PREDICTING PARATRANSIT CRASH SEVERITY USING MACHINE LEARNING: A CASE STUDY OF GAZIPUR, BANGLADESH

Joydeep Banik*[1], Arman Hossain[2], Md. Hossain Manik[3], Md Sifat Bin Siraj[4], Armana Sabiha Huq[5]

[1] *Graduate Student, Department of Civil Engineering, Bangladesh University of Engineering and Technology (BUET), Bangladesh, e-mail: joydeepbanik1999@gmail.com*
[2] *Graduate Student, Department of Civil Engineering, Bangladesh University of Engineering and Technology (BUET), Bangladesh, e-mail: armanhossain5616@gmail.com*
[3] *Senior Lecturer, Department of Civil Engineering, Chittagong Institute of Engineering & Technology (CIET), Affiliated with Chittagong University, Muradpur- Chattogram, Bangladesh, e-mail: hossainmanik.ce@gmail.com*
[4] *Graduate Student, School of Transportation and Logistics, Southwest Jiaotong University, Sichuan, China, e-mail: sifatbinsiraj@gmail.com*
[5] *Assistant Professor, Accident Research Institute (ARI), Bangladesh University of Engineering and Technology (BUET), Bangladesh, e-mail: ahuq002@fiu.edu*

***Corresponding Author**

ABSTRACT

In Bangladesh, transportation is primarily operated by paratransit providers who manage their services independently due to a lack of sufficient formal transportation options. People frequently use the paratransit system extensively for their daily travel needs. Consequently, the increasing number of crashes related to paratransit has become a significant safety concern. This study evaluates the safety concerns related to paratransit, with a primary focus on determining the crash severity associated with these systems. A survey was conducted of 507 drivers of several paratransits in 10 locations in Gazipur City. The survey included information on drivers' sociodemographic, driving-related and crash experience characteristics, including the severity of past crash involvement. To predict the crash severity, a set of single and ensemble machine learning algorithms has been employed. These are Logistic Regression, K-Nearest Neighbor (KNN), Decision Tree, Random Forest, Gradient Boosting and Extreme Gradient Boosting. These models' performances have been compared based on prediction accuracy, precision, recall, F1-score and area under the receiver operator characteristic (AUROC). From the single and ensemble models, the best model for predicting crash severity was found to be Extreme Gradient Boosting (XGBoost), which achieved a prediction accuracy of 78% with an F1-score of 0.75. Most crashes resulted in non-injury (75.84%) and non-grievous injury (16.37%). The causes of these crashes were primarily a non-expert driver (80.63%) and driving at higher speeds (17.16%). The majority of these paratransits were Easybike (60.67%) and the Rickshaw (28.54%). These findings can provide valuable insights and help the government to take targeted measures to address paratransit crashes.

Keywords: *Paratransit safety, crash severity, machine learning, single classifier, ensemble classifier*

1. INTRODUCTION

Paratransit is one of the most demandable transportation modes among today's urban transportation systems (M. T. Ahmed et al., 2023). Its services are increasing day by day due to their availability, which saves travel time, costs, and provides comfort and security for goods, among other benefits. Around 72% of city households use paratransit for their daily travels in Bangladesh (Bin Siraj et al., 2021). Paratransit services differ between developing and developed countries. In developing countries, it is one of the best solutions for transport service to passengers with lower fare prices, flexible service schedules, and comparatively door-to-door service (Bin Siraj, Farjana, et al., 2023). Commonly known paratransit vehicles are CNG, rickshaw, laguna, tempo, shared Uber, shared bike service, pedicabs, motorbikes, van-type minibuses and so on (Bin Siraj et al., 2021). In contrast, frequent availability and easy modes make it more susceptible to road accidents (Bin Siraj, Hasan, et al., 2023). Every year, more than 1.19 million people tragically lose their lives, and 20 to 50 million sustain non-fatal injuries from traffic crashes (*Road Traffic Injuries*, 2023). Southeast Asia has fatality rates twice as high as those in the developed countries of America or Europe (*Global Status Report on Road Safety ,WHO*, 2015). In the case of Bangladesh, the fatality rate of road crashes, more than 160 deaths per 10,000 vehicles, is quite alarming compared to other developing countries. Most of the paratransit drivers are illiterate, unskilled, don't have driving licenses, do not follow traffic rules and regulations that cause traffic congestion, and most of the accidents (Md. H. Rahman et al., 2021).

Traffic crashes pose significant safety concerns worldwide. Countries and international organizations have designed technologies, systems, and policies to prevent accidents. Whilst several measures have been adopted to make the roads safe for users, there is no such real-time warning system available to guide the user about the probability of an accident (S. Ahmed et al., 2021). Though a large number of studies were conducted on motorcycle (paratransit) crash severity in the context of developed countries, there is still a lack of literature in developing countries, especially in Bangladesh, which has one of the highest motorcycle crash rates in the world (Md. H. Rahman et al., 2021). In recent years, traffic accident analysis has garnered considerable attention from researchers as they seek to identify the factors that significantly contribute to traffic accidents. However, most research methods are based on statistical records or on conducting some simple surveys based on interviews or questionnaires (Labib et al., 2019). Researchers have employed several approaches to identifying the responsible factors, including econometric models, traditional statistical approaches, and data mining frameworks (Bhuiyan et al., 2022). Studies show that traditional methods cannot fully capture the unpredictable and spontaneous nature of traffic accidents, making accurate data prediction quite difficult. Machine Learning algorithms can be a better solution for tackling data unpredictability and spontaneity (Labib et al., 2019). These algorithms can explore the relationships (non-linear) between crash severity injury categories and contributing factors such as driver behaviour, vehicle characteristics, roadway geometry, significantly unveiling the complex situation (S. Ahmed et al., 2021).

