Final Project Report

Predicting Pedestrian Injury Severity in Crashes Using Machine Learning

ABSTRACT

This project explores the application of supervised machine learning models to predict pedestrian injury severity in traffic crashes using data from North Carolina between 2020 and 2024. The primary focus is on binary classification of serious or fatal injuries versus less severe outcomes, with an initial investigation into multiclass prediction and model interpretability. To address the significant class imbalance in the data, SMOTE was applied to oversample serious/fatal cases during model training. Several models were evaluated, including Decision Tree, Random Forest, XGBoost, and Multilayer Perceptron (MLP), with hyperparameter tuning performed via grid search using recall as the primary metric. An ensemble model combining Random Forest, XGBoost, and Decision Tree achieved the best overall balance between recall, F1-score, and accuracy. While binary classification performance was encouraging, multiclass prediction remains challenging due to extreme class imbalance and overlapping feature patterns. Marginal effect analysis provided additional interpretability by estimating the impact of key categorical variables on injury severity probabilities. These results demonstrate the promise of machine learning for pedestrian injury severity analysis while highlighting the need for careful handling of imbalance, model interpretability, and specialized techniques to further improve multiclass outcomes in transportation safety research.

1. Background

Traffic crashes continue to pose a major public health and safety concern in the United States, resulting in thousands of fatalities and millions of injuries each year. Among all road users, pedestrians represent a particularly vulnerable group due to their lack of physical protection and exposure to motorized traffic. According to national statistics, pedestrian-involved crashes are far more likely to result in severe or fatal injuries compared to vehicle-only incidents. Accurately predicting the severity of pedestrian injuries is critical for informing timely emergency response, designing safer transportation systems, and shaping effective traffic safety policies.

Historically, crash injury severity has been analyzed using statistical models such as binary or multinomial logistic regression. These models are valued for their interpretability and ability to test for statistical significance and have been widely used to identify risk factors associated with crash severity (Abrari Vajari et al., 2020; Kaplan and Prato, 2012; Gray et al., 2008; Yuan et al., 2021). However, traditional models are constrained by assumptions such as linearity and independence, and often struggle to capture complex, nonlinear interactions among crash variables—especially when dealing with high-dimensional datasets or interdependent features.

To address these limitations, researchers have increasingly turned to machine learning (ML) techniques to improve predictive accuracy and reveal more nuanced relationships between crash conditions and injury outcomes. A wide range of ML models have been applied in crash severity prediction, including artificial neural networks (Arhin and Gatiba, 2019; Wang et al., 2019), support vector machines (Hegde and Rokseth, 2020; Mohamed Radzi et al., 2017), decision trees (Abellán et al., 2013; Chang and Chien, 2013), random forests, gradient-boosted trees, and association rule mining (Jiang et al., 2020; Samerei et al., 2021). While these methods offer strong predictive power and flexibility, many existing studies rely on a single algorithm or focus exclusively on vehicle-occupant crashes, leaving a gap in understanding pedestrian-specific injury mechanisms.

This project aims to leverage crash data from North Carolina (2020–2024) to predict injury severity outcomes specifically for pedestrian-involved crashes. Using a rich set of crash-related features—including roadway characteristics, environmental conditions, and pedestrian and driver factors—we implement and compare several machine learning models: decision trees, random forests, XGBoost (gradient-boosted decision trees), and multilayer perceptrons (neural networks). In addition, we explore ensemble learning techniques to aggregate predictions from multiple models, with the goal of improving generalizability and robustness. Model performance is evaluated for both binary (severe vs. non-severe) and multiclass (ordinal injury level) classification tasks, offering insights into the strengths and limitations of different machine learning approaches for pedestrian injury risk analysis.

2. Data description

This project extracts pedestrian-involved crash data from the state of North Carolina, covering the years 2020 to 2024. The dataset was obtained from a bicycle-pedestrian crash database maintained by the North Carolina Department of Transportation. Each record in the dataset corresponds to a single pedestrian crash event and contains a variety of coded categorical and numerical features related to crash conditions, road environment, driver and pedestrian characteristics, and injury outcomes. The primary target variable, PedInjury, denotes the level of injury sustained by the pedestrian and follows the KABCO injury classification system. It includes the following categories: O: No Injury; C: Possible Injury; B: Suspected Minor Injury; A: Suspected Serious Injury; K: Killed. Figure 1 presents the distribution of the pedestrian injury severity.

For modeling purposes, this variable is used in both binary and multiclass formats. In the binary classification setup, injuries are grouped as either non-severe (O, C, B) or severe (A, K). In the multiclass setting, the categories are retained as an ordinal variable encoded from 0 (No Injury) to 4 (Killed). The dataset contains a rich set of crash-level and individual-level features, including crash type; alcohol or drug involvement; vehicle type; pedestrian characteristics; roadway and environmental features. All categorical variables were transformed using one-hot encoding. The cleaned and preprocessed dataset consists of 9,689 records after removing entries with missing or unknown injury severity.

