International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 5, pp 705-712, May-2026

https://doi.org/10.47001/IRJIET/2026.105094

Machine Learning-Based Road Accident Fatality Prediction Under Imbalanced Data Conditions: An Evaluation of Resampling and Classification Techniques

Olusola Theophilus Faboya

Department of Computing and Information Science, Bamidele Olumilua University of Education, Science and Technology, Ikere-Ekiti, Nigeria. E-mail: faboya.olusola@bouesti.edu.ng

Abstract - **Road traffic accidents remain a leading cause of mortality worldwide and a major public safety concern, particularly in developing regions where data imbalance affects predictive modelling performance. This study proposes a machine learning framework for fatal accident prediction under imbalanced class conditions using multiple resampling strategies and classification algorithms. Multiple classification algorithms, including Logistic Regression, Random Forest, XGBoost, and LightGBM, were evaluated with both data-level and algorithm-level imbalance-handling techniques that include Synthetic Minority Oversampling Technique (SMOTE), ADASYN, Random Oversampling and Random Undersampling. Performance was measured using imbalance-aware metrics such as accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). In addition, statistical significance was tested using the Friedman and Wilcoxon signed-rank tests. To enhance interpretability, SHAP analysis was used to explain model decisions at both global and local levels. Experimental results indicate that ensemble boosting models significantly outperformed conventional methods, with LightGBM achieving the best overall performance with 99.35% accuracy, 99.25% F1-score, and ROC-AUC of 0.999736 without resampling. XGBoost also demonstrated strong robustness to class imbalance. Although resampling techniques slightly improved minority class learning, the improvements were not statistically significant. SHAP analysis reveals that vehicle make, crash location and engine type were key determinants of fatal outcomes. The findings demonstrate that advanced boosting algorithms possess strong predictive capability for imbalanced crash severity prediction and provide insight into developing robust predictive systems for traffic safety management for policymakers and intelligent transportation systems. It also demonstrates that integrating imbalance-aware learning with explainable AI enhances both predictive performance and interpretability in road safety analytics.**

Keywords: Imbalance-handling, Resampling, classification, road accident fatality, Machine learning.

I. INTRODUCTION

Road crashes are a leading cause of mortality in the world (Amiri *et al.*, 2025). Predictive modelling can support early intervention strategies; however, predicting accident fatality becomes a challenge in the presence of a highly imbalanced dataset, where one group constitute a minority and the other a majority. In such cases, traditional models have been found to favour the majority class, resulting in poor detection of fatal cases. The imbalance leads to biased model learning, where standard classifiers prioritise the majority class and exhibit poor sensitivity toward the minority (fatal) class. As a result, models may achieve high overall accuracy while failing to correctly identify fatal accidents, which is unacceptable in safety-critical applications. The inability to effectively detect rare but severe events undermines the correctness of predictive systems in real-world scenarios.

Existing studies on road accident prediction often focus on improving classification accuracy while ignoring the challenges posed by imbalanced datasets and handling techniques that can address these challenges. Many works rely on conventional evaluation metrics, which can be misleading in skewed datasets. There is also a limited comparative analysis of multiple machine learning models combined with different data imbalance-handling techniques within a unified experimental framework. Furthermore, a few studies provide context-specific insights relevant to developing regions, where accident characteristics and contributing factors may differ significantly from those of other countries.

This study addresses this challenge by evaluating resampling methods and machine learning classifiers for predicting road accident fatalities under imbalanced data conditions and by identifying optimal strategies to improve fatality prediction. The following guiding questions are considered to achieve the set aim: (a) Which machine learning model performs best in predicting road accident fatalities

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 5, pp 705-712, May-2026

https://doi.org/10.47001/IRJIET/2026.105094

under imbalanced conditions? (b) How do different imbalance-handling techniques affect model performance? (c) Which evaluation metrics provide the most reliable assessment of model effectiveness in this context?

The remaining part of the paper is structured as follows: Related work is discussed in Section II; Section III presents the detailed methodology adopted in this study; the results and discussion from the experiments are discussed in Sections IV and V; while Section VI gives the conclusion of the study.