Prior studies on crash severity using machine learning in Bangladesh are limited. A study using BUET/ARI records (2001–2015) showed that ML classifiers, particularly AdaBoost (≈80% accuracy), can effectively predict crash severity (Labib et al., 2019). Rabbani et.al (2023) studied comparing eight ML algorithms and found Random Forest (RF) performed best, showing RF's strong potential for crash severity prediction and traffic safety management in Bangladesh (Rabbani & Anik, 2023). A study by Siraj et al.(2023) investigated paratransit service quality using Multinomial Logit Modeling, identifying accident type, vehicle type, speed, and driver licensing as key factors, and providing insights for improving safety and planning (Bin Siraj, Hasan, et al., 2023). A study measuring paratransit service quality using Structural Equation Modeling (SEM) found that passenger security, seating comfort, and riding safety had the greatest influence on passengers' perceived quality (F. Rahman et al., 2023). Another study on unconventional vehicles (UVOs), including rickshaws and motorized vehicles, found that mid-block crashes were the severe crashes in Dhaka (Saha et al., 2022). There is a noticeable gap in understanding paratransit crash severity and in identifying the significant crash predictors. This gap is partly due to limited resources and the lack of comprehensive data.

From a strategic standpoint, the ML algorithms have a distinguishing capability to predict which scenarios are likely to result in a crash (Rabbani & Anik, 2023). For any machine learning algorithm, the maximum achievable predictive accuracy depends on the quality of the data (Yahaya et al., 2020). As there are relatively lower numbers of works on prediction models for paratransit crashes and their associated severities, this study has mainly two objectives which are:

1. To develop a paratransit crash prediction model
2. To unfold the contributing factors leading to the respective crash severities.

A questionnaire survey was conducted, involving 507 drivers from 10 locations in Gazipur, where most paratransit trips occur. The questionnaire collected data on demographic and economic characteristics, driving behavior, past crash experiences and characteristics. Crash severity in the questionnaire was categorized into 4 types: minor, non-grievous, grievous, and fatal. To predict crash severity, a combined total of 6 single and ensemble machine learning algorithms are used which include Logistic Regression, K-Nearest Neighbor, Decision Tree, Random Forest, Gradient Boosting, and Extreme Gradient Boosting.

2. METHODOLOGY

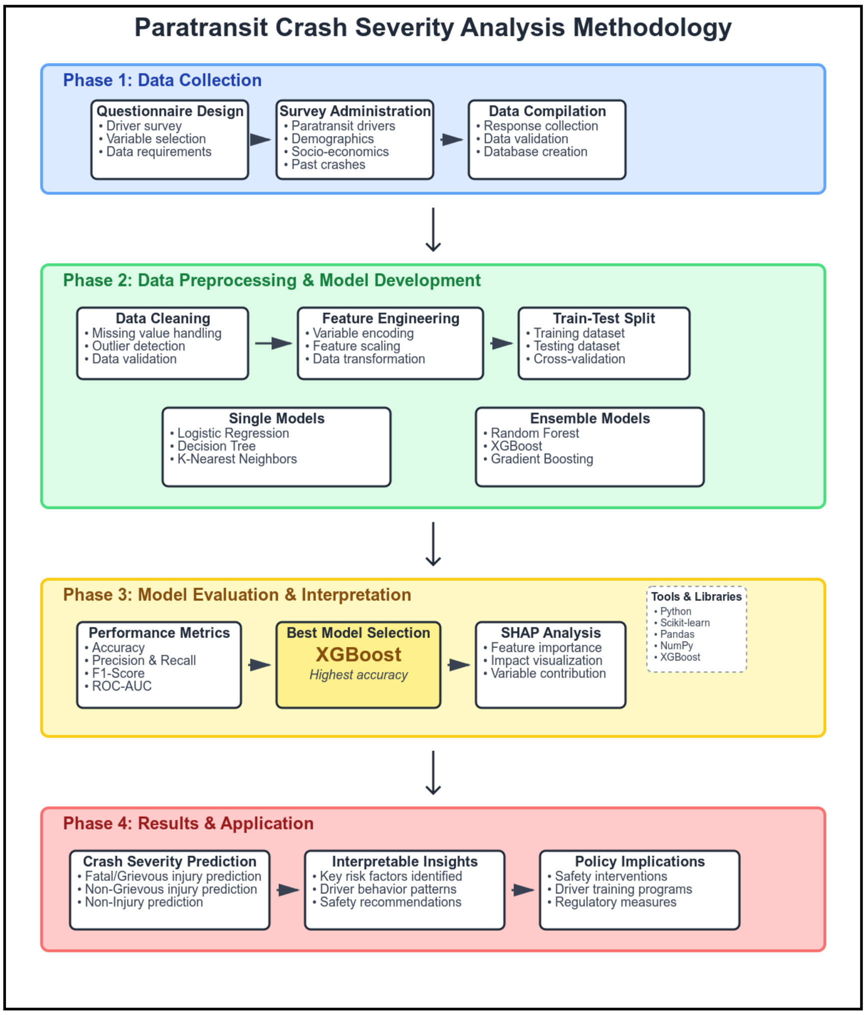

Figure 1: Flow Diagram of the Study

Figure 1 above shows the comprehensive flow diagram of the methodology of this study. The overall study is divided into four phases: data collection, data preprocessing and model development, model evaluation, and results interpretation. The four phases are described below.

