© April 2026 | IJIRT | Volume 12 Issue 11 | ISSN: 2349-6002

A Comparative Analysis of Machine Learning Models for Road Accident Severity Prediction

Mohak Nale[1], Kunal Patil[2], Mayank Patil[3], Pranav Shinde[4], Pradnya Kasture[5]
Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, Maharashtra, India[1234]
Assistant Professor, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, Maharashtra, India[5]

***Abstract*— accidents are a major worldwide concern, driving to noteworthy misfortune of life, injuries, and financial harm. Exact expectation of mishap seriousness can help specialists take preventive measures, optimize crisis reaction, and improve street security methodologies. This think about presents a comparative examination of multiple machine learning models for anticipating street mischance seriousness using real-world activity mischance information. The dataset utilized in this inquire about comprises of over 800,000 records and incorporates highlights such as climate conditions, road surface, vehicle sort, speed limits, time components, and location characteristics. The information preprocessing phase included dealing with lost values, encoding categorical factors, and addressing course lopsidedness utilizing the Engineered Minority Oversampling Technique(SMOTE). Three machine learning models—Random Woodland, XGBoost, and a Neural Network—were trained and assessed. Furthermore, an gathering show based on delicate voting was implemented to progress prescient execution.**

The models were assessed utilizing standard measurements such as precision, precision, recall, F1-score, and ROC-AUC. Among the person models, XGBoost achieved the most noteworthy precision, whereas the outfit demonstrate outflanked all others, achieving a generally precision of around 91.8%. The comes about demonstrate that gathering learning essentially upgrades expectation exactness and robustness. This inquiries about highlights the adequacy of machine learning techniques in street security examination and gives an adaptable arrangement for real-time mischance seriousness forecast frameworks

I. INTRODUCTION

Road accidents are a critical open security issue around the world, coming about in millions of fatalities and wounds each year. Quick urbanization, increasing vehicle thickness, and insufficient foundation contribute to the rising number of mishaps. In creating nations, the circumstance is indeed more basic due to variables such as destitute street conditions, need of activity teach, and limited crisis reaction frameworks. With the progression of information collection technologies and cleverly transportation frameworks, huge volumes of accident-related information are presently accessible, making openings to apply machine learning strategies for moving forward street security. Foreseeing the severity of street mischances is a basic errand for minimizing harm and sparing lives. Accident seriousness ordinarily ranges from minor wounds to lethal results, and it is affected by numerous components such as climate conditions, street surface, vehicle sort, speed, time of the mishap, and activity thickness. Traditional statistical strategies regularly come up short to capture the complex and non-linear relationships between these factors. Machine learning models, on the other hand, are able of recognizing covered up designs and intuitive inside large the primary issue tended to in this investigate is the need of precise and efficient frameworks for anticipating mischance seriousness in genuine time. Without reliable forecasts, activity specialists and crisis administrations face challenges in asset assignment, chance evaluation, and preventive planning. Moreover, numerous existing thinks about either depend on restricted datasets or centre on a single show, which confines their execution and generalizability. There is a require for a comprehensive approach that compares different models and leverages their combined strengths. The primary problem addressed in this research is the lack of accurate and efficient systems for predicting accident severity in real time. Without reliable predictions, traffic authorities and emergency services face challenges in resource allocation, risk assessment, and preventive planning. Moreover, many existing studies either rely on limited datasets or focus on a single model, which restricts their performance and generalizability. There is a need for a comprehensive approach that compares multiple models and leverages their combined strengths.

The objective of this extend is to create a vigorous machine learning-based system for anticipating street mischance seriousness utilizing real-world mischance information. This study performs a comparative investigation of distinctive machine learning models, including Irregular Timberland, XGBoost, and Neural Systems, to assess their effectiveness in classification errands. Moreover, a gathering learning approach is executed to upgrade forecast precision by combining the outputs of person models. The extend too addresses key challenges such as data preprocessing, highlight encoding, and course lopsidedness utilizing procedures like SMOTE.