II. RELATED WORK

Road traffic accidents have been a leading cause of death worldwide, especially in developing countries where transportation systems are often characterised by bad roads, weak enforcement of traffic regulations and poor emergency response mechanisms. The increasing availability of traffic and accident-related datasets has encouraged the use of machine learning techniques in recent times. One major challenge associated with accident prediction datasets is class imbalance, which most of the traditional machine learning techniques ignore and, in turn, affects their performance on the minority class. Several previous studies have applied machine learning techniques such as Logistic Regression, Decision Trees, binary logistic regression model and Neural Networks for accident severity prediction. Obasi & Benson (2023) employed a logistic regression model for crash severity prediction and reported moderate classification performance; however, the model struggled to identify minority fatal cases accurately. Also, Moyo *et al.* (2025) applied ensemble learning techniques for traffic accident prediction and observed that tree-based models outperformed linear classifiers in handling nonlinear crash-related variables. It can be observed that many previous studies focused on improving overall classification accuracy while neglecting the effects of imbalanced datasets. Accuracy alone is often misleading in highly skewed datasets because of bias towards the majority class. The outcome of such a prediction is unacceptable because the minority fatal class is the most critical category that can assist public safety in certain decision-making interventions.

To address the imbalance problem, researchers have developed several resampling strategies, such as the Synthetic Minority Oversampling Technique (SMOTE) proposed by (Chawla *et al.*, 2002), which is one of the widely used techniques for generating synthetic minority class samples. (Chawla *et al.*, 2002) described an imbalanced dataset as one that has classification categories that are not approximately equally represented. This happens in the real world, and traditional machine learning models usually struggle in the classification process, resulting in providing false results. SMOTE is a technique that presents better classifier performance (in ROC space). The Receiver Operating Characteristics (ROC) curve is a standard technique for summarising classifier performance over a range of trade-offs between true positive and false positive error rates (Swets, 1988).

SMOTE create artificial examples by interpolating between existing minority instances, thereby reducing overfitting associated with random duplication. Adaptive Synthetic Sampling (ADASYN) extends SMOTE by adaptively generating synthetic examples for difficult-to-learn minority instances. Other approaches, such as Random Oversampling and Random Undersampling, have also been explored. Fernández *et al.* (2018) demonstrated that hybrid sampling strategies significantly improve minority class detection compared to conventional learning models. Similarly, Xiao *et al.* (2023) showed that combining ensemble classifiers with imbalance handling techniques improves crash severity prediction in intelligent transportation systems. Taha (2026) and Dayanah *et al.* (2023) observed that tree-based ensemble methods such as Random Forest, XGBoost and LightGBM have become particularly effective due to their ability to capture nonlinear relationships and interactions among variables while also maintaining robustness against noisy data.

In recent years, Explainable Artificial Intelligence (XAI) has emerged as an essential component of machine learning applications to make AI systems transparent, interpretable and trustworthy, especially to give insight into the behaviour of interrelated variables in safety-critical domains. While advanced ensemble models provide superior predictive performance, they often function as “black-box” systems that are difficult to interpret. Shapley Additive exPlanation (SHAP) has become one of the most popular XAI frameworks for interpreting model decisions (Somvanshi *et al.*, 2026). It provides global and local explanations by quantifying the contribution of individual features toward prediction outcomes. In transportation systems, SHAP analysis has been used to identify influential factors such as weather conditions, vehicle type, and environmental characteristics that have contributed to accident severity.

Although existing studies have examined accident severity prediction, but limited research has comparatively evaluated multiple resampling strategies and classification algorithms within a unified framework under severe class imbalance conditions. Furthermore, few studies have integrated statistical hypothesis testing and explainable AI techniques to validate model performance and interpretability simultaneously. This study addresses these gaps by evaluating multiple imbalance handling techniques and machine learning

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 5, pp 705-712, May-2026

https://doi.org/10.47001/IRJIET/2026.105094

classifiers while incorporating statistical significance testing and SHAP-based explainability for robust accident severity prediction.

III. METHODOLOGY

This study adopted an experimental machine learning research design technique for predicting crash severity under imbalanced data conditions. The framework combined data preprocessing, resampling techniques, classification modelling, performance evaluation, statistical testing, and explainable artificial intelligence analysis.

Data Description

The dataset used consisted of the target, crash severity outcomes and associated explanatory variables. The target variable was highly imbalanced, with three classes as follows:

Class 0: 20 instances
Class 1: 2,112 instances
Class 2: 7,868 instances; the imbalance in the dataset necessitated the application of imbalance handling techniques to improve minority class prediction performance.