A summary of the distribution of key variables is presented in Table 1. As shown, most crashes resulted in "Possible Injury" or "Suspected Minor Injury," with fatal pedestrian injuries accounting for roughly 10% of the total. Most crashes involved passenger cars and occurred during daylight hours. A significant share of pedestrians was struck while walking in travel lanes or crosswalks. The dataset also shows considerable variation in roadway configuration, lighting conditions, and traffic control environments.

Table 1 Descriptive statistics of model variables (N = 9,689)

Variable	Description	Frequency	Percent
Pedestrian injury severity level	No Injury(O)	473	4.88%
	Possible Injury(C)	2,915	30.09%
	Suspected Minor Injury(B)	3,767	38.88%
	Suspected Serious Injury(A)	1,390	14.34%
	Killed(K)	1,144	11.81%
Primary pre-crash behavior (crash type)	Pedestrian Failed to Yield	1,594	16.5%
	Pedestrian dash	939	9.7%
	Others	7,156	73.9%
Motorist intoxicated (alcohol or drug)	No	8,399	86.7%
	Yes	322	3.3%
	Others	968	10.0%
Vehicle Type	Passenger Car	4,812	49.7%
	Sport Utility	1,891	19.5%
	Pickup	1,295	13.4%
	Others	1,691	17.5%
Pedestrian's age	below 16	800	8.3%
	16 to 19	655	6.8%
	20 to 29	2,074	21.4%
	30 to 39	1,941	20.0%
	40 to 49	1,383	14.3%
	50 to 59	1,420	14.7%
	above 60	1,416	14.6%
Pedestrian intoxicated (alcohol or drug)	No	8,174	84.4%
	Yes	1,147	11.8%
	Others	368	3.8%
Pedestrian's sex	Female	3,325	34.3%
	Not female	6,364	65.7%

Pedestrian position	Travel Lane	6,085	62.8%
	Crosswalk Area	1,970	20.3%
	Sidewalk / Shared Use Path / Driveway Crossing	312	3.2%
	Paved Shoulder / Bike Lane / Parking Lane	530	5.5%
	Others	792	8.2%
Light condition	Dark - Lighted Roadway	2,390	24.7%
	Dark - Roadway Not Lighted	2,393	24.7%
	Dawn-Dusk	412	4.3%
	Daylight	4,440	45.8%
	Others	54	0.6%
Total Number of through lanes	2	4,915	50.7%
	3	754	7.8%
	4	1,955	20.2%
	5	984	10.2%
	6	431	4.4%
	Others	650	6.7%
Road alignment	Straight - Level	7,796	80.5%
	Straight - Grade	1,010	10.4%
	Straight - Hillcrest	295	3.0%
	Curve - Grade	157	1.6%
	Curve - Level	303	3.1%
	Others	128	1.3%
Speed limit	Below 25 MPH	1,791	18.5%
	30 - 35 MPH	3,742	38.6%
	40 - 45 MPH	2,482	25.6%
	50 - 55 MPH	1,284	13.3%
	60 - 75 MPH	390	4.0%
Traffic control	Double Yellow Line, No Passing Zone	1,331	13.7%
	No Control Present	5,180	53.5%
	Stop Sign	770	7.9%
	Stop And Go Signal	1,900	19.6%
	Others	508	5.2%

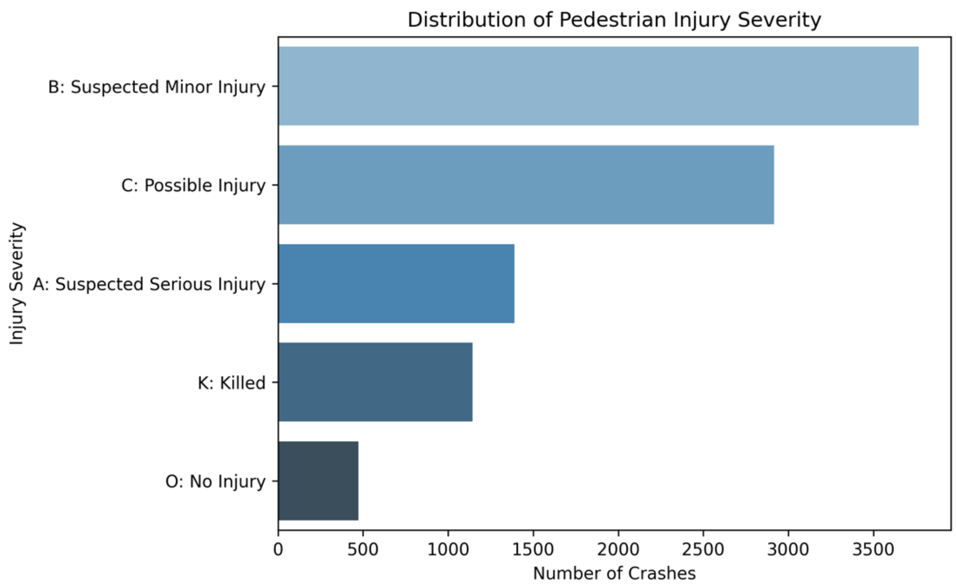

Figure 1. Distribution of pedestrian injury severity.

