International Journal of Multidisciplinary
Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 8, Issue 8, August 2025

CRIMERATE-AI - PREICTIVE ANALYTICS FOR CRIME TRENDS IN INDIA

Dr. Vidya Pol, Syed Sheesh Abul Hussain

Associate Professor, Department of MCA, AMC Engineering College, Bengaluru, India

Student, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: Crime forecasting is a crucial tool for governments and law enforcement agencies to anticipate crime patterns and plan effective interventions. CrimeRate-AI is a web-based application that predicts future crime rates for Indian states and union territories using advanced statistical and machine learning models. Built with Python, Flask, and modern visualization libraries, the system supports multiple forecasting techniques including SARIMA, Prophet, XGBoost, and LSTM. It covers five key crime categories—Murder, Rape, Kidnapping, Robbery, and Theft—based on historical data from 2001 to 2012. The application features an interactive dashboard for visualization, downloadable reports, clustering-based state grouping, and scenario simulations. By integrating interpretability tools, the platform not only provides predictions but also insights into factors influencing crime trends.

KEYWORDS: Crime Prediction, SARIMA, Prophet, LSTM, XGBoost, Flask, Visualization, Crime Analytics, India

I. INTRODUCTION

Rising crime rates pose a significant challenge for policymakers and law enforcement agencies. Traditional crime analysis methods rely heavily on past records and manual assessment, limiting their ability to identify emerging trends. The CrimeRate-AI system is designed as an end-to-end predictive analytics platform that uses historical data to generate actionable forecasts. It provides state-wise and crime-wise predictions, enabling stakeholders to plan targeted interventions.

II. LITERATURE SURVEY

[1] Time-series models such as ARIMA and SARIMA have been extensively used in crime forecasting due to their ability to model seasonality and trends.
[2]Facebook Prophet has gained popularity for its flexibility in handling multiple seasonalities and missing data.
[3]Machine learning models such as XGBoost and Random Forest leverage feature engineering to improve short-term forecast accuracy.
[4] LSTM networks capture non-linear temporal dependencies and outperform traditional models when sufficient data is available
[5]Several studies highlight that ensemble models combining multiple approaches often yield more robust predictions than single-model solutions.

EXISTING SYSTEM
Current crime analytics platforms are primarily descriptive, focusing on visualizing historical data rather than forecasting future trends. Many lack multi-model forecasting capabilities, advanced interpretability, or user-friendly visualization tools. Furthermore, they often do not allow users to compare models, export reports, or perform scenario analysis.

PROPOSED SYSTEM
The proposed CrimeRate-AI system addresses these limitations by integrating:
- Multiple forecasting models (SARIMA, Prophet, XGBoost, LSTM) for model comparison and selection.
- Interactive visualizations using Chart.js for trend plots and clustering maps.
- Downloadable reports in PDF and CSV formats.
- State clustering based on similar crime patterns.

- Interpretability features like feature importance and SHAP values for machine learning models.

The system predicts future crime counts for each state and crime category, enabling more informed policy-making.

III. SYSTEM ARCHITECTURE

The system consists of:
- Frontend: HTML/CSS/JS interface with Chart.js visualizations.
- Backend: Flask API serving predictions from pre-trained models.
- Database: Historical crime data stored locally or in the cloud.
- Model Layer: Time-series models (SARIMA, Prophet) and machine learning models (XGBoost, LSTM) trained per crime type and state.
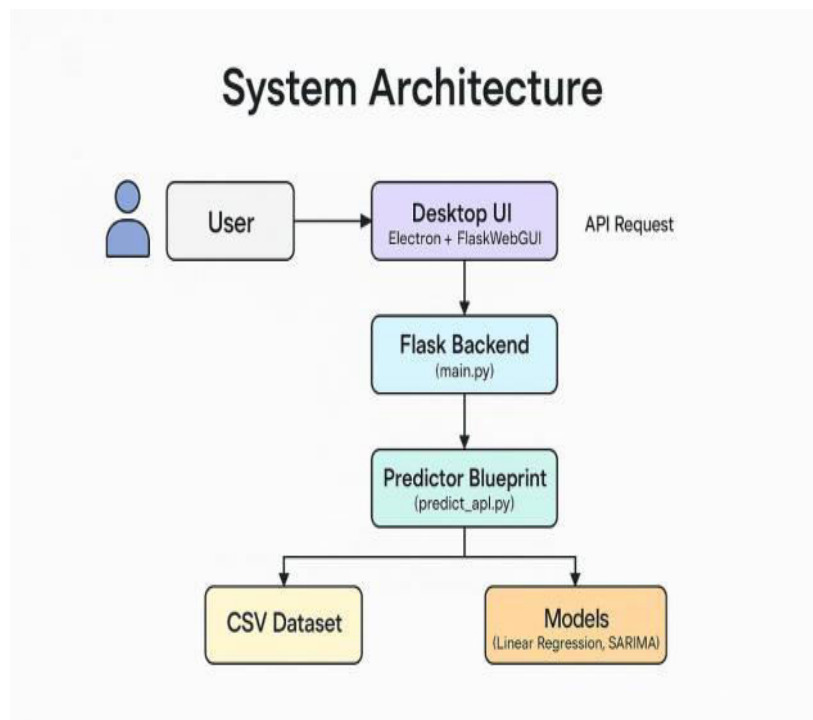- Export & Reporting: CSV and PDF report generation for forecasts and analysis.

