CRIME RATE PREDICTION AND ANALYSIS USING MACHINE LEARNING

[1]Students: K. Yogitha, MD. Salma, G. Adithya, S. Akash
[2]Assistant Professor: Mr. B. Veeru
[1,2] Department of Computer Science Engineering,
Balaji Institute of Technology and Science, Warangal, Telangana

ABSTRACT

Crime analysis and prediction is a methodical approach for identifying the crime. This system can predict region which have high chance for crime occurrences and visualize crime prone area. Using the concept of data mining we can extract previously unknown, useful information from an unstructured data. The extraction of new information is predicted using the existing datasets. Crimes are treacherous and common social problem faced worldwide. Crimes affect the quality of life, economic growth and reputation of nation. The aim of securing the society from crimes, there is a need for advanced systems and new approaches for improving the crime analytic for protecting their communities.

In recent years, report points out that the crimes in India have seen a spike. The report adds that the cases of murder, rapes, and kidnapping have seen a rise. Most of countries in the world have seen a remarkable increase in the crime rate. There is no particular reason for any trouble for criminal activities. To prevent this problem in police sector, have to predict crime rate using machine learning techniques. mainly we used Random Forest classification technique.

Doi:10.48047/ijiee.2025.15.5.14

1. INTRODUCTION

Crime is an act that is prohibited by law and is punishable by a fine, imprisonment, or other legal action. Crime prediction is significant to determine increase or decrease in crime rate from preceding years. Complex and advanced ML algorithms can be used to identify patterns in data to detect and predict criminal activities.

It can also be used to create predictive models that identify the likelihood of a certain crime occurring in a particular area or time frame. Using machine learning, law enforcement professionals can gain insight into criminal behavior, enabling them to better detect and prevent crime

Crime rate prediction is a multifaceted endeavor aimed at anticipating the likelihood of criminal activity within specific locations and time frames. Its primary objective is to empower proactive crime prevention strategies, facilitate efficient resource allocation for law enforcement agencies, and inform evidence-based policy decisions concerning public safety. This predictive capability is achieved through the meticulous analysis of diverse datasets, encompassing historical crime records, socioeconomic indicators, demographic distributions, and other pertinent variables.

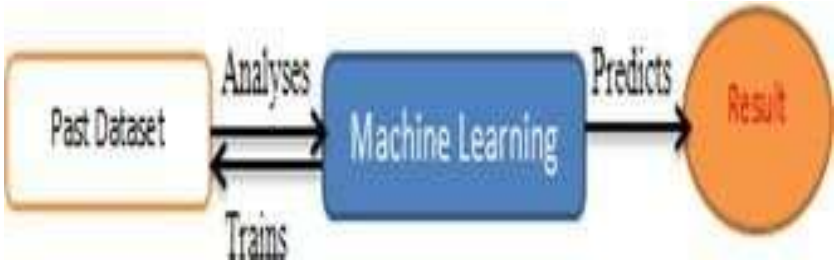
Doi:10.48047/ijiee.2025.15.5.14

Methodologies employed in this analysis range from traditional statistical modelling to advanced machine learning algorithms and sophisticated geospatial analysis techniques.

Machine learning algorithms, such as neural networks, decision trees, and support vector machines, are proving highly effective in discerning complex patterns and generating accurate predictions. Geospatial analysis, through the utilization of mapping and spatial statistical techniques, allows for the precise identification of crime hotspots. The integration of big data analytics, enabling the processing and analysis of vast datasets from diverse sources, further enhances the precision and scope of crime rate predictions.

crime rate prediction shifts the focus from past occurrences to future probabilities. This predictive endeavor leverages analytical models, often incorporating sophisticated machine learning and statistical techniques, to forecast the likelihood of criminal events. By assimilating historical crime data, demographic information, socioeconomic indicators, and other relevant variables, predictive models aim to identify areas and time frames where crime is more prone to occur. This proactive approach equips law enforcement with the capacity to anticipate criminal activity, enabling them to strategically allocate resources and implement preventive measures. In essence, crime analysis serves to elucidate "what has happened," while crime rate prediction endeavors to anticipate "what might happen." Both methodologies are indispensable tools for fostering safer communities, with crime analysis providing the foundation for understanding criminal behavior and crime rate prediction facilitating proactive intervention.