3. Models

This project applies supervised machine learning techniques to predict the severity of pedestrian injuries in crashes. The problem is framed as both a binary classification task—distinguishing serious or fatal injuries (A/K) from less severe outcomes (B/C/O)—and a multiclass ordinal classification task, capturing the full five-level KABCO severity scale.

3.1 Model Selection

We implemented and compared four widely used machine learning algorithms, each representing a different modeling paradigm:

Decision Tree Classifier: A simple, interpretable model that recursively splits the data into subsets based on feature thresholds. Decision trees provide transparent decision logic, making them useful for understanding which conditions most strongly relate to injury severity. However, they are prone to overfitting when used alone.

Random Forest Classifier: An ensemble method that builds multiple decision trees on bootstrapped subsets of the data and aggregates their predictions. Random forests mitigate overfitting by reducing variance and introducing randomness in feature selection. They are effective for handling high-dimensional, categorical data, and offer variable importance scores that aid in model interpretation.

XGBoost (Extreme Gradient Boosting): A gradient-boosted decision tree model that builds trees sequentially, where each new tree corrects errors made by the previous ones. XGBoost is known for its high performance in structured tabular data, robustness to outliers, and ability to handle feature interactions. It also includes built-in regularization to prevent overfitting and often outperforms other tree-based models in predictive accuracy.

Multilayer Perceptron (MLP): A type of feedforward artificial neural network composed of multiple fully connected layers and nonlinear activation functions. MLPs can model complex, nonlinear relationships and interactions between features that may not be easily captured by decision trees. However, they require more careful tuning and are less interpretable than tree-based models.

In addition to binary classification, we also explored the use of the same models for multiclass classification, predicting all five levels of pedestrian injury severity as defined in the KABCO scale. While the model architecture and tuning approach remained consistent, multiclass performance was notably weaker, particularly in detecting rare but critical outcomes such as fatal injuries. We include these results in Section 4 for completeness and discussion but place primary emphasis on the binary classification results.

3.2 Hyperparameter Tuning

To optimize each model, we performed grid hyperparameter search using 3-fold cross-validation on the training data. The search spaces were defined based on best practices and prior research. Table 2 shows the chosen hyperparameter for each model.

Decision Tree:

max_depth: maximum depth of the tree, which limits how deeply the tree can grow

min_samples_split: minimum number of samples required to split an internal node

min_samples_leaf: minimum number of samples required to be at a leaf node, providing additional regularization

criterion: function used to measure the quality of a split (gini impurity or information gain via entropy)

Random Forest:

n_estimators (number of trees): controls model complexity and variance reduction

max_depth: limits tree depth to prevent overfitting

min_samples_split: controls when a node should be split further

XGBoost:

n_estimators: number of boosting rounds

max_depth: depth of individual trees

learning_rate: step size shrinkage to control overfitting

subsample: fraction of data sampled for each tree, introducing regularization

MLP:

hidden_layer_sizes: number of neurons and layers in the architecture

activation: nonlinear function applied at each node (ReLU or tanh)

alpha: L2 regularization term to prevent overfitting

learning_rate_init: controls weight update step size

We tuned all models with Grid Search and a 5-fold cross validation. In addition to evaluating individual models, we implemented an ensemble model that combines the predictions of the top-performing classifiers. Ensemble learning is often used to improve prediction accuracy and robustness by aggregating outputs from diverse models. Specifically, we used a soft voting ensemble, where the predicted class probabilities from Random Forest, XGBoost, and MLP are averaged to generate the final prediction. This approach leverages the strengths of each model while reducing their individual weaknesses.

3.3 Handling Class Imbalance

Crash datasets often exhibit significant class imbalance, particularly when modeling rare but critical outcomes such as serious or fatal pedestrian injuries (as shown in Figure1). To mitigate this issue and improve the model's ability to learn from minority class instances, we applied the Synthetic Minority Oversampling Technique (SMOTE) to the training data prior to model fitting. SMOTE generates synthetic examples for the minority class by interpolating between existing samples and their nearest neighbors in feature space. This approach helps balance the class distribution while preserving the original data structure. Importantly, SMOTE was applied only to the training data to avoid information leakage. The test set retained the original class distribution to reflect real-world deployment conditions and to evaluate the model's true generalization ability.

By addressing class imbalance during training, SMOTE allows machine learning models to better detect serious or fatal injuries without becoming biased toward the majority (non-severe) class.

