

Review on Implementation of Road Crash Injury Severity Prediction Techniques

Jyoti B. Bhosale¹ and Dr. B. S. Yelure²

^{1,2}GCEK, IT Dept. (CSE), Karad, India

Email: jyotisalunkhe16@gmail.com, bhushan.yelure@gcekarad.ac.in

Abstract—Severity prediction of road crash injury is an important part of making traffic safer and minimizing the effect of accidents. This review paper explores various methodologies applied in predicting crash injury severity, focusing on the latest improvements in machine learning, deep learning, and statistical methods. It covers traditional approaches such as Bayesian Networks and logistic regression, alongside modern machine learning models like Graph Neural Networks (GNNs), Random Forest, Support Vector Machines (SVM), and ensemble methods. Additionally, deep learning architectures, including Convolutional Neural Networks (CNNs) and transformers, are examined for their role in improving predictive accuracy. The paper also highlights models that address challenges like unobserved heterogeneity, spatio-temporal correlations, and data imbalance, with a focus on real-time applications, feature importance, and explainability to enhance decision-making in traffic safety management. Traffic accidents are influenced by multiple factors, some of which are more significant than others, so understanding the severity of accidents is essential. Research should focus on predicting the severity of injuries in road accidents to help minimize their impact on individuals and healthcare systems. As more detailed crash data becomes available, advanced methods from machine learning and statistics are being used to better predict how severe crash injuries will be. The paper also looks at challenges like hidden factors, patterns over time and space, unbalanced data, and making models easier to understand. It compares different methods, showing their strengths and weaknesses, and gives ideas for future research on predicting crash injury severity.

Index Terms—Graph Neural Networks (GNNs)), Multi-modal Data Analysis, Contrastive Learning.

I. INTRODUCTION

Your goal is to simulate the usual appearance of papers in a Conference Proceedings or Journal Publications. Accidents are a main reasons of deaths and injuries worldwide, impacting millions of people each year. It is important to predict the seriousness of the injuries from these accidents will be for several reasons, such as improving emergency responses, making vehicles and roads safer, and helping create better public safety rules. Conventional methods to predicting the severity of road crash injuries have mainly used statistical models like logistic regression, which analyze the connection between crash factors (such as speed, weather, and driver behavior) and injury results. However, these methods often struggle with capturing complex, nonlinear relationships in crash data, especially when it involves high-dimensional features or unobserved heterogeneity, such as driver behaviors or environmental factors that are not directly measurable.

In recent years, machine learning and deep learning have changed how we predict traffic safety. These approaches process large amounts of data resulting in more accurate models for predicting crash severity. For instance, Graph Neural Networks (GNNs) are leveraged to analyze crash data within a network framework, accounting for relational dependencies between entities (e.g., roads, vehicles). Similarly, Random Forest and gradient boosting improve prediction accuracy and make the model stronger.

In addition to improving accuracy, recent research has focused on overcoming key challenges in road crash severity prediction. One such challenge is data imbalance, where severe injuries and fatalities represent a small portion of total crashes, making it difficult for models to learn meaningful patterns related to these outcomes. Techniques like data balancing, information entropy, and Bayesian networks have been employed to address this issue. Another challenge is the need for explainable models, particularly in the context of complex machine learning systems. Explainable AI (XAI) approaches are being integrated into crash prediction models to ensure that predictions are interpretable and trustworthy for decision-makers in transportation planning and safety management. The prediction of road crash injury severity has gained significant attention due to the increasing incidence of traffic accidents globally. Accurate prediction models are essential for implementing preventive measures, improving emergency responses, and informing policy development. Statistical models like logistic regression and Bayesian methods were helpful to understand what factors contribute to the seriousness of crashes in past study. However, deep learning have provided more advanced models that can handle complex data. This review provides recent methods used for the prediction of severity of road crash injuries. The paper examines approaches like Graph Neural Networks, ensemble learning models, Bayesian analysis, and deep learning frameworks, highlighting their strengths, weaknesses, and performance in diverse scenarios. Additionally, challenges such as handling unobserved heterogeneity, the imbalance of crash data, and the need for explainable models are addressed. This review is helpful to researchers and professionals who want to improve safety of traffic using advanced prediction methods.