2.1 Data Collection

This study was conducted to analyze the severity of paratransit crashes in Bangladesh. Therefore, Gazipur City Corporation was selected as the study area. Data were collected through a questionnaire survey in 10 locations, as shown in Figure 2, where most paratransit trips occur. The questionnaire covered a variety of characteristics, including demographic information, crash occurrence, driving experience, and other relevant factors. The sample of minimum people required for this study has been determined by the equation below:

$$n \geq N\left[1+\frac{N-1}{P(1-P)}\left(\frac{d}{z_{\alpha/2}}\right)^2\right]^{-1} \tag{1}$$

Where, n is the minimum sample size to be considered, N is the population of the city, P is the quality characteristic that is to be measured. As per (Johnson & Wichern, 2002), for neutral cases or where no previous experience exists, the value of P is taken as 0.5, d is the margin of error, which is taken as 5%, $z_{\alpha/2}$ = 1.96 for 95% confidence interval.

From the equation, for a 2.67 million (*Bangladesh: Gazipur City Corporation (City Districts and Wards) - Population Statistics, Charts and Map*, 2022) population size, the minimum sample size is n ≅ 385.

In this study, a total of 507 active paratransit drivers were surveyed using the questionnaire.

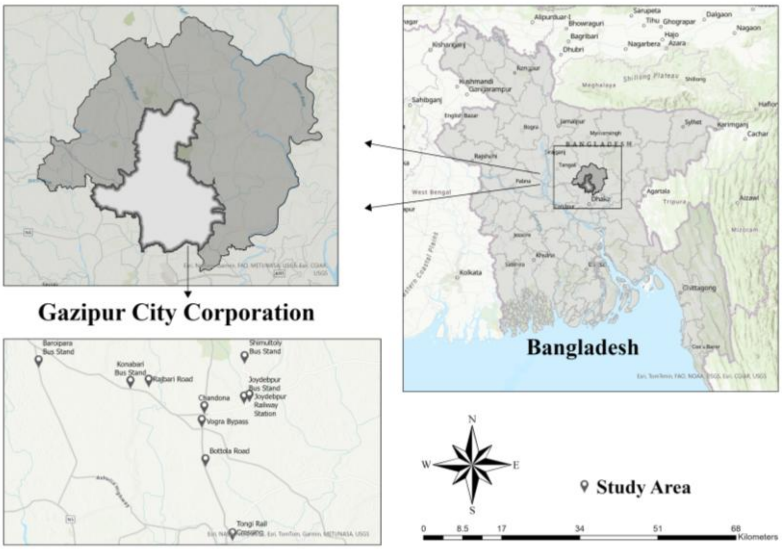

Figure 2: Study Area of This Research

2.2 Descriptive Statistics

Table 1: Demographic Characteristics

Occupation of Driver	No	Percentage
Business	9	1.78
Care Taker	1	0.20
Day Labour	1	0.20
Driving	491	96.84
Farmer	2	0.39
Student	3	0.59
Age		
13-30	152	29.98
31-40	182	35.90
41-50	106	20.91
>50	67	13.21
Subject Vehicle		
CNG	40	7.89
Easy bike	310	61.14
Leguna	12	2.37
Rickshaw	143	28.21
Tempo	2	0.39
Education Level		
Graduate	1	0.20

H.Secondary	1	0.20
Illiterate	111	21.89
Primary	333	65.68
Secondary	61	12.03
Driving Experience		
0-5 years	293	57.79
6-15 years	162	31.95
16-25 years	37	7.30
26-35 years	15	2.96
Institutional Driving Training		
No	499	98.42
Yes	8	1.58
Driving Duration (Hous)		
0-5	14	2.76
6-10	343	67.65
15-20	150	29.59
Driver's Income (Daily)		
0-500	79	15.58
500-1000	382	75.35
1000-1500	44	8.68
1500-2000	2	0.39

Table 1 illustrates the demographic characteristics of paratransit drivers. The majority, 96.84%, are reported driving as their main occupation, with the highest age groups being 31–40 years (35.90%) and 13–30 years old (29.98%). Easy Bikes (61.14%) and Rickshaws (28.21%) were the most common as paratransit vehicles in this study. Most drivers had only a primary education, with 65.68% holding a secondary education and 21.89% being illiterate. The driving experience was concentrated within the 0–5 years range, at 57.79%. Notably, 98.42% of the participants had no institutional training. On average, 67.65% drove 6–10 hours a day, with 75.35% earning between 500 and 1000 BDT per day.

Table 2: Crash Characteristics

Crash Severity	No	Percentage
Fatality	6	1.18
Grievous	33	6.51
Non-Injury	386	76.13
Non-Grievous	82	16.17
Vehicle Speed		
High (40-60)	47	9.27
Medium (25-40)	320	63.12
Low (15-25)	140	27.61
Cause of Crash		
Driven Faster	86	16.96
Environmental Effect	1	0.20
Non-expert driver	410	80.87
Talking on phone	10	1.97
Crash Partner		
Bike	49	9.66
Bus	11	2.17
Car	7	1.38
CNG	4	0.79
Easy bike	277	54.64
Pedestrian	86	16.96
Rickshaw	72	14.20
Truck	1	0.20

Incident Time		
Night	56	11.05
Day	451	88.95
Awareness of Traffic Rules		
Yes	5	0.99
No	502	99.01

Table 2 illustrates the crash characteristics of paratransit. The maximum crash severity found was non-injury at 76.13% and non-grievous at 16.17%. Most crashes occurred at a medium speed of vehicles, accounting for 63.12%, and were primarily caused by non-expert drivers, at 80.87%. Easy Bikes (54.64%) and pedestrians (16.96%) are the most common crash partners, with incidents predominantly occurring during the daytime (88.95%). Notably, 99.01% of drivers reported having no awareness of traffic rules.