3.4 Evaluation Strategy

Model performance was assessed using a variety of standard classification metrics: Accuracy: overall proportion of correct predictions; Precision: proportion of predicted positive cases that are actually positive. Recall: proportion of actual positive cases that are correctly identified; F1-score: harmonic mean of precision and recall, emphasizing balanced performance; ROC AUC (Area Under the Receiver Operating Characteristic Curve): evaluates the model's ability to distinguish between classes across all classification thresholds, offering a threshold-independent performance measure.

Special attention was given to the F1-score and recall for the serious/fatal injury class, given the critical importance of correctly identifying high-severity outcomes. Confusion matrices and ROC curves were also generated to visualize performance trade-offs, and side-by-side comparisons were made across models and the ensemble.

4. Results

4.1 Baseline comparison: effect of SMOTE on Class 1 Performance

To evaluate the impact of class imbalance handling, we trained two baseline Random Forest models—one using the original, imbalanced training data and the other using a SMOTE-balanced training set. Both models were evaluated on the same test data to ensure a fair comparison. As shown in Figure 2, the model trained with SMOTE achieved notably higher recall (**0.48** vs. **0.35**) and a modest gain in F1-score (**0.49** vs. **0.42**) compared to the model trained without oversampling. While precision slightly decreased (**0.51** vs. **0.54**), the improvement in recall—indicating better detection of serious or fatal pedestrian injuries—justifies the trade-off.

These findings highlight the importance of balancing the training data when predicting rare but critical outcomes such as serious or fatal injury. SMOTE allows the model to learn more representative patterns for minority classes, leading to more responsive predictions in high-risk scenarios. The limited performance of both versions—particularly in terms of recall—motivates the need for hyperparameter tuning and the application of more advanced or specialized models, such as XGBoost and neural networks, which are evaluated in subsequent sections.

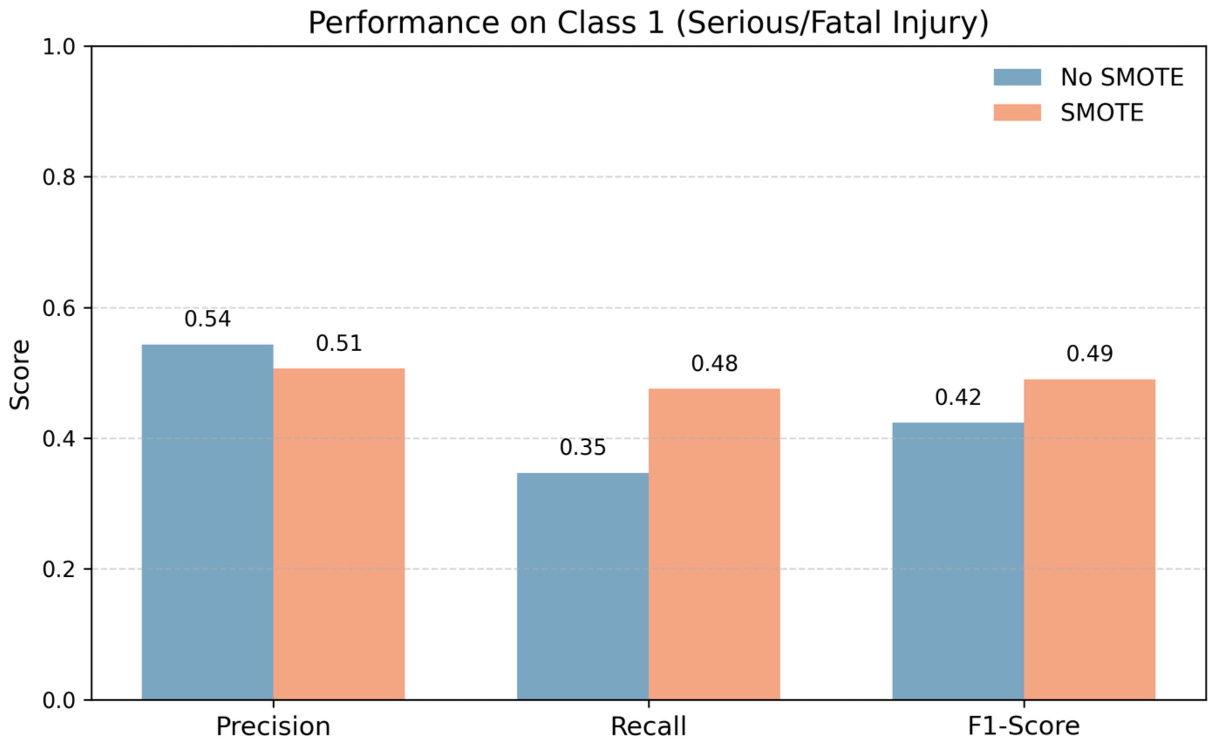

Figure 2. Precision, recall, and F1-score for class 1 (serious/fatal injury) using Random Forest models trained with and without SMOTE.