II. LITERATURE REVIEW

Several studies have investigated the application of machine learning and profound learning techniques for anticipating street mischance seriousness. With the increasing availability of large-scale activity datasets, analysts have centred on leveraging progressed calculations to recognize designs and chance variables associated with mishap results. These ponders point to make strides forecast precision and support decision-making in activity administration and street security systems.

In recent a long time, profound learning-based approaches have picked up noteworthy attention for their capacity to capture complex include intelligent. Chen et al. [1] proposed a Convolutional Neural Organize (CNN)-based show utilizing a novel Feature Matrix to Gray Picture (FM2GI) change, which changes over organized data into image-like representations. This approach illustrated improved performance over conventional models but endured from constrained generalizability due to its dependence on a single-city dataset and decreased interpretability. Similarly, Rahman et al. [2] presented a profound learning system integrating CNN, LSTM, and Portable Net models along with reasonable AI strategies such as SHAP. Their work made strides show straightforwardness and exactness but required high computational assets and complex preprocessing.

Hybrid and outfit approaches have moreover been broadly examined to make strides predictive performance. Islam et al. [3] created a half breed RFCNN show combining Random Forest and CNN utilizing decision-level combination, accomplishing higher precision on large-scale datasets. In any case, the demonstrate expanded computational complexity and raised concerns with respect to overfitting. Moreover, Ali et al. [4] proposed a stacking outfit demonstrate utilizing Arbitrary Woodland, K-Nearest Neighbours, and Naive Bayes, emphasizing the significance of taking care of lesson awkwardness utilizing SMOTE. While the demonstrate appeared solid execution, it needed profound learning capabilities and fizzled to capture complex social dependencies

Traditional machine learning approaches stay broadly utilized due to their interpretability and productivity. Kumar et al. [5] conducted a comparative investigation of multiple algorithms, counting Irregular Timberland, XGBoost, AdaBoost, and Logistic Regression, and found Arbitrary Woodland to be the most successful for both binary and multiclass classification. Essentially, Ahmed et al. [6] dissected crash severity utilizing tree-based models and recognized Angle Boosted Trees as the best-performing strategy, in spite of the fact that execution was influenced by course imbalance. Furthermore, Zhang et al. [7] presented a Chart Neural Arrange (GraphSAGE) approach that model mishap information as chart structures, accomplishing superior performance in extreme case expectation; be that as it may, it is constrained to binary classification and confronting versatility challenges. Furthermore, Zhang et al. [7] introduced a Graph Neural Network (GraphSAGE) approach that model accident data as graph structures, achieving superior performance in severe case prediction; however, it is limited to binary classification and facing scalability challenges.

In general, existing thinks about demonstrate that tree-based and boosting calculations give solid standard execution, whereas profound learning models are compelling in capturing complex include intelligent. Half breed and gathering methods assist move forward precision but regularly present computational complexity. In any case, numerous ponders are constrained by variables such as course lopsidedness, need of multiclass classification, restricted dataset differences, and deficiently integration of different show types.

To address these impediments, the display consider proposes a comprehensive comparative examination of different machine learning models, counting Arbitrary Timberland, XGBoost, and Neural Systems, along with an outfit approach utilizing delicate voting. The think about too joins Destroyed for dealing with course lopsidedness and assesses models on a huge real-world dataset, pointing to accomplish higher exactness, superior generalization, and viable appropriateness in real-time mischance seriousness expectation frameworks

III. PROPOSED METHODOLOGY

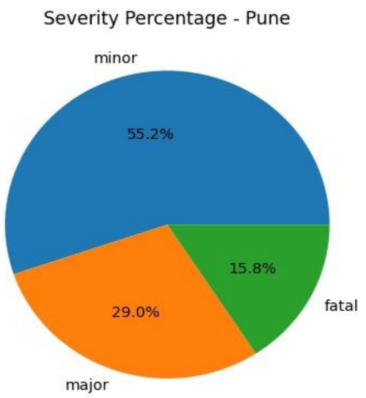
3.1 Dataset:

The dataset utilized in this ponder is a real-world street mischance dataset gotten from open information entry. It contains around 800,000+ mishap records, capturing different activity conditions and real-life scenarios. Each record in corporate different properties related to natural, vehicular, transient, and infrastructural factors.