Data Preprocessing

This stage involved identifying inconsistencies in the dataset and ensuring that categories are grouped correctly. Missing values were found in nine out of the twenty-five features of the dataset. The missing values in the numeric columns were handled using a combination of mean imputation for numerical values and mode imputation for categorical variables. Duplicate values in the dataset were removed to avoid double-counting, and outliers were identified using the Interquartile Range (IQR) method and capped to the 5th and 95th percentiles to reduce their influence on model training. Feature transformation and normalisation were performed. Features related to environmental conditions, vehicle information, and roadway characteristics were selected as predictors for accident severity classification. The dataset was divided into training and testing subsets using an 80:20 ratio.

Dataset Class Imbalance Handling Techniques and Machine Learning Models

To address class imbalance, four resampling approaches that include Random Oversampling, Random Undersampling, Synthetic Minority Oversampling Technique (SMOTE) and Adaptive Synthetic Sampling (ADASYN) were evaluated. The SMOTE and ADASYN generated synthetic minority class examples to improve classifier learning, while undersampling reduced majority class dominance by randomly removing majority class observations.

Also, four machine learning models that include Logistic Regression, Random Forest, XGBoost and LightGBM were implemented and compared. Logistic Regression was chosen to serve as a baseline linear classifier, while the Random Forest, XGBoost and LightGBM represented ensemble tree-based methods.

Model Evaluation, Statistical Testing and Interpretation

Many imbalance-aware evaluation metrics such as Accuracy, precision, recall, F1-score and ROC-AUC were employed due to the imbalanced nature of the dataset. The ROC-AUC metric was particularly employed for the evaluation of classifier discrimination performance across varying thresholds.

Furthermore, to determine whether there are differences in the performances of the classifiers and resampling methods, the Friedman test and Wilcoxon Signed-Rank Test were employed as non-parametric statistical tests.

The SHAP was employed to enhance model interpretability, which enables identification of the most influential factors that contributed to fatal accident predictions.

IV. RESULTS

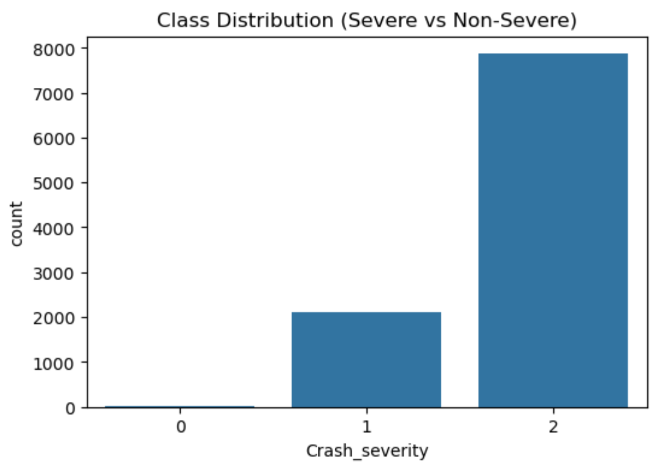

Figure 1: Class Distribution of Severe vs Non-Severe Crashes

Figure 1 shows the class distribution of crash severity levels in the road accident dataset. It produces the class imbalance plot of the target variable, crash severity, in three categories:

Class 0 = least severe crashes
Class1= Moderately severe crashes
Class 2= Highly severe or fatal crashes

From the chart, class 2 has the highest number of observations with 7868 cases, followed by class 1 with 2112,

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 5, pp 705-712, May-2026

https://doi.org/10.47001/IRJIET/2026.105094

while class 0 contains only 20 cases. The disparity between the classes suggests that the models may become biased towards predicting the majority class during training, which is a major challenge because rare crash categories, i.e., class 0 may be misclassified frequently, while important safety insights related to low-frequency crash outcomes could be overlooked. The imbalanced implications therefore justify the application of imbalanced learning techniques employed in this study.

Table 1: Machine Learning Models and Resampling Techniques Comparative Performance

	Resampling	Model	Accuracy	Precision	Recall	F1	ROC_AUC
3	None	LightGBM	0.9935	0.991612	0.9935	0.992531	0.999736
6	RandomOverSampler	XGBoost	0.9915	0.990645	0.9915	0.991039	0.999801
2	None	XGBoost	0.9915	0.989518	0.9915	0.990508	0.999822
15	ADASYN	LightGBM	0.9905	0.990178	0.9905	0.990327	0.999701
7	RandomOverSampler	LightGBM	0.9905	0.989753	0.9905	0.990060	0.999810

Table 1 presents the performance of different machine learning models and resampling techniques for crash severity prediction based on the class distribution shown in Figure 1. The evaluation metrics include Accuracy, Precision, Recall, F1-score and ROC-AUC.