4.2 Performance comparison of tuned models

To improve predictive accuracy for serious and fatal pedestrian injuries, we performed hyperparameter tuning for each model using grid search with cross-validation. The current best hyperparameter configurations for Decision Tree, Random Forest, XGBoost, and MLP are summarized in Table 2. These results reflect the best configurations identified as of the current stage of experimentation. Given the scope of the hyperparameter space and computational constraints, tuning is still in progress and additional improvements may be achieved with further refinement.

Table 2. Best hyperparameters identified for each model (subject to update as tuning continues).

Model	Best Hyperparameter
Decision Tree	{'criterion': 'gini', 'max_depth': 5, 'min_samples_leaf': 1, 'min_samples_split': 10}
Random Forest	{'max_depth': 15, 'min_samples_split': 2, 'n_estimators': 200}
XGBoost	{'learning_rate': 0.05, 'max_depth': 5, 'n_estimators': 100, 'subsample': 0.7}
MLP	{'activation': 'relu', 'alpha': 0.0001, 'hidden_layer_sizes': (128, 64), 'learning_rate_init': 0.001}

Figure 3 presents the performance of the tuned models on the test set using four key evaluation metrics: F1-score, precision, recall, and accuracy. Among individual models, Random Forest and XGBoost achieved the highest overall accuracy (0.75 and 0.74, respectively), while Decision Tree attained the highest recall (0.67), indicating its relative strength in identifying serious or fatal injury cases. However, its lower precision suggests a tendency toward over-predicting the positive class. The MLP underperformed across all metrics, with noticeably lower scores in recall (0.44) and F1-score (0.49). Due to its limited contribution, MLP was excluded from the final ensemble model. The ensemble model, which combines the predictions of top individual classifiers, achieved a balanced profile, matching the highest accuracy (0.75) and improving recall (0.62) while maintaining a acceptable F1-score (0.56). This supports the use of ensemble methods to stabilize predictions and leverage complementary strengths across model types. In addition to standard classification metrics, we also compared the Receiver Operating Characteristic (ROC) curves of all models to evaluate their performance across different classification thresholds. As shown in Figure 4, XGBoost achieved the highest AUC score (0.77), indicating the strongest overall ability to distinguish between serious/fatal and non-serious injuries across varying thresholds. Random Forest and the ensemble model followed closely (both at 0.76), while Decision Tree achieved a slightly lower AUC of 0.74. Consistent with earlier findings, the MLP model underperformed (AUC = 0.69) and was excluded from the ensemble. The similarity between ensemble and Random Forest performance in ROC space suggests overlapping strengths, though ensembling may offer gains in stability and recall. These results reinforce the robustness of tree-based models for this classification task. We also report the confusion matrix of ensemble matrix.

Overall, while the tuned models outperform baseline versions, F1-scores remain modest, suggesting further improvement may be possible through expanded tuning.