This involves meticulously dissecting crime types, pinpointing their geographic distribution, and charting their temporal occurrence. Through this process, crime analysts identify hotspots, discern recurring crime patterns, and strive to comprehend the multifaceted factors that contribute to criminal behavior. The ultimate objective of crime analysis is to furnish law enforcement agencies with actionable intelligence, empowering them to refine crime prevention strategies and enhance investigative capabilities.



Machine Learning:

Allows organizations to make predictions about possible outcomes based on historical data.

Analysis:

The process of inspecting, cleansing, and modelling data to discover useful information.

Doi:10.48047/ijiee.2025.15.5.14

2. LITERATURE SURVEY

- C. S. Ogbodo, "Law, human rights, crime and society," Afr. Hum. Rights Law J., vol. 8, no. 1, pp. 1–15, Jan.2024.
- G. Garcia-Zanabria et al., "CriPAV: Street-level crime patterns analysis and visualization," IEEE Trans.Vis. Compute. Graph., vol. 28, no. 12, pp. 4000–4015, Dec. 2021.Leverage advanced technologies without compromising the security and privacy of individuals.
- Y. Xie, S. Shekhar, and Y. Li, "Statistically-robust clustering techniques for mapping spatial hotspots: A survey," ACM Compute. Surv., vol. 55, no. 2, pp. 1–38, Feb. 2022.
- Intelligent Crime Anomaly Detection in Smart Cities, Authors: Sharmila Chackravarthy, Steven Schmitt, Li Yang, Issued on: 2018.
- By considering above surveys I can develop a project that is based on the previous crime rate.by developing this we can identify the upcoming and next generation of crime rate and which area can be effected by the crime and which crime they are effected by people.

By using Machine learning techniques are:

Supervised learning: used labeled data to train algorithms to predict outcomes and recognize patterns.

Linear Regression: this algorithm is used to predict a continuous outcomes variable

Doi:10.48047/ijiee.2025.15.5.14

Random Forests: This is an ensemble method that combines multiple decision trees

3. PROBLEM STATEMENT

- **Crime as a Social Concern** – Crime affects public safety, economic growth, and overall quality of life.
- **Limitations of Traditional Methods** – Manual crime analysis is time-consuming and less effective in identifying patterns.
- **Need for an Intelligent System** – A data-driven approach can improve crime prediction and prevention.
- **Objective of the Project** – Develop a machine learning- based system to analyze crime trends and predict future occurrences.
- **Crime Pattern Recognition** – Identify trends and correlations in crime data.
- **Crime Hotspot Prediction** – Forecast areas with high crime probabilities.
- **Real-time Crime Monitoring** – Provide alerts and updates on potential criminal activities
- the development of machine learning models capable of accurately predicting crime hotspots and temporal trends; the identification and integration of relevant features from diverse data sources, including historical crime records, census data, and environmental factors; the evaluation and optimization of model performance through rigorous validation and testing; and the development of a user- friendly interface that enables law enforcement agencies to access and utilize the predictive insights generated by the system.

Doi:10.48047/ijiee.2025.15.5.14

4. METHODOLOGIES

1.Data Collection

Sources: Gather crime data from reliable sources such as national crime bureaus, police records, or social media platforms.

Types of Data: Collect historical crime data, including types of crimes, locations, and timestamps.

Data Cleaning: Ensure data is accurate, complete, and consistent.

2.Data Preprocessing

Feature Engineering: Extract relevant features from the data, such as time of day, day of the week, month, year, location (latitude and longitude), and crime type.

Normalization/Scaling: Scale numerical features to ensure they are on the same scale, which can improve model performance.

3.Model Selection and Training

Machine Learning Algorithms: Choose appropriate algorithms based on the nature of the data and the problem. Commonly used algorithms include:

Random Forest Regression: Useful for predicting continuous values like crime rates.

Linear Regression: Can be used for predicting crime rates over time.

Decision Trees and Random Forest Classification: Useful for predicting the likelihood of specific types of crimes.

Doi:10.48047/ijiee.2025.15.5.14

K-Nearest Neighbors (KNN): Can be used for both classification and regression tasks.

Training and Validation: Split the data into training and testing sets to evaluate model performance.

4. Model Evaluation

Metrics: Use metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared for regression models, and accuracy, precision, recall, and F1-score for classification models.

Hyper parameter Tuning: Adjust model parameters to optimize performance.

5. Prediction and Visualization

Prediction: Use the trained model to predict future crime rates or types.

Visualization: Employ visualization tools to display crime trends, hotspots, and patterns over time and space.

6. Deployment

Application Development: Create a user-friendly interface for law enforcement agencies to input parameters and receive predictions.

Integration with Existing Systems: Integrate the model with existing crime management systems for seamless data flow and decision- making.

.

Doi:10.48047/ijiee.2025.15.5.14

1. Random Forest Classification

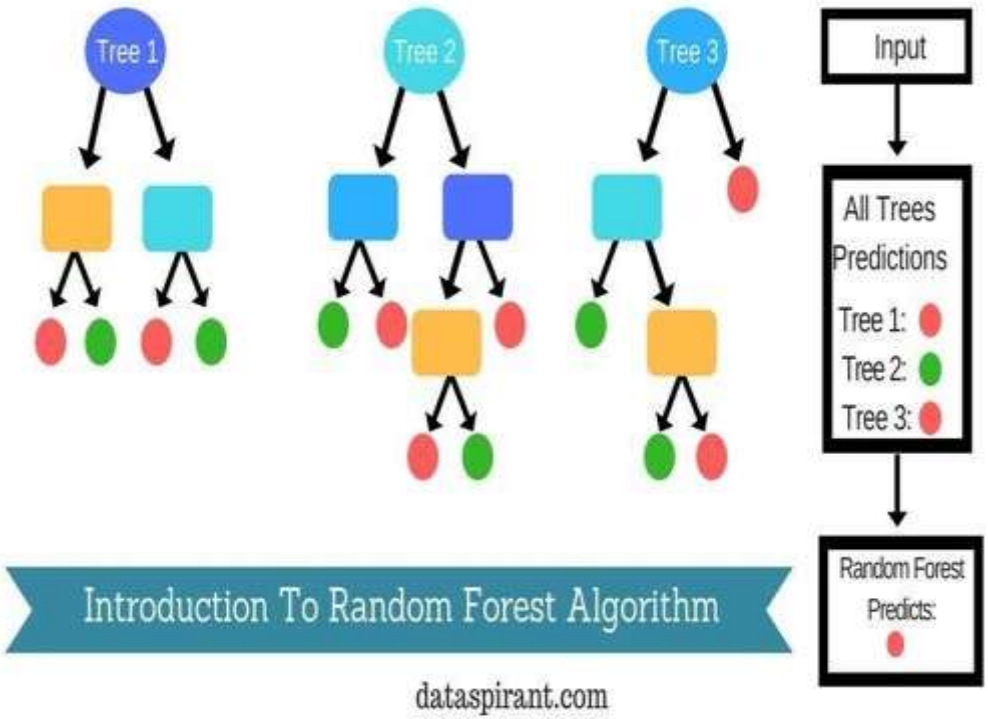
Random forest classification is an ensemble machine learning algorithm that combines multiple decision trees to classify data. It is widely used for its ability to handle both classification and regression tasks effectively.

How a Random Forest Classification Works

1. **Decision Trees:** Random forest creates multiple decision trees, each trained on a random subset of the data and features. This process is known as "bagging" or bootstrap aggregating, which helps reduce over fitting by introducing randomness.

2. **Feature Selection:** At each node of a decision tree, instead of considering all features, only a random subset is evaluated for splitting. This further enhances the randomness and diversity among trees, reducing correlation between them.

3. **Prediction:** For classification tasks, each decision tree predicts a class, and the final prediction is determined by a majority vote among all trees. For regression tasks, the predictions from all trees are averaged to obtain the final result.



Doi:10.48047/ijiee.2025.15.5.14

5.CONCULSION

In this paper focused on building predictive models for crime frequencies per crime type per month.

The crime rates in India are increasing day by day due to many factors such as increase in poverty, implementation, corruption, etc.

The proposed model id very useful for both the investigating agencies and the police official in taking necessary steps to reduce crime.

The project helps the crime analysis these crime networks by means of various interactive visualization.

Future enhancement of this research work on training bots to predict the crime prone areas by using machine learning techniques.

Since, machine learning is similar to data mining advanced concept of machine learning can be used for better prediction.

The data privacy, reliability, accuracy can be improved for enhanced prediction.

Doi:10.48047/ijiee.2025.15.5.14

REFERENCE

1. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima" A Brain Tumor Identification and Classification Using Deep Learning based on CNN-LSTM Method" Computers and Electrical Engineering , 101 (2022) 107960
2. Ramdas Vankdothu,.Mohd Abdul Hameed "Adaptive features selection and EDNN based brain image recognition on the internet of medical things", Computers and Electrical Engineering , 103 (2022) 108338.
3. Ramdas Vankdothu,.Mohd Abdul Hameed,Ayesha Ameen,Raheem,Unnisa " Brain image identification and classification on Internet of Medical Things in healthcare system using support value based deep neural network" Computers and Electrical Engineering,102(2022) 108196.
4. Ramdas Vankdothu,.Mohd Abdul Hameed" Brain tumor segmentation of MR images using SVM and fuzzy classifier in machine learning" Measurement: Sensors Journal,Volume 24, 2022, 100440 .
5. Ramdas Vankdothu,.Mohd Abdul Hameed" Brain tumor MRI images identification and classification based on the recurrent convolutional neural network" Measurement: Sensors Journal,Volume 24, 2022, 100412 .
6. Bhukya Madhu, M.Venu Gopala Chari, Ramdas Vankdothu,.Arun Kumar Silivery,Veerender Aerranagula " Intrusion detection models for IOT networks via deep learning approaches " Measurement: Sensors Journal,Volume 25, 2022, 100641
7. Mohd Thousif Ahemad ,Mohd Abdul Hameed, Ramdas Vankdothu" COVID-19 detection and classification for machine learning methods using human genomic data" Measurement: Sensors Journal,Volume 24, 2022, 100537
8. S. Rakesh [a], NagaratnaP. Hegde [b], M. VenuGopalachari [c], D. Jayaram [c], Bhukya Madhu [d], MohdAbdul Hameed [a], Ramdas Vankdothu [e], L.K. Suresh Kumar "Moving object detection using modified GMM based background subtraction" Measurement: Sensors ,Journal,Volume 30, 2023, 100898

Doi:10.48047/ijiee.2025.15.5.14

9. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Efficient Detectionof Brain Tumor Using Unsupervised Modified Deep Belief Network in Big Data" Journal of Adv Research in Dynamical & Control Systems, Vol. 12, 2020.

10. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Internet of Medical Things of Brain Image Recognition Algorithm and High Performance Computing by Convolutional Neural Network" International Journal of Advanced Science and Technology, Vol. 29, No. 6, (2020), pp. 2875 – 2881

11. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Convolutional Neural Network-Based Brain Image Recognition Algorithm And High-Performance Computing", Journal Of Critical Reviews,Vol 7, Issue 08, 2020(Scopus Indexed)

12. Ramdas Vankdothu, Dr.Mohd Abdul Hameed "A Security Applicable with Deep Learning Algorithm for Big Data Analysis",Test Engineering & Management Journal,January-February 2020

13. Ramdas Vankdothu, G. Shyama Chandra Prasad " A Study on Privacy Applicable Deep Learning Schemes for Big Data" Complexity International Journal, Volume 23, Issue 2, July-August 2019

14. Ramdas Vankdothu, Dr.Mohd Abdul Hameed, Husnah Fatima " Brain Image Recognition using Internet of Medical Things based Support Value based Adaptive Deep Neural Network" The International journal of analytical and experimental modal analysis, Volume XII, Issue IV, April/2020

15. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima" Adaptive Features Selection and EDNN based Brain Image Recognition In Internet Of Medical Things " Journal of Engineering Sciences, Vol 11,Issue 4 , April/ 2020(UGC Care Journal)

16. Ramdas Vankdothu, Dr.Mohd Abdul Hameed " Implementation of a Privacy based Deep Learning Algorithm for Big Data Analytics", Complexity International Journal , Volume 24, Issue 01, Jan 2020

17. Ramdas Vankdothu, G. Shyama Chandra Prasad" A Survey On Big Data Analytics: Challenges, Open Research Issues and Tools" International Journal For Innovative Engineering and Management Research,Vol 08 Issue08, Aug 2019.

18. Vankdothu, R., Hameed, M.A. "An Effective Congestion and Interference Secure Routing Protocol for Internet of Things Applications in Wireless Sensor Network " Wireless Personal Communication Journal 140, 143–161 (2025)

Doi:10.48047/ijiee.2025.15.5.14

19. Vankdothu, R., Bhukya, H. & Bhukya, R.R. “Hybrid TDR-MI Based Wireless Sensor Network for Underground Water Pipeline Leakage Detection and Localization Using Pressure Residuals and Classifiers Wireless Personal Communications 139, 803–823 (2024).

20. Vankdothu, R., Cheng, X. “Energy Efficient TDMA and Secure Based MAC Protocol for WSN Using AQL Coding and ASGWI Clustering”. Wireless Personal Communications 136, 2125–2143 (2024)

21. Vankdothu, R., Hameed, M.A., Fatima, H. *et al.* Multicast Scaling in Heterogeneous Wireless Sensor Networks for Security and Time Efficiency. Wireless Personal Communications (2025).

22. Vankdothu, R., Hameed, M.A., Fatima, H. *et al.* Multicast Scaling in Heterogeneous Wireless Sensor Networks for Security and Time Efficiency. Wireless Personal Communications (2025)

23. Ramdas Vankdothu, Mohd Abdul Hameed” Brain MRI Images for Tumor Detection using Storage Optimization Technique”,Mobile Radio Communications and 5G Networks,Lecture Notes in Networks and Systems,425-437, Springer .

24. Bandi Krishna , Ramdas Vankdothu , Varun Revuri and B. Prashanth” A brain tumor identification using convolution neural network in the deep learning” MATEC Web of Conferences 392, 01131 (2024) ,https://doi.org/10.1051/matecconf/202439201131 ICMED 2024

25. .S. Ogbodo, “Law human rights, crime and society,” Afr. Hum. Rights Law J., vol .8, no.1, pp. 1-15, Jan.2024.

26. G. Garcia -Zanabria et al.,”CriPAV: Street -level crime patterns analysis and visualization,” IEEE Trans.Vis.pp.4000-4015,Dec. 2021.Leverage advanced technologies without compromising the security and privacy of individuals.

27. Y. Xie, S. Shaker, and Y. Li,”Statistically-robust clustering techniques for mapping spatial hotspots: A survey,”ACM Compute. Surv., vol. 55, no. 2, pp. 1-38, Feb.2022.

Doi:10.48047/ijiee.2025.15.5.14

7.BIBILOGRAPHY

I'm Yogitha. I am currently in my 8th semester of Computer Science in the Bachelor's Degree at Balaji Institute of Technology and Science. My research interest is done based on "**CRIME RATE PREDICTION AND ANALYSIS USING MACHINE LEARNING"**

I'm Adithya. I am currently in my 8th semester of Computer Science in the Bachelor's Degree at Balaji Institute of Technology and Science. My research interest is done based on **"CRIME RATE PREDICTION AND ANALYSIS USING MACHINE LEARNING"**

I'm Akash. I am currently in my 8th semester of Computer Science in the Bachelor's Degree at Balaji Institute of Technology and Science. My research interest is done based on **"CRIME RATE PREDICTION AND ANALYSIS USING MACHINE LEARNING"**

I'm Salma. I am currently in my 8th semester of Computer Science in the Bachelor's Degree at Balaji Institute of Technology and Science. My research interest is done based on **"CRIME RATE PRDICTION AND ANALYSIS USING MACHINE LEARNING"**

Doi:10.48047/ijiee.2025.15.5.14

Doi:10.48047/ijiee.2025.15.5.14