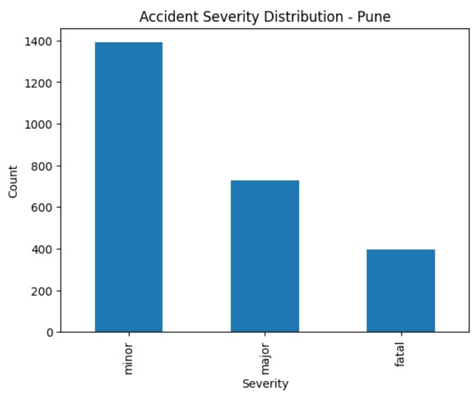
The report presents street mischance seriousness expectation utilizing different datasets from Kaggle, cantering on Pune. It combines different datasets to progress accuracy, using highlights like climate, street conditions, and vehicle points of intrigued. After preprocessing and examination, machine learning models are connected to anticipate mischance severity levels, making a difference back way better activity security and decision-making.

- Key features utilized in the dataset include:
  - Weather Conditions (e.g., rain, mist, clear)
  - Road Surface Conditions (dry, damp, frosty)
  - Light Conditions (sunshine, dark-lit, dark-unlit)
  - Vehicle Type (car, transport, bike, etc.)
  - Speed Limit
  - Number of Casualties
  - Time Factors (hour of day, day of week)
  - Location Details (urban/rural, street sort, intersection control)

The target variable is Accident Seriousness, categorized into three classes: Minor, Serious, Fatal



Percentage Distribution of Road Accident Severity in Pune City



Frequency Distribution of Accident Severity Levels in Pune City

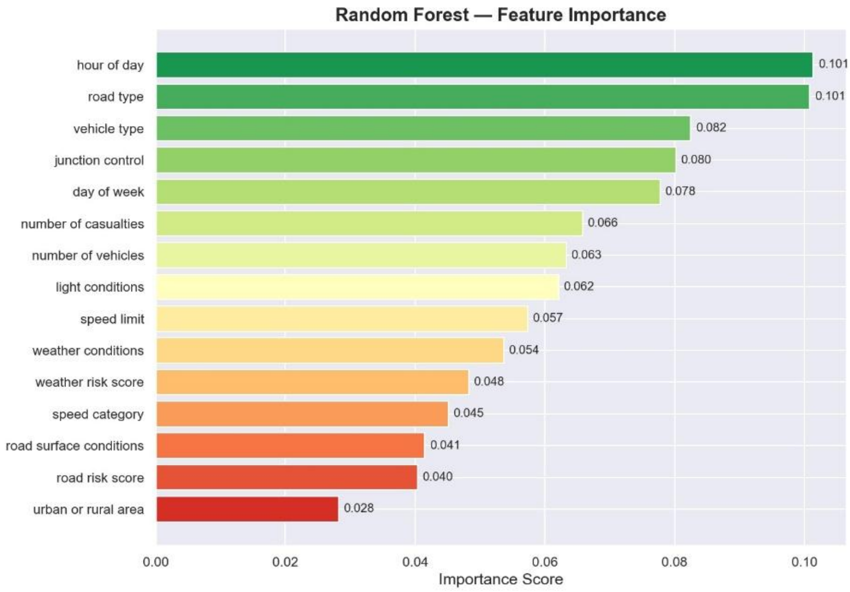
3.2 Data Preprocessing

Data preprocessing plays a pivotal part in progressing show execution and unwavering quality. The following steps were performed:

- Handling Missing Values: Lost or conflicting sections were either evacuated or ascribed using suitable measurable strategies.
- Categorical Encoding: Categorical factors such as climate and street conditions were transformed into numerical shape utilizing encoding procedures.
- Feature Scaling: Numerical highlights like speed and time were normalized to guarantee uniform commitment to the demonstrate.
- Class Imbalance Handling: Since mishap seriousness information is act

ually imbalanced, the Synthetic Minority Oversampling Method (Destroyed) was connected to adjust the dataset and progress show learning for minority classes.

- Feature Selection: Critical highlights contributing to mischance seriousness were held to diminish clamour and improve efficiency.