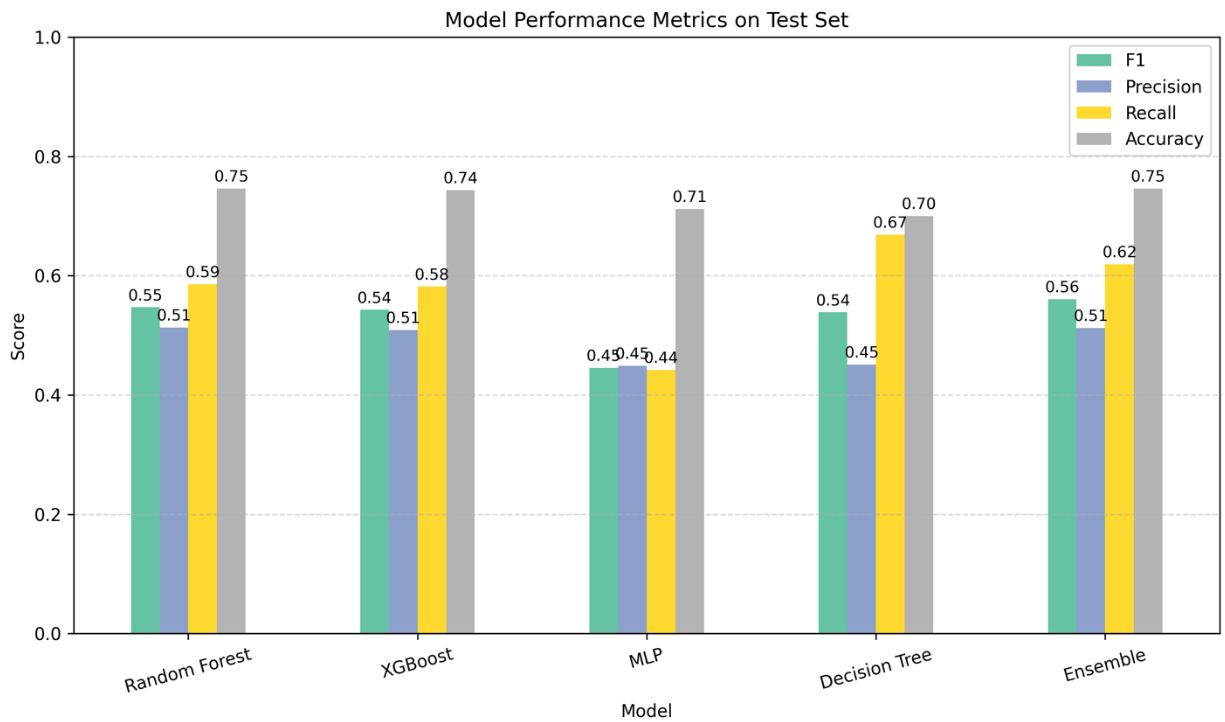

Figure 3. Comparison of model performance on the test set

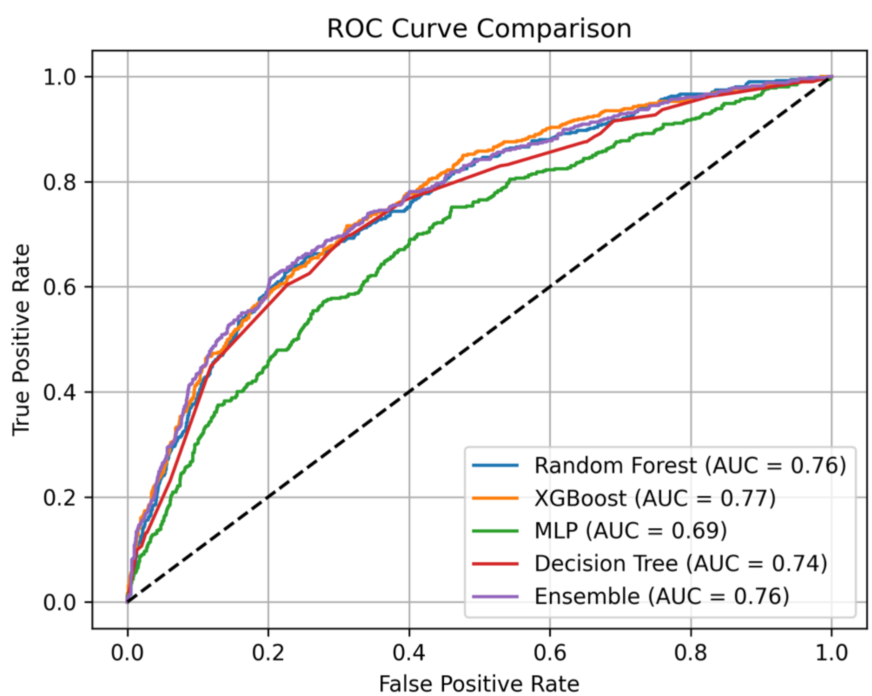

Figure 4. ROC curve comparison and AUC values for all models.

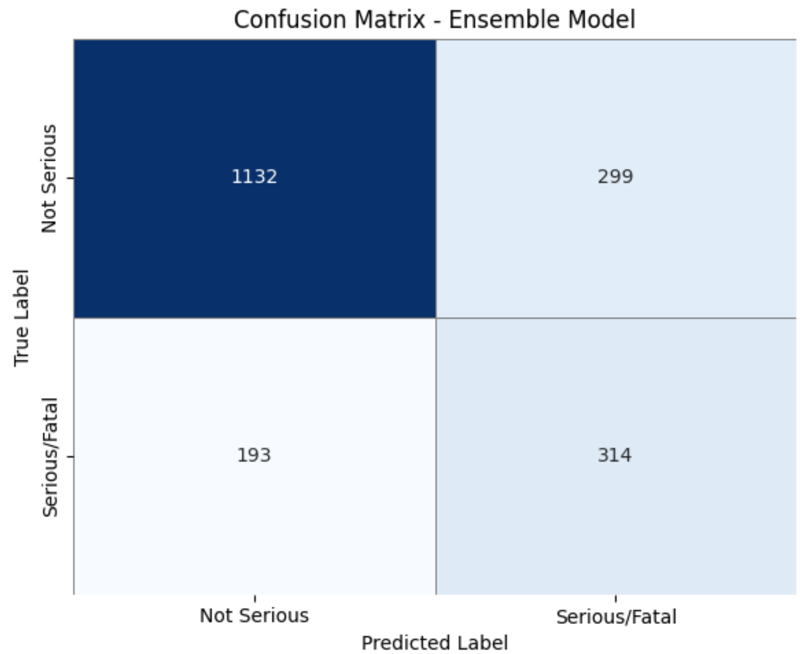

Figure 5. **Confusion matrix of ensemble model.**

4.3 Multiclass injury severity prediction

In addition to binary classification, we attempted to predict pedestrian injury severity across five discrete classes following the KABCO injury scale: K (Killed), A (Suspected Serious Injury), B (Suspected Minor Injury), C (Possible Injury), and O (No Injury). The goal was to evaluate the model's ability to differentiate between various severity levels, providing a more nuanced understanding of injury outcomes.