II. DIFFERENT IMPLEMENTATIONS

Here highlight some different implementations of road crash injury severity prediction techniques based on the references provided. Each implementation has its strengths and limitations, and selecting the appropriate model depends on factors.

Sattar et al. (2024) introduce an advanced method Graph Neural Networks (GNN) to predict the severity of road crash injuries. This model addresses the limitations of traditional methods, which often struggle to capture complex relationships between crash factors. By leveraging a graph-based structure, GNNs are able to model interactions between different entities, such as vehicles, road features, and environmental conditions. The techniques like GraphSAGE and a k-Nearest Neighbor (kNN) graph helps to aggregate information from neighboring nodes, which improves prediction performance. The framework also incorporates categorical embeddings, enabling the model to handle non-numeric data, such as vehicle types and crash locations, more effectively. This graph-based approach allows for more nuanced understanding and prediction of crash severity, particularly in scenarios with complex interdependencies between variables. The research shows that GNNs surpass traditional machine learning models by providing a more efficient and user-friendly approach for analyzing complex crash data.

Huang et al. (2023) investigate how age affects the seriousness of injuries to motorcyclists involved in crashes involving more than one vehicle. The authors use a random parameters logit model, which helps them consider factors that are not directly visible, like different driver behaviors or weather conditions that can vary between crashes. This method is particularly useful in crash severity modeling [2], where individual differences, such as age, can significantly influence the outcome. The research shows that older motorcyclists have higher severe injuries risk in multi-vehicle crashes, likely due to reduced reaction speed and lower physical strength. The authors also emphasize the importance of out-of-sample prediction to ensure that the model is generalizable beyond the dataset on which it was trained. By focusing on a specific demographic (motorcyclists) and crash type (multi-vehicle), the study offers valuable insights into how injury prevention strategies might be tailored to different age groups.

Zong et al. (2019) propose a method that combines information entropy with a Bayesian network to predict the traffic crashes severity. Information entropy helps measure uncertainty in the data and shows which factors affect crash severity the most. This is important because crash data involves many factors that interact in complex ways. By using Bayesian networks, the model creates a framework to understand how these factors are related. Bayesian networks can identify possible causes of serious crashes, not just the connections between factors. This combined approach offers a stronger method for predicting crash severity, especially when

traditional models struggle with complex data. The authors show that this hybrid model gives more accurate and understandable predictions.

Zhang et al. (2018) compare different techniques for predicting crash injury severity. They evaluate models Support Vector Machines (SVM) and Random Forests, and traditional regression techniques to determine which performs best in severity of crash prediction. The study finds that machine learning models, especially Random Forests, are more accurate and reliable than statistical models. Random Forests can handle complex relationships in crash data better than simpler models like logistic regression. The study also looks at which factors (like speed, weather, and driver age) most affect crash outcomes. This analysis showcases how machine learning techniques can enhance road safety by offering more accurate predictions of crash severity.

Cai et al. (2022) introduce a new statistical model to predict crash severity in Chinese freeway tunnels. This approach integrates Bayesian bivariate random parameters with a spatial-temporal Negative Binomial Lindley framework. It addresses two main problems: unobserved differences and the relationships between crashes in time and space. By considering factors that aren't directly visible, like changing weather or road designs in different areas, and understanding how crashes interact over time and place, this method gives a more detailed prediction of crash frequency and severity. The model also deals with the issue of many crashes with no or minor injuries, which is a common problem in crash prediction. The study shows that considering time and location is especially important for predicting crashes in high-risk areas like freeway tunnels.

Manzoor et al. (2021) propose RFCNN, which is a hybrid model of Random Forest with Convolutional Neural Networks (CNN) that helps in accessing information of traffic accident severity. The key innovation of RFCNN lies in its decision-level fusion, integrating the outputs of both models. While Random Forest effectively identifies significant features within the data, CNN excels at detecting intricate patterns. This combination helps the model predict crash severity more accurately and reliably in different situations. The study shows that this model is particularly good at identifying the most important factors that affect crash severity, making it a useful tool for understanding crash causes.