2.3 Data Analysis

In the questionnaire, crash severity was categorized into four levels: minor, non-grievous, grievous, and fatal.

Minor crashes are those in which participants suffer little injury, which doesn't require or requires little medical support. Non-grievous injuries that participants may require moderate medical support. Grievous injuries, which involve major damage, include admission to the hospital. And fatal are where deaths occur. This case is particularly a multiclass classification, where the dependent variable has four classes. Therefore, multiclass classification algorithms have been employed for this analysis. Multiclass classification algorithms can be categorized into two main types: Single Classifier Model and Ensemble Classifier Model. These two are described below:

Single Model: In a single model classifier, only one model is used for the prediction of the output. In the training stage, the model is trained on the data to find the values of coefficients $F(w; x)$ for which the model's loss function gets minimized (Breiman et al., 2017). Here, w is the coefficient and x is the variable.

$$F = \arg\min_w \ Loss(y; F(\mathrm{w}; \mathrm{x})) \tag{2}$$

A total of three single models have been used. There are Logistic Regression (LR), K-Nearest Neighbor (KNN) and Decision Tree (DT).

Logistic Regression: Logistic Regression is a basic multiclass classification algorithm that assigns a separate linear Z score to each class j, which is then converted into probabilities for each class using the softmax function.

$$y = \frac{e^{zi}}{\sum_{j=1}^{k} e^{zj}} \tag{3}$$

Where Z_i is the score for the target class i, k is the total number of classes, and y is the predicted probability for class

K-Nearest Neighbors (KNN): KNN is a popular machine learning algorithm used for classification. It operates on the principle that similar data points tend to cluster together in feature space. The algorithm begins by selecting a value for k, representing the number of nearest neighbors to consider. Then the distance between data points is calculated, commonly using Euclidean distance.

$$d(x_i, \hat{x}_i) = \sqrt{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2} \tag{4}$$

Decision Tree: A decision tree is a supervised learning algorithm used for both classification and regression. The algorithm works by recursively splitting the dataset into subsets based on feature values. Each internal node represents a test on a feature, each branch corresponds to an outcome of the test, and each leaf node holds a class label. The splitting is guided by impurity measures, such as the Gini index or entropy, that quantify how mixed the classes are at a node. In this study, it's entropy,

$$IH = -\sum_{j=1}^{c} p_j log_2(p_j) \quad (5)$$

where p_j is the proportion of class j in the node, and C is the number of classes.

Ensemble Model: An ensemble model is one that utilizes multiple base models to achieve the lowest loss function *L(y; F(w; x))* and find the most suitable predicted output from the model (Dietterich, 2000).

$$F = arg\ min_w\ Ex;y[Loss(y; F(w; x)] \quad (6)$$

Three ensemble models have been used for analysis in this study: Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Gradient Boosting (GB).

Random Forest: Random Forest is an ensemble learning algorithm that constructs multiple decision trees using the bagging technique. It randomly samples the training data with replacement (bootstrapping) and builds each tree on a different subset. During prediction, all trees vote, and the majority class is selected. RF reduces variance and overfitting, and its performance depends on parameters such as the number of trees, the number of features considered at each split, and the minimum samples required to split a node.

Gradient Boosting: Gradient Boosting is a sequential ensemble method that builds decision trees iteratively, where each new tree is trained to correct the residual errors of the previous ensemble. This algorithm minimizes a specified loss function using gradient descent. GB is sensitive to hyperparameters such as the learning rate, number of estimators, and tree depth, and requires regularization techniques to prevent overfitting.

Extreme Gradient Boosting: XGBoost is an advanced implementation of gradient boosting that incorporates regularization (L1 and L2), second-order gradient optimization, and parallelized tree construction. This algorithm handles missing values and supports column subsampling, making it highly efficient for large-scale structured data. XGBoost is known for its scalability, speed, and predictive accuracy, and the performance is influenced by parameters such as learning rate, maximum depth, number of estimators, and regularization coefficients.

2.4 Model Metrics

These models were trained on 80% of the data, and the remaining 20% was used to evaluate the predictive accuracy. A confusion matrix consisting of 4 values (True Positive, False Positive, True Negative, False Negative) from which accuracy, precision, recall score, and F1 score are evaluated to validate each model's predictive performance.

$$Accuracy = \frac{True\ Positive}{True\ Positive\ +\ False\ Positive\ +\ True\ Negative\ +\ False\ Negative} \quad (7)$$

$$Precision = \frac{True\ Positive}{True\ Positive\ +\ False\ Positive} \quad (8)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (9)$$

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (10)$$

AUROC: It's known as the area under the receiver operating characteristic curve, a measure of how well a classifier distinguishes between different classes.

2.5 Shapley Additive Explanations (SHAP)

SHAP is a unified framework for interpreting machine learning model predictions based on cooperative game theory. It assigns each feature an importance value by computing its marginal contribution across all possible combinations of features. SHAP values are consistent and locally accurate, making them suitable for both global and individual-level explanations. The equation for SHAP follows below:

$$y_j = y_{base} + f(X_{j1}) + f(X_{j2}) + f(X_{j3}) + \ldots + f(X_{jk}) \quad (11)$$

Here, y_j represent predicted value for instance j, y_{base} is the average model output for the entire dataset, $f(X_{jk})$ refers to the SHAP value for feature k in instance j.