Results indicate that multiclass prediction remains significantly more challenging than binary classification. Although overall accuracy was moderate, performance on individual classes, especially the rarest categories such as "Killed" (K), was poor. The confusion matrices revealed substantial misclassification between adjacent injury categories (e.g., between serious injury and minor injury), suggesting that the models struggle to distinguish severity levels that are inherently close and subjective in nature.

Future work will explore advanced modeling strategies, such as ordinal classification techniques or cost-sensitive learning frameworks, which may better handle the natural ordering and imbalance of the severity levels. Additional feature engineering and resampling strategies will also be considered to improve multiclass model performance.

At this stage, multiclass modeling remains an important but unresolved component of the study, and further refinement is necessary to produce actionable, interpretable results for pedestrian injury severity prediction across multiple outcome levels.

4.4 Model interpretation

Model interpretability is an important aspect of injury severity modeling, particularly when informing policy or engineering interventions. We used marginal effect analysis, to better understand which features most influence predictions. Marginal effect quantifies the relationship between an input variable and the predicted probability of an injury outcome. The marginal effect shows how a change in the level of a categorical variable impacts the probability of the outcome, while holding all other variables constant. Marginal effect shows the direction and magnitude of change for each category of a variable relative to a reference (or base) category. The partial dependence calculation for a specific variable *S* is given by:

$$\hat{f}_s(x_s) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}\left(x_s, x_c^i\right) \tag{9}$$

where, x_s is the specific value of the variable being analyzed. x_c^i are the values of all other variables c in the i-th observation. $\hat{f}$ is the prediction function. n is the number of observations. This equation calculates the average predicted probability of police injury for a given value of x_s . Then the marginal effect of switching from a base category to a target category is calculated as:

$$ME = \hat{f}_{Starget}\left(x_{Starget}\right) - \hat{f}_{Sbase}\left(x_{Sbase}\right) \tag{10}$$

where, $x_{Starget}$ is the target category of the variable and x_{Sbase} is the base category of the variable.

Table 3. Marginal Effects for binary classification ensemble model

Variable	**Description**	**Marginal effect**
Primary pre-crash behavior (crash type)	Pedestrian Failed to Yield	0.4%
	Pedestrian dash (base)	
	Others	-2.2%
Motorist intoxicated (alcohol or drug)	No(base)	
	Yes	4.9%
	Others	-8.2%
Vehicle Type	Passenger Car (base)	
	Sport Utility	1.1%
	Pickup	0.9%
	Others	0.05%
Pedestrian intoxicated (alcohol or drug)	No (base)	
	Yes	4.7%
	Others	12.5%
Pedestrian position	Travel Lane (base)	
	Crosswalk Area	-7.9%
	Sidewalk / Shared Use Path / Driveway Crossing	-2.1%
	Paved Shoulder / Bike Lane / Parking Lane	-4.5%
	Others	-4.8%
Light condition	Dark - Lighted Roadway (base)	
	Dark - Roadway Not Lighted	4.4%
	Dawn-Dusk	-4.4%
	Daylight	-13.7%
	Others	5.2%
Road alignment	Straight – Level (base)	
	Straight - Grade	2.9%

	Straight - Hillcrest	2.5%
	Curve - Grade	3.8%
	Curve - Level	3.7%
	Others	2.8%
Speed limit	Below 25 MPH (base)	
	30 - 35 MPH	16.4%
	40 - 45 MPH	30.1%
	50 - 55 MPH	39.3%
	60 - 75 MPH	52.2%
Traffic control	Double Yellow Line, No Passing Zone	2.5%
	No Control Present (base)	
	Stop Sign	-2.6%
	Stop And Go Signal	-2.0%
	Others	-0.01%

Given that all variables were one-hot encoded, marginal effects were calculated by setting all categories to the base level and selectively activating one target category to measure the resulting change in the predicted probability of serious or fatal injury.