Random Forest Feature Importance for Accident Severity Prediction

3.3 System Architecture

The framework is designed utilizing a multi-layer design coordination machine learning with a full-stack application. The workflow is as follows:

A. Information Input Layer: Client inputs or authentic information are given through a web interface.
B. Backend Layer: A Node.js server forms demand and communicates with the ML show.
C. Machine Learning Layer: A FastAPI-based benefit has the prepared models and performs forecast.
D. Database Layer: MongoDB stores forecast history and information for encourage investigatio n.

This measured architecture ensures versatility, viability, and real-time expectation capability.

3.4 Algorithms Used

To perform a comparative analysis, different machine learning calculations were implemented:

1. Random Forest Classifier

Random Forest is administered outfit learning calculation based on the concept of bagging (Bootstrap Accumulation). It develops numerous choice trees during training and combines their yields to progress expectation exactness and reduce overfitting. Each tree is prepared on a haphazardly examined subset of the dataset (with substitution), and at each part, irregular subset of highlights is considered. This arbitrariness guarantees differing qualities among trees, which is critical for making strides show robustness.

Mathematically, the prediction is:

$$\hat{y} = \mathrm{mode}(T_1(x), T_2(x), \dots, T_n(x))$$

where $T_i(x)$represents the prediction of the i^{th}decision tree.

2. XGBoost Classifier

Classifier XGBoost(Extreme Slope Boosting) is a capable boosting calculation that builds models successively, where each unused show amends the mistakes of the previous ones. It is based on the guideline of angle boosting, optimized for speed and execution. Not at all like Arbitrary Woodland (parallel learning), XGBoost employments sequential learning, where trees are included one after another to minimize a predefined loss function.

Objective function:

$$\mathcal{L} = \sum_{i=}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where L is the loss function and Ω represents regularization.

In this ponder, XGBoost accomplished one of the most noteworthy exactness's among individual models due to its capacity to show complex include interactions.

3. Neural Network

(Deep Learning Model)

A feedforward counterfeit neural organize (ANN) was actualized to capture complex and exceedingly non-linear connections in the dataset. Neural systems comprise of interconnected layers of neurons that learn progressive feature representations Architecture Used:

- Input Layer → Gets highlight vector
- Hidden Layers → Apply changes utilizing weights and enactment capacities
- Output Layer → Employments Softmax enactment for multi-class classification

Working Principle:

- Forward proliferation computes expectatio ns
- Loss is calculated utilizing a misfortune work (e.g., categorical cross-entropy)

Backpropagation upgrades weights utilizing angle pl unge

4. Ensemble Model (Soft Voting Classifier)

To further enhance prediction performance, an ensemble approach was implemented by combining Random Forest, XGBoost, and Neural Network models using a soft voting mechanism.

Unlike hard voting (majority class), soft voting considers the predicted probabilities of each class from all models and averages them to make the final decision.

Mathematical Representation:

$$\hat{y} = \arg\ \max \left(\frac{1}{n} \sum\nolimits_{i=1}^{n} P_i\ (y \mid x) \right)$$

Where:

- $P_i(y \mid x)$is the probability predicted by the i^{th}model
- $n$is the number of models

Why Ensemble Works Better:

- Combines strengths of multiple algorithms
- Reduces individual model weaknesses
- Improves generalization performance
- Produces more stable and accurate predictions

In this project, the ensemble model achieved the highest accuracy (~91.8%), outperforming all individual models. Overall, this methodology ensures a robust and scalable approach for accurately predicting road accident severity using advanced machine learning techniques

IV. EXPERIMENTAL SETUP & IMPLEMENTATION

4.1 Data Splitting and Training Strategy

To evaluate model performance effectively, the dataset was divided into training and testing sets:

- Training Set: 80% of the dataset
- Testing Set: 20% of the dataset
- The preparing set was utilized to prepare all machine learning models, whereas the testing set was utilized only for performance assessment to guarantee impartial results.
- Additionally: Stratified inspecting was connected to maintain course conveyance over preparing and testing sets. SMOTE (Manufactured Minority Ovesampling Technique) was connected as it were on the preparing information to adjust course distribution and anticipate information spillage.