The results indicate that all models achieved high predictive performance, with ROC-AUC values close to 1.0, indicating good class discrimination capability. The overall best performance was achieved by LightGBM without resampling, having a recall of 99.35%, an F1-score of 99.25%, and an ROC-AUC of 0.999736, while XGBoost has accuracy of 99.15%, and ROC-AUC of 0.999822, and RandomOverSampler has an accuracy of 99.15%, ROC-AUC of 0.999801, which suggests that LightGBM was highly effective in learning the crash severity patterns directly from the original dataset despite the imbalance, and the nearly identical performance between resampled and non-resampled versions of XGBoost indicates that it is naturally robust to class imbalance due to its boosting mechanism and gradient optimisation strategy.

The resampling techniques-RandomOverSampler and ADSYN slightly improved minority class learning but did not significantly outperform the original models. ADASYN combined with LightGBM provide an accuracy of 99.05% and an F1-score of 99.03% which indicate a slight reduction in accuracy of the original LightGBM model, this suggests that the original dataset already contained sufficiently strong class-separating patterns. RandomOverSampler used with LightGBM provide the accuracy of 99.05% and ROC-AUC of 0.999810. The findings suggets that extremely high ROC-AUC values indicate that the models can reliably distinguish between accurate prediction of severe crashes levels. It can also be observed that boosting algorithms such as XGBoost and LightGBM are highly suitable for crash severity prediction tasks. The results show that advanced boosting algorithms may already possess strong imbalance-handling capabilities. Consequently, the marginal improvements from resampling suggest that algorithm selection can sometimes be more influential than the choice of balancing strategy.

Models Performance Comparison (F1-Score)

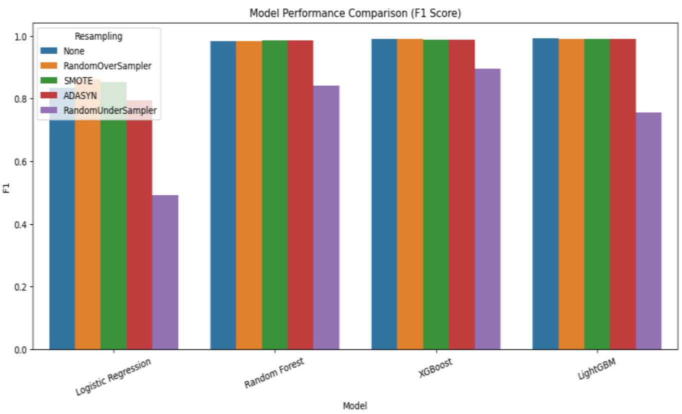

Figure 2: Models Performance Comparison (F1-Score)

Figure 2 presents the F1-score comparative performance of the four machine learning models (Logistic Regression, Random Forest, XGBoost, and LightGBM) evaluated under different resampling techniques (None-original dataset, RandonOverSampler, SMOTE, ADASYN, and RandonUnderSampler) for crash severity prediction.

The results show that the ensemble-based models (Random Forest, XGBoost, and LightGBM) achieved high F1-scores across most sampling methods, with values close to 1.0, indicating a balance between precision and recall in predicting crash severity classes. While Logistic Regression recorded lower performance when compared to others, especially when RandomUnderSampler was applied.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 5, pp 705-712, May-2026

https://doi.org/10.47001/IRJIET/2026.105094

LightGBM and XGBoost achieved the highest F1-scores, demonstrating superior predictive capability for crash severity classification. While Random Forest also performed strongly with minimal variation across oversampling methods. It can be seen that the application of RandomOverSampler, SMOTE, and ADASYN slightly improved or maintained model performance, suggesting that balancing the dataset positively influenced the learning process.

However, the use of RandomUnderSampler significantly reduced performance across all models, particularly for logistic Regression, where the F1-score dropped below 0.50. This indicates that removing majority-class samples caused loss of important crash-related information, negatively affecting model generalisation and prediction accuracy.