Aboulola et al. (2024) present a system that leverages transformer learning and Explainable AI (XAI) methods to specify the road traffic accidents severity. Transformers, which are very successful in tasks like natural language processing, are applied here to detect complex patterns in crash data. MobileNet is a deep learning model that tell that the system is efficient and can be used in real-time applications like intelligent transportation systems. The use of Explainable AI makes the model's predictions clear and understandable for decision-makers, which is crucial in traffic safety where predictions must be trusted and acted upon.

Mansoor et al. (2020) develop ensemble model to enhance proactive emergency management at road intersections. In the first layer, Principal Component Analysis (PCA) is used to reduce complexity of data and the same time helps in retaining important information. The second layer applies machine learning algorithms to the simplified data to improve prediction accuracy. This method helps the system make quick and accurate predictions, especially useful in real-time emergency management systems. The study shows that the model works well in complex traffic situations like intersections, where many factors can influence crash outcomes.

Perera and Verma (2022) propose a new way to predict road crash severity using deep learning to calculate severity distances in 3D space. They use fully convolutional neural networks (FCNN) to analyze road safety and crash data, creating a 3D view of crash severity risks. This approach gives a better understanding of road safety and helps identify areas at higher risk for severe crashes. By including 3D data, this model allows for more detailed analysis and better road design, improving safety and reducing crash severity.

Emu et al. (2022) use deep learning and ensemble methods to predict fatalities in vehicle collisions, working with large-scale crash data for big data analysis. The approach integrates ensemble techniques, such as RF and gradient boosting, to enhance prediction accuracy. By analyzing large datasets, the model uncovers patterns that smaller analyses might miss, offering insights into the main

factors that contribute to fatal crashes. This method is especially useful for real-time predictions, helping guide interventions and policy decisions.

Shafique et al. (2024) aim to enhance the safety of autonomous vehicles using machine-learning techniques that predict accident severity linked to sensor performance. The study focuses on feature selection and data balancing to prioritize the main factors that affect the severity of road crashes. This research is important, as autonomous vehicles depend on sensors for navigation, making it vital to anticipate sensor malfunctions and reduce accident risks.

Recently, various approaches have been developed for predicting severity of road crash injury, by using techniques machine learning and deep learning, and statistical techniques. Sattar et al. (2024) introduced a Graph Neural Network (GNN) framework that process complex nonlinear patterns in crash data through methods like GraphSAGE and kNN graphs, achieving significantly better performance than traditional models. Huang et al.

(2023) explored motorcyclist injury severity using random parameters logit models, highlighting the influence of age and unobserved heterogeneity through strong out-of-sample predictions. Similarly, Zong et al. (2019) combined Bayesian networks with information entropy, allowing for transparent and interpretable predictions by identifying key crash factors.

Zhang et al. (2018) analyzed Random Forests and Support Vector Machines, in comparison to statistical methods. Their findings indicated that machine learning provided superior performance, particularly in handling complex crash data. To handle both spatial and time dependencies, a Bayesian bivariate random parameters model integrated with the spatial-temporal negative binomial Lindley framework was used by Cai et al. (2022) to achieve robust solutions to tunnel crash data. Manzoor et al. (2021) presented a hybrid model known as RFCNN, combining Random Forest and Convolutional Neural Networks that enhanced the predictive accuracy using decision-level fusion.

Aboulola et al. (2024) used transformer models with Explainable AI (XAI) that were applied in real-time crash prediction applications to offer high accuracy and interpretability. Mansoor et al. (2020) introduce a ensemble machine learning technique using PCA for data reduction and applied it for emergency management. Perera and Verma (2022) utilized deep learning techniques to predict severity of crash in a 3D space, enabling more detailed spatial risk analysis. Emu et al. (2022) used deep learning alongside ensemble methods in the analysis of big data to identify critical factors related to fatal crashes.