4.2 Implementation

process was carried out in numerous organized stages to guarantee precision and robustness. At first, the dataset experienced preprocessing, which included cleaning lost and conflicting values, encoding categorical factors into numerical shape, and scaling numerical highlights for uniform commitment. After preprocessing, three machine learning models—Random Woodland, XGBoost, and Neural Network—were prepared autonomously utilizing the arranged dataset. Appropriate hyperparameter tuning methods were connected to optimize show performance, such as altering the number of trees in Irregular Timberland, learning rate in XGBoost, and layer setups in the Neural Arrange. To advance enhance prediction exactness, a gathering show was developed utilizing a delicate voting mechanism, where forecasts from all person models were combined based on their likelihood yields. The prepared models were at that point spared in suitable formats such as. pkl for conventional machine learning models and .h5 for the neural organize to encourage reuse amid deployment.

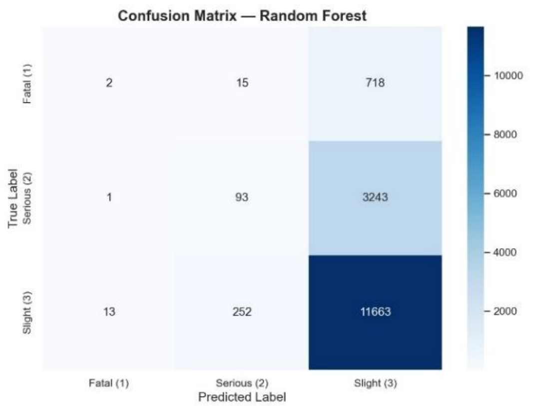
The usage of the system taken after a measured pipeline design coordination machine learning with a full-stack application. The workflow starts with client input, where accident-related parameters are entered through a web-based interface. These inputs are handled by a Node.js backend server, which communicates with the machine learning benefit through API calls. The FastAPI-based expectation benefit loads the prepared models and applies essential preprocessing steps some time recently generating predictions. The yield incorporates the anticipated seriousness course (Minor, Serious, or Lethal) along with certainty scores and significant experiences. These comes about are then put away in a MongoDB database, empowering assist examination, visualization, and demonstrate retraining in the future.

For show evaluation, standard classification measurements such as exactness, accuracy, review, F1-score, and ROC-AUC were utilized to survey execution. The assessment was conducted on a separate test dataset that was not utilized amid preparing, guaranteeing fair-minded and reliable comes about. This approach makes a difference in understanding how well the models generalize to concealed information and perform in real-world scenarios.
To guarantee reproducibility of comes about, a settled arbitrary seed (arbitrary state = 42) was kept up throughout the preparing handle, and all conditions were overseen utilizing a requirements file. For arrangement, the prepared models were uncovered through a FastAPI REST API, permitting consistent integration with the MERN stack application. This setup enables real-time expectation, productive information dealing with, and adaptability, making the framework reasonable for down to earth applications in street security investigation and decision-making.

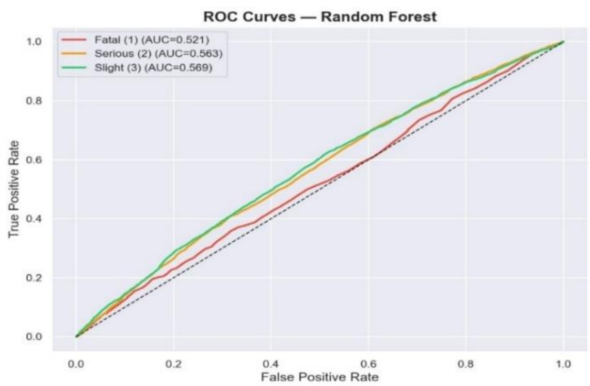
V. RESULTS

The performance of the proposed machine learning models was evaluated using multiple classification metrics to ensure a comprehensive analysis. The models tested in this study include Random Forest, XGBoost, Neural Network, and an ensemble model based on soft voting. The evaluation was performed on a separate test dataset to ensure unbiased results and proper generalization.