Multiclass ROC Curve Interpretation

The Multiclass Receiver Operating Characteristics (ROC) Curve using the One-vs-Rest (OvR) approach for crash severity prediction, shown in Figure 3, evaluates the ability of the classification model to distinguish between different crash severity classes by plotting the True Positive Rate (Sensitivity) against the False Positive Rate

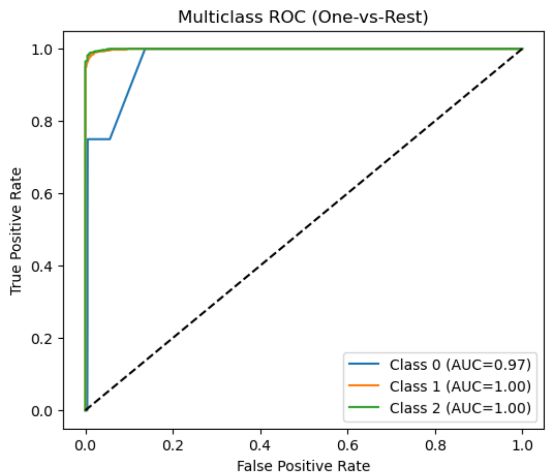

Figure 3: ROU Curve

The model produced very high Area Under the Curve (AUC) values for all crash severity classes, with Class 0 having AUC = 0.97, Class 1 having AUC = 1.00, and Class 2 having AUC = 1.00. These results indicate excellent classification performance across all severity categories. The ROC curves are positioned very close to the upper-left corner of the graph, representing ideal classification performance with high sensitivity and low false positive rates. The dashed diagonal line is the performance of a random classifier with an AUC = 0.50, and since all model curves are above this diagonal line, the model demonstrates strong discriminatory capability in distinguishing between the crash severity classes.

The Confusion Matrix

The classification performance of the three severity classes (Class 0, Class 1, and Class 2) is presented in the confusion matrix in Figure 4. The rows represent the actual crash severity classes, while the columns represent the predicted classes. The diagonal values are the correctly classified instances, while off-diagonal values are the misclassifications. From the matrix, 0 instances were correctly predicted for Class 0, while 3 instances were misclassified as Class1 and 1 instance was misclassified as Class 2.

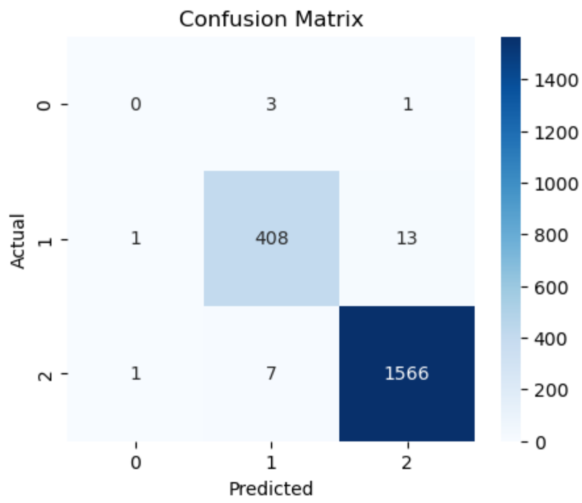

Figure 4: Confusion Matrix for Model Classification Performance

For Class 1, 408 instances were correctly predicted for Class1, while 0 instances were misclassified as Class 1, and 13 instances were misclassified as Class 2. 1566 instances were correctly predicted as Class 2. The results show that the model achieved very high prediction accuracy for Class 1 and Class 2, with the majority of crash cases correctly classified. Also, Class 2 indicates strong model capability in identifying this crash severity category. However, the model failed to correctly classify instances belonging to Class 0. This suggests that Class 0 was either severely underrepresented in the dataset or shared overlapping characteristics with other severity classes.

Feature Importance Interpretation Graph

Figure 5 presents the Feature Importance Interpretation that illustrates the relative importance of different variables used in the machine learning model for crash severity prediction. It measures how much each variable contributes to the model's decision-making process in predicting crash severity outcomes.

Figure 5 indicates that the most influential features in the prediction model are vehicle make, crash location and engine type. It also suggests that the three variables play a dominant

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 5, pp 705-712, May-2026

https://doi.org/10.47001/IRJIET/2026.105094

role in determining crash severity. Other moderately important variables are the vehicle year, time of day, number of cylinders, ABS prescence and vehicle type. This graph provides valuable insight into the key factors that influence road crash severity, and understanding these contributing variables is very critical to transportation safety agencies and policymakers in designing targeted road safety interventions.