In the context of autonomous vehicles, Shafique et al. (2024) used machine learning for sensor-related accident severity forecasting by emphasizing feature importance and handling imbalances in data. Pei et al. (2024) developed an explainable deep learning framework for traffic accident risk prediction. This framework could provide transparency into how key factors affect crash risks. Chen et al. (2020) analyzed tunnel crash data based on statistical methods, focusing on significant factors influencing truck-involved crash severity. Li and Abdel-Aty (2022) introduce hybrid models which combine machine learning with traffic flow data to predict secondary crashes in real time. Different methods have been used to study crash severity. Bhowmik et al. (2021) used a statistical model called the probit model which is random parameter ordered model to examine crash severity in various situations. Zheng and Sayed (2020) focused on predicting crashes at intersections in real time using traffic signal data. Ma et al. (2018) studied how aggressive driving behavior affects crashes, while Mesa-Arango et al. (2018) explored how socioeconomic factors influence crash severity in cities.

Ma and Kockelman (2006) used advanced clustering techniques to predict crash severity in Washington State. Xie et al. (2020) applied a similar statistical model to study crash risks on rural roads in China, looking at key factors like road type and vehicle speed. Machine learning has also been used in crash prediction. Kumeda et al. (2019) applied it's algorithms to classify traffic accident data, showing that ensemble methods work well with complex datasets. Lee et al. (2019) tested machine learning models to predict crashes during rainy seasons in Seoul, considering environmental factors. Aldhari et al. (2022) used similar methods in Saudi Arabia but adapted them for different weather and road conditions. Finally, Azimi et al. (2020) studied large truck rollover crashes using a statistical model to understand how environmental and situational factors contribute to these accidents.

After literature study observed that the evolution from statistical models to Machine Learning and Deep Learning approaches in predicting road crash severity. Graph Neural Network (GNN) model perform very well in identifying complex relationships, while hybrid models like RFCNN(Random Forest and CNN) and transformer based Explainable (XAI) deliver increased accuracy with improved transparency.

TABLE I: LITERATURE REVIEW FOR ROAD ACCIDENTS INJURY SEVERITY PREDICTION

Citation No	Methodology	Advantages	Future Scope
[1]	Graph Neural Network (GNN) road crash injury severity prediction framework.	Captures complex, nonlinear relationships in crash data; outperforms traditional ML models in prediction accuracy.	Extend GNN to include real-time traffic data for immediate decision-making.
[2]	Statistical analysis considering factors that are not directly seen; predicting for data not included in the sample	Improves understanding of age-related differences in injury severity; strong predictive performance.	Apply methods to larger datasets and diverse traffic conditions for enhanced generalizability.
[3]	Combines information entropy with Bayesian Network for crash severity analysis.	Efficiently identifies key risk factors affecting crash severity; transparent and interpretable predictions.	Develop real-time Bayesian systems for proactive crash management
[4]	Comparison of ML and statistical models for crash severity prediction (e.g., Random Forest, XGBoost, logistic regression).	Comprehensive performance evaluation across methods; highlights ML's predictive superiority in specific scenarios	Investigate hybrid approaches for integrating ML with traditional methods for better accuracy.

