International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 08 | Aug - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Advanced Crime Prediction Using Machine Learning Along with Crime Forecasting and Categorization

RANGISETTI HARSHA BHARADWAJ
Post Graduate Student, M.C.A. Department of Information Technology, Jawaharlal Nehru Technological University, Hyderabad,
harshabharadwajrangisetti@gmail.com

DR. V. UMARANI
Professor, Department of Information Technology,
Jawaharlal Nehru Technological University Hyderabad,
umarani@jntuh.ac.in

DR.SUNITHA VANAMALA
Lecturer, Department Of Computer Science ,TSWRDCW, Warangal East, Warangal,Telangana,India
sunithavanamala@gmail.com

ABSTRACT
Crime has become one of the most pressing concerns in modern societies, with urbanization and population growth contributing to its rising complexity. Traditional crime analysis methods are largely reactive and fail to capture the hidden trends required for proactive policing. To address this challenge, this research introduces a machine learning-based crime prediction system that utilizes Decision Tree and Bagging Classifier algorithms. A dataset comprising 505,063 records from Portland, Oregon (2015–2023) was preprocessed and analyzed. The models achieved strong performance, recording 98% accuracy on training data and 95% accuracy on testing data. The system classifies crimes into 20 distinct categories and provides predictive insights that support law enforcement in hotspot identification and preventive action. By employing ensemble-based methods, the framework demonstrates both robustness and interpretability, thereby contributing to the development of intelligent data-driven solutions for public safety.

KEYWORDS
Crime Prediction, Machine Learning, Decision Tree, Bagging Classifier, Ensemble Learning, Predictive Policing, Flask Application

1.INTRODUCTION

Public safety is a vital concern for every nation, and crime continues to present significant challenges to law enforcement agencies. Conventional approaches rely heavily on manual data analysis and reactive measures, which are inadequate when faced with the growing scale and complexity of criminal activity. Predictive crime analysis provides a transformative solution by leveraging machine learning to identify hidden patterns, classify crime types, and forecast trends based on historical records.

In this work, crime prediction is studied using data from Portland, Oregon, covering more than half a million crime incidents collected over an eight-year span. The proposed model applies Decision Tree and Bagging Classifier algorithms to categorize crimes into 20 classes, enabling finer granularity in predictions. The integration of spatial and temporal attributes allows for better identification of crime hotspots and trends. Furthermore, the development of a web-based interface ensures that the predictive system is not only accurate but also accessible and usable in real-time by agencies and policymakers.

2 .OBJECTIVES

The primary objectives of this research are:

- To evaluate the efficiency of Decision Tree and Bagging Classifier algorithms in crime prediction.
- To analyze a large publicly available dataset and extract features relevant for accurate classification.
- To categorize crimes into 20 different classes for detailed crime pattern analysis.
- To measure model performance using accuracy, precision, recall, and confusion matrices.
- To implement a web-based Flask application that enables real-time prediction and visualization.

3. METHODOLOGY

The methodology followed in this study involves dataset acquisition, preprocessing, model development, evaluation, and deployment. The dataset, sourced from Kaggle, contains 505,063 crime records from Portland recorded between 2015 and 2023. It includes 15 features such as case number, crime category, date, time, and location. Data preprocessing steps involved managing missing values, normalizing attributes, and splitting the dataset into training (80%) and testing (20%) subsets.

Model Architecture

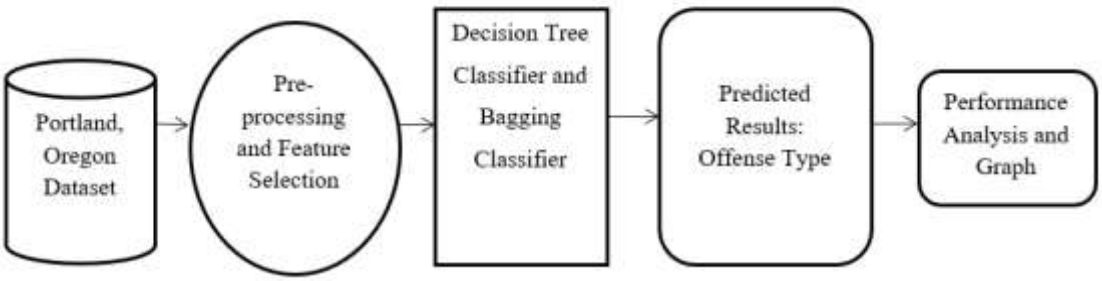

Fig. 1: System Architecture

Two algorithms were implemented. The Decision Tree Classifier was chosen for its interpretability and efficiency in handling categorical data. The Bagging Classifier, an ensemble approach combining multiple decision trees, was employed to improve robustness, reduce overfitting, and enhance prediction accuracy. Both models were assessed using accuracy scores, confusion matrices, and feature importance analysis.

To facilitate usability, the predictive model was deployed through a Flask-based web application. The interface allows users to upload datasets, input crime details, and obtain predictions, with visual outputs such as confusion matrices and accuracy graphs for interpretability.

4. RESULTS

The evaluation results confirmed the effectiveness of the proposed models. The Decision Tree Classifier achieved 98% accuracy on the training set and 95% on the testing set, while the Bagging Classifier obtained similar accuracy with greater resilience against overfitting. Compared with traditional Random Forest approaches, the models offered improved interpretability and computational efficiency.

The system was capable of classifying crimes into 20 categories, including theft, assault, burglary, fraud, and homicide. These results demonstrate the model's potential to assist law enforcement in detailed crime analysis and hotspot identification. Graphical results, such as confusion matrices, further validated the accuracy and reliability of the predictions.

5. OUTPUT SCREENS

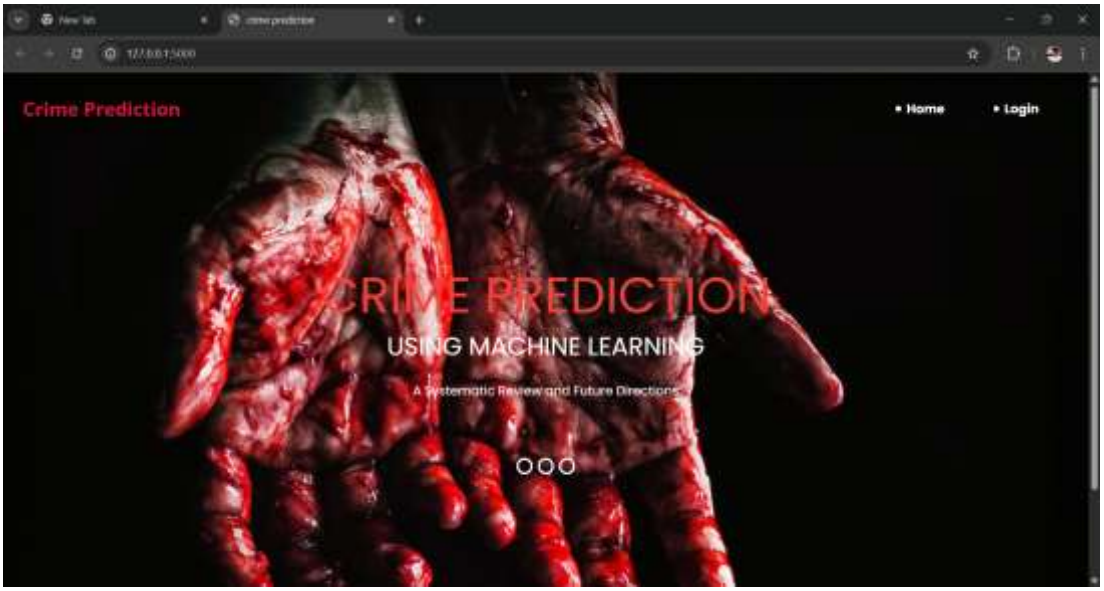

Fig. 2: Home Screen Interface

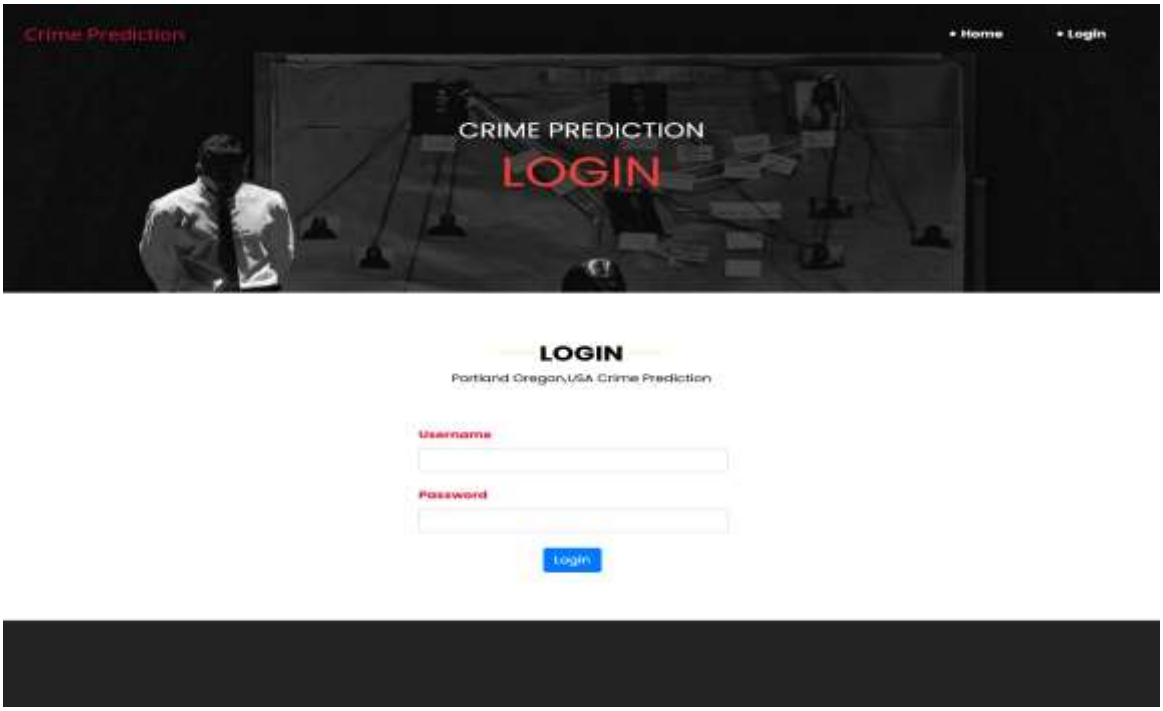

Fig. 3: Login Page

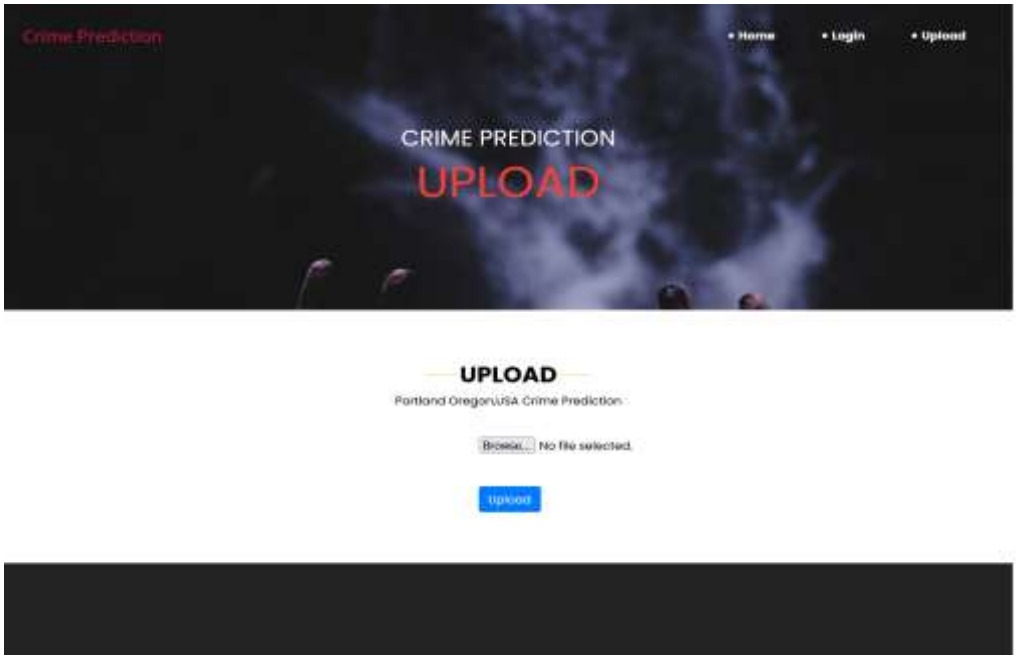

Fig. 4: File Upload Page

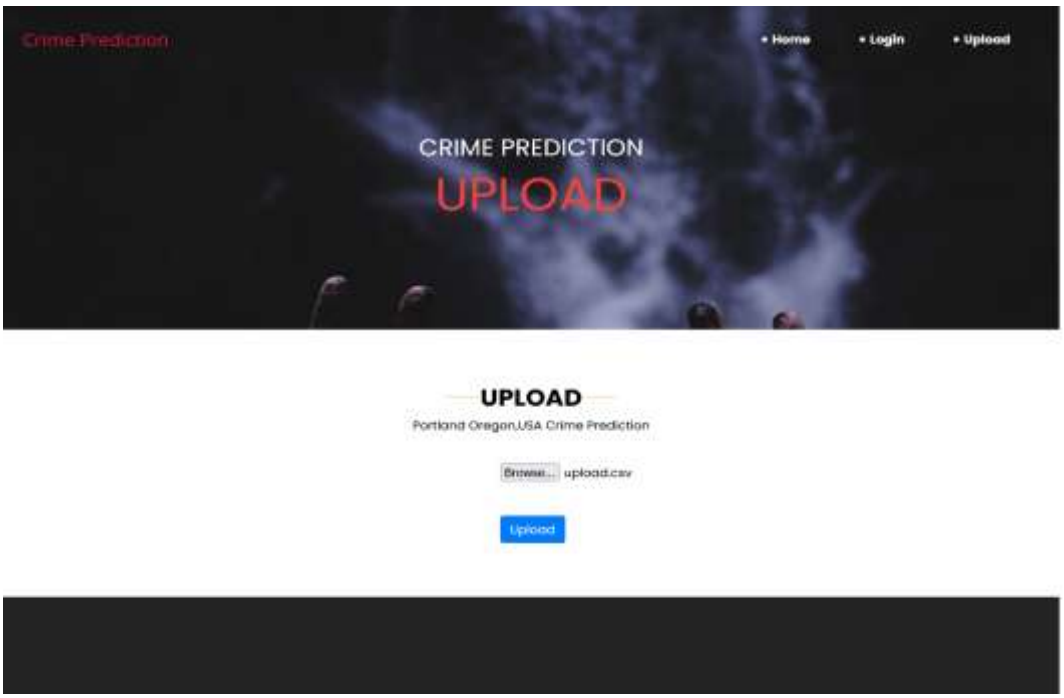

Fig. 5: Owner Uploaded Files

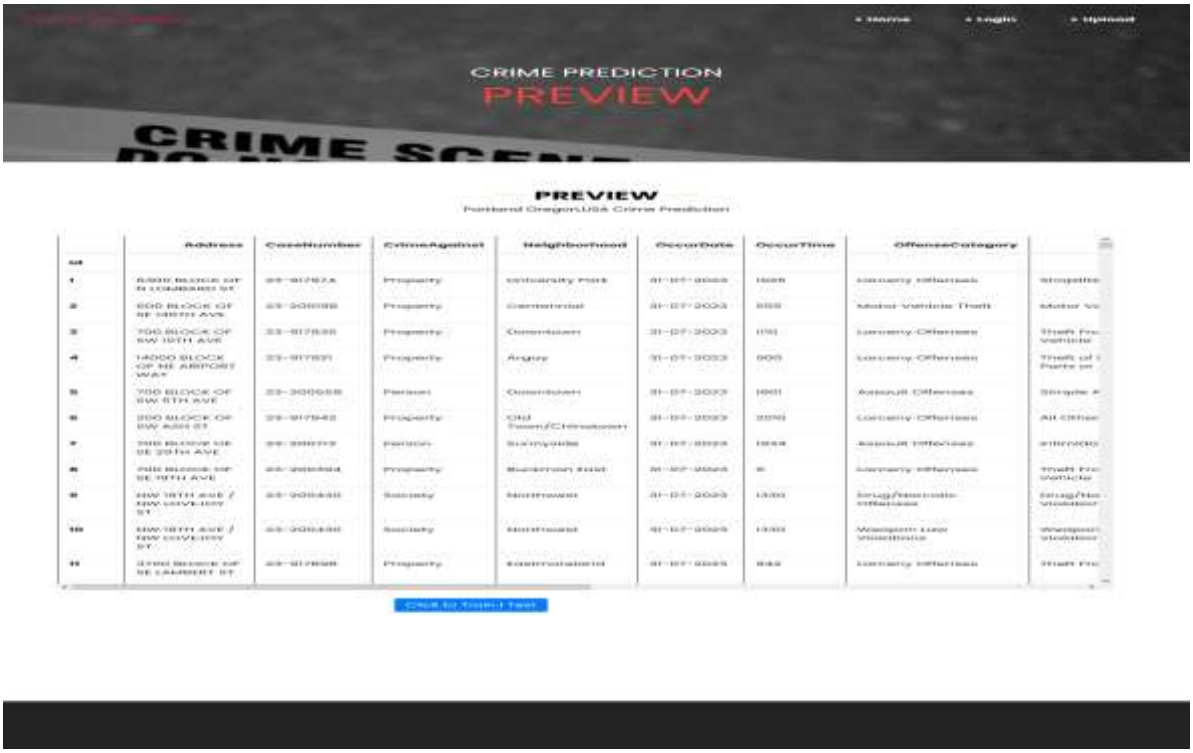

Fig. 6: User Dashboard –Files View Page

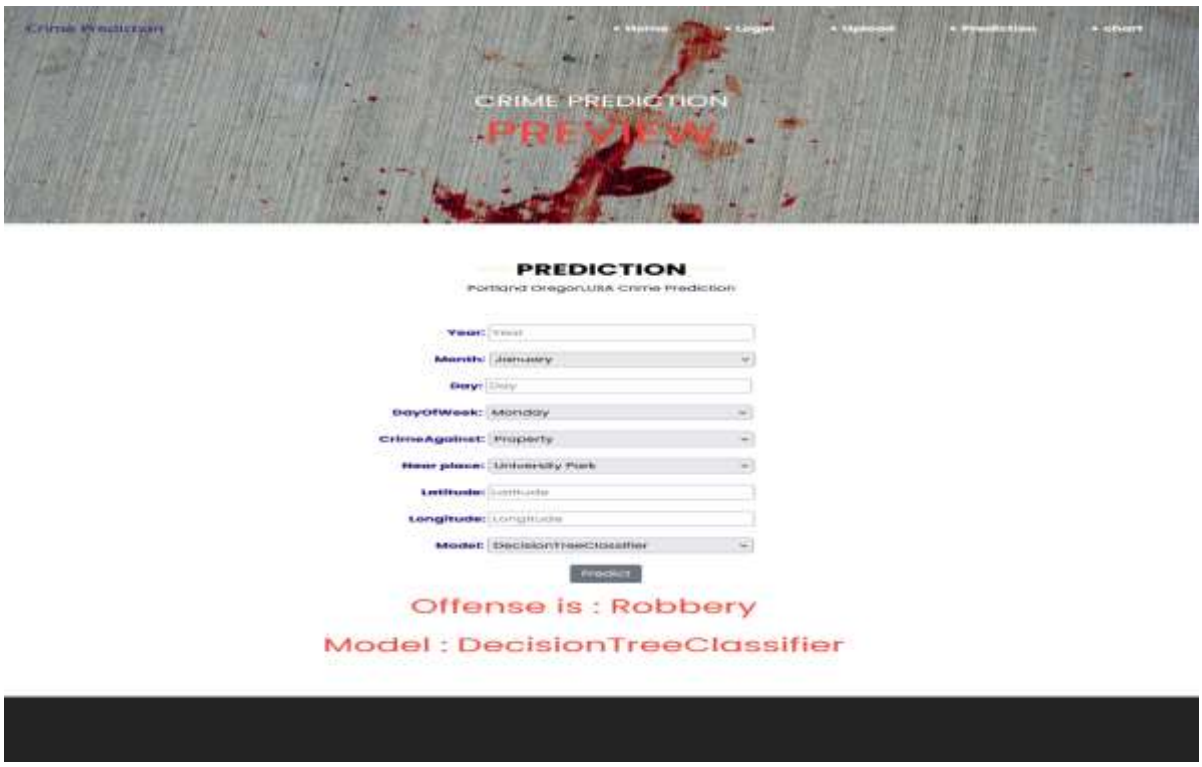

Fig. 7: Server Requests View Page

6. CONCLUSION

This study presents a machine learning-based crime prediction framework using Decision Tree and Bagging Classifier algorithms. Achieving 95% accuracy on testing data, the models successfully predict and classify crimes into 20 categories, offering actionable insights to law enforcement agencies. The system combines robust ensemble techniques with a web-based interface, enabling practical usability and data-driven decision making.

Key contributions include the integration of interpretable machine learning models, the application of ensemble methods for improved prediction stability, and the development of a scalable web platform. Future work will focus on incorporating advanced deep learning approaches, addressing class imbalance, integrating external datasets such as socioeconomic indicators, and enabling real-time predictive capabilities. These enhancements will further strengthen the system's role in supporting smart city initiatives and predictive policing.

REFERENCES

[1] Safat, W., Asghar, S., & Gillani, S. A. (2020). Comparative study of machine learning and deep learning for crime forecasting. *International Journal of Computer Applications.*

[2] Chun, S. A., Paturu, V. A., Yuan, S., Pathak, R., Atluri, V., & Adam, N. R. (2019). Deep neural networks for crime prediction. *IEEE Transactions on Big Data.*

[3] Raza, D. M., & Victor, D. B. (2018). Regional crime prediction using random forest. *International Journal of Advanced Computer Science and Applications.*

[4] Kshatri, S. S., Singh, D., Narain, B., Bhatia, S., Quasim, M. T., & Sinha, G. R. (2019). A stacked ensemble approach for crime prediction using machine learning. *Journal of Data Mining and Knowledge Discovery.*

[5] Shah, N., Bhagat, N., & Shah, M. (2018). Crime forecasting through machine learning and computer vision. *International Journal of Computer Science Research.*