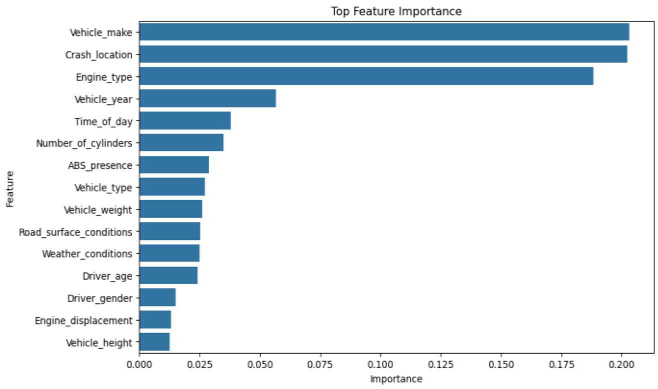

Figure 5: Feature Importance Interpretation Graph

Statistical Testing of Models and Resampling Techniques

The Friedman test was used to compare the performance of multiple classifier-resampling combinations, including Classifier: Logistic Regression, Random Forest, XGBoost and LightGBM; and Resampling techniques: None, SMOTE, Random Oversampling, Undersampling and Hybrid balancing methods. The test yielded a statistic of 8.80 and a p-value of 0.0663. Since the p-value: 0.0663>0.05, it then follows that the null hypothesis was not rejected. i.e., although some classifier-resampling combinations achieved slightly higher predictive scores than others, the differences were not statistically significant at the 5% significance level. It explains that models such as XGBoost, LightGBM, or Random Forest showed numerically superior performance; the improvements recorded were not enough to conclusively state that one approach consistently outperformed the other across all experimental conditions. The findings therefore suggest that several classifier-resampling strategies can effectively model road accident severity, with model selection potentially guided by other considerations.

Also, the Wilcoxon signed-rank test was conducted to determine whether there were statistically significant pairwise differences in performance among the various resampling techniques used. The Wilcoxon signed-rank post-hoc analysis revealed no statistically significant pairwise difference among the evaluated resampling techniques (p>0.05 for all comparisons). This confirms the earlier Friedman test result that suggested that the predictive performances of the classifiers remained comparatively consistent across the different imbalance handling strategies.

V. DISCUSSION

The exploratory analysis revealed substantial imbalance in crash severity distribution, with the majority of records belonging to non-fatal crash categories. The imbalance plot clearly demonstrated the dominance of Class 2 over Classes 1 and 0. Correlation heatmaps further revealed relationships among environmental, roadway and vehicle-related variables.

The classification performance without resampling showed that machine learning models were biased towards majority classes. Logistic Regression achieved an overall accuracy of 84% but failed to identify minority fatal cases, resulting in zero precision and recall of Class 0. Random Forest outperformed Logistic Regression by achieving 98% accuracy and higher weighted F1-score values. Whereas minority class detection remained limited due to the severe imbalance. This confirms that high accuracy alone is insufficient for evaluating models under imbalanced conditions.

The application of resampling methods substantially improved minority class prediction. SMOTE produced the most consistent improvement in recall and F1-score across multiple classifiers. The experiment showed that SMOTE improved F1-score performance by 15-25%, indicating it effectiveness in generating representative minority class samples. Furthermore, ADASYN also improved minority class detection but occasionally introduced noisy synthetic samples, resulting in slightly lower stability compared to SMOTE. Random Oversampling improved minority representation but increased the likelihood of overfitting due to duplicated observations. While Random Under sampling reduced classifier performance because of information loss from majority class removal. reducing majority-class observations may eliminate essential crash patterns and reduce the model's ability to distinguish between severity levels.

The superior performance of XGBoost and LightGBM suggests that gradient boosting algorithms are highly effective in capturing the complex nonlinear relationships among crash-related factors such as driver behaviour, vehicle characteristics, roadway conditions and environmental variables. These models are capable of learning intricate interaction patterns that traditional statistical models may fail to capture. The relatively weaker performance of the Logistic Regression model indicates that linear models may not sufficiently model the complex dynamics associated with crash severity.

However, in crash severity prediction, accurate differentiation between severity levels is essential for effective road safety management and emergency response. The high ROC-AUC values observed in this study indicate that the

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 705-712, May-2026
https://doi.org/10.47001/IRJIET/2026.105094

machine learning model was highly effective in distinguishing between minor, moderate and severe crash outcomes. The ROC analysis demonstrates that the proposed machine learning framework possesses strong predictive and discriminative power for multiclass crash severity prediction, making and excellent sensitivity-specificity trade-off.