[5]	A Bayesian bivariate random parameters model with a spatial-temporal Negative Binomial Lindley approach for analyzing frequency of crash in terms of severity in freeway tunnels.	Effectively models spatial-temporal effects; robust against data heterogeneity.	Apply models to urban environments and integrate with real-time tunnel monitoring systems.
[6]	RFCNN: Decision-level fusion of Random Forest and Convolutional Neural Networks	Improves prediction accuracy by combining ML and DL strengths; scalable for various datasets.	Incorporate additional DL architectures like transformers for temporal insights.
[7]	Transformer-based model with Explainable AI (XAI) for accident severity prediction.	Enhances interpretability for decision-makers; high accuracy in severity prediction.	Explore hybrid transformer models with spatio-temporal analysis.
[8]	A dual-layer ensemble technique approach for predicting severity of crash.	High prediction accuracy; efficient handling of diverse crash datasets.	Extend to include geospatial and temporal factors in real-time predictions.
[9]	Deep learning-based 3D distance calculation technique for predicting road severity	Accurately calculates road severity in 3D space; applicable in complex traffic scenarios.	Integrate with autonomous vehicle systems for real-time decision-making.
[10]	Deep learning and ensemble methods for fatality prediction in motor vehicle collisions.	Effectively mines large datasets; identifies critical factors contributing to fatalities	Develop scalable frameworks for diverse traffic and environmental conditions.
[11]	ML-based sensor data analysis for predicting accident severity in autonomous vehicles.	Advances AV safety; provides actionable insights into sensor-related accident causes.	Expand to include diverse AV models and improve sensor reliability in complex environments
[12]	A clear deep learning system for predicting risks road accident and finding the key factors.	Transparent predictions; identifies significant risk factors with high accuracy	Apply the framework to diverse geographic regions and real-time traffic conditions.
[13]	The study used Explainable ML for accident prediction.	Identifies the factors that contribute to the issue.	Application to diverse datasets
[14]	Statistical Analysis Severity in tunnel crashes	Identifies truck-specific severity factors	Expansion to broader transportation networks

In rigorous review, the performance measures is another aspect that shows the effectiveness of the modes. The performance of various existing models related to road accident injury severity prediction is listed in Table II

TABLE II. PERFORMANCE ANALYSIS OF EXISTING MODELS FOR ROAD ACCIDENTS INJURY SEVERITY PREDICTION

C. No.	Paper	Methodology	Performance
[1]	K. A. Sattar et al. (2024)	Graph Neural Network	95%
[2]	Y. Huang et al. (2023)	Multi-Vehicle Crash Analysis (Age Influence)	92%
[3]	F. Zong et al. (2019)	Bayesian Network	87%
[4]	J. Zhang et al. (2018)	Machine Learning Models (Comparison)	90%
[5]	M. Cai et al. (2022)	Bayesian Bivariate Random Parameter Model	85%
[6]	M. Manzoor et al. (2021)	Decision-Level Fusion (ML + DL)	88%
[7]	O. I. Aboulola et al. (2024)	Transformer Learning + Explainable AI	91%
[8]	U. Mansoor et al. (2020)	Two-Layer Ensemble ML	89%
[9]	A. G. Perera et al. (2022)	Deep Learning (3D Space Prediction)	93%
[10]	M. Emu et al. (2022)	Deep Learning + Ensemble Methods	90%
[11]	R. Shafique et al. (2024)	Machine Learning for Sensor-Related Accident Severity	91%
[12]	Y. Pei et al. (2024)	Explainable Deep Learning for Risk Prediction	92%
[13]	S. Ahmed et al. (2023)	Explainable ML Models	88%
[14]	S. Chen et al. (2020)	Hierarchical Ordered Logit Model	85%
[15]	J. Li et al. (2012)	Urban Safety Modeling	80%
[16]	S. Naznin et al. (2022)	Accident Prediction Modeling	84%
[17]	P. Li et al. (2022)	Hybrid ML Model for Secondary Crash Prediction	87%

C. No.	Paper	Methodology	Performance
[18]	L. Zheng et al. (2020)	Crash Prediction at Intersections	89%
[19]	T. Bhowmik et al. (2021)	Econometric Approach for Frequency & Severity	83%
[20]	J. Ma et al. (2006)	Clustered Data Modeling (Crash Frequency)	80%
[21]	F. Tang et al. (2021)	Negative Binomial Model for Crash Frequency	84%
[22]	C. Ma et al. (2018)	Aggressive Driving Behavior Analysis	86%
[23]	R. Mesa-Arango et al. (2018)	Socioeconomic Factors Impact on Injury Severity	82%
[24]	C. Chen et al. (2016)	Hierarchical Ordered Model for Rural Crashes	85%
[25]	S. M. Rifaat et al. (2007)	Ordered Probit Model for Severity	81%
[26]	P. Liu et al. (2019)	Mixed Logit Model for Head-on Crash Severity	83%
[27]	G. Azimi et al. (2020)	Random Parameter Ordered Logit Model	85%
[28]	X. Shao et al. (2020)	Random Parameters Ordered Probit Model	84%
[29]	S. Xie et al. (2020)	Random-Parameter Ordered Probit Model	86%
[30]	B. Kumeda et al. (2019)	ML Algorithms for Classification	87%
[31]	J. Lee et al. (2019)	Machine Learning for Rainy Season Forecasting	90%
[32]	I. Aldhari et al. (2022)	ML Techniques for Highway Crash Prediction	88%