3. RESULTS & DISCUSSION

3.1 Correlation Matrix

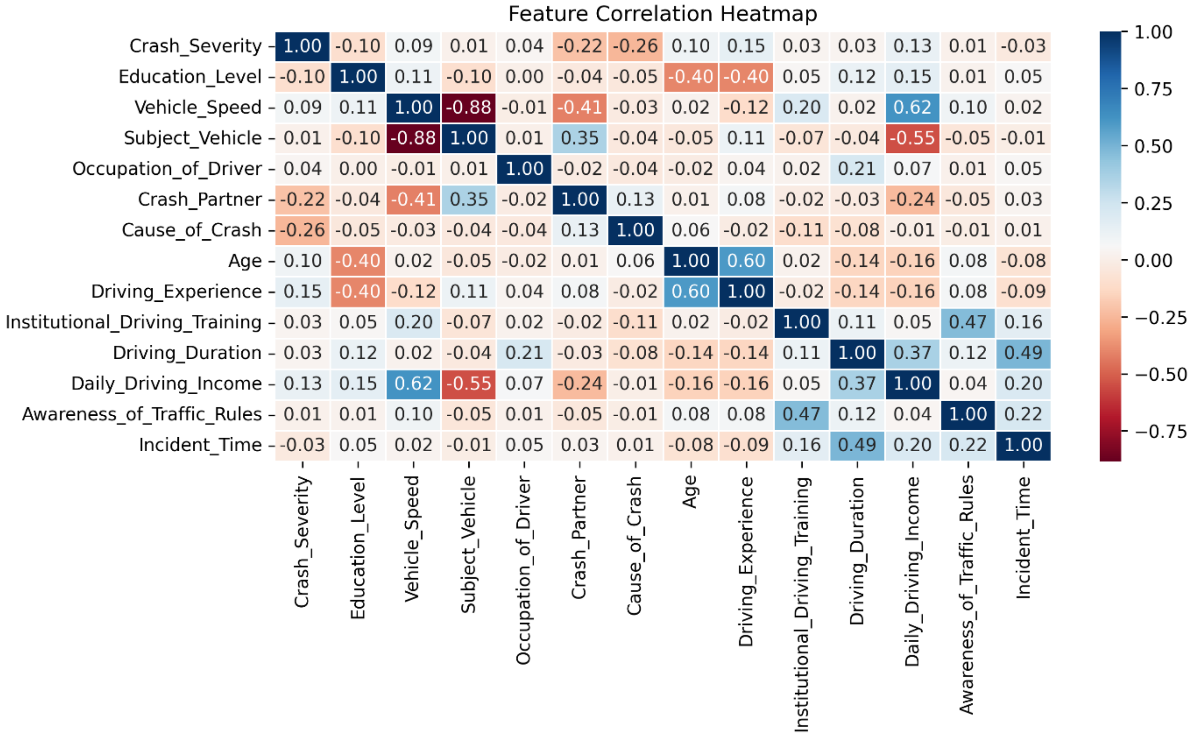

Figure 3: Correlation of Matrix of Variables

Figure 3 presents the correlation matrix of the variables, computed using Spearman's correlation. Spearman's correlation assesses the strength and direction of monotonic relationships between variables, making it suitable for ordinal or rank-transformed data. Before calculation, categorical variables were encoded into numerical form to enable correlation analysis. As observed in Figure 3, the variable "Subject Vehicle" exhibited a non-monotonic association with other features, therefore, it was excluded from the correlation analysis to avoid misinterpretation.

3.2 Model Evaluation

Table 3: Model Metrics for Machine Learning Models

Metrics	RF	XGB	GB	DT	KNN	LR
Accuracy	0.7451	0.7843	0.7157	0.5686	0.7647	0.7647
Precision	0.7412	0.7388	0.6787	0.7074	0.7288	0.5848
Recall	0.7451	0.7843	0.7157	0.5686	0.7647	0.7647
F1 Score	0.7412	0.7501	0.688	0.6124	0.7063	0.6627

Table 3 presents the four metrics used to compare the different models. To evaluate the predictive performance of the machine learning models, standard classification metrics including accuracy, precision, recall, and F1-score were computed using a weighted averaging scheme. This approach ensures that performance evaluation remains robust despite class imbalance. Given the disproportionately small representation of severe crash outcomes (fatal: 1.18%, grievous: 6.51%), all models were trained with class-balancing strategies either through built-in *class_weight='balanced'* parameters or custom sample weighting where applicable (e.g., XGBoost). Furthermore, hyperparameter tuning was conducted for each model using grid search with cross-validation to optimize generalization performance. Among the six ensemble and single models, Extreme Gradient Boost (XGBoost) demonstrated the best accuracy, as well as the highest F1 score. It denotes that XGBoost performed well in terms of model training and predictions. Further, the AUROC value can be a decisive factor in demonstrating model validation. In this study, the One vs. Rest strategy has been employed to determine which model predicts more effectively, distinguishing the classes.