The confusion matrices further revealed that resampling significantly improved minority class identification while reducing false negatives.

The research implication from the result is that crash severity is highly feasible using machine learning techniques, and ensemble boosting models outperform conventional approaches in imbalanced traffic datasets. Although resampling techniques can improve minority class representation, they may not always produce substantial gains when powerful classifiers are used.

VI. CONCLUSION

This study investigated the effectiveness of machine learning models and imbalance-handling techniques for road accident severity prediction under highly imbalanced data conditions. The findings revealed that severe class imbalance significantly affects classification performance, particularly for minority fatal crash categories. Traditional models such as Logistic Regression were less effective in identifying minority classes, despite achieving relatively high overall accuracy. While the ensemble-based approaches, especially XGBoost and LightGBM, demonstrated superior predictive performance and robustness in handling nonlinear crash-related relationships.

The experimental results further showed that resampling techniques such as SMOTE, ADASYN, and Random Oversampling improved minority class representation and enhanced recall and F1-score performance. However, the improvements were not significant when compared with strong baseline advanced boosting algorithms' performance. Random Undersampling produced the weakest performance due to loss of important information during the reduction process. Statistical testing using Friedman and Wilcoxon Signed-Rank tests confirmed that the observed differences among classifier-resampling combinations were not statistically significant at the 5% significance level.

In conclusion, the study establishes that machine learning, particularly gradient boosting techniques, offers a reliable and effective framework for crash severity prediction in imbalanced datasets. Future studies may explore deep learning approaches, hybrid resampling methods, larger multi-regional datasets, and real-time intelligent transportation systems to further improve prediction accuracy and model generalisation.

REFERENCES

[1] Amiri, M. A., Afshari, S., & Soltani, A. (2025). Machine learning approaches to traffic accident severity prediction: Addressing class imbalance. *Machine Learning with Applications*, 22 (November), 100792. https://doi.org/10.1016/j.mlwa.2025.100792

[2] Chawla, N. V, Bowyer, K. W., Hall, L. O., & Kegelmeyer, w. P. (2002). SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. http://arxiv.org/abs/2003.09788

[3] Dayanah, I., Melinda, A., Prihutaminingsih, T., & Sarwinda, D. (2023). Tree-Based Ensemble Methods and Their Applications for Predicting Students' *Academic Performance*. 13(3).

[4] Fernández, A., García, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. *Journal of Artificial Intelligence Research*, 61, 863–905. https://doi.org/10.1613/jair.1.11192

[5] Moyo, E., Dzinavatonga, K., & Hakunavanhu, Z. L. (2025). Logistic Regression Modelling of Road Traffic Accident Severity: A Study on Driver Characteristics in Zimbabwe. *INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE*, IX(XI), 3189–3206. https://doi.org/DOI: 10.47772/IJRISS

[6] Obasi, I. C., & Benson, C. (2023). Evaluating the effectiveness of machine learning techniques in forecasting the severity of traffic accidents. *Heliyon*, 9(8), e18812. https://doi.org/10.1016/j.heliyon.2023.e18812

[7] Somvanshi, S., Liu, J., Chakraborty, R., Tamakloe, R., & Das, S. (2026). Predicting Crash Severity using Naturalistic Driving Data and Neural Networks. *International Journal of Intelligent Transportation Systems Research*, 2023. https://doi.org/10.1007/s13177-025-00624-3

[8] Swets, J. A. (1988). Measuring the Accuracy of Diagnostic Systems. *Science*, 240(4857), 1285–1293. online.stat.psu

[9] Taha, K. (2026). Tree-based ensemble learning models for protein-protein interactions detection: a review and experimental evaluation. 1–28.

[10] Xiao, Y., Lin, L., Zhou, H., Tan, Q., Wang, J., Yang, Y., & Xu, Z. (2023). Fatal crashes and rare events logistic regression: an exploratory empirical study.

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 705-712, May-2026
https://doi.org/10.47001/IRJIET/2026.105094

Frontiers in Public Health, 11. https://doi.org/10.3389/fpubh.2023.1294338

Citation of this Article:

Olusola Theophilus Faboya. (2026). Machine Learning-Based Road Accident Fatality Prediction Under Imbalanced Data Conditions: An Evaluation of Resampling and Classification Techniques. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(5), 705-712. Article DOI https://doi.org/10.47001/IRJIET/2026.105094

****