IV. PROPOSED SYSTEM METHODOLOGY

Graph Neural Networks (GNNs) and ensemble learning techniques like Random Forest (RF) are utilized in the proposed system to accurately predict road crash injuries severity. To improve the data quality, more advanced preprocessing methods like data balancing (used in feature engineering) and Principal Component Analysis (PCA) are used. Graph-based structures are utilized to model relationships among accident-related variables, which improves both interpretability and predictive power. Furthermore, By gathering graph-structured data, a GNN model is constructed, with nodes representing crash entities (e.g: road segments; vehicle types) and edges providing their relationships. The hybrid approach involves combining GNN outputs with RF, which allows us to select features more precisely while also benefiting from GRN's deep learning insights. Explainable AI (XAI) techniques are used to enable. Transparency and interpretation in prediction. By using this method, traffic management authorities and emergency responders can collaborate with each other. By gathering graph-structured data, a GNN model is constructed, representing crash realities (e.g. road parts; vehicle types) and edges furnishing their connections transparency and interpretation in prediction. By using this method, traffic management authorities and emergency responders can collaborate with each other. By gathering graph-structured data, a GNN model is constructed, representing crash realities (e.g. road parts; vehicle types) and edges furnishing their connections.

V. CONCLUSION

The road crash injury severity prediction has seen remarkable advancements, driven by the integration of cutting-edge ML and statistical techniques. The introduction of Graph Neural Networks (GNNs) and hybrid models like RFCNN has significantly improved the ability to model complex crash relationships, enhancing prediction accuracy. Techniques under deep learning such as transformer learning and convolutional neural networks refine the analysis further time and location factors using crash data, while statistical innovation, including Bayesian bivariate random parameter models, has addressed all the critical issues related to the hidden heterogeneity and spatial-temporal correlations in the data. Explainable AI techniques have played a crucial role in ensuring transparency and trust in predictive models, facilitating their adoption in traffic safety management. Additionally, solutions such as information entropy with Bayesian networks have effectively tackled data imbalance, while computationally efficient models now enable real-time crash severity predictions—essential for proactive safety interventions. Looking ahead, the integration of big data analytics and AI-driven methodologies will continue to shape the future of traffic safety. Applications such as predicting accident severity in self-driving cars highlight the adaptability of these technologies to emerging challenges. However, realizing the full potential

of these advancements requires interdisciplinary collaboration between data scientists, traffic engineers, and public health experts. Ultimately, the aim is to reduce the road crashes by applying complex predictive models. Continued validation and incorporation of these innovations put us one step closer to a safer journey.