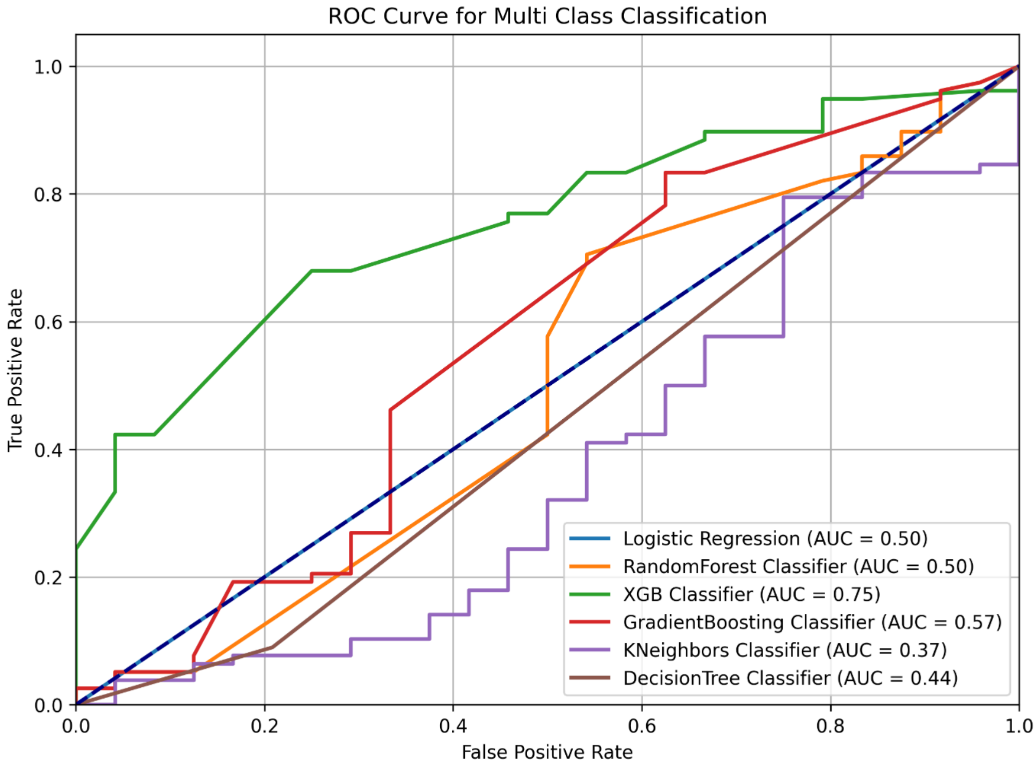

Figure 4: AUROC comparison across six algorithms

Figure 4 illustrates that XGBoost clearly dominated the rest of the models. XGBoost showed an AUROC value of 0.75, which is the highest among the six models, indicating that XGBoost is the most suitable model for predicting crash severity. Therefore, XGBoost will be used for further analysis. While the dataset size is modest, the use of interpretable tree-based models and careful validation protocols supports the reliability of the reported metrics. These results are presented not as definitive predictors, but as exploratory benchmarks to assess the feasibility of data-driven crash severity modeling. To further enhance interpretability, SHAP summary plots and waterfall plots were employed, providing transparent insights into the contribution of individual features to model predictions.

3.3 Inferential Statistics

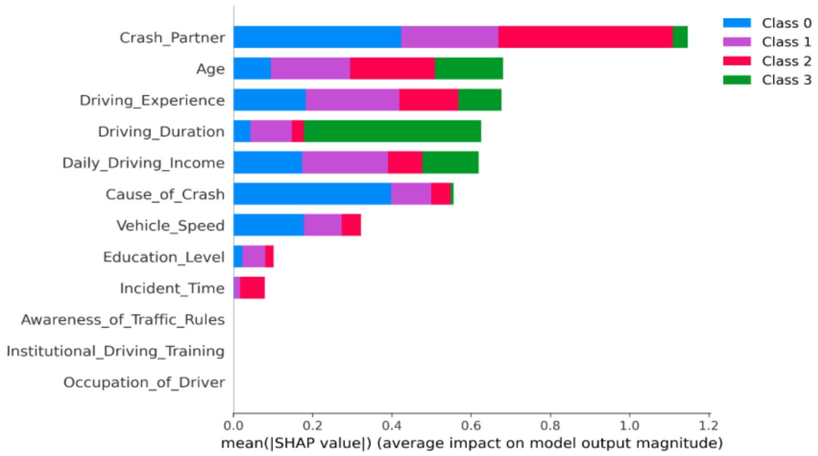

Figure 5: Feature Importance Across Different Classes

Figure 5 depicts the overall significant crash predictors for multiple classes in descending order. Crash_Partner has the highest influence on resulting in non-injury, non-grevious, and grievous crashes. Age and driving experience also have significant influences on all types of crashes. Driving duration is the most influential crash predictor for fatal crashes (class 3).

