash locations and time, challenges in data linkages (for instance, with traffic data) because of database discrepancies, intensity misclassification, errors and incompleteness of affected users' demographics, and wrong identification of accident contributory determinants are just a few of these issues [8]. Furthermore, it is challenging to identify and assess factors influencing pedestrian crash deaths because of the heterogeneity intrinsic in pedestrian crash data, which results from unobservable characteristics that are not recorded by police and cannot be collected from crash reports. As a result of this heterogeneity, parameter estimation may be skewed, leading to possibly inaccurate findings [9–11].

To study the crash data, traditional broadly utilised discrete choice modelling approaches, including mixed logit models, multinomial logit models, ordered logit/probit models, and partial proportional odds logit models have been utilised. Most of the solutions mentioned above, however, rely heavily on pre-existing assumptions. Machine learning (ML) techniques have more flexibility than traditional statistical models in that they can analyse noisy data, outliers, and missing data, without or with minimal previous assumptions about inputs [12–18]. In addition, ML methods are notable instances of data-driven techniques that strive to improve the efficiency and precision of accident data processing and forecasting. Early research employed multiple ML methods, including support vector machines, decision trees, artificial neural network, and ensemble learning, to forecast the severity and frequency of pedestrian-involved crashes, and their findings show that these techniques are very flexible and can outperform conventional methods. Hence, this study selected the ML-based approach in a Bayesian neural network (BNN) to analyse data associated with pedestrian deaths due to road crashes (PDRC).

Due to advancements in computer methods, Bayesian computing approaches are becoming more prominent. Bayesian models offer the privilege of dealing with extremely complicated models, particularly those with difficult-to-calculate probability functions. On the other hand, standard NN models have been criticized for their inability to fit training data accurately, and they may generate forecasted results with undesirable variances [19–22]. Overfitting is among the main causes of this issue. Even if the standard NN model has stronger linear and nonlinear estimation capabilities than traditional statistical approaches, this technique, being vulnerable to the overfitting issue, has poor generalization, which restricts its utility for crash severity and frequency forecasts [19]. In various domains, several earlier studies have shown that applying the Bayesian algorithm in NN models can significantly lessen overfitting while maintaining the NN's excellent nonlinear approximation ability (e.g., [23–25]). However, the combination mentioned above has rarely been used in the domain of crash prediction (e.g., [19]), especially for predicting PDRC.

The main aim of this study was to see how effective the BNN model is for the prediction of PDRC. Furthermore, our study contributes to the area of pedestrian road crash fatality modelling in the following ways: (1) building a combination of architectures to assess the model's performance; (2) evaluating a variety of characteristics that might help with pedestrian fatality classification and forecasting; (3) evaluating BNN in comparison to other machine learning models.

Utilizing data obtained on road transport crash fatalities in Australia, different BNN structures were evaluated to achieve the study's goal. The authors estimated 16 BNN structures and compared their performances utilizing several performance criteria. The authors compared the performance of the best model with other ML models. In addition, the influences of predictors were evaluated using a different approach in which various types of factors were combined to determine the best variable set for predicting PDRC.

The rest of this paper is layed out as follows. A literature review on the methods used for analysing pedestrian crash data is presented in Section 2. Moreover, the necessity of using advanced ML techniques for analysing pedestrian crash data is highlighted in this section. Section 3 presents brief explanations of the methods, performance criteria, and dataset. The feature selection, model development process, selection of the best BNN structure, comparison of the selected BNN model with other ML models, significance of the influential factors, and study limitations are presented in Section 4. The last section presents a summary of the paper and some recommendations for future studies.

2. Literature Review

Traditional statistical methods have been employed in the majority of pedestrian-involved crash forecast studies. These models included the ordered probit model [26–30], binary logit model [31], and multinomial logit model (MNL) [29,32–36]. MNL was widely used to study pedestrian crashes; nevertheless, it was criticised since it relies on the assumption that independent variables have the same impacts across instances, which could be contradicted if there are unobserved data heterogeneities. This is a concern because of the incompleteness of the data on road crashes, which means that the impacts may change in different circumstances. Therefore, the mixed logit model was utilised to circumvent the restriction imposed by the independence of irrelevant alternatives (IIA) property by randomly distributing the parameters among individual observations [32,36–39]. Along with the mixed logit model that overcomes the drawbacks of MNL, other models, including partial proportional odds (PPO), also were applied to examine the pedestrian-involved crashes [40–44]. The PPO allows some of the parameter estimates to have different effects on a dependent variable, which is suitable for modelling the pedestrian crash injury severities.

Traditional statistical methods for predicting pedestrian-related crashes are widely used; however, they may become out-of-date if efficacy and accuracy are taken into account. Furthermore, the majority of traditional approaches are regression-based, which include drawbacks such as assuming linear or nonlinear correlations between exploratory factors and the target variable. When such requirements are not satisfied, the models may inadvertently lead to incorrect conclusions [45]. Abreast with the fast evolution of ML techniques and the growing amount of data available, it is becoming increasingly popular to use ML to solve transportation-related issues. In comparison to traditional statistical methods, ML techniques, as non-parametric approaches, have fewer restrictions on pre-existing assumptions regarding the correlations between road accident fatality outcomes and major contributors [46].

Neural networks (NN), random forest, support vector machines (SVMs), decision trees (DTs), and gradient boosting (GB) are among the most frequently used ML techniques for crash data analysis. A list of some studies that have employed the ML techniques for analysing pedestrian crash data is provided in Table 1. It should be noted identifying contributing elements in road crashes is basically a multiclass or binary class problem. Among all ML techniques utilized for the pedestrian crash data, DT-based models, including classification and regression trees (CART), XGBoost, and random forest (RF), were the

most frequently used techniques. Instead, powerful models such as NN were rarely used for analysis of pedestrian-related crash data (e.g., [47]). The standard NN models have been criticized for their vulnerability to overfitting and poor generalization [19]. Consequently, some solutions have been proposed, including combining the Bayesian inference method with the NN algorithm [24,48–50]. This combination allows the neural network to choose hidden neurons and input variables with greater freedom. The BNN model has been successfully applied in many fields of research. However, to date, a very limited number of studies in traffic safety have adopted this model (e.g., [19,51]). These studies mostly attempted to predict motor vehicle collisions and estimate the energy equivalent speed. Given the benefits of NN models over traditional statistical models, along with the gains made by consolidating Bayesian inference into NN, it is worth looking into whether the BNN model can be utilized to model PDRC effectively and whether it outperforms other ML techniques.

Table 1. Some studies on the prediction of crashes related to pedestrians using ML techniques.

Study	Study Aim	ML Technique Employed
Pour et al. [52]	To determine the impact of temporal, geographical, and personal variables on the severity of vehicle-pedestrian collisions.	DT, KDE
Ding et al. [53]	To provide a different perspective on the effects of pedestrian collisions.	MAPRT
Mokhtarimousavi [54]	To predict the severity of injuries in pedestrian collisions.	SVM, MNL
Das et al. [55]	To create a framework for classifying crash kinds from unstructured textual input using ML algorithms.	RF, SVM, XGBoost
Rahimi et al. [56]	To identify death patterns in heavy truck-related pedestrian/bike collisions.	RF, DT
Guo et al. [57]	To simulate the issue of categorizing three levels of severity in older pedestrian traffic crashes.	XGBoost
Saha and Dumbaugh [58]	To assess the characteristics of the relationships between built environment variables and pedestrian crash frequency at the census block group level.	GB, DT, GAM
Zhu [47]	To look into the elements that contribute to the intensity of vehicle-pedestrian collisions at crossings.	CART, GB, RF, ANN, SVM

Support vector machines = SVM; artificial neural network = ANN; random forest = RF; decision tree = DT; classification and regression trees = CART; kernel density estimation = KDE; multiple additive Poisson regression trees = MAPRT; multinomial logit model = MNL; extreme gradient boosting = XGBoost; generalized additive model = GAM; gradient boosting = GB.

3. Methodology

This study primarily aimed at predicting and classifying PDRC using a dataset from Australia. This study employed the BNN algorithm to achieve the objective mentioned above. The flowchart of the investigation is shown in Figure 1. The following sections provide more in-depth descriptions of the stages.

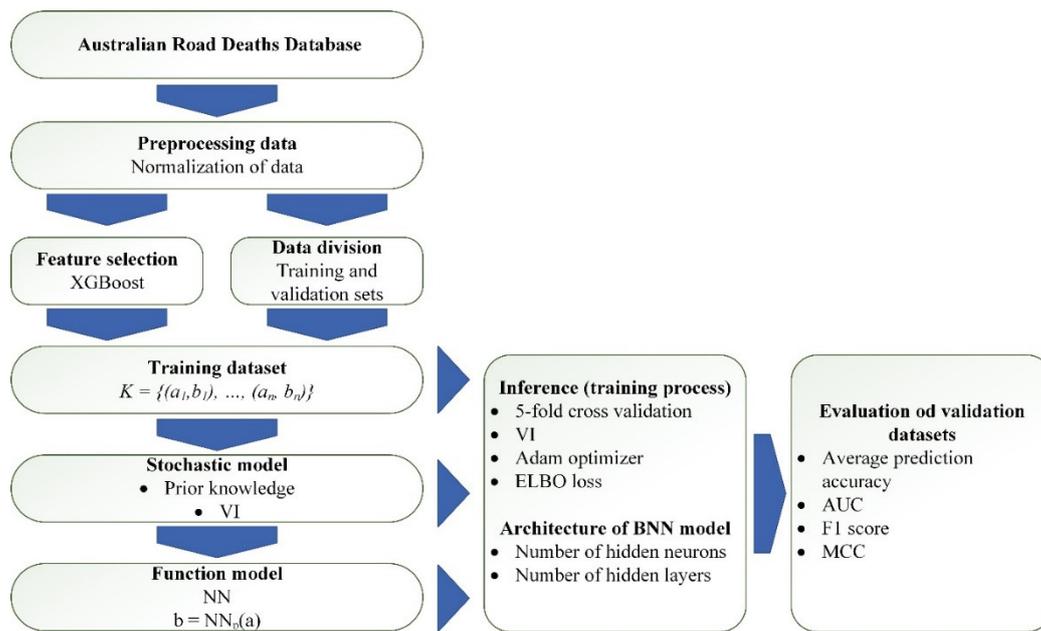


Figure 1. A flowchart of this study.

3.1. A Basic Understanding of the Bayesian Neural Network and Bayesian Inference

This study utilized a Bayesian method to forecast and classify PDRC. Employing Bayes' theory, Bayesian models attempt to derive and determine characteristics regarding a likelihood distribution from collected data (Equation (1)).

$$P(\alpha|K) = \frac{P(K|\alpha)}{P(K)} \quad (1)$$

where α is a collection of uncalibrated model parameters, which must be calibrated with dataset K . Posterior distribution on α is indicated by $P(\alpha)$, and it reflects our understanding of how data are produced prior to observing them. The posterior distribution, abbreviated as $P(\alpha|K)$, represents the uncertainty levels of attribute values that accurately describe observed data. The probability function $P(K|\alpha)$ denotes how likely distinct values of α are to produce the observed dataset K . $P(K)$ uses a proper probability density to normalize the posterior distribution.

The use of Bayesian inference in NN has gotten a great deal of interest. This study focuses on expanding the BNN's usage for forecasting and classifying PDRC. A BNN is a NN that has been trained to fit measured values utilizing Bayesian inference, with the assumption that the network's parameters are arbitrary based on a prior probability distribution [49]. In the training stage, various sorts of NN use different approaches to learn from the data and adjust network weights [59]. The weights of a standard NN are regarded deterministic, and then when the model is trained, a single data point approximation is achieved. Contrastingly, instead of assuming a singular point estimation following training, the BNN's weights are expressed as likelihood distributions across feasible data points. The variance of the weights' network distribution reveals the BNN's performance uncertainty. The distinction between a BNN and a deterministic NN is shown in Figure 2.

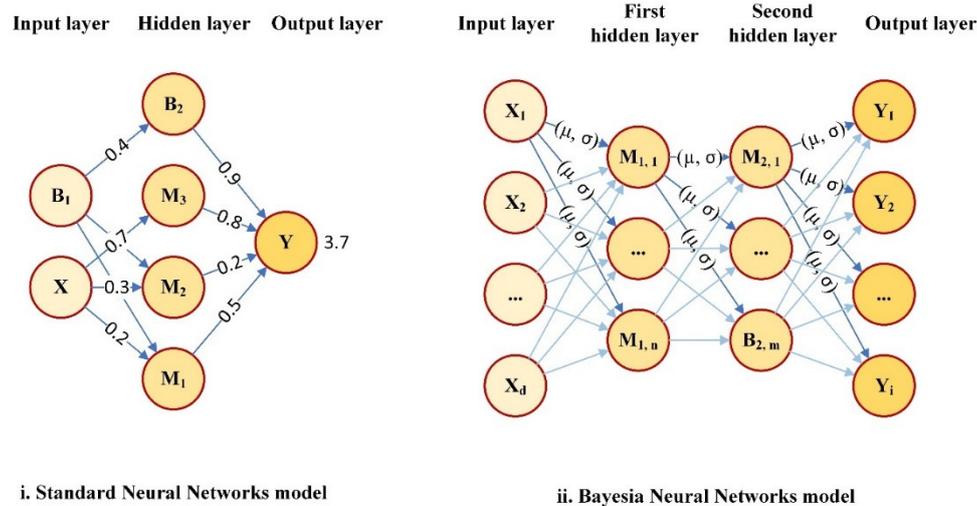


Figure 2. Typical structures of NN and BNN.

3.2. Bayesian Neural Network

The authors employed a BNN in this research to perform a binary classification between the two tasks—0 = non-pedestrian death and 1 = pedestrian death—while considering data uncertainty. The authors utilize variational inference (VI) to train the BNN, an optimization algorithm for approximating likelihood densities. VI is different from other traditional approaches, such as Markov chain and Monte Carlo, as it determines the parameters of these distributions rather than the weights directly.

The BNN used in this study can be regarded as a probabilistic model $P(b|a, \gamma)$. Here, b is a collection of our categories— $b = 0$ or 1 ; a is a collection of attributes; γ is the weight parameter; $P(b|a, \gamma)$ is a categorical probability. The likelihood function (LF) that is a function of the parameter Y could be generated using the training dataset K . The following is the LF:

$$P(K|\gamma) = \prod P(b|a, \gamma) \quad (2)$$

The maximum likelihood estimate (MLE) of γ can be obtained via maximisation of the LF, with the objective function being negative log-likelihood. Based on the Bayes theory, the posterior distribution is proportionate to the outcome of the prior distribution, $P(\gamma)$ and the probability $P(K|\gamma)$. MLE, on the other hand, uses point calculations for parameters; therefore, the uncertainty in the weights is not represented. As a result, a BNN averages forecasts from a number of NN that are weighted according to the posterior distribution of the γ . The following is the mathematical equation for the posterior predictive distribution:

$$P(b|a, K) = \int P(b|a, \gamma)P(\gamma|K)d\gamma \quad (3)$$

A BNN can employ a variational distribution $S(\gamma|\vartheta)$ of established functional form to estimate the correct posterior distribution because determining the posterior distribution, $P(\gamma|K)$, is complicated. To accomplish this, the Kullback–Leibler (KL) divergence between the correct posterior $P(\gamma|K)$ and $S(\gamma|\vartheta)$ concerning ϑ is reduced [60]. The following is the relevant objective function:

$$KL(S(\gamma|\vartheta)||P(\gamma|K)) = E[\log S(\gamma|\vartheta)] - E[\log P(\gamma)] - E[\log P(K|\gamma)] + \log P(K) \quad (4)$$

Since the KL cannot be determined, this study employs the evidence lower bound (ELBO) that does not comprise the component $\log P(K)$ and is the inverse of the KL divergence function. Since $\log p(K)$ is a constant, it may be ignored, making maximization of the ELBO function equal to minimization of the KL divergence. The adaptive moment

estimation (Adam) optimizer is employed to calibrate the variational parameters γ , which can be modified adaptively. The ELBO function's mathematical form is given below.

$$ELBO(S) = E[\log P(\gamma)] + E[\log P(K|\gamma)] - E[\log S(K|\gamma)] \quad (5)$$

3.3. Evaluation of Various Models' Performances

This work used the k-fold cross-validation method to arbitrarily divide a whole dataset into five distinct subdivisions with nearly equivalent numbers of data points to avoid biases and overfitting throughout model training. The performances of BNN models in classifying and forecasting pedestrian fatalities due to traffic crashes were assessed using the set of criteria:

- Average training accuracy (ATA): Prediction accuracy in this study's binary class case is defined as the total number of correct forecasts over two classes divided by the total number of forecasts.
- Average F1-score: In the binary-class forecasting study, the average F1-score was employed to approximate criteria for each classification, and the average was calculated by estimating the number of correctly predicted occurrences.
- The area under the ROC curve (AUC): the area under the receiver operating characteristic curve (AUC) was utilised to estimate a scoring classifier at multiple cutoffs in this investigation. The AUC measures a model's ability to distinguish between positive and negative classifications.
- Matthew's correlation coefficient (MCC): The MCC was employed to assess the quality of binary classifications in this investigation. The MCC is a balanced measure that can be utilized even if the categories are of significantly distinct sizes, since it considers true and inaccurate positives and negatives. This criterion is a correlation coefficient that produces a number between -1 and $+1$ for actual and forecasted binary classes.

This study also used some other common criteria to assess the performance of various BNN architectures. However, the final evaluations and comparisons were based on the four metrics mentioned above. These additional criteria included false discovery rate, false negative rate, false positive rate, negative predictive value, precision, sensitivity, and specificity.

3.4. Dataset

The Australian Road Deaths Database (ARDD) provided the data for this research [2]. This database contains information on deaths in road transport crashes in Australia, as provided by the police to state/local road safety bodies monthly. The ARDD collects demographic and crashes information for individuals who died in car accidents in Australia. A road death, often known as a fatality, occurs when an individual dies because of injuries sustained in a car accident within 30 days of the accident. In this dataset, a pedestrian crash is defined as any collision in which a pedestrian is killed, regardless of the number of cars involved. The ARDD includes 24 columns/variables, and 13 of these variables are suitable for predicting pedestrian crashes. It is worth noting that the data utilized in this study were the most up to date, having been collected between 1989 and 2021. This dataset has a sample size of 52,843, and it was used in its entirety to forecast pedestrian fatalities. Table 2 provides a summary of the variables used in this research. This dataset includes basic information about the PDRC. These variables allowed us to achieve the objective of this study, which was applying the combination of Bayesian theory and neural network to pedestrian crash data. Future studies can extend this study by employing datasets with a higher number of variables.

Table 2. Summary of the variables employed in this present research.

Variable and Sub-variable	Description	Value	Mean/Mode
<i>Individual characteristics (IC)</i>			
Age	Age of the individual who was killed (years)	1–101	39.662
Gender	Person's sex	Female, male	Male
<i>Time and occasions (TO)</i>			
Month	Month of crash	1–12	12
Day of week	Specifies whether the crash happened on a weekday or on a weekend.	Weekend; weekday	Weekday
Time of day	Specifies whether the crash happened during the day or at night.	Day, Night	Day
Christmas Period	Specifies whether the crash happened in the 12 days starting on 23 December.	Yes, no	No
Easter Period	Specifies whether the crash happened within the five days leading up to Good Friday.	Yes, no	No
<i>Road characteristics (RC)</i>			
Speed limit	The designated speed limit at the location of the crash.	10–130 km	82.10
National Road Type		Access Road, Arterial Road, Collector Road, Local Road, National or State Highway, Pedestrian Thoroughfare, Sub-arterial Road	National or State Highway
<i>Crash attributes (CA)</i>			
Crash Type	If a pedestrian was died in a collision, it is marked as a pedestrian crash; else, the vehicles engaged is recorded.	Multiple, single	Single
Bus involvement	Shows that a bus was involved in the accident.	Yes, no	No
Heavy Rigid Truck Involvement	Shows that a heavy rigid truck was involved in the collision.	Yes, no	No
Road User (target variable)	Road user type of killed person.	Non-pedestrian, pedestrian	Non-pedestrian

It is worth mentioning that input variables were normalized and transformed as follows:

- The order of nominal variables was rearranged, with the smallest category appearing first and the largest category appearing last.
- In continuous variables, missing values were substituted with the mean.
- The mode was used to substitute missing values in nominal variables.
- The median was utilized to substitute missing values in ordinal variables.
- The target variable (road user) was initially nominal, and its values included driver, motorcycle pillion passenger, motorcycle rider, passenger, pedal cyclist, pedestrian. The road user was transformed into a binary variable. This new variable included two classes: non-pedestrian death and pedestrian death.

4. Results and Discussions

4.1. Determination of Significant Variables

This study applied the advanced XGBoost technique to refine irrelevant inputs for a Bayesian-inferred pedestrian death model. It has been proven that the XGBoost method is superior to other non-linear classification methods; however, few studies have applied this technique for feature selection in pedestrian crash prediction and classification (e.g., [57,61]). XGBoost adopts the F-score to determine the significance score (weight) of each variable. A greater F-score is assigned to a variable that embodies more information for classification.

The F-score is calculation using the number of occasions an input is employed for dividing, weighted through the squared enhancement of the model as a consequence of every division, and averaged over all probabilities [62]. This criterion is capable of treating both categorical and continuous inputs fairly to evaluate and rank the inputs. The authors applied the XGBoost technique on 12 variables. Figure 3 illustrates the input rank outcomes organised by their influence. This algorithm selected the ten most important inputs, including speed limit, crash type, age, time of day, bus involvement, gender, day of the week, month, Christmas period, and national road type.

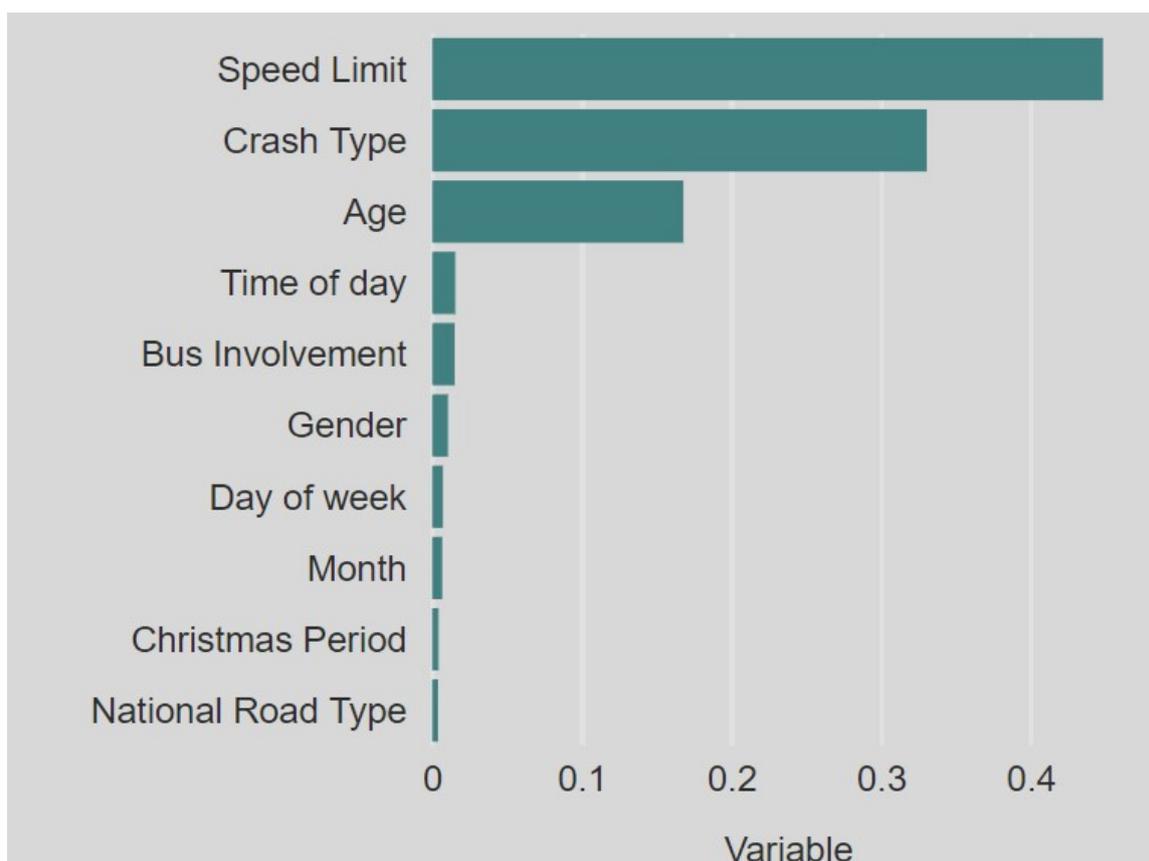


Figure 3. Importance of variables analysed by XGBoost technique.

4.2. Development and Performance Assessment of the BNN Model

Creating a proper neural network structure is reliant on problems and data. Initially, the authors used a rectified linear unit (ReLU) as the activation function between the consecutive hidden layers to induce non-linearity in the neuron's output. To calculate the error gradient, a batch size of 64 samples from the training dataset was employed. In order to detect the error gradient of the model optimization during the learning stage, various learning rates (LRs) for the Adam optimizer operation were evaluated (10-3, 10-2, 10-1). Then, ELBO loss was observed on validation and training sets. In the prediction of PDRC, Figure 4 shows in what way LRs affected model convergence utilising a BNN model with a single hidden layer (hidden units = 16). Figure 4a illustrates a desirable match, as the validation and training losses rapidly climb to the established position, with little divergence between the two ultimate loss rates. Figure 4b,c shows noisy fluctuations around the training and validation loss, with every iteration moving ahead at an excessively large step size thanks to the high LR. The authors tuned the BNN model utilizing the Adam optimizer's LR of 0.001 to determine the best number of hidden layers and neurons.

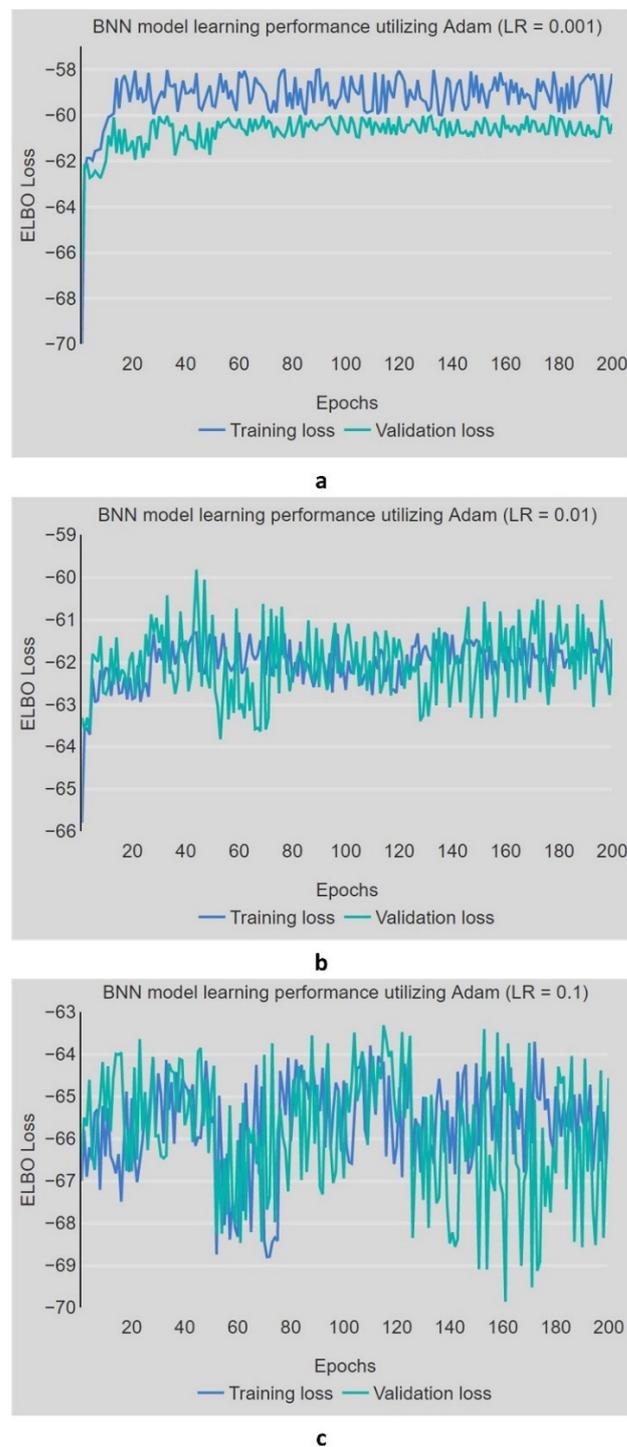


Figure 4. Convergence of the BNN model with varying LR. (a): LR = 0.001, (b): LR = 0.01, (c): LR = 0.1.

Various structures of BNN were trained 200 times. Table 3 presents the Bayesian-inferred PDRC model's forecasting performance. The authors evaluated the forecasting performances of several model structures with eleven performance criteria.

Table 3. Performances of several BNN designs.

NS	HL	ATA	Sensitivity	Specificity	Precision	NPV	FPR	FDR	FNR	F1 Score	AUC	MCC
NS ₁	1	0.802	0.940	0.421	0.817	0.721	0.578	0.182	0.059	0.874	0.808	0.442
NS ₂	1	0.834	0.844	0.171	0.986	0.015	0.828	0.013	0.155	0.909	0.748	0.005
NS ₃	1	0.879	0.890	0.747	0.974	0.345	0.252	0.021	0.110	0.932	0.776	0.454
NS ₄	1	0.887	0.900	0.752	0.974	0.413	0.247	0.025	0.100	0.939	0.780	0.503
NS ₅	2	0.607	0.967	0.271	0.553	0.900	0.728	0.446	0.032	0.704	0.811	0.329
NS ₆	2	0.570	0.972	0.256	0.504	0.922	0.743	0.4951	0.027	0.664	0.811	0.312
NS ₇	2	0.670	0.9605	0.303	0.636	0.858	0.696	0.363	0.039	0.765	0.804	0.361
NS ₈	2	0.767	0.947	0.379	0.766	0.771	0.620	0.233	0.052	0.847	0.796	0.419
NS ₉	2	0.784	0.945	0.399	0.791	0.750	0.600	0.208	0.055	0.861	0.822	0.432
NS ₁₀	3	0.867	0.879	0.687	0.976	0.274	0.313	0.023	0.120	0.925	0.824	0.377
NS₁₁	3	0.894	0.906	0.774	0.975	0.453	0.225	0.0245	0.093	0.939	0.844	0.540
NS ₁₂	3	0.792	0.928	0.399	0.816	0.661	0.600	0.183	0.071	0.868	0.788	0.396
NS ₁₃	3	0.781	0.944	0.394	0.786	0.750	0.605	0.213	0.055	0.858	0.734	0.426
NS ₁₄	3	0.717	0.936	0.334	0.711	0.750	0.665	0.289	0.063	0.808	0.719	0.353
NS ₁₅	3	0.707	0.949	0.334	0.687	0.811	0.665	0.312	0.050	0.797	0.717	0.376
NS ₁₆	3	0.668	0.946	0.304	0.640	0.811	0.696	0.359	0.053	0.764	0.719	0.336

NS = network structure; NS1 = 16; NS2 = 32; NS3 = 64; NS4 = 128; NS5 = (16, 8); NS6 = (16, 16); NS7 = (32, 8); NS8 = (32, 16); NS9 = (32, 32); NS10 = (8, 8, 8); NS11 = (16, 8, 8); NS12 = (16, 16, 8); NS13 = (32, 8, 8); NS14 = (32, 16, 8); NS15 = (32, 32, 16); NS16 = (64, 32, 16). HL = hidden layers; AAT = average training accuracy; NPV = negative predictive value; FPR = false positive rate; FDR = false discovery rate; FNR = false negative rate; AUC = area under curve; MCC = Matthews's correlation coefficient.

Concerning ATA, the BNN with three hidden neuron layers (NS11) had the best results (ATA = 0.894). The second best ATA belonged to a BNN architecture including a hidden layer of 128 elements (namely, NS4). NS5 and NS6 were the two poorest network architectures. Regarding AUC, F1 score, and MCC, NS11 also outperformed other BNN structures, which indicated the model's success in classifying PDRC.

The BNN design with three hidden layers (NS11) performed reasonably well, with sixteen hidden neurons in the first layer and eight hidden neurons each in the second and third layers. As a result, this research focuses on this BNN model in the subsequent sections to see how the model's classification uncertainties affect the forecasts of PDRC.

4.3. Quantification of Ambiguity in the Forecast and Classifying Probability

A Sankey plot was built to depict the relationship between actual and forecasted labels to understand the classification errors of the BNN model (Figure 5). The actual classes are represented by the left nodes on the Sankey plot, whereas the forecasted classes are displayed by the right-hand points. The thicknesses of the color connections and streams are proportional to the amounts of data. As seen in Figure 5, non-pedestrian deaths (class 0) were mainly predicted to be non-pedestrian deaths, with only a few being misclassified as pedestrian deaths (class 1). However, more than half of pedestrian deaths (54.6%) were incorrectly predicted as non-pedestrian deaths. The proposed BNN's classification of the "non-pedestrian death" class is superior to that of the "pedestrian death" class with forecast rates of 97.5% and 45.37% accuracy, according to the comparison of forecast performance across each category.

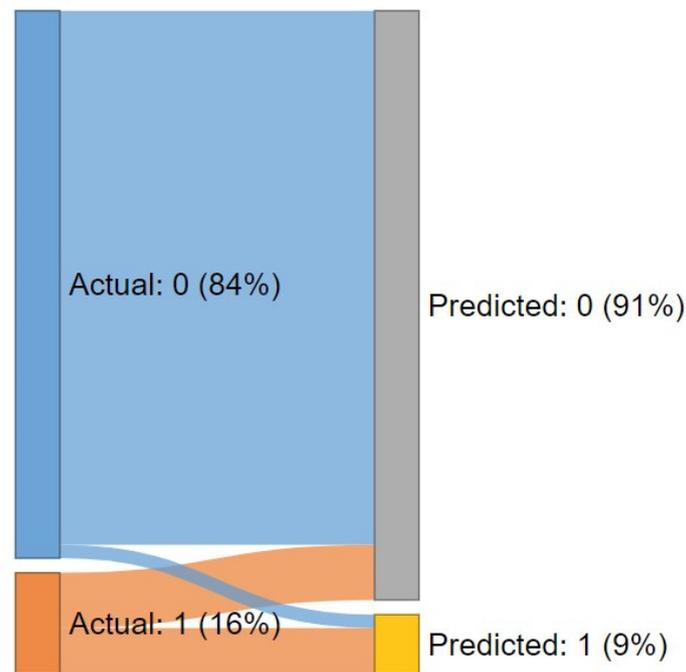


Figure 5. Correct and wrong forecasts classified by the BNN model.

The Bayesian technique has two notable features: (1) it yields predictive class probabilities rather than deterministic class label forecasts, and (2) it produces the standard deviation of the posterior prediction to indicate the level of uncertainty. The findings are shown as a raincloud graphic, which mixes a data distribution depiction and box plots overlaid on jittered raw data. For two death categories, Figure 6 depicts the range of the predictive probabilities and the forecast uncertainty. As can be seen in thick regions, the probability values for both classes are predominantly concentrated in the great probability zones that are in the range of 0.8–1.0. Both classes' prediction uncertainties are highly aggregated in the range of 0.0 and 0.1, indicating a low level of ambiguity. Overall, the BNN had a great level of confidence in classifying both death classes.

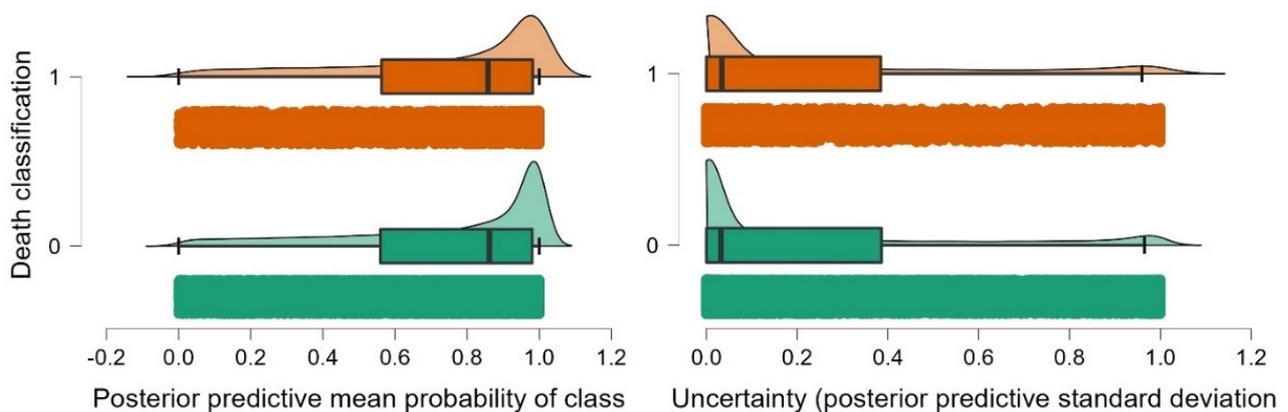


Figure 6. All class labels' posterior predictive mean probabilities and uncertainties.

4.4. Variable Significance

When performing field research, knowing the impacts of variables on a model's predictive ability can lower the cost of gathering data on PDRC. Assessing the significance of all specified traits and their conceivable combinations, on the other hand, is time-intensive and computationally costly. In this investigation, ten XGBoost-selected variables were categorised according to the kinds to which they related. This study built eleven combinations

in which different types of factors were combined to identify the best variable combination. Simultaneously, the model's performance was analysed in order to determine the smallest number of variables that must be collected while maintaining reliable prediction performance. Table 4 presents the outcomes of the models' executions. The outcomes of this analysis showed that ARR8 (TO + RC) was the weakest combination. In contrast, ARR7 (IC + TO + RC + CA) was the best combination, followed by ARR6 (IC + RC + CA). These findings imply that the combination of factors related to the time, occasions, and road characteristics is not able to predict the PDRC accurately alone. The predictions based on these two types of data should be improved using other factors, such as individual characteristics and crash attributes. The findings of this study are in line with those of Onieva-García et al. [63], Toran Pour et al. [64], Park and Ko [65], Li and Fan [44], and Kim et al. [66], who confirmed the significant roles of age and gender in pedestrian-related crashes and deaths. Several studies also confirmed the effects of bus involvement on the risks of injury and death of pedestrians (e.g., [67–69]), which indicates the significant role of crash attributes in the prediction of PDRC. Overall, when personal characteristics and crash features are factored in, this model appears to be successful and accurate.

Table 4. The performance of the BNN model with various variable arrangements.

Arrangement	Combination of Variables	ATA	F1 Score	AUC	MMC
ARR ₁	IC + TO	0.845	0.915	0.591	0.1375
ARR ₂	IC + RC	0.855	0.918	0.698	0.324
ARR ₃	IC + CA	0.858	0.920	0.761	0.317
ARR ₄	IC + TO + RC	0.860	0.921	0.754	0.3527
ARR ₅	IC + TO + CA	0.863	0.923	0.762	0.361
ARR ₆	IC + RC + CA	0.890	0.937	0.788	0.526
ARR ₇	IC + TO + RC + CA	0.894	0.939	0.844	0.540
ARR ₈	TO + RC	0.844	0.915	0.500	0
ARR ₉	TO + CA	0.846	0.916	0.695	0.113
ARR ₁₀	TO + RC + CA	0.866	0.924	0.774	0.387
ARR ₁₁	RC + CA	0.847	0.916	0.762	0.127

ARR = arrangement; AAT = average training accuracy; AUC = area under curve; MCC = Matthews's correlation coefficient; IC = individuals' characteristics; TO = time and occasions; RC = road characteristics; CA = crash attributes.

4.5. Comparison of BNN Modes with Other ML Models

The authors of this study compared the BNN model with various ML models, including a random forest (RF), a standard Bayesian network (BN), and a standard neural network (NN). This comparison helps with determining which machine learning algorithm has the highest prediction accuracy. It will help to reduce future work spent on selecting acceptable methods for PDRC data analysis. The advantages of using BNN for PDRC prediction are further confirmed by this comparison. The outcomes of this comparison are presented in Table 5. This comparison shows that the BNN model outperformed the other models, especially the standard NN model. Additionally, the standard BN model showed a poor prediction performance compared with the other models. The RF model showed a desirable performance that can be rooted in its capabilities for ensembling weak learners [70]. This study's Bayesian-inferred pedestrian fatality model performs well in prediction and classification based on the presented results.

Table 5. Comparisons of the BNN model with other ML models used to predict the PDRC.

Model	ATA	F1 Score	AUC	MMC
BNN	0.89	0.94	0.84	0.54
RF	0.87	0.92	0.79	0.60
BN	0.85	0.91	0.80	0.36
NN	0.84	0.91	0.53	0.07

4.6. Limitations and Future Enhancements

The Australian Road Deaths Database (ARDD) was employed to create and test Bayesian inference with NN for forecasting and classifying road-related pedestrian deaths. However, there are a few drawbacks to be aware of, and potential enhancements for the future. Even though the Bayesian-inferred pedestrian fatality model outperformed traditional ML models, BNN, like many other ML techniques, is a data-driven modelling approach, and the ARDD contains little variety and skewed distributions. This suggests that in certain severe circumstances, the model would be unstable. Future investigations are required to improve this model by consolidating a more varying set of environmental factors, built environment factors, and road characteristics (e.g., weather conditions, use patterns, and road widths), as past studies have confirmed their usefulness (e.g., [71–73]).

While this study was effective at using a BNN model to predict PDRC, it is important to remember that the performances of ML models vary depending on the data. If the data are within the range of the current study's data, the results of this study can be replicated. Future research could use this technique, possibly with some tweaks, to analyse other datasets and present their findings. It enables a valid assessment of the BNN's ability to forecast PDRC.

Several prior studies also have found that walking behaviors can have a role in pedestrian fatalities as a result of road crashes. When it comes to pedestrian-involved collisions, the pedestrian crossing pattern is one of the most essential features of walking behaviour [74]. Pedestrians who were tragically wounded or admitted to hospital were typically crossing unlawfully and/or at fault, according to prior research (e.g., [33,75]). However, ARDD does not capture pedestrian activities at the moment of a collision, such as crossing and use of a mobile phone. The ARDD must include a wide variety of characteristics of both sides, vehicles and pedestrians, to gain a deep understanding of the reasons behind pedestrian-involved crashes.

5. Conclusions

The Australian Road Deaths Database was employed to train the BNN model to generate sound pedestrian death forecasts based on individuals' characteristics, time, occasions, road characteristics, and crash attributes in this study. For every road crash fatality class, this study created BNN models, including various structures, to assess their performances and to examine their corresponding predictive ambiguities. Below is a summary of this study's findings:

- The BNN model, which consists of three hidden neuron layers with sixteen hidden nodes in the first layer, provided the best training accuracy of 0.894. Its posterior predictive probabilities are in the great probability range, and the predictive ambiguity is tightly concentrated in the 0–0.1 range.
- BNN model outperformed RF, BN, and NN models.
- Personal characteristics and time and occasion factor groups are clearly essential, greatly boosting the performance of the model if they are used as inputs.
- Individually, the most important parameters in PDRC prediction were the speed limit, collision type, and age.

The following are some practical implications of the major results that may be of interest to both academics and practitioners in the domain.

For pedestrian safety on special occasions, such as Christmas and Easter, specific effective pedestrian safety strategies should be implemented. These policies may assist pedestrians in using roads safely and developing sustainable commute habits.

The speed limit has emerged as the most important factor for predicting pedestrian fatalities due to road crashes. It is obvious that increasing vehicle speed raises the collision risk exponentially [76]. According to Australian and worldwide case reports, lowering the posted maximum speed on rural roads by 10 km/h reduces the chance of an accident by 20–25%. Furthermore, after the removal of unrestricted speeds in some highways, the Australian road mortality database reveals that there was a 3.4 per year decline in fatalities

on highways with speed restrictions of 110 km/h and above. In Australia, for every person killed on the road, another 23 persons are injured as a result of an accident, highlighting the social benefit of any speed restriction lowering [77].

Another key factor in predicting pedestrian fatalities due to traffic collisions was age. Several prior studies have found that senior pedestrians are more prone to pedestrian crossing collisions. Elder pedestrian crashes are more likely to occur in congested metropolitan locations, and older pedestrians are more likely to be at fault because of their incapacity to manage complicated traffic scenarios, such as crossing roads [78]. These problems can be avoided if government agencies and licensing departments enhance crossing safety by reducing intersection ambiguity, increasing visibility, increasing conspicuity, and eliminating right-hand turns that require gap acceptance decisions. In addition, they can install or retrofit systems that defend pedestrians in locations where there is a significant risk of pedestrian fatality, such as high-pedestrian-activity places.

Our BNN model is capable of predicting future PDRC accurately, and it has a low level of predictive uncertainty. Although further research is needed in this area, the methods utilised in this study could be employed as a starting point for finding pedestrian risk determinants and developing appropriate legislation.

Author Contributions: Conceptualization, M.A. (Mahdi Aghaabbasi) and W.T.; methodology, M.A. (Mahdi Aghaabbasi); software, M.A. (Mahdi Aghaabbasi) and M.A. (Mujahid Ali); validation, M.A. (Mahdi Aghaabbasi); formal analysis, M.A. (Mahdi Aghaabbasi); investigation, M.A. (Mahdi Aghaabbasi) and M.A. (Mujahid Ali); resources, A.H.A., A.A.A., and E.E.H.; writing—original draft preparation, M.A. (Mahdi Aghaabbasi) and M.A. (Mujahid Ali); writing—review and editing, W.T., M.A. (Mahdi Aghaabbasi), M.A. (Mujahid Ali), A.H.A., R.Z., A.A.A. and E.E.H.; supervision, R.Z.; funding acquisition, A.H.A. and E.E.H. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by Taif University researchers support project, number TURSP-2020/252, Taif University, Taif, Saudi Arabia.

Acknowledgments: The authors would like to acknowledge financial support from the Taif University researchers support project, number TURSP-2020/252, Taif University, Taif, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Australian Transport Council (ATC). *National Road Safety Strategy 2011–2020*; Australian Transport Council: Canberra, Australia, 2011.
2. Department of Infrastructure Regional Development and Cities. Australian Road Deaths Database. 2021. Available online: https://www.bitre.gov.au/statistics/safety/fatal_road_crash_database (accessed on 13 November 2021).
3. Bureau of Infrastructure, Transport and Regional Economics. *Road Trauma Involving Heavy Vehicles 2018 Crash Statistical Summary*; BITRE: Canberra, Australia, 2020.
4. Zegeer, C.V.; Bushell, M. Pedestrian crash trends and potential countermeasures from around the world. *Accid. Anal. Prev.* **2012**, *44*, 3–11. [[CrossRef](#)] [[PubMed](#)]
5. Anderson, R.; Ponte, G.; Doecke, S. *A Survey of Bullbar Prevalence at Pedestrian Crash Sites in Adelaide, South Australia*; Centre for Automotive Safety Research: Adelaide, Australia, 2008.
6. Samerei, S.A.; Aghabayk, K.; Shiwakoti, N.; Karimi, S. Modelling bus-pedestrian crash severity in the state of Victoria, Australia. *Int. J. Inj. Control Saf. Promot.* **2021**, *28*, 233–242. [[CrossRef](#)] [[PubMed](#)]
7. Arnold, P.; Rosman, D.; Thornett, M. Pedestrian crash risk in Western Australia for both pedestrians and drivers. *Road Transp. Res.* **1992**, *1*, 60–75.
8. Imprialou, M.; Quddus, M. Crash data quality for road safety research: Current state and future directions. *Accid. Anal. Prev.* **2019**, *130*, 84–90. [[CrossRef](#)]
9. Mannering, F.L.; Bhat, C.R. Analytic methods in accident research: Methodological frontier and future directions. *Anal. Methods Accid. Res.* **2014**, *1*, 1–22. [[CrossRef](#)]
10. Shaheed, M.S.; Gkritza, K. A latent class analysis of single-vehicle motorcycle crash severity outcomes. *Anal. Methods Accid. Res.* **2014**, *2*, 30–38. [[CrossRef](#)]
11. Sun, M.; Sun, X.; Shan, D. Pedestrian crash analysis with latent class clustering method. *Accid. Anal. Prev.* **2019**, *124*, 50–57. [[CrossRef](#)]

12. Aghaabbasi, M.; Shekari, Z.A.; Shah, M.Z.; Olakunle, O.; Armaghani, D.J.; Moeinaddini, M. Predicting the use frequency of ride-sourcing by off-campus university students through random forest and Bayesian network techniques. *Transp. Res. Part A Policy Pract.* **2020**, *136*, 262–281. [[CrossRef](#)]
13. Qian, Y.; Aghaabbasi, M.; Ali, M.; Alqurashi, M.; Salah, B.; Zainol, R.; Moeinaddini, M.; Hussein, E.E. Classification of Imbalanced Travel Mode Choice to Work Data Using Adjustable SVM Model. *Appl. Sci.* **2021**, *11*, 11916. [[CrossRef](#)]
14. Aghaabbasi, M.; Shah, M.Z.; Zainol, R. Investigating the Use of Active Transportation Modes Among University Employees Through an Advanced Decision Tree Algorithm. *Civ. Sustain. Urban Eng.* **2021**, *1*, 26–49. [[CrossRef](#)]
15. Ali, M.; de Azevedo, A.R.G.; Marvila, M.T.; Khan, M.I.; Memon, A.M.; Masood, F.; Almahbashi, N.M.Y.; Shad, M.K.; Khan, M.A.; Fediuk, R.; et al. The Influence of COVID-19-Induced Daily Activities on Health Parameters—A Case Study in Malaysia. *Sustainability* **2021**, *13*, 7465. [[CrossRef](#)]
16. Ali, M.; Dharmowijoyo, D.B.; Harahap, I.S.; Puri, A.; Tanjung, L.E. Travel behaviour and health: Interaction of Activity-Travel Pattern, Travel Parameter and Physical Intensity. *Solid State Technol.* **2020**, *63*, 4026–4039.
17. Ali, M.; Dharmowijoyo, D.B.E.; de Azevedo, A.R.G.; Fediuk, R.; Ahmad, H.; Salah, B. Time-Use and Spatio-Temporal Variables Influence on Physical Activity Intensity, Physical and Social Health of Travelers. *Sustainability* **2021**, *13*, 12226. [[CrossRef](#)]
18. Chen, Y.; Aghaabbasi, M.; Ali, M.; Anciferov, S.; Sabitov, L.; Chebotarev, S.; Nabiullina, K.; Sychev, E.; Fediuk, R.; Zainol, R. Hybrid Bayesian Network Models to Investigate the Impact of Built Environment Experience before Adulthood on Students' Tolerable Travel Time to Campus: Towards Sustainable Commute Behavior. *Sustainability* **2022**, *14*, 325. [[CrossRef](#)]
19. Xie, Y.; Lord, D.; Zhang, Y. Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. *Accid. Anal. Prev.* **2007**, *39*, 922–933. [[CrossRef](#)]
20. Ali, M.; Abbas, S.; Salah, B.; Akhter, J.; Saleem, W.; Haruna, S.; Room, S.; Abdulkadir, I. Investigating Optimal Confinement Behaviour of Low-Strength Concrete through Quantitative and Analytical Approaches. *Materials* **2021**, *14*, 4675. [[CrossRef](#)] [[PubMed](#)]
21. Ali, M.; Room, S.; Khan, M.I.; Masood, F.; Ali Memon, R.; Khan, R.; Memon, A.M. Assessment of local earthen bricks in perspective of physical and mechanical properties using Geographical Information System in Peshawar, Pakistan. *Structures* **2020**, *28*, 2549–2561. [[CrossRef](#)]
22. De Azevedo, A.R.G.; Marvila, M.T.; Ali, M.; Khan, M.I.; Masood, F.; Vieira, C.M.F. Effect of the addition and processing of glass polishing waste on the durability of geopolymeric mortars. *Case Stud. Constr. Mater.* **2021**, *15*, e00662. [[CrossRef](#)]
23. Liu, T.; Liu, Y.; Liu, J.; Wang, L.; Xu, L.; Qiu, G.; Gao, H. A Bayesian learning based scheme for online dynamic security assessment and preventive control. *IEEE Trans. Power Syst.* **2020**, *35*, 4088–4099. [[CrossRef](#)]
24. Marzban, C.; Witt, A. A Bayesian neural network for severe-hail size prediction. *Weather Forecast.* **2001**, *16*, 600–610. [[CrossRef](#)]
25. Yan, D.; Zhou, Q.; Wang, J.; Zhang, N. Bayesian regularisation neural network based on artificial intelligence optimisation. *Int. J. Prod. Res.* **2017**, *55*, 2266–2287. [[CrossRef](#)]
26. Zajac, S.S.; Ivan, J.N. Factors influencing injury severity of motor vehicle–crossing pedestrian crashes in rural Connecticut. *Accid. Anal. Prev.* **2003**, *35*, 369–379. [[CrossRef](#)]
27. Rifaat, S.M.; Chin, H.C. Accident severity analysis using ordered probit model. *J. Adv. Transp.* **2007**, *41*, 91–114. [[CrossRef](#)]
28. Obeng, K.; Rokonuzzaman, M. Pedestrian injury severity in automobile crashes. *Open J. Saf. Sci. Technol.* **2013**, *3*, 33341. [[CrossRef](#)]
29. Kwigizile, V.; Sando, T.; Chimba, D. Inconsistencies of ordered and unordered probability models for pedestrian injury severity. *Transp. Res. Rec.* **2011**, *2264*, 110–118. [[CrossRef](#)]
30. Yasmin, S.; Eluru, N. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. *Accid. Anal. Prev.* **2013**, *59*, 506–521. [[CrossRef](#)]
31. Sze, N.-N.; Wong, S. Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes. *Accid. Anal. Prev.* **2007**, *39*, 1267–1278. [[CrossRef](#)]
32. Kim, S.; Ulfarsson, G.F. Traffic safety in an aging society: Analysis of older pedestrian crashes. *J. Transp. Saf. Secur.* **2019**, *11*, 323–332. [[CrossRef](#)]
33. Ulfarsson, G.F.; Kim, S.; Booth, K.M. Analyzing fault in pedestrian–motor vehicle crashes in North Carolina. *Accid. Anal. Prev.* **2010**, *42*, 1805–1813. [[CrossRef](#)]
34. Tay, R.; Choi, J.; Kattan, L.; Khan, A. A multinomial logit model of pedestrian–vehicle crash severity. *Int. J. Sustain. Transp.* **2011**, *5*, 233–249. [[CrossRef](#)]
35. Zhou, Z.-P.; Liu, Y.-S.; Wang, W.; Zhang, Y. Multinomial logit model of pedestrian crossing behaviors at signalized intersections. *Discret. Dyn. Nat. Soc.* **2013**, *2013*, 172726. [[CrossRef](#)]
36. Chen, Z.; Fan, W. Modeling pedestrian injury severity in pedestrian-vehicle crashes in rural and urban areas: Mixed logit model approach. *Transp. Res. Rec.* **2019**, *2673*, 1023–1034. [[CrossRef](#)]
37. Kim, J.-K.; Ulfarsson, G.F.; Shankar, V.N.; Mannering, F.L. A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accid. Anal. Prev.* **2010**, *42*, 1751–1758. [[CrossRef](#)]
38. Haleem, K.; Alluri, P.; Gan, A. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accid. Anal. Prev.* **2015**, *81*, 14–23. [[CrossRef](#)]
39. Tulu, G.S.; Washington, S.; Haque, M.M.; King, M.J. Injury severity of pedestrians involved in road traffic crashes in Addis Ababa, Ethiopia. *J. Transp. Saf. Secur.* **2017**, *9*, 47–66. [[CrossRef](#)]
40. Rifaat, S.M.; Tay, R.; de Barros, A. Urban street pattern and pedestrian traffic safety. *J. Urban Des.* **2012**, *17*, 337–352. [[CrossRef](#)]

41. Sasidharan, L.; Menéndez, M. Partial proportional odds model—An alternate choice for analyzing pedestrian crash injury severities. *Accid. Anal. Prev.* **2014**, *72*, 330–340. [[CrossRef](#)]
42. Pour, A.T.; Moridpour, S.; Tay, R.; Rajabifard, A. A partial proportional odds model for pedestrian crashes at mid-blocks in Melbourne metropolitan area. *MATEC Web Conf.* **2016**, *81*, 02020. [[CrossRef](#)]
43. Li, Y.; Fan, W.D. Modelling severity of pedestrian-injury in pedestrian-vehicle crashes with latent class clustering and partial proportional odds model: A case study of North Carolina. *Accid. Anal. Prev.* **2019**, *131*, 284–296. [[CrossRef](#)]
44. Li, Y.; Fan, W. Pedestrian injury severities in pedestrian-vehicle crashes and the partial proportional odds logit model: Accounting for age difference. *Transp. Res. Rec.* **2019**, *2673*, 731–746. [[CrossRef](#)]
45. Chang, L.-Y.; Chen, W.-C. Data mining of tree-based models to analyze freeway accident frequency. *J. Saf. Res.* **2005**, *36*, 365–375. [[CrossRef](#)]
46. Gong, Y.; Abdel-Aty, M.; Cai, Q.; Rahman, M.S. A decentralized network level adaptive signal control algorithm by deep reinforcement learning. In Proceedings of the Transportation Research Board 98th Annual Meeting, Washington, DC, USA, 13–17 January 2019.
47. Zhu, S.Y. Analyse vehicle-pedestrian crash severity at intersection with data mining techniques. *Int. J. Crashworthiness* **2021**, 1–9. [[CrossRef](#)]
48. Mackay, D.J.C. Bayesian Methods for Adaptive Models. Ph.D. Thesis, California Institute of Technology, Pasadena, CA, USA, 1992.
49. Neal, R.M. *Bayesian Learning for Neural Networks*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; Volume 118.
50. Liang, F. Bayesian neural networks for nonlinear time series forecasting. *Stat. Comput.* **2005**, *15*, 13–29. [[CrossRef](#)]
51. Riviere, C.; Lauret, P.; Ramsamy, J.M.; Page, Y. A Bayesian neural network approach to estimating the energy equivalent speed. *Accid. Anal. Prev.* **2006**, *38*, 248–259. [[CrossRef](#)]
52. Pour, A.T.; Moridpour, S.; Rajabifard, A.; Tay, R. Spatial and temporal distribution of pedestrian crashes in Melbourne metropolitan area. *Road Transp. Res.* **2017**, *26*, 4–20.
53. Ding, C.; Chen, P.; Jiao, J.F. Non-linear effects of the built environment on automobile-involved pedestrian crash frequency: A machine learning approach. *Accid. Anal. Prev.* **2018**, *112*, 116–126. [[CrossRef](#)]
54. Mokhtarimousavi, S. A Time of Day Analysis of Pedestrian-Involved Crashes in California: Investigation of Injury Severity, a Logistic Regression and Machine Learning Approach Using HSIS Data. *ITE J.-Inst. Transp. Eng.* **2019**, *89*, 25–33.
55. Das, S.; Le, M.; Dai, B.Y. Application of machine learning tools in classifying pedestrian crash types: A case study. *Transp. Saf. Environ.* **2020**, *2*, 106–119. [[CrossRef](#)]
56. Rahimi, A.; Azimi, G.; Asgari, H.; Jin, X. Injury Severity of Pedestrian and Bicyclist Crashes Involving Large Trucks. In Proceedings of the ASCE International Conference on Transportation and Development (ASCE ICTD), Seattle, WA, USA, 26–29 May 2020; pp. 110–122.
57. Guo, M.; Yuan, Z.; Janson, B.; Peng, Y.; Yang, Y.; Wang, W. Older pedestrian traffic crashes severity analysis based on an emerging machine learning XGBoost. *Sustainability* **2021**, *13*, 926. [[CrossRef](#)]
58. Saha, D.; Dumbaugh, E. Use of a model-based gradient boosting framework to assess spatial and non-linear effects of variables on pedestrian crash frequency at macro-level. *J. Transp. Saf. Secur.* **2021**, *13*, 32. [[CrossRef](#)]
59. Wang, H.; Yeung, D.-Y. Towards Bayesian deep learning: A framework and some existing methods. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 3395–3408. [[CrossRef](#)]
60. Blei, D.M.; Kucukelbir, A.; McAuliffe, J.D. Variational inference: A review for statisticians. *J. Am. Stat. Assoc.* **2017**, *112*, 859–877. [[CrossRef](#)]
61. Wu, S.; Yuan, Q.; Yan, Z.; Xu, Q. Analyzing Accident Injury Severity via an Extreme Gradient Boosting (XGBoost) Model. *J. Adv. Transp.* **2021**, *2021*, 3771640. [[CrossRef](#)]
62. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
63. Onieva-García, M.Á.; Martínez-Ruiz, V.; Lardelli-Claret, P.; Jiménez-Moleón, J.J.; Amezcu-Prieto, C.; de Dios Luna-del-Castillo, J.; Jiménez-Mejías, E. Gender and age differences in components of traffic-related pedestrian death rates: Exposure, risk of crash and fatality rate. *Inj. Epidemiol.* **2016**, *3*, 14. [[CrossRef](#)]
64. Toran Pour, A.; Moridpour, S.; Tay, R.; Rajabifard, A. Influence of pedestrian age and gender on spatial and temporal distribution of pedestrian crashes. *Traffic Inj. Prev.* **2018**, *19*, 81–87. [[CrossRef](#)]
65. Park, S.; Ko, D. Investigating the Factors Influencing Pedestrian–Vehicle Crashes by Age Group in Seoul, South Korea: A Hierarchical Model. *Sustainability* **2020**, *12*, 4239. [[CrossRef](#)]
66. Kim, J.-K.; Ulfarsson, G.F.; Shankar, V.N.; Kim, S. Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic logit analysis. *Accid. Anal. Prev.* **2008**, *40*, 1695–1702. [[CrossRef](#)]
67. Park, H.-C.; Joo, Y.-J.; Kho, S.-Y.; Kim, D.-K.; Park, B.-J. Injury severity of bus–pedestrian crashes in South Korea considering the effects of regional and company factors. *Sustainability* **2019**, *11*, 3169. [[CrossRef](#)]
68. Li, P.; Abdel-Aty, M.; Yuan, J. Using bus critical driving events as surrogate safety measures for pedestrian and bicycle crashes based on GPS trajectory data. *Accid. Anal. Prev.* **2021**, *150*, 105924. [[CrossRef](#)]

69. Sivasankaran, S.K.; Balasubramanian, V. Severity of Pedestrians in Pedestrian-Bus Crashes: An Investigation of Pedestrian, Driver and Environmental Characteristics Using Random Forest Approach. In Proceedings of the Congress of the International Ergonomics Association, Online, 13–18 June 2021; pp. 825–833.
70. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
71. Shawky, M.; Garib, A.; Al-Harthei, H. The impact of road and site characteristics on the crash-injury severity of pedestrian crashes. *Adv. Transp. Stud.* **2014**, *1*, 27–36.
72. Zhai, X.; Huang, H.; Sze, N.; Song, Z.; Hon, K.K. Diagnostic analysis of the effects of weather condition on pedestrian crash severity. *Accid. Anal. Prev.* **2019**, *122*, 318–324. [[CrossRef](#)] [[PubMed](#)]
73. Ukkusuri, S.; Miranda-Moreno, L.F.; Ramadurai, G.; Isa-Tavarez, J. The role of built environment on pedestrian crash frequency. *Saf. Sci.* **2012**, *50*, 1141–1151. [[CrossRef](#)]
74. Jain, A.; Gupta, A.; Rastogi, R. Pedestrian crossing behaviour analysis at intersections. *Int. J. Traffic Transp. Eng.* **2014**, *4*, 103–116. [[CrossRef](#)]
75. Kucerová, I.; Sulíková, B.; Polácková, K. Pedestrian accidents-actual trend in the Czech Republic. *Trans. Transp. Sci.* **2013**, *6*, 145. [[CrossRef](#)]
76. Aarts, L.; Van Schagen, I. Driving speed and the risk of road crashes: A review. *Accid. Anal. Prev.* **2006**, *38*, 215–224. [[CrossRef](#)]
77. Read, D.J. Open speeds on Northern Territory roads: Not so fast. *Med. J. Aust.* **2015**, *203*, 14–15. [[CrossRef](#)]
78. Oxley, J.; Whelan, M. It cannot be all about safety: The benefits of prolonged mobility. *Traffic Inj. Prev.* **2008**, *9*, 367–378. [[CrossRef](#)]