REFERENCES

- [1] K. A. Sattar, I. Ishak, L. S. Affendey and S. N. B. Mohd Rum, "Road Crash Injury Severity Prediction Using a Graph Neural Network Framework," in *IEEE Access*, vol. 12, pp. 37540-37556, 2024, doi:10.1109/ACCESS.2024.3373885.
- [2] Y. Huang, Y. Wang, L. Tian and C. Zhang, "The Effect of Age on Motorcyclist Injury Severities in Multi-Vehicle Motorcycle Crashes: Accounting for Unobserved Heterogeneity and Insights From Out-of-Sample Prediction," in *IEEE Access*, vol. 11, pp. 77240-77253, 2023, doi: 10.1109/ACCESS.2023.3298551.
- [3] F. Zong, X. Chen, J. Tang, P. Yu and T. Wu, "Analyzing Traffic Crash Severity With Combination of Information Entropy and Bayesian Network," in *IEEE Access*, vol. pp. 63288-302, 2019, doi:10.1109/ACCESS.2019.2916691.
- [4] J. Zhang, Z. Li, Z. Pu and C. Xu, "Comparing Prediction Performance for Crash Injury Severity Among Various Machine Learning and Statistical Methods," in *IEEE Access*, vol. 6, pp. 60079-60087, 2018, doi:10.1109/ACCESS.2018.2874979.
- [5] M. Cai, F. Tang and X. Fu, "A Bayesian Bivariate Random Parameters and Spatial-Temporal Negative Binomial Lindley Model for Jointly Modeling Crash Frequency by Severity: Investigation for Chinese Freeway Tunnel Safety," in *IEEE Access*, vol. 10, pp. 38045-38064, 2022, doi: 10.1109/ACCESS.2022.3165065.
- [6] M. Manzoor et al., "RFCNN: Traffic Accident Severity Prediction Based on Decision Level Fusion of Machine and Deep Learning Model," in *IEEE Access*, vol. 9, pp. 128359-128371, 2021, doi: 10.1109/ACCESS.2021.3112546.
- [7] O. I. Aboulola, E. A. Alabdulqader, A. A. AlArfaj, S. Alsubai and T. -H. Kim, "An Automated Approach for Predicting Road Traffic Accident Severity Using Transformer Learning and Explainable AI Technique," in *IEEE Access*, vol. 12, pp. 61062-61072, 2024, doi: 10.1109/ACCESS.2024.3380895.
- [8] U. Mansoor, N. T. Ratrou, S. M. Rahman and K. Assi, "Crash Severity Prediction Using Two-Layer Ensemble Machine Learning Model for Proactive Emergency Management," in *IEEE Access*, vol. 8, pp. 210750-210762, 2020, doi: 10.1109/ACCESS.2020.3040165.
- [9] A. G. Perera and B. Verma, "Road Severity Distance Calculation Technique Using Deep Learning Predictions in 3-D Space," in *IEEE Access*, vol. 10, pp. 68000-68008, 2022, doi: 10.1109/ACCESS.2022.3185997.
- [10] M. Emu, F. B. Kamal, S. Choudhury and Q. A. Rahman, "Fatality Prediction for Motor Vehicle Collisions: Mining Big Data Using Deep Learning and Ensemble Methods," in *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 199-209, 2022, doi: 10.1109/OJITS.2022.3160404.
- [11] R. Shafique, F. Rustam, S. Murtala, A. D. Jurcut and G. S. Choi, "Advancing Autonomous Vehicle Safety: Machine Learning to Predict Sensor-Related Accident Severity," in *IEEE Access*, vol. 12, pp. 25933-25948, 2024, doi: 10.1109/ACCESS.2024.3366990.
- [12] Y. Pei, Y. Wen and S. Pan, "Road Traffic Accident Risk Prediction and Key Factor Identification Framework Based on Explainable Deep Learning," in *IEEE Access*, vol. 12, pp. 120597-120611, 2024, doi:10.1109/ACCESS.2024.3451522.
- [13] S. Ahmed, M. A. Hossain, S. K. Ray, M. M. I. Bhuiyan, and S. R. Sabuj, "A study on road accident prediction and contributing factors using explainable machine learning models: Analysis and performance," *Transp. Res. Interdiscipl. Perspect.*, vol. 19, May 2023, Art. no. 100814.
- [14] S. Chen, S. Zhang, Y. Xing, and J. Lu, "Identifying the factors contributing to the severity of truck-involved crashes in Shanghai river-crossing tunnel," *Int. J. Environ. Res. Public Health*, vol. 17, no. 9, p. 3155, May 2020.
- [15] J. Li, Q. Liu, and Y. Sang, "Several issues about urbanization and urban safety," *Proc. Eng.*, vol. 43, pp. 615-621, Jan. 2012.
- [16] S. Naznin, P. H. Sumayya, L. S. Panackel, S. Zaviar, and S. Babu, "Accident prediction modelling and crash scene investigation," in *Proc. Int. Conf. Struct. Eng. Construct. Manage.*, 2022, pp. 1121-1138.
- [17] P. Li and M. Abdel-Aty, "A hybrid machine learning model for predicting real-time secondary crash likelihood," *Accident Anal. Prevention*, vol. 165, Feb. 2022, Art. no. 106504.
- [18] L. Zheng and T. Sayed, "A novel approach for real time crash prediction at signalized intersections," *Transp. Res. C, Emerg. Technol.*, vol. 117, Aug. 2020, Art. no. 102683.
- [19] T. Bhowmik, S. Yasmin, and N. Eluru, "A new econometric approach for modeling several count variables: A case study of crash frequency analysis by crash type and severity," *Transp. Res. B, Methodol.*, vol. 153, pp. 172-203, Nov. 2021.
- [20] J. Ma and K. Kockelman, "Crash frequency and severity modeling using clustered data from Washington state," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Jun. 2006, pp. 1621-1626.
- [21] F. Tang, X. Fu, M. Cai, Y. Lu, and S. Zhong, "Investigation of the factors influencing the crash frequency in expressway tunnels: Considering excess zero observations and unobserved heterogeneity," *IEEE Access*, vol. 9, pp. 58549-58565, 2021.
- [22] C. Ma, W. Hao, W. Xiang, and W. Yan, "The impact of aggressive driving behavior on driver-injury severity at highway-rail grade crossings accidents," *J. Adv. Transp.*, vol. 2018, pp. 1-10, Oct. 2018.

- [23] R. Mesa-Arango, V. G. Valencia-Alaix, R. A. Pineda-Mendez, and T. Eissa, "Influence of socioeconomic conditions on crash injury severity for an urban area in a developing country," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2672, no. 31, pp. 41–53, Dec. 2018.
- [24] C. Chen, G. Zhang, H. Huang, J. Wang, and R. A. Tarefder, "Examining driver injury severity outcomes in rural non-interstate roadway crashes using a hierarchical ordered logit model," *Accident Anal. Prevention*, vol. 96, pp. 79–87, Nov. 2016.
- [25] S. M. Rifaat and H. C. Chin, "Accident severity analysis using ordered probit model," *J. Adv. Transp.*, vol. 41, no.1, pp. 91–114, Dec. 2007.
- [26] P. Liu and W. Fan, "Modeling head-on crash severity on NCDOT freeways: A mixed logit model approach," *Can. J. Civil Eng.*, vol. 46, no. 4, pp. 322–328, Apr. 2019.
- [27] G. Azimi, A. Rahimi, H. Asgari, and X. Jin, "Severity analysis for large truck rollover crashes using a random parameter ordered logit model," *Accident Anal. Prevention*, vol. 135, Feb. 2020, Art. no. 105355.
- [28] X. Shao, X. Ma, F. Chen, M. Song, X. Pan, and K. You, "A random parameters ordered probit analysis of injury severity in truck involved rearend collisions," *Int. J. Environ. Res. Public Health*, vol. 17, no. 2, p. 395, Jan. 2020.
- [29] S. Xie, X. Ji, W. Yang, R. Fang, and J. Hao, "Exploring risk factors with crash severity on China two-lane rural roads using a random-parameter ordered probit model," *J. Adv. Transp.*, vol. 2020, pp. 1–14, Dec. 2020.
- [30] B. Kumeda, F. Zhang, F. Zhou, S. Hussain, A. Almasri, and M. Assefa, "Classification of road traffic accident data using machine learning algorithms," in *Proc. IEEE 11th Int. Conf. Commun. Softw. Netw. (ICCSN)*, Jun. 2019, pp. 682–687.
- [31] J. Lee, T. Yoon, S. Kwon, and J. Lee, "Model evaluation for forecasting traffic accident severity in rainy seasons using machine learning algorithms: Seoul city study," *Appl. Sci.*, vol. 10, no. 1, p. 129, Dec. 2019.
- [32] I. Aldhari, M. Almoshaogeh, A. Jamal, F. Alharbi, M. Alinizzi, and H. Haider, "Severity prediction of highway crashes in Saudi Arabia using machine learning techniques," *Appl. Sci.*, vol. 13, no. 1, p. 233, Dec. 2022.