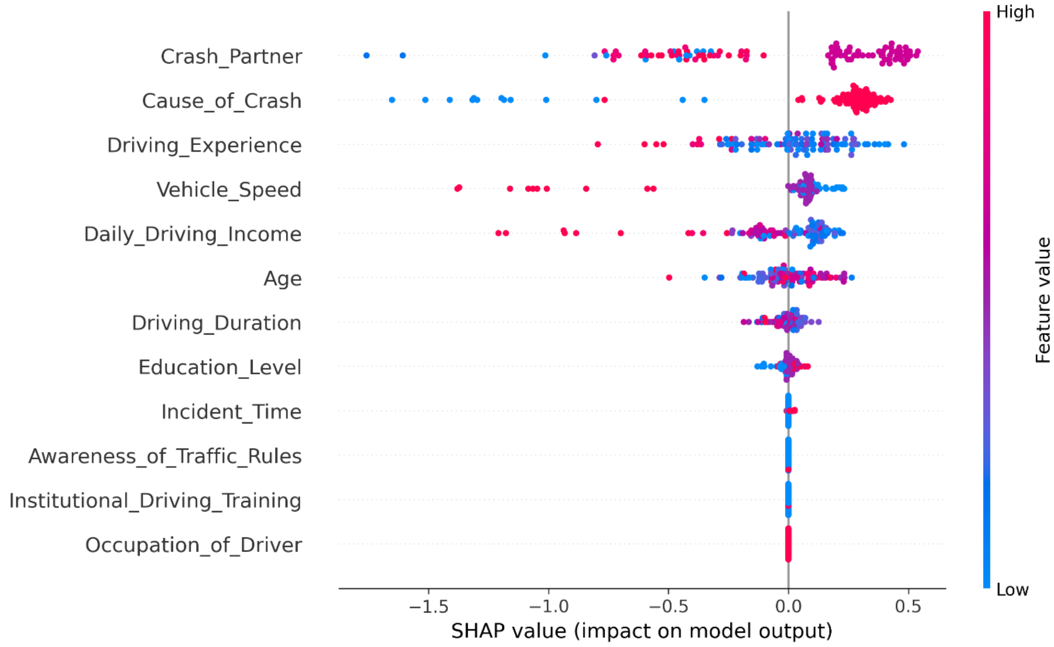

Figure 6: SHAP Summary Plot for Non-Injury Crash

According to questionnaire data, the majority of crashes resulted in non-injury. The SHAP summary plot has been utilized to uncover the underlying patterns of crashes, specifically those resulting in Non-Injury Crashes. The red value indicates the higher value for features, and blue denotes lower values, while the more red points in the right denote the increasing value for that feature, which increases the log odds for that given class. Figure 6 shows that the mid-level values of crash partners in our cases, which are (easybike, pedestrians, CNG, rickshaw), increase the probability of predicting a non-injury crash. A higher value of causes, in our case, such as talking on mobile phones and driving inexperience, mostly result in these non-injury crashes. Additionally, a noticeable pattern exists in vehicle speed, low vehicle speeds mostly increase the odds of resulting in non-injury crashes. The rest of the features are centred around Shapley value 0, so they are not significant predictors for the mentioned class.

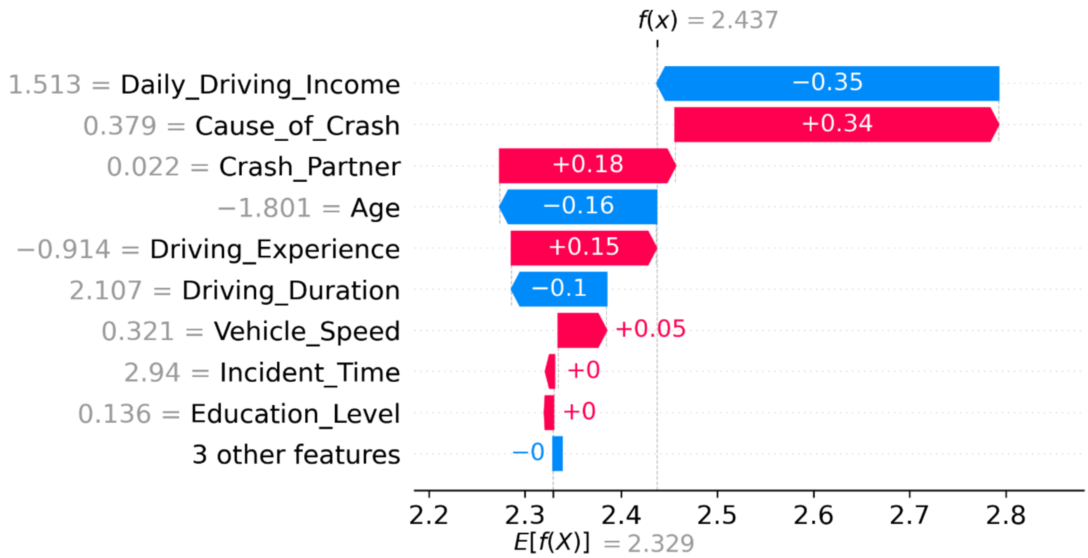

Figure 7: Waterfall Plot for a Single Instance

The waterfall plot in Figure 6 illustrates how each feature influences and impacts a single prediction. For this specific prediction, the model's expected value, E[f(x)], is 2.329. However, some features pushed it upwards, and ultimately, the output f(x) is 2.437. The driver had less driving experience, which contributed to the increase in models predicting it as a non-injury crash. Moderate value of crash partner, high value of cause of crash, contributed positively to the model's prediction of a non-injury crash.

4. CONCLUSIONS

Evaluating the severity of paratransit crashes is crucial, especially as the number of crashes increases daily. To reduce the overall crash outcomes, the root of those crashes needs to be assessed. This study aimed to develop a crash prediction model and identify the risks and crash outcomes associated with paratransit services in a densely populated city. These paratransits are the dominant transportation choice for most residents, so it is necessary to evaluate the crash outcomes associated with them.

From the analysis, the most reliable model was XGBoost, which performed well in distinguishing between every class in our case, specifically the different levels of crash severity. The model identified the most significant features across all crash severities as crash partner, age, driving experience, driving duration, and the cause of the crash. The paratransits mostly resulted in minor or non-injury crashes, where easybikes, CNG, and rickshaws were the main causes of these crashes, in which pedestrians were the victims of those incidents. Pedestrians are vulnerable while walking on the roads, as most light vehicles increase the probability of being involved in a crash with pedestrians. Talking on the phone, inadequate driving experiences, and low-speed vehicles on the busiest roads often result in crashes. There is a lack of institutional training among the drivers, as well as a lack of driving knowledge. To improve the overall crash scenario in Bangladesh, the root cause of these daily incidents should be given more focus. The easybike, CNG and rickshaw drivers should be given proper knowledge about driving, and these light vehicles should have clear restrictions and strict regulations while operating on the roads.

This crash severity prediction model was trained on only 507 data points from the questionnaire survey. The relatively small sample size and class imbalance constrain the generalizability of the findings. If the study had included crash databases from official departments and covered a vast area, the results would have been more reliable for predicting crashes across a wider region and would also have improved its ability to capture crash outcomes.

Declaration of Use of AI

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript, including in the writing, data analysis, or research methodology.

REFERENCES

Ahmed, M. T., Bin Siraj, M. S., & Campisi, T. (2023). Paratransit Safety as a Key Resource for Sustainable Mobility in Developing Countries. *Communications - Scientific Letters of the University of Zilina*, *25*. https://doi.org/10.26552/com.C.2023.018

Ahmed, S., Hossain, M. A., Bhuiyan, M. M. I., & Ray, S. K. (2021). A Comparative Study of Machine Learning Algorithms to Predict Road Accident Severity. *2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS)*, 390–397. https://doi.org/10.1109/IUCC-CIT-DSCI-SmartCNS55181.2021.00069

Bangladesh: Gazipur City Corporation (City Districts and Wards)—Population Statistics, Charts and Map. (2022). https://www.citypopulation.de/en/bangladesh/gazipurcity/admin/

Bhuiyan, H., Ara, J., Hasib, K. M., Sourav, M. I. H., Karim, F. B., Sik-Lanyi, C., Governatori, G., Rakotonirainy, A., & Yasmin, S. (2022). Crash severity analysis and risk factors identification based on an alternate data source: A case study of developing country. *Scientific Reports*, *12*(1), 21243. https://doi.org/10.1038/s41598-022-25361-5

Bin Siraj, M. S., Farjana, M., & Ullah, M. (2023). *Paratransit Safety Evaluation Using Poisson Regression: A Case Study Based on Driver Perception*.

Bin Siraj, M. S., Hasan, M., Farjana, M., Ratna, K., Asma, Mossa. S., Islam, Md. J., Rabbi, F., & Rakib, R. (2023). *Realize Service Quality of Paratransit in a Developing Country: A Multinomial Logit Approach*.

Bin Siraj, M. S., Hossain, M., Hasan, M., Chowdhury, M. M. H., Uddin, M. H., Islam, M., Ratna, K., & Asma, Mossa. S. (2021). *Mode Choice Behavior of Chittagong City Dwellers in Bangladesh*. *6*, 10.

Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (2017). *Classification and Regression Trees*. Chapman and Hall/CRC. https://doi.org/10.1201/9781315139470

Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. *Multiple Classifier Systems*, 1–15. https://doi.org/10.1007/3-540-45014-9_1

Global status report on road safety ,WHO. (2015). https://www.afro.who.int/publications/global-status-report-road-safety-2015

Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis*. https://scholar.google.com/scholar_lookup?title=Applied%20Multivariate%20Statistical%20Analysis&author=R.A.%20Johnson&publication_year=2002

Labib, Md. F., Rifat, A. S., Hossain, Md. M., Das, A. K., & Nawrine, F. (2019). Road Accident Analysis and Prediction of Accident Severity by Using Machine Learning in Bangladesh. *2019 7th International Conference on Smart Computing & Communications (ICSCC)*, 1–5. https://doi.org/10.1109/ICSCC.2019.8843640

Rabbani, B. U., & Anik, B. M. T. H. (2023). *A Comparative Study on Machine Learning Algorithms for Crash Severity Prediction: A Case Study for Bangladesh* (SSRN Scholarly Paper No. 4611094). Social Science Research Network. https://doi.org/10.2139/ssrn.4611094

Rahman, F., Islam, Md. A., & Hadiuzzaman, Md. (2023). Paratransit service quality modeling reflecting users' perception-A case study in Dhaka, Bangladesh. *IATSS Research*, *47*(3), 335–348. https://doi.org/10.1016/j.iatssr.2023.07.001

Rahman, Md. H., Zafri, N. M., Akter, T., & Pervaz, S. (2021). Identification of factors influencing severity of motorcycle crashes in Dhaka, Bangladesh using binary logistic regression model. *International Journal of Injury Control and Safety Promotion*, *28*(2), 141–152. https://doi.org/10.1080/17457300.2021.1878230

Road traffic injuries. (2023). https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

Saha, B., Fatmi, M. R., & Rahman, Md. M. (2022). Modeling Injury Severity of Unconventional Vehicle Occupants: Hybrid of Latent Segments and Random Parameters Logit Models. *Transportation Research Record*, *2676*(6), 35–47. https://doi.org/10.1177/03611981211069949

Yahaya, M., Fan, W., Fu, C., Li, X., Su, Y., & Jiang, X. (2020). A machine-learning method for improving crash injury severity analysis: A case study of work zone crashes in Cairo, Egypt. *International Journal of Injury Control and Safety Promotion*, *27*(3), 266–275. https://doi.org/10.1080/17457300.2020.1746814