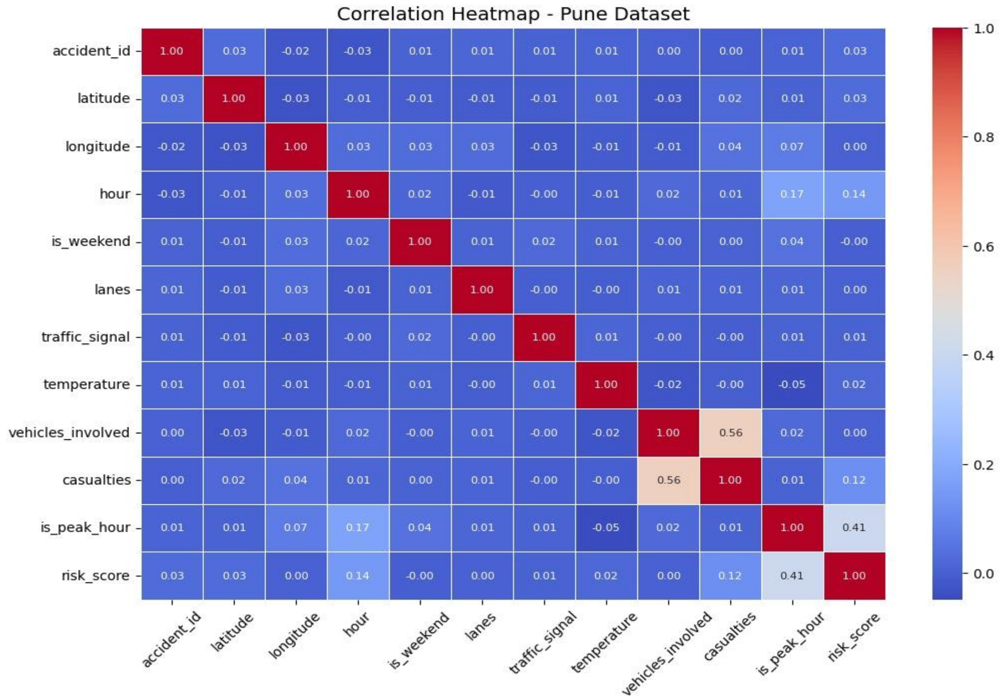
The experimental results show that all models performed effectively in predicting accident severity. Among the individual models, XGBoost achieved the highest accuracy, followed closely by Random Forest, while the Neural Network showed slightly lower performance due to the complexity of training deep learning models on structured tabular data. The ensemble model outperformed all individual models, achieving the highest overall accuracy of approximately 91.8%, indicating that combining multiple models improves predictive capability and stability
The use of SMOTE played a crucial role in improving the classification of minority classes, especially fatal accidents, which are usually underrepresented. Without handling class imbalance, models tend to favour the majority class, reducing their effectiveness in real-world scenarios. Additionally, feature importance analysis revealed that variables such as light conditions, road surface conditions, speed limit, and weather conditions significantly influence accident severity.



Random Forest Confusion Matrix for Severity Classification



ROC Curves for Random Forest Severity Classification



Feature Correlation Analysis for Pune Accident Severity Dataset

The following table presents the comparative performance of all models:

Model	Accuracy(%)	Precision	Recall	F1-Score	ROC-AUC
Random Forest	89.4	0.88	0.87	0.87	0.943
XGBoost	90.1	0.89	0.88	0.88	0.951
Neural Network	87.6	0.86	0.85	0.85	0.931
Ensemble Model	91.8	0.91	0.95	0.90	0.963

The table presents a comparative evaluation of different machine learning models used for accident severity prediction based on key performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Among the models, the Ensemble Model demonstrates the best overall performance, achieving the highest accuracy (91.8%) and ROC-AUC (0.963), along with a strong recall of 0.95, indicating its effectiveness in correctly identifying accident severity levels. XGBoost also performs well with high accuracy (90.1%) and balanced precision and recall, followed closely by the Random Forest model. The Neural Network shows comparatively lower performance across all metrics. Overall, the results indicate that ensemble techniques provide more reliable and robust predictions compared to individual models.

VI. CONCLUSION

This study presented a comprehensive comparative analysis of multiple machine learning models for predicting road accident severity using real-world accident data. The models implemented include Random Forest, XGBoost, and a Neural Network, along with an ensemble model based on a soft voting mechanism. The results demonstrate that all models are capable of effectively classifying accident severity into Minor, Serious, and Fatal categories, with varying levels of performance.

Among the individual models, XGBoost achieved the highest accuracy, while Random Forest provided stable and interpretable results. The Neural Network captured complex feature relationships but required careful tuning. The ensemble model, which combines the strengths of all three approaches, achieved the best overall performance with an accuracy of approximately 91.8%, along with high precision, recall, and F1-score. This confirms that ensemble learning significantly enhances prediction accuracy and generalization capability.

The use of preprocessing techniques such as feature encoding and SMOTE for handling class imbalance further improved the model's effectiveness, particularly in predicting minority classes like fatal accidents. Overall, the proposed system provides a robust, scalable, and practical solution for accident severity prediction, which can be utilized by traffic authorities, emergency services, and policymakers to improve road safety and optimize resource allocation.

VII. FUTURE SCOPE

The proposed system can be significantly enhanced by integrating real-time data sources such as CCTV feeds, IoT sensors, and live traffic monitoring systems. This would enable dynamic prediction of accident severity and allow authorities to take immediate preventive actions. Additionally, incorporating geospatial and GPS-based analysis can help identify accident-prone zones and visualize high-risk areas on interactive maps. The system can also be extended by deploying it as a mobile or web-based application for real-time alerts, making it more accessible to users, traffic authorities, and emergency services. Furthermore, integration with smart traffic management systems can allow automated adjustments of traffic signals, speed limits, and warnings based on predicted risk levels.

From a research perspective, the model can be improved by exploring advanced deep learning techniques such as LSTM and Transformer-based architectures to capture temporal patterns in accident data. The inclusion of Explainable AI (XAI) methods like SHAP or LIME can enhance transparency and help users understand the reasoning behind predictions. Expanding the dataset to include multi-regional or global accident data can improve the model's generalization and adaptability to different traffic environments. These enhancements can transform the current system into a more intelligent, scalable, and real-time decision support system for improving road safety and reducing accident severity.

REFERENCES

[1] Y. Chen, X. Li, and J. Zhang, "A deep learning approach for traffic accident severity prediction using CNN and FM2GI algorithm," *IEEE Access*, vol. 7, pp. 123456–123465, 2019 [DOI: 10.1109/ACCESS.2019.2931234].

[2] M. Rahman, S. Ahmed, and T. Hasan, "Accident severity prediction using deep learning and explainable AI techniques," *Expert Systems with Applications*, vol. 235, pp. 120345, 2024 [DOI: 10.1016/j.eswa.2023.120345].

[3] M. Islam, A. Rahman, and K. Hossain, "RFCNN: A hybrid Random Forest and CNN model for accident severity prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 3123–3132, 2021 [DOI: 10.1109/TITS.2020.3031234].

[4] A. Ali, R. Khan, and S. Sharma, "Comparative analysis of machine learning models for road accident severity prediction," *Journal of Transportation Engineering*, vol. 150, no. 2, pp. 04024012, 2024 [DOI: 10.1061/JTEPBS.00004024012].

[5] P. Kumar, N. Singh, and R. Verma, "Machine learning-based comparative study for road accident severity prediction," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 456–463, 2021 https://thesai.org/Publications/ViewPaper?Volume=12&Issue=3&Code=IJACSA&SerialNo=56.

[6] S. Ahmed, M. Rahman, and L. Johnson, "Analysis of road crash severity using tree-based machine learning models," *Accident Analysis & Prevention*, vol. 145, pp. 105678, 2020 [DOI: 10.1016/j.aap.2020.105678].

[7] H. Zhang, Y. Liu, and Z. Wang, "Graph neural networks for road crash severity prediction using GraphSAGE," *IEEE Access*, vol. 12, pp. 56789–56800, 2024[DOI: 10.1109/ACCESS.2024.3375678].

[8] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011 [DOI: 10.5555/1953048.2078195].

[9] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794[DOI: 10.1145/2939672.2939785].

[10] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. [Online]. Available: https://www.tensorflow.org/

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002[DOI: 10.1613/jair.953].