Fig 3.1 System Architecture

IV. METHODOLOGY

1. Data Preprocessing: Handle missing values, normalize counts, generate lag features, and split data for training/testing.
2. Model Training:
- SARIMA: Seasonal differencing and parameter tuning.
- Prophet: Automatic changepoint detection and seasonality fitting.
- XGBoost: Gradient boosting on lagged features and external regressors.
- LSTM: Sequence modeling with dropout to prevent overfitting.
3. Evaluation: RMSE, MAE, and MAPE metrics computed for each model.
4. Deployment: Save trained models and serve them through Flask API endpoints.

V. DESIGN & IMPLEMENTATION

- User Interface: Simple navigation with tabs for forecasts, visualizations, clustering, and downloads.
- Model Storage: Serialized model files for fast prediction loading.
- Real-Time Visualization: Forecast plots with confidence intervals.
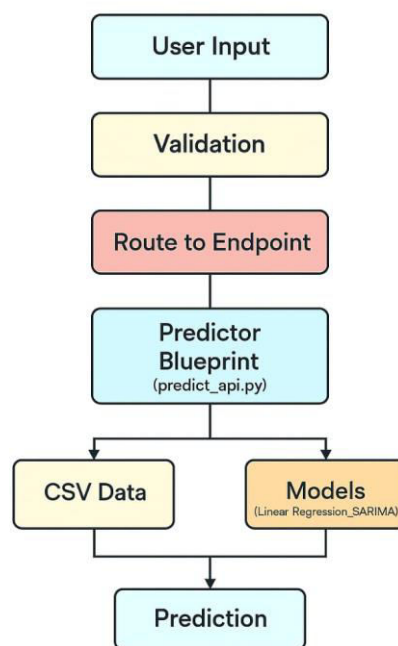- Export Features: Generate downloadable CSV/PDF files.

Fig 5.1 Flowchart Working

VI. OUTCOME OF RESEARCH

The system successfully forecasts multi-year crime trends for different states. Example: Predicted Murder cases in Maharashtra for 2026: 357.73. SARIMA and Prophet models provide stable long-term forecasts, while XGBoost and LSTM offer improved short-term accuracy. Clustering reveals states with similar crime trends, supporting collaborative strategies.

VII. RESULT AND DISCUSSION

CrimeRate-AI provides a practical, scalable solution for state-level crime prediction. The ensemble of models ensures robustness, while the interactive interface enhances usability for non-technical users. Limitations include the coarse yearly granularity of the dataset and potential missing socio-economic features.

VIII. CONCLUSION

The CrimeRate-AI system demonstrates the potential of AI-driven forecasting in crime prevention and resource allocation. Future work will expand the dataset to monthly resolution, integrate socio-economic indicators, add automated model selection (AutoML), and develop real-time alert systems.

REFERENCES

[1] Box, G. E. P., et al., Time Series Analysis: Forecasting and Control.
[2] Hyndman, R. J., & Athanasopoulos, G., Forecasting: Principles and Practice.
[3] Taylor, S. J., & Letham, B., "Prophet: Forecasting at Scale".
[4] Chen, T., & Guestrin, C., "XGBoost: A Scalable Tree Boosting System".
[5] Hochreiter, S., & Schmidhuber, J., "Long Short-Term Memory".

INTERNATIONAL JOURNAL OF

MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |