

Article

Analyzing Crash Severity: Human Injury Severity Prediction Method Based on Transformer Model

Yalan Jiang ¹, Xianguo Qu ^{2,*}, Weiwei Zhang ^{3,4}, Wenfeng Guo ⁴, Jiejie Xu ³, Wangpengfei Yu ³ and Yang Chen ³

¹ School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; jiangyl@sues.edu.cn

² SAMR Defective Product Recall Technical Center, Beijing 100088, China

³ Shanghai Smart Vehicle Cooperating Innovation Center, Shanghai 201805, China; zhangww.2021@tsinghua.org.cn (W.Z.); xujie@smarvcte.com (J.X.); pengfeixiaoyu@smarvcte.com (W.Y.); chenyang@smarvcte.com (Y.C.)

⁴ School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China; guowenfeng@smarvcte.com

* Correspondence: quxg@samrdprc.org.cn

Abstract: Traffic accident-related injuries and fatalities are a serious global public health and social development challenge. The accurate prediction of crash severity improves road safety and reduces casualties, as well as serving road managers and policy makers. Prediction models need to learn and analyze the various characteristic factors of traffic accidents and capture from them the inherent complex relationship between accident characteristics and the severity of traffic accidents. However, most accident prediction studies lack analytical predictions of injury severity, and predictive models rely on the content and quality of accident datasets. To increase the robustness and accuracy of prediction models, this paper leverages a Transformer-based architecture for the severity prediction of traffic collisions from human injury severity. This framework learns both text and sequence data from accident datasets. After comparative analysis, the framework can achieve the prediction of human injury severity under different data categories and show good prediction performance at low injury severity levels using only textual data or sequence data.



Academic Editors: Deogratias Eustace, Bhaven Naik, Heng Wei and Parth Bhavsar

Received: 2 December 2024

Revised: 5 January 2025

Accepted: 14 January 2025

Published: 15 January 2025

Citation: Jiang, Y.; Qu, X.; Zhang, W.; Guo, W.; Xu, J.; Yu, W.; Chen, Y. Analyzing Crash Severity: Human Injury Severity Prediction Method Based on Transformer Model. *Vehicles* **2025**, *7*, 5. <https://doi.org/10.3390/vehicles7010005>

Copyright: © 2025 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/>).

Keywords: traffic safety; crash severity prediction; road safety; injury severity; Transformer

1. Introduction

The modern transportation system plays a vital role in meeting the demands of daily traffic and logistics transportation while profoundly influencing socio-economic development and enhancing people's living standards. With the rapid increase in the number of motor vehicles, traffic pressure has significantly risen and the frequency of traffic collisions has steadily increased. The issue of human casualties resulting from traffic accidents is becoming increasingly prevalent, exerting a significant economic burden and psychological impact on society and families. This growing crisis has heightened awareness across society regarding the seriousness of road safety issues and the urgent need for improvement. In response, governments worldwide unanimously committed—through UN General Assembly Resolution 74/299—to the Decade of Action for Road Safety 2021–2030. This initiative aims to reduce road traffic deaths and injuries by at least 50% during this period. The Global Status Report on Road Safety 2023 indicates that annual road traffic deaths have slightly decreased to 1.19 million. However, achieving the global goal of reducing road traffic injuries and fatalities by at least half by 2030 still requires urgent and coordinated efforts [1].

Traffic accident prediction has emerged as a critical research area for enhancing road traffic safety. Through comprehensive data analysis and the construction of robust accident prediction models, high-frequency, high-risk, and high-casualty accidents can be effectively identified and mitigated. This approach enables the implementation of proactive preventive measures, fosters advancements in road traffic safety technologies and traffic safety management systems, and ultimately reduces casualties. Most prior research on predicting accident severity has adopted an accident-focused perspective, utilizing various classification systems. These include binary categorizations based on the presence or absence of fatalities [2] and multilevel classifications reflecting the severity of injuries sustained by individuals [3,4]. Accident severity categorization is typically determined by the information content of the dataset or the predefined classification schemes established by researchers, resulting in a degree of uniqueness in prediction algorithms. This study argues that accident severity categorization should be based on injury severity and that a more standardized and consistent categorization framework should be employed. Such an approach not only facilitates the development of a unified predictive framework for traffic accident severity but also directly enhances rescue efficiency and reduces casualties from the perspective of the injured. In this study, the Abbreviated Injury Scale (AIS) from the Injury Severity Rating Scale for trauma patients was adopted as the prediction model's output, with the Maximum Abbreviated Injury Scale (MAIS) used to represent the severity of traffic accident outcomes.

A systematic analysis of the factors influencing crash severity (e.g., crash velocity, environmental conditions, driver reaction time, and vehicle safety performance) can elucidate the relative impact of these factors on driver injury risk. On the one hand, it directly influences overall road safety within a traffic system, offering guidance for road infrastructure improvements and providing data support to road traffic management departments in formulating more comprehensive road safety regulations. On the other hand, it impacts drivers' safe driving habits, encourages protective behaviors, and supports the development of intelligent vehicles through enhanced automatic driving intervention strategies and automatic emergency avoidance systems, further mitigating the adverse effects of traffic accidents on individual safety. Safe driving is no longer solely reliant on reaction but is increasingly rooted in proactive prediction. Therefore, studying the severity of human injuries in traffic accidents contributes to advancing the theoretical framework of traffic safety management and offers actionable insights for future traffic accident prevention and personal injury risk mitigation.

The primary goal of this study was to develop a Transformer-based architecture capable of capturing and processing textual and sequential data related to traffic accidents. The focus was on predicting the severity of driver injuries. This approach facilitates accurate injury severity predictions, even with limited available accident information.

The main contributions of this study are threefold:

- (1) This study establishes an attention-based framework for predicting the severity of human injuries in traffic accidents. The framework can process and analyze textual data from accident descriptions and sequential data from accident datasets to accurately predict accident severity.
- (2) Traditional severity analysis and prediction approaches are typically constrained by the severity levels pre-defined within datasets, resulting in a limited scope for covering a broader range of accident severities or enabling finer-grained classifications. This study proposes a novel prediction model for driver injury severity in traffic accidents that specifically focuses on human injury outcomes.
- (3) A generalized traffic accident severity prediction model that is not restricted by external factors such as accident location, road structure, and similar variables is developed.

2. Research Methodology

2.1. Extraction of Key Factors Affecting Accident Severity

Understanding the factors contributing to road crashes is critical for advancing researchers' insights into traffic crash safety. Factors influencing the severity of traffic accidents are typically examined and extracted through logic models and their variants, such as multinomial logit (MNL), nested logit (NL), and ordered logit (OL). Ghandour et al. proposed a model employing a hybrid ensemble machine learning classifier, which integrates sequential minimal optimization and decision trees, to identify risk factors associated with fatal road injuries [5]. In addition to investigating factors such as driving behavior and vehicle conditions, Xu et al. utilized association rule analysis (ARA) to delve deeper into the interdependencies among causal factors in mass casualty accidents [6].

2.2. Traffic Accident Severity Prediction

Recent studies have categorized accident severity prediction models into three primary types: probabilistic models, machine learning, and natural language processing (NLP).

2.2.1. Probabilistic Models

Injury severity prediction is commonly framed as a classification task, typically defined by the severity categories recorded in datasets or user-defined classifications. Probability-based models represent the traditional approach for injury severity prediction. These models reduce severity prediction to a probabilistic solution, often relying on logistic regression or probabilistic frameworks with a dependent variable.

Considering the bidirectional nature of collisions, Fu et al. utilized a bivariate generalized ordered probability model in their collision severity prediction study to parameterize collision thresholds and intra-collision correlations, classifying collision severity across different risk modes based on characteristic data [7]. Christoforou et al. applied a stochastic parameter-ordered probability model to examine the effects of traffic flow and speed on driver injury severity [8]. Ordered probit models are frequently employed to analyze the impact of diverse factors on injury severity in road traffic accidents, including vehicle types, road conditions, environmental variables, and accident types [9,10]. Do et al. developed a stochastic parametric negative binomial (RPNB) model based on a binary logit framework to evaluate the influence of geographic and event-specific variables on crash severity involving vulnerable traffic participants, categorizing severity into three primary levels: fatalities, injuries, and property damage [11].

2.2.2. Machine Learning Models

Logistic models are widely applied in forecasting tasks due to their simplicity and efficient predictive performance. However, their sensitivity to noise and outliers makes logistic models prone to underfitting. Consequently, traditional machine learning algorithms have gained popularity in injury severity analysis and prediction. Katanalp et al. utilized a decision tree model to examine the impact of injury severity classification and driving-related parameters on cyclist injury severity. They also employed a fuzzy decision-making approach to evaluate the classification performance [12].

The performance of decision tree models is often limited by feature selection and parameter tuning, which makes their predictive accuracy inferior to that of more advanced machine learning methods. Alkheder et al. applied an artificial neural network (ANN) to predict traffic accident severity, demonstrating its effectiveness in processing and predicting accident outcomes [13]. Sameen et al. employed a recurrent neural network (RNN) to model collisions in highway scenarios, analyzing temporal interactions and dependencies

in sequential data. Their findings suggest that RNNs, within deep learning frameworks, hold significant promise for predicting injury severity in traffic accidents [14].

Artificial neural networks (ANNs) and deep neural networks (DNNs) exhibit developmental challenges, including training complexity, overfitting issues, and limited generalization. Consequently, more advanced machine learning models are being adopted for injury severity prediction. Jamal et al. employed the Extreme Gradient Boosting (XGBoost) model to analyze collision damage severity, demonstrating superior predictive performance compared with traditional algorithms such as decision trees and random forests [15]. Niyogisubizo et al. introduced a novel hybrid stacking framework that integrates gradient boosting (GB), adaptive boosting (AdaBoost), and multi-layer perceptron (MLP) by leveraging the strengths and mitigating the weaknesses of various machine learning techniques. This framework enhanced the severity analysis and prediction accuracy for property-damage-only and injury-related collisions [16].

In addressing the significant imbalance between collision and non-collision data, Cai et al. utilized a deep convolutional generative adversarial network (DCGAN) model to analyze and generate synthetic collision data, providing a robust dataset for severity prediction [17]. Rahim et al. developed a deep learning algorithm based on convolutional neural networks (CNNs) to predict traffic collision severity. Their findings validated that deep learning methods outperformed traditional machine learning approaches, such as support vector machines, particularly in predicting the severity of fatal and injury crashes [18]. Islam et al. combined deep learning with integrative learning strategies to design a model capable of the real-time prediction of collisions across various severity levels, effectively handling large-scale collision datasets and severity measurements [19].

2.2.3. NLP

Previous research on crash severity prediction has primarily relied on the content and quality of crash datasets, with specialized data collection being both time-consuming and labor-intensive. To overcome this limitation, a new research direction has emerged that focuses on extracting key factors from collision accidents and predicting severity by applying natural language processing (NLP) algorithms to accident text data. Fan et al. converted structured crash dataset information into textual data, leveraging the textual reasoning capabilities of large language models (LLMs) to predict crash severity. The authors framed crash feature learning as a novel textual reasoning task and used various large language models (LLMs) to predict detailed crash outcomes, including severity [20]. Li et al. employed the structural topic model (STM) and an interpretable XGBoost model from NLP to examine the relationship between latent topics and crash severity in vehicle crash narratives [21].

In most research on crash severity prediction, the analysis of predictive models and key factors is shaped by the content of the dataset records, which are geographically and road structure-dependent, limiting coverage to specific geographic regions. Moreover, crash data collection is often subject to the discretion of investigators and experts, undermining the objectivity of data-driven algorithms and the accuracy of predictive models. Therefore, this study aimed to construct a generalized accident severity prediction model that is not limited by the region where the accident occurred or the content of the accident dataset.

3. Analysis Methodologies

This section provides a detailed description of the attention framework proposed in this paper for predicting injury severity in crash accidents. The model framework is based on the Transformer architecture, a learning framework solely reliant on attention mechanisms that was initially developed for machine translation [22]. The success of the

Transformer in natural language tasks has led to new research directions and applications in other domains. Currently, the Transformer framework has demonstrated outstanding performance in natural language processing (NLP) [23], computer vision (CV) [24–26], and other fields. It has also gradually evolved into a flexible, scalable, and replicable general-purpose backbone network architecture for various downstream tasks. This architecture offers fresh avenues for exploration and development in traditional fields.

3.1. Framework

This study closely followed the original Transformer framework's architectural design, aiming to create a scalable and efficient attention framework for predicting injury severity in crash accidents. Figure 1 illustrates the overall architecture, which consists of an encoder and a decoder. The encoder processes a series of inputs, represented as $(x_1, \dots, x_n)$, consisting of symbols and transforms them into a series of consecutive representations denoted as $z = (z_1, \dots, z_n)$. Subsequently, the decoder generates the output sequence $(y_1, \dots, y_n)$ step by step. In exceptional cases, the model is autoregressive, where the outputs generated based on the input sequence information serve as new additional inputs into the model through the masking process.

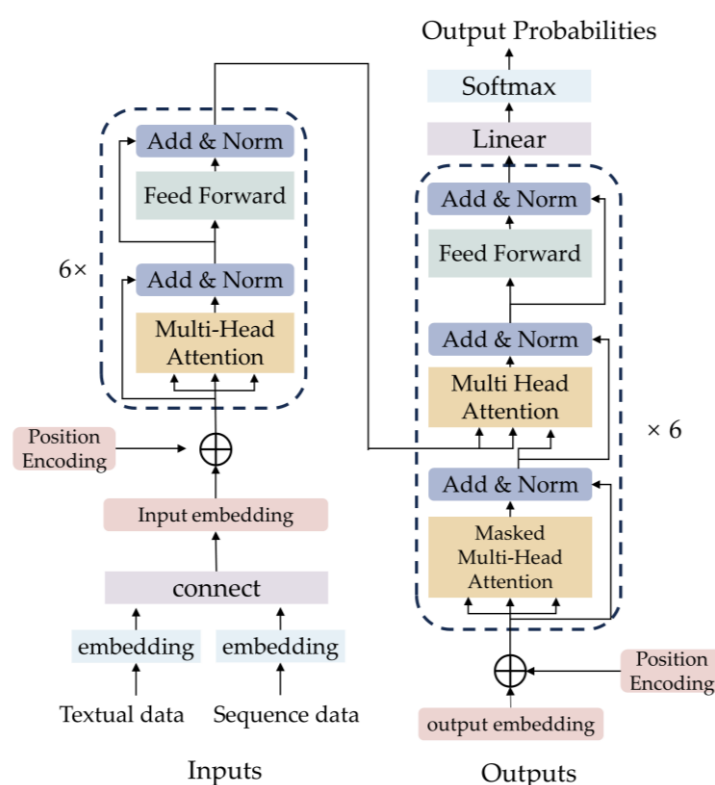


Figure 1. Attentional framework structure.

The input textual data and sequence data are spliced after embedding processing, followed by common position encoding to obtain the input sequence containing positional relations. For text data processing, the input embedding module converts and maps the discrete tokens of the input text into continuous, fixed-size vectors. Since the architecture does not include convolution or recursion to capture sequence positions, positional encoding is introduced to capture sequence order information. The positional encoding module employs different combinations of sine and cosine functions to generate position vectors

for positional encoding, as shown in Equations (1) and (2), where i denotes dimension and d denotes the dimension size of the input vector.

$$p_{p,2i} = \sin\left(\frac{p}{1000^{\frac{2i}{d}}}\right) \quad (1)$$

$$p_{p,2i+1} = \cos\left(\frac{p}{1000^{\frac{2i}{d}}}\right) \quad (2)$$

3.1.1. Encoder–Decoder Architectures

As shown in Figure 1, the encoder and decoder each consist of a stack of six identical layers, including a Multi-Head Attention sublayer and a feed-forward network sublayer. Each of these sublayers is connected through residual links and normalization. One key difference is the presence of a masked Multi-Head Attention sublayer in the decoder that is designed to mask the unpredicted information. The Transformer framework's main feature is its use of an attention mechanism, rather than traditional convolutional networks, to process global information. The entire sequence is trained simultaneously, effectively avoiding issues like gradient vanishing and gradient explosion. The attention mechanism enables the model to capture the global dependencies among sequence elements by processing each element while considering all others in the sequence.

The encoder receives the word vectors after the embedding and positional encoding processes and generates three identical vectors, which are multiplied by different weight matrices W^Q , W^K , and W^V to obtain the Query matrix, Key matrix, and Value matrix, respectively. Their vector length and word length are 64 and 512, respectively. The attention function is shown in Equation (3), where $(Q,K,V) \in \mathbb{R}^{t \times d}$ denotes the vector matrix, with t representing the number of vectors and d the dimension of the vector. The weighted sum, known as Self-Attention, is obtained through the dot product matrix calculation and normalization.

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

Both the encoder and decoder architectures utilize a Multi-Head Attention module to enhance the Self-Attention layer's ability to focus on different locations. Multi-Head Attention employs 8 attention heads to project the encoder input vector information into 8 subspaces, allowing the spatial information of multiple locations in each sublayer to be captured simultaneously. The formula for Multi-Head Attention is shown in Equation (4).

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_8)W_0 \quad (4)$$

where $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$, $W_i^Q \in \mathbb{R}^{d \times d_Q}$, $W_i^K \in \mathbb{R}^{d \times d_K}$, $W_i^V \in \mathbb{R}^{d \times d_V}$, and $W^0 \in \mathbb{R}^{8d \times d}$. Multi-Head Attention is a concatenation and transformation of the Self-Attention structure. As shown in Figure 2, the outputs of the Self-Attention results for each subspace are concatenated and passed through a linear layer to obtain the original spatial dimension size.

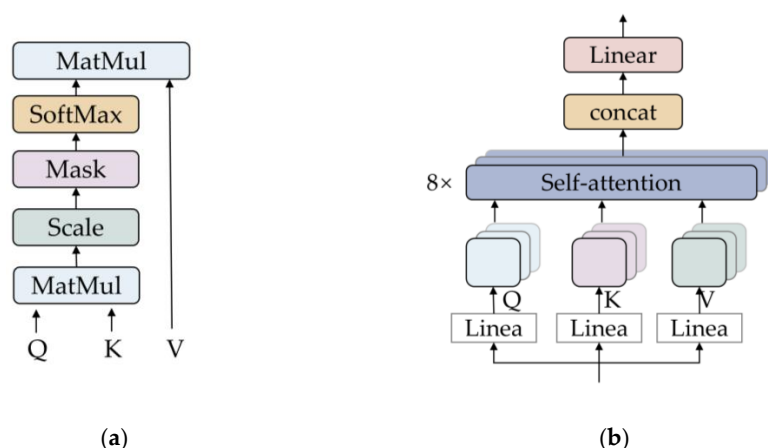


Figure 2. Two attention mechanisms in architecture: (a) Self-Attention and (b) Multi-Head Attention.

The Transformer framework assigns different weights to each position in the input sequence by calculating the degree of association between each position and others, which is the essence of Multi-Head Attention. This enables the model to learn complex, unstructured data by assigning varying levels of attention, allowing it to identify important and relevant information in the input sequence. There are two main structural differences between the encoder and decoder. First, the input to the Multi-Head Attention sublayer in the decoder consists of both the current input of the decoder and the output from the encoder, allowing the model to focus on the original input information. The autoregressive property of the Transformer framework is realized through the masked Multi-Head Attention sublayer. Secondly, the masked Multi-Head Attention sublayer in the decoder includes one of the two masking operations in the architecture to ensure that the generated sequence can only depend on previous sequence information, not future information. Therefore, the masking operation on future positions is applied to maintain causality in the autoregressive generation. The other masking operation is explained in Section 3.2.2.

3.1.2. Feed-Forward Networks

The encoder–decoder structure also incorporates a feed-forward network module, which performs additional non-linear processing of the output from the Multi-Head Attention mechanism, enhancing the model's expressiveness. The feed-forward network structure, comprising two fully connected layers, is illustrated in Figure 3. The first fully connected layer performs a linear transformation, mapping the dimension d of the output vectors to a higher dimension $d' = 2048$, while the second layer remaps the activated vectors to the original dimension. This study uses the ReLU function as the non-linear activation function, as shown in Equation (5).

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (5)$$

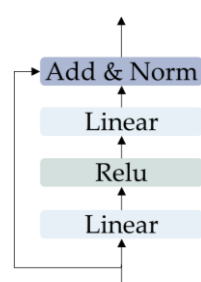


Figure 3. Structure of feed-forward network.

Each sublayer in the encoder and decoder modules is connected using skip connections and layer normalization, which aims to address the training issues of gradient explosion and vanishing gradients. The residual connections and normalization are shown in Equations (6) and (7).

$$\text{output} = \text{LayerNorm}(x + \text{SubLayer}(x)) \quad (6)$$

$$\text{LayerNorm}(x) = \frac{x - \mu}{\delta} \quad (7)$$

Skip connections address the issue of gradient vanishing by propagating gradients through layers, adding the inputs to the outputs of the sublayers. Layer normalization stabilizes the distribution of activation values by reducing internal covariate shift, which helps stabilize the training process and mitigate gradient vanishing and explosion. Together, these mechanisms enhance the effectiveness of Transformer model training.

3.2. Data

The traffic accident data were obtained from the Crash Report Sampling System (CRSS) of the National Highway Traffic Safety Administration (NHTSA) in the United States, the German In-Depth Accident Study (GIDAS) in German, and the Institute of Traffic Accident Research and Data Analysis (ITARDA) in Japan. There may be incomplete records or obvious errors in the dataset. Therefore, we first cleaned the dataset by eliminating accident cases with obvious errors or incomplete records. After compiling and filtering, we selected 45,373 vehicle-to-vehicle collisions cases from those recorded between 2017 and 2022, including 90,746 participants. The collisions selected for this study were those in which two parties were involved. The severity of multi-party collisions is influenced by information about the collision characteristics of both parties. For instance, under the same environmental and crash conditions, the different parties involved may experience varying driver injuries, implying heterogeneity within each crash case. Each participant represents a distinct data point, which not only improves the efficiency of data usage but also enables a more precise understanding and prediction of injury severity.

The accident data contain different variables, including basic accident information, vehicle information, driver information, environmental factors, and the severity of injuries sustained by crash victims. These variables provide a detailed record of crash characteristics from various perspectives.

3.2.1. Processing of Textual Data

Each accident's textual description consisted of approximately 100 words. Accident investigations and reports must describe, as clearly and simply as possible, the time and location of the collision, the type of road, the nature of the vehicle impact, the injuries sustained, and the specific circumstances of the accident. There is no fixed format for incident descriptions, which often contain non-standardized information. As a result, it was essential to filter and clean the data. Figure 4 shows the processing of the textual data in the dataset.

Text cleaning began with a standardization process that included converting text to a uniform case, removing punctuation, and eliminating irrelevant symbols. Words in the accident descriptions that did not provide meaningful information about the incident, such as the location or the car manufacturer, also had to be removed. This helped to reduce computational complexity and improve model training efficiency. Notably, the textual data in accident descriptions do not include the severity level of the driver's injury, defined in the sequential data. After stemming and word reduction, static word embeddings were generated using the pre-trained Word2Vec model [27] to create context-independent vector

representations. The pre-processed text data were then encoded into a vector of size (256, 512). Crash feature learning was framed as a new textual inference problem.

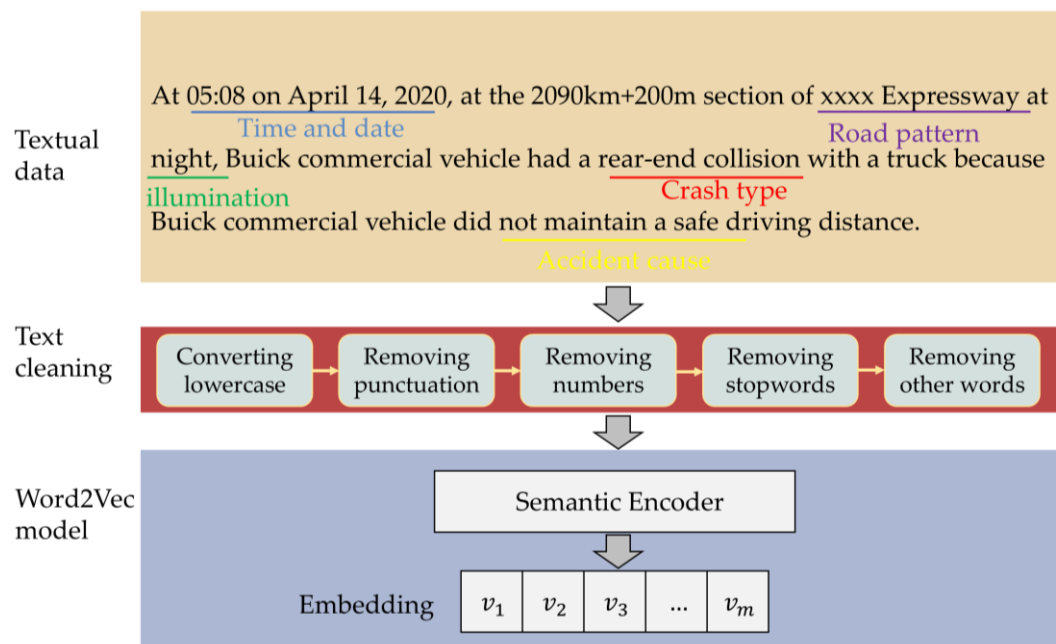


Figure 4. Vector embedding of accident textual data.

3.2.2. Encoding of Mixed Data

The Transformer model was designed to extract characteristic variables from sequence data. The dataset used in this study consisted of both textual information describing accidents and sequential data containing various feature types. Both data types needed to be transformed and mapped into continuous tensors of fixed size within the embedding module. These fixed tensors were then integrated into the encoder–decoder structure to construct the feature-factor mapping relationship.

The structured sequence data and unstructured textual data were fed into the embedding module, resulting in an output tensor of the form (seq_length, embedding_size), where seq_length is 256 and embedding_size is 512. The continuous variables within the characteristic variables were transformed into discrete values, as illustrated in Table 1 and encoded using unique heat coding. The embedded sequence and text data were then directly concatenated and processed to achieve a tensor size of (512 × 512).

Table 1. Crash percentages and characteristics by varying crash severity.

Characteristic	Category	Number	Percentage of Traffic Crash Severity (%)						
			AIS0	AIS1	AIS2	AIS3	AIS4	AIS5	AIS6
Crash Velocity	0–20	51,858	81.46	12.14	4.6	1.53	0.11	0.11	0.05
	20–40	25,699	74.34	14.12	8.19	2.9	0.11	0.25	0.09
	40–60	10,833	66.63	15.76	10.67	6.09	0.6	0.17	0.08
	60–80	2255	75.3	10.47	7.05	3.81	1.06	1.46	0.85
	80–150	85	63.53	11.76	5.88	4.71	3.53	4.71	5.88
Gender	Male	47,616	81.56	10.5	5.35	2.18	0.21	0.11	0.09
	Female	43,130	73.08	15.95	7.56	2.91	0.17	0.2	0.13
Age	0–25	22,713	80.4	11.35	5.9	2.12	0.1	0.1	0.03
	26–55	48,115	77.31	13.63	6.22	2.43	0.17	0.16	0.08
	56–100	19,918	74.92	13.77	7.41	3.21	0.36	0.21	0.12
Illumination	Daytime	69,892	78.47	12.79	6.07	2.25	0.15	0.15	0.12
	Night (with streetlight)	12,924	75.28	14.05	7.2	3.03	0.2	0.19	0.05
	Night (without streetlight)	3774	69.9	13.86	9.7	5.38	1.06	0.1	0
	Dawn/Dusk	4156	75.35	14.26	6.47	2.97	0.51	0.23	0.21

Table 1. Cont.

Characteristic	Category	Number	Percentage of Traffic Crash Severity (%)						
			AIS0	AIS1	AIS2	AIS3	AIS4	AIS5	AIS6
Alcohol	Yes	942	77.7	13.1	6.34	2.44	0.18	0.15	0.09
	No	89,804	63.02	12.21	11.89	11.04	1.27	0.42	0.15
Airbag	Yes	13,257	18.43	21.54	22.54	12.89	6.39	7.45	10.76
	No	77,489	85.72	8.07	3.96	1.73	0.07	0.33	0.12
Crash Type	Side Impact Collision	29,724	68.54	16.82	10.13	4.01	0.25	0.18	0.07
	Side Rear-End Collision	210	92.38	6.19	0.95	0.48	0	0	0
	Side Collision	2010	87.75	5.72	4.03	1.89	0.05	0.4	0.16
	Parallel Side Collision	6400	90.28	6.5	2.48	0.64	0.02	0.08	0
	Same-Direction Collision	5722	93.8	4.67	1.1	0.38	0.02	0.03	0
	Head-On Side Collision	235	81.62	8.51	5.11	2.98	0	1.28	0.5
	Head-On Collision	595	84.95	7.23	4.2	2.86	0.34	0.34	0.08
	Frontal Collision	4448	52.43	17.65	16.03	11.6	1.89	0.4	0
	Rear-End Collision	40,554	81.67	12.77	4.28	1.12	0.03	0.13	0
Expressway	Other	848	94.1	4.6	0.83	0.47	0	0	0
Expressway	Yes	83,246	77.14	13.21	6.57	2.63	0.19	0.17	0.09
	No	7500	81.99	11.79	4.56	1.45	0.19	0.02	0
Weather	Sunny	70,047	77.54	13.13	6.32	2.55	0.19	0.17	0.1
	Foggy	204	74.02	14.22	7.35	3.92	0.49	0	0
	Cloudy	12,356	77.19	13.29	6.68	2.54	0.25	0.05	0
	Other	152	71.23	18.42	5.26	3.29	0.66	0.66	0.48
	Snowy	814	80.95	9.34	7.25	1.84	0.25	0.25	0.12
	Rainy	7173	77.99	12.67	6.62	2.36	0.15	0.14	0.07

The position encoding module is responsible for adding positional information to the fixed-size sequence data within the input embedding module. This enables the model to understand sequence positions and facilitates global learning. After adjusting the shape and processing the data to a fixed length, the resulting tensor, with a shape of (batch_size × seq_length × embedding_size), is input into the Transformer architecture, where batch_size represents the size of the data batch, seq_length represents the sequence length, and embedding_size represents the embedding size.

3.3. Experimental Methodology

The principal objective of this study was to develop a framework for forecasting the severity of traffic accidents based on human injuries. The prediction of crash accident severity is divided into seven categories, AIS0–AIS6, according to the level of injury sustained by the driver. The existing accident severity prediction models rely on a single data structure, and their efficacy heavily depends on the dataset’s quality. In response to these limitations, this study proposes an attention framework for predicting human injury severity in traffic accidents that can simultaneously analyze and process both types of data: text data and sequence data.

This study used 90% of the dataset for model training, while the remaining 10% was reserved for test validation. The proposed model was trained on the original dataset, and its effectiveness was validated using three different types of incident data: mixed data, sequence-only data, and text-only data.

Indicators for Model Evaluation

Experiments were conducted to assess the effectiveness of the proposed prediction model. The accident data were divided into a test set and a validation set, with prediction accuracy used as the evaluation metric, as shown in Equation (8).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (8)$$

where the true positive (TP) is the number of positive cases correctly identified as such, the true negative (TN) is the number of negative cases correctly identified as such, the false

negative (FN) is the number of positive cases incorrectly identified as negative, and the false positive (FP) is the number of negative cases incorrectly identified as positive.

A common metric for evaluating classification models is accuracy, which measures the percentage of correct predictions in a sample. However, in cases where data quality is poor, such as when the dataset is imbalanced or the sample size is inappropriate, using accuracy as the sole evaluation metric may lead to misinterpretations of the model's performance. In this study, we conducted a comprehensive assessment of the model's predictive performance by incorporating additional metrics: precision, recall, and F1 score. Precision is defined as the proportion of correct positive predictions among all positive results. The recall metric reflects the model's ability to identify positive class samples, quantifying the proportion of true positives that are correctly identified. The F1 score is the harmonic mean of precision and recall, providing a balanced measure of performance. This makes it an effective metric for assessing models with unbalanced datasets, offering a more holistic view of the model's predictive capability. The formulas are as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

4. Results and Discussion

4.1. Correlation Analysis of Sequence Data

In the statistical analysis and modeling of collisions, it is essential to conduct a correlation analysis of individual characteristic factors due to the potential for multicollinearity issues to arise among them. High correlations between multiple characteristic variables in a model can lead to instability and inaccuracies in model estimation, as well as compromise the simplicity of the model.

Figure 5 presents the correlation coefficient matrix for each factor involved in an accident. The darker the color, the higher the correlation between the two characteristic factors. The correlation coefficient matrix not only reveals the relationship between each characteristic factor and the severity of the accident but also shows the correlations among the characteristic factors themselves.

It can be concluded that, with regard to accident severity, the variable representing whether the airbag is activated (K) has the most significant influence, with a correlation coefficient of 0.26. Next, the magnitude of the speed at which the accident occurred (A) was found to have a correlation of 0.1 with accident severity. Other characteristic factors, such as alcohol consumption (L), collision mode (M, N), and lighting conditions (O), also showed a significant correlation with injury severity. Additionally, road structure features, such as the number of lanes and lane direction, as well as personal characteristics including the driver's age and gender, exhibited a slight correlation with injury severity.

It is evident that there is a high correlation between several characteristic factors, including vehicle mass (R) and vehicle size and shape characteristics (S, T, U, V, W), and a strong correlation between collision velocity (A), lane type (B), number of lanes (C), lane speed limit (D), lane direction (E), lane gradient (F), collision location (G), and driving environment (Q). This is because a vehicle's velocity is directly influenced and constrained by the driving environment and road structure. Therefore, features with high correlations are avoided. Based on the correlation coefficient matrix of each factor, these are the critical factors in efficiency and accuracy.

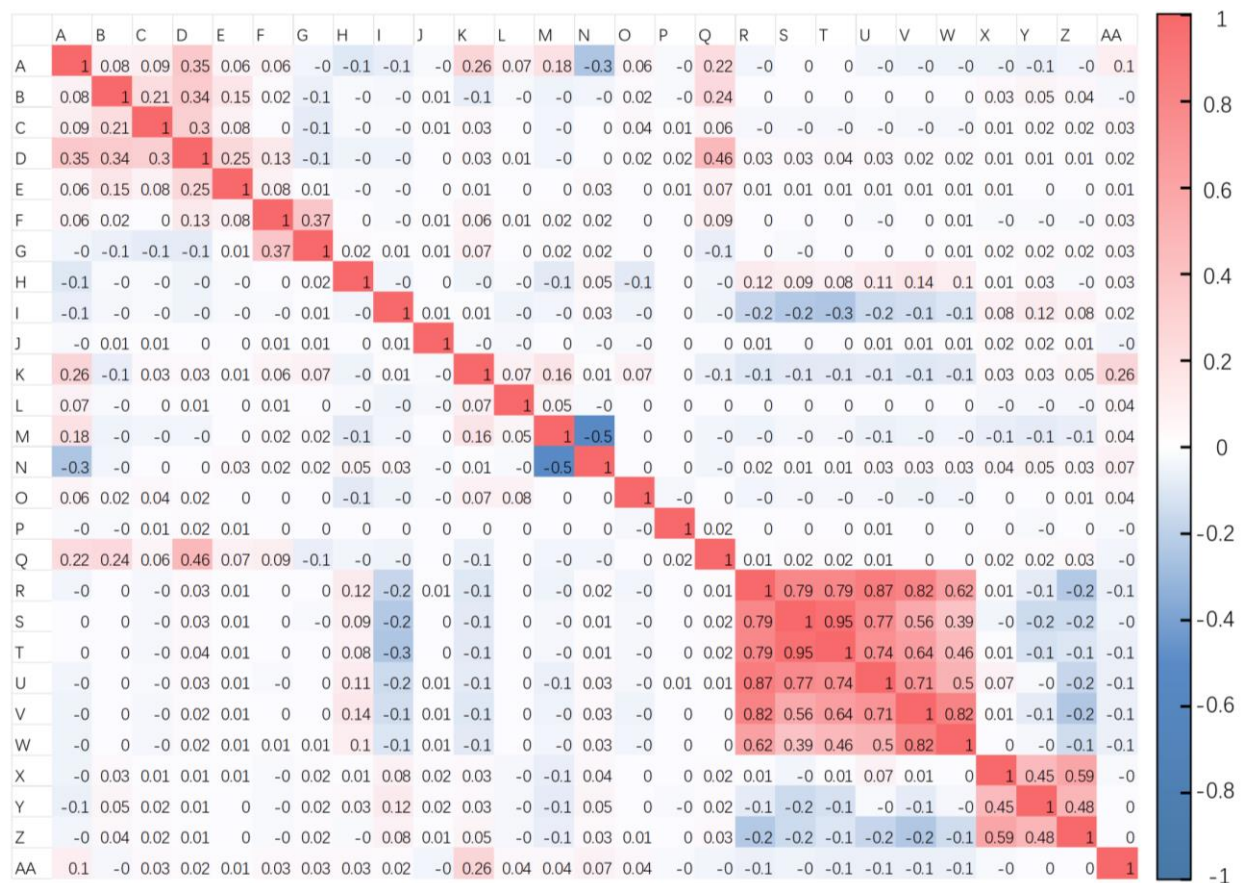


Figure 5. Matrix of correlation coefficients for crash parameters. A—Crash velocity; B—Lane type; C—Number of lanes; D—Lane speed limit value; E—Direction of lane; F—Gradient of lane; G—Crash location; H—Age; I—Sex; J—Seat belts; K—Airbags; L—Drunk driving; M—Crash site of own vehicle; N—Crash site of other vehicle; O—Lighting conditions; P—Weather; Q—Expressway; R—Overall mass; S—Length of the vehicle; T—Wheelbase; U—Width of the vehicle; V—Height of the vehicle; W—Minimum ground clearance; X—Side air curtains; Y—Emergency braking; Z—Front side airbags; and AA—AIS.

The key variables were selected based on their correlation with other factors and their relationship to injury severity. Nine critical factors were identified: the driver's age and gender, vehicle velocity, crash mode, airbag activation, drunk driving, ambient lighting levels, weather conditions, and driving environment. The dataset on human injury severity resulting from accidents was categorized into seven levels corresponding to the AIS scale. Table 1 provides descriptive statistics for the key determinants of injury severity among crash participants. It highlights each key factor's percentage distribution of injury severity levels in collisions. The table also shows the proportion of individuals experiencing different levels of injury severity, categorized by the crash characteristics.

As Table 1 shows, the severity of human injuries was concentrated in the minor injury classes AIS0 and AIS1, with the number of hours spent on the different injury classes showing an extremely uneven distribution.

Crash velocity is related to the severity of injuries to the driver in a collision. As the velocity of the vehicle increases, the probability that the driver will be injured in the crash will become bigger.

In terms of gender, there was no significant difference in the distribution of severity injuries, as shown in Figure 6. However, Table 1 shows that females were observed to be more susceptible to both minor and severe injuries during accidents. This difference is often attributed to lower average physical fitness levels compared with males. However,

the proportion of females involved in severe injury incidents was found to be lower than that of males. Figure 5 reveals a negative correlation between gender and collision velocity, indicating that females tend to drive at lower speeds compared with males due to differing driving habits.

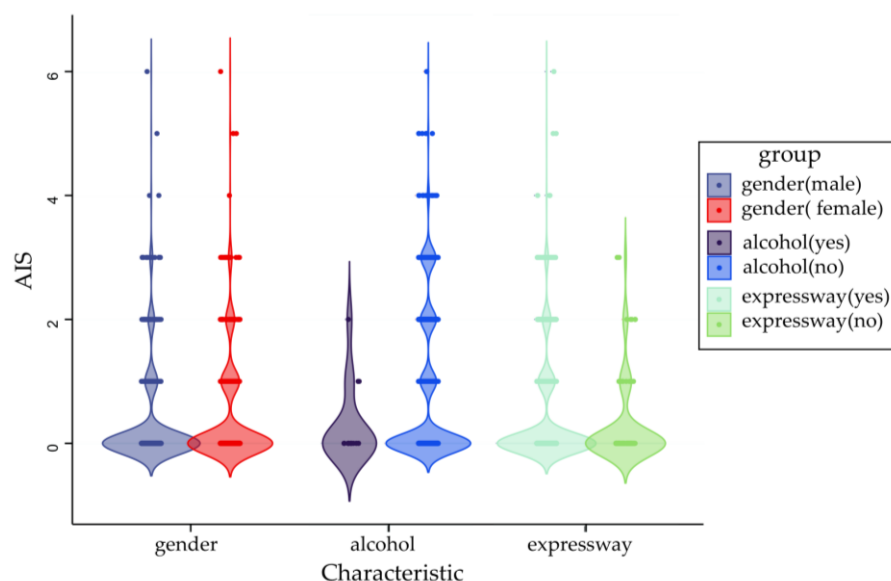


Figure 6. Distribution of accident severity under different characteristics.

To be clear, drivers who have consumed alcohol are significantly more likely to sustain serious injuries compared with those who have not. Moreover, the severity of injuries is also greater in such cases. As shown in Figure 5, alcohol consumption influences both the speed of the vehicle and the collision type. Figure 6 shows that drivers who consume alcohol cause less severe accidents than those who do not. This is due to the fact that the dataset records a smaller number of drink-driving accidents than regular accidents, which led to unbalanced data and a greater bias in the results.

An analysis of collision types and accident severity indicates that frontal collisions are the most fatal, primarily due to the high relative speeds involved. Secondly, although there is no significant relative speed in the same-direction side collisions, excessive driving speed increases the risk of loss of vehicle control. This can contribute to post-crash secondary collisions with surrounding vehicles or infrastructure.

Lighting conditions also play a critical role in accident severity. Accidents occurring under good lighting conditions tend to have lower severity levels compared with those in poor lighting conditions. This difference is attributed to the driver's improved ability to make timely judgments in well-lit environments. In contrast, low visibility at night increases the likelihood of inappropriate collision avoidance maneuvers.

Lastly, the analysis of the accident scenarios shows that a smaller proportion of low-severity crashes occur in highway scenarios compared with non-highway scenarios. This is likely due to better road conditions, more expansive fields of vision, improved traffic management, and fewer instances of sudden crossing incidents on highways. Furthermore, rear-end collisions are more common on highways, typically occurring at lower relative speeds, which contributes to the lower severity of injuries. However, Figure 6 shows that crashes on expressways tend to have a high severity that cannot be separated from the speed of the collision. This can be seen from the strong correlation between motorway and crash speed in Figure 5.

4.2. Predictive Performance Analysis

To evaluate the effectiveness of the proposed methodology, we compared the experimental results of this model with those of a model utilizing a single data source. A multilayer perceptron (MLP) was chosen to classify and predict the baseline model. The results are presented in Table 2. The prediction framework proposed in this study demonstrated superior performance in predicting human injury severity on both training and test sets, achieving a training accuracy of 0.94 and a test accuracy of 0.93, both of which outperformed the baseline model.

Table 2. Validation results for different data types.

	Training Accuracy	Test Accuracy
Mixed Data Validation		0.93
Only Textual Data Validation	0.94	0.92
Only Sequence Data Validation		0.89
Base	0.87	0.85

The hybrid data validation model, which integrates textual and sequential data, exhibited better prediction performance on the test set compared with models only using textual data or sequential data. Among the single-source models, the model utilizing sequential data significantly outperformed the one based on textual data. This can be attributed to the high correlation between accident sequence data and injury severity, which enhances the accuracy of predictions. Conversely, the textual dataset contains non-standardized content, leading to incomplete feature representation. As a result, the prediction accuracy of the model using only textual data was notably lower than that of the other models.

Since the classification of human injury severity is divided into seven categories, the accuracy metric presented in Table 2 may only partially capture the model's performance across different severity levels.

To address this, the prediction performance of the models was further evaluated using additional metrics, including precision, recall, and F1 score. Table 3 summarizes the prediction performance for each model across various injury severity grades. As shown in Table 3, the classification performance exhibited a clear trend: the lower the severity grade of human injury, the better the model's prediction performance, and vice versa.

Table 3. Model prediction performance for various injury severities.

Title 1	Injury Severity	Precision	Recall	F1-Score
Mixed Data Validation	AIS0	0.953	0.956	0.954
	AIS1	0.964	0.951	0.957
	AIS2	0.941	0.892	0.916
	AIS3	0.793	0.792	0.792
	AIS4	0.689	0.714	0.701
	AIS5	0.584	0.627	0.605
	AIS6	0.563	0.547	0.555
Only Textual Data Validation	AIS0	0.964	0.953	0.958
	AIS1	0.956	0.964	0.960
	AIS2	0.926	0.895	0.910
	AIS3	0.725	0.785	0.754
	AIS4	0.653	0.704	0.678
	AIS5	0.582	0.649	0.614
	AIS6	0.551	0.582	0.566
Only Sequence Data Validation	AIS0	0.935	0.927	0.931
	AIS1	0.912	0.895	0.903
	AIS2	0.713	0.732	0.722
	AIS3	0.568	0.624	0.595
	AIS4	0.437	0.536	0.481
	AIS5	0.349	0.438	0.388
	AIS6	0.324	0.452	0.377

4.3. Applicability and Limitations of the Proposed Method

The model proposed in this paper is a severity prediction of traffic accidents model with practical applications. It can provide government managers with proactive strategies in preventing traffic accidents, thus enhancing road safety and improving the current situation of traffic injuries and fatalities. Most existing crash severity prediction studies primarily focus on analyzing the severity of crashes from the perspective of the crash events themselves. Accident severity classification is basically based on the three levels of property damage accident, personal injury accident, and fatal accident. Undoubtedly, this method of accident severity definition has certain advantages in the acquisition of traffic accident data, the applicability of research methods, and the development of preventive measures. However, this severity classification index cannot accurately reflect the actual injury situation of the injured, nor can it fully take into account the differences between the severity of individual injuries and accident severity. The severity prediction of traffic accidents model proposed in this paper can predict the severity of driver injuries through the driving characteristics of vehicles involved in collisions. Most severity prediction of traffic accidents models use accident datasets with sequential accident feature content or textual data containing accident features. This reduces the biased features of structured data and the lack of semantic understanding of textual data. It also increases the accuracy of the prediction and the robustness of the model.

However, the proposed model has certain constraints and limitations. As with other predictive models, the main limitation of the model is the quality of the input dataset. Accident data, due to human collection, commonly suffer from inaccurate, incomplete and unstandardized records. The resulting problems of bias, noise, and missing values directly limit the performance of the prediction model. Secondly, the collection of human injury severity information is different from the processing of standardized information, which inevitably increases the difficulty of collection. These limit the research on severity prediction of traffic accidents models. Future research will focus on adapting the proposed framework to diverse traffic accident datasets, particularly those that lack AIS score information, to enhance the generalizability of the prediction model. One key challenge is addressing the issue of low dataset quality, which is primarily reflected in data imbalances. For instance, there is a significant disparity between the number of accidents with severity ratings of MAIS = 5 or MAIS = 6 and those with minor injury severity ratings. This imbalance negatively impacts the prediction accuracy of the model, especially for higher injury severity categories.

5. Conclusions

This study was based on the Transformer model for severity prediction of traffic collisions. The model was trained using Chinese road traffic accident data to learn the mapping relationship between crash characteristic factors and human injury severity. The dataset contains both structured sequence data and unstructured text data of accidents. This paper used AIS to quantify the severity of human injuries as a way of reflecting the severity of a collision. This allows the predictive model to more accurately reflect the impact of crashes on the human body, and it also helps to develop response strategies that are directly targeted at reducing injuries and fatalities.

The results of the study show that the prediction model proposed in this paper performed better than the base model and achieved better prediction accuracy when only textual or sequential data were used. Notably, the textual descriptions of accidents in this dataset were inconsistently formatted, demonstrating the capability of this framework to predict traffic accident severity, even with incomplete accident information. This facilitates the enhancement of the robustness and generalization of the prediction model. However,

the prediction performance at various injury severity grades showed extreme imbalance, which was limited by the uneven distribution of different damage severity levels in the accident dataset. Future efforts could focus on mitigating data imbalance issues to improve the prediction accuracy of the algorithm, particularly for severe injury ratings.

Author Contributions: The authors' contributions can be summarized as follows: conceptualization, Y.J., W.Y. and W.G.; methodology, Y.J., X.Q. and W.Y.; software, Y.J. and W.Z.; validation, Y.J. and Y.C.; formal analysis, Y.J. and Y.C.; resources, X.Q., J.X. and W.Z.; investigation, X.Q., W.Z. and J.X.; data curation, Y.J., Y.C. and J.X.; writing—original draft, Y.J. and W.G.; writing—review and editing, Y.J. and W.Y.; visualization, Y.J. and W.Y.; supervision, Y.J. and W.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to the Confidentiality and Protection of Personal Information Act.

Conflicts of Interest: Author Xianguo Qu was employed by the Defective Product Recall Technology Center of the State Administration of Market Supervision. Authors Weiwei Zhang, Jiejie Xu, Wangpengfei Yu and Yang Chen are employed by the Shanghai Intelligent Vehicle cooperating Innovation Center. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. World Health Organization. Global Status Report on Road Safety 2023, 1st ed. World Health Organization: Geneva, Switzerland, 2023.
2. Cicek, E.; Akin, M.; Uysal, F.; Topcu Aytas, R. Comparison of Traffic Accident Injury Severity Prediction Models with Explainable Machine Learning. *Transp. Lett.* **2023**, *15*, 1043–1054. [\[CrossRef\]](#)
3. Mussone, L.; Bassani, M.; Masci, P. Analysis of Factors Affecting the Severity of Crashes in Urban Road Intersections. *Accid. Anal. Prev.* **2017**, *103*, 112–122. [\[CrossRef\]](#)
4. Luan, S.; Jiang, Z.; Qu, D.; Yang, X.; Meng, F. Investigating Risk Factors Associated with Injury Severity in Highway Crashes: A Hybrid Approach Integrating Two-Step Cluster Analysis and Latent Class Ordered Regression Model with Covariates. *Accid. Anal. Prev.* **2024**, *208*, 107805. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Ghandour, A.J.; Hammoud, H.; Al-Hajj, S. Analyzing Factors Associated with Fatal Road Crashes: A Machine Learning Approach. *IJERPH* **2020**, *17*, 4111. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Xu, C.; Bao, J.; Wang, C.; Liu, P. Association Rule Analysis of Factors Contributing to Extraordinarily Severe Traffic Crashes in China. *J. Saf. Res.* **2018**, *67*, 65–75. [\[CrossRef\]](#)
7. Fu, C.; Tu, H.-T. Investigating Vehicle-Vehicle and Vehicle–Pedestrian Crash Severity at Street Intersections with the Latent Class Parameterized Correlation Bivariate Generalized Ordered Probit. *Accid. Anal. Prev.* **2024**, *207*, 107745. [\[CrossRef\]](#)
8. Christoforou, Z.; Cohen, S.; Karlaftis, M.G. Vehicle Occupant Injury Severity on Highways: An Empirical Investigation. *Accid. Anal. Prev.* **2010**, *42*, 1606–1620. [\[CrossRef\]](#)
9. Fountas, G.; Rye, T. A Note on Accounting for Underlying Injury-Severity States in Statistical Modeling of Injury Accident Data. *Procedia Comput. Sci.* **2019**, *151*, 202–209. [\[CrossRef\]](#)
10. Garrido, R.; Bastos, A.; De Almeida, A.; Elvas, J.P. Prediction of Road Accident Severity Using the Ordered Probit Model. *Transp. Res. Procedia* **2014**, *3*, 214–223. [\[CrossRef\]](#)
11. Du, B.; Zhang, C.; Sarkar, A.; Shen, J.; Telikani, A.; Hu, H. Identifying Factors Related to Pedestrian and Cyclist Crashes in ACT, Australia with an Extended Crash Dataset. *Accid. Anal. Prev.* **2024**, *207*, 107742. [\[CrossRef\]](#)
12. Katanalp, B.Y.; Eren, E. The Novel Approaches to Classify Cyclist Accident Injury-Severity: Hybrid Fuzzy Decision Mechanisms. *Accid. Anal. Prev.* **2020**, *144*, 105590. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Alkheder, S.; Taamneh, M.; Taamneh, S. Severity Prediction of Traffic Accident Using an Artificial Neural Network. *J. Forecast.* **2017**, *36*, 100–108. [\[CrossRef\]](#)
14. Sameen, M.; Pradhan, B. Severity Prediction of Traffic Accidents with Recurrent Neural Networks. *Appl. Sci.* **2017**, *7*, 476. [\[CrossRef\]](#)

15. Jamal, A.; Zahid, M.; Tauhidur Rahman, M.; Al-Ahmadi, H.M.; Almoshaogeh, M.; Farooq, D.; Ahmad, M. Injury Severity Prediction of Traffic Crashes with Ensemble Machine Learning Techniques: A Comparative Study. *Int. J. Inj. Control Saf. Promot.* **2021**, *28*, 408–427. [\[CrossRef\]](#)
16. Niyogisubizo, J.; Liao, L.; Lin, Y.; Luo, L.; Nziyumva, E.; Murwanashyaka, E. A Novel Stacking Framework Based On Hybrid of Gradient Boosting-Adaptive Boosting-Multilayer Perceptron for Crash Injury Severity Prediction and Analysis. In Proceedings of the 2021 IEEE 4th International Conference on Electronics and Communication Engineering (ICECE), Xi'an, China, 17–19 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 352–356. [\[CrossRef\]](#)
17. Cai, Q.; Abdel-Aty, M.; Yuan, J.; Lee, J.; Wu, Y. Real-Time Crash Prediction on Expressways Using Deep Generative Models. *Transp. Res. Part C Emerg. Technol.* **2020**, *117*, 102697. [\[CrossRef\]](#)
18. Rahim, M.A.; Hassan, H.M. A Deep Learning Based Traffic Crash Severity Prediction Framework. *Accid. Anal. Prev.* **2021**, *154*, 106090. [\[CrossRef\]](#)
19. Islam, M.R.; Wang, D.; Abdel-Aty, M. Calibrated Confidence Learning for Large-Scale Real-Time Crash and Severity Prediction. *npj. Sustain. Mobil. Transp.* **2024**, *1*, 1. [\[CrossRef\]](#)
20. Fan, Z.; Wang, P.; Zhao, Y.; Zhao, Y.; Ivanovic, B.; Wang, Z.; Pavone, M.; Yang, H.F. Learning Traffic Crashes as Language: Datasets, Benchmarks, and What-If Causal Analyses. *arXiv* **2024**, arXiv:2406.10789. [\[CrossRef\]](#)
21. Li, P.; Chen, S.; Yue, L.; Xu, Y.; Noyce, D.A. Analyzing Relationships between Latent Topics in Autonomous Vehicle Crash Narratives and Crash Severity Using Natural Language Processing Techniques and Explainable XGBoost. *Accid. Anal. Prev.* **2024**, *203*, 107605. [\[CrossRef\]](#)
22. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2023**, arXiv:1706.03762. [\[CrossRef\]](#)
23. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2019**, arXiv:1810.04805. [\[CrossRef\]](#)
24. Carion, N.; Massa, F.; Synnaeve, G.; Usunier, N.; Kirillov, A.; Zagoruyko, S. End-to-End Object Detection with Transformers. *arXiv* **2020**, arXiv:2005.12872. [\[CrossRef\]](#)
25. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv* **2021**, arXiv:2010.11929. [\[CrossRef\]](#)
26. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. *arXiv* **2021**, arXiv:2103.14030. [\[CrossRef\]](#)
27. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.; Dean, J. Distributed Representations of Words and Phrases and Their Compositionality. *arXiv* **2013**, arXiv:1310.4546. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.