Table 3 presents the marginal effects for several important variables. The results indicate several meaningful patterns. For primary pre-crash behavior, crashes involving "Pedestrian Failed to Yield" slightly increase the probability of serious/fatal outcomes by 0.4% relative to the "Pedestrian Dash" baseline. Motorist intoxication increases the predicted risk by 4.9%, whereas pedestrian intoxication is associated with a 4.7% increase. Pedestrian position plays a significant role: compared to standing in a travel lane, being in a crosswalk area or on a sidewalk reduces the likelihood of serious injury by 2.1% and 4.5%, respectively. Lighting conditions also substantially affect injury severity, with crashes occurring in daylight reducing the serious/fatal injury probability by 13.7% relative to crashes on dark, lighted roadways. Roadway geometry shows smaller but consistent effects, where curves (both grade and level) increase severity probabilities by approximately 3%. In contrast, speed limits have a dramatic influence: moving from areas below 25 MPH to higher-speed zones increases the probability of serious injury by 16.4% (30–35 MPH), 30.1% (40–45 MPH), 39.3% (50–55 MPH), and 52.2% (60–75 MPH). Finally, the presence of traffic control devices such as stop signs and double yellow lines slightly mitigates injury severity risk compared to areas with no control present.

Overall, the marginal effects analysis supports the broader findings from model performance evaluation and highlights the critical influence of roadway speed, pedestrian positioning, and lighting conditions on pedestrian injury outcomes. These interpretable insights can inform targeted infrastructure interventions and policy strategies aimed at improving pedestrian safety.

5. Conclusion

This project investigates the use of supervised machine learning models to predict pedestrian injury severity in traffic crashes using North Carolina data from 2020 to 2024. The primary focus has been on binary classification of serious/fatal injuries versus less severe outcomes, with an initial exploration into multiclass prediction and model interpretability.

Our results demonstrate that addressing class imbalance using SMOTE significantly improves model performance on the minority class. Among the models tested, XGBoost and

Random Forest achieved the strongest overall performance in terms of accuracy, F1-score, and AUC. The ensemble model, which combines the predictions of Random Forest, XGBoost, and Decision Tree, achieved the best balance between recall and stability, making it a promising approach for practical applications. Conversely, the MLP underperformed and was excluded from the final ensemble. Despite these encouraging results, the models still struggle with relatively low recall and F1-scores for serious injury cases, indicating that even after tuning and oversampling, a significant number of high-severity outcomes are missed. The complexity of injury severity prediction, influenced by overlapping feature patterns and class imbalance, suggests that further improvements may be gained through more extensive hyperparameter optimization, enhanced feature engineering, and alternative model structures.

Multiclass injury severity prediction was also explored but remains challenging, with poor performance especially in the rarest outcome classes. Future work will investigate ordinal classification techniques and cost-sensitive learning approaches to better capture the ordered nature of injury severity outcomes.

Finally, to improve model interpretation, we performed a marginal effect analysis to interpret how key categorical factors such as speed limit, pedestrian position, and lighting conditions influence the probability of severe injuries.

Overall, results reinforce the value of machine learning for pedestrian safety analysis, while emphasizing the importance of thoughtful preprocessing, class balancing, interpretability, and targeted evaluation metrics in addressing real-world public safety challenges.

References

[1] Vajari, M. A., Aghabayk, K., Sadeghian, M., & Shiwakoti, N. (2020). A multinomial logit model of motorcycle crash severity at Australian intersections. Journal of safety research, 73, 17-24.

[2] Kaplan, S., & Prato, C. G. (2012). Risk factors associated with bus accident severity in the United States: A generalized ordered logit model. Journal of safety research, 43(3), 171-180.

[3] Gray, R. C., Quddus, M. A., & Evans, A. (2008). Injury severity analysis of accidents involving young male drivers in Great Britain. Journal of safety research, 39(5), 483-495.

[4] Yuan, Y., Yang, M., Guo, Y., Rasouli, S., Gan, Z., & Ren, Y. (2021). Risk factors associated with truck-involved fatal crash severity: Analyzing their impact for different groups of truck drivers. Journal of safety research, 76, 154-165.

[5] Arhin, S. A., & Gatiba, A. (2019). Predicting Injury Severity of Angle Crashes Involving Two Vehicles at Unsignalized Intersections Using Artificial Neural Networks. Engineering, Technology & Applied Science Research, 9(2).

[6] Wang, J., Kong, Y., & Fu, T. (2019). Expressway crash risk prediction using back propagation neural network: A brief investigation on safety resilience. Accident Analysis & Prevention, 124, 180-192.

[7] Hegde, J., & Rokseth, B. (2020). Applications of machine learning methods for engineering risk assessment–A review. Safety science, 122, 104492.

[8] Abellán, J., López, G., & De OñA, J. (2013). Analysis of traffic accident severity using decision rules via decision trees. Expert Systems with Applications, 40(15), 6047-6054.

[9] Chang, L. Y., & Chien, J. T. (2013). Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. Safety science, 51(1), 17-22.

[10] Jiang, F., Yuen, K. K. R., & Lee, E. W. M. (2020). Analysis of motorcycle accidents using association rule mining-based framework with parameter optimization and GIS technology. Journal of safety research, 75, 292-309.

[11] Samerei, S. A., Aghabayk, K., Mohammadi, A., & Shiwakoti, N. (2021). Data mining approach to model bus crash severity in Australia. Journal of safety research, 76, 73-82.