Hybrid Crime Prediction using GRU and ARIMAX Models

A. Sudhakar
Lakireddy Balireddy College of Engineering,Dept.of CSE Mylavaram
sudhakar.atchala@gmail.com

Dirisala Chaitanya Nandini
Lakireddy Balireddy College of Engineering,Dept.of CSE Mylavaram
dirisalachaitanyanandini@gmail.com

Pingala Ganga Bhavani
Lakireddy Balireddy College of Engineering,Dept.of CSE Mylavaram
pingalagangabhavani@gmail.com

Lamu Lakshmi Priyanka
Lakireddy Balireddy College of Engineering,Dept.of CSE Mylavaram
lamulakshmipriyanka@gmail.com

2025 8th International Conference on Trends in Electronics and Informatics (ICOEI) | 979-8-3315-4460-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICOEI65986.2025.11013209

Abstract*— *Crime prediction and analysis are crucial for enhancing public safety and optimizing law enforcement strategies. This paper proposes a hybrid crime forecasting framework that integrates deep learning and statistical modeling to analyze crime patterns across Indian states and districts. The approach combines a Bidirectional Gated Recurrent Unit (GRU) model and Seasonal Auto-Regressive Integrated Moving Average with Exogenous Variables (SARIMAX) to leverage historical crime data spanning 12 years. The GRU model effectively captures long-term temporal dependencies and complex patterns, while SARIMAX incorporates historical trends and seasonal variations. Their predictions are merged through a weighted hybrid model to enhance forecasting accuracy. To improve predictive performance, feature engineering techniques such as cyclical encoding and MinMax scaling are applied to preprocess temporal and numerical data. The model demonstrates its effectiveness in predicting crime distributions across 30+ categories, including violent crimes, property crimes, and other offenses. Performance is evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 Score, ensuring robust validation. Additionally, visualization tools provide intuitive insights into predicted crime trends, aiding policymakers and law enforcement agencies in resource allocation and preventive measures. This study introduces a novel integration of deep learning and statistical models for crime prediction, offering a data-driven decision-support system to address a critical societal challenge in predictive analytics.

***Keywords*— Crime Prediction, Crime Analysis, Hybrid Model, Bidirectional GRU, SARIMAX, Machine Learning, Predictive Analytics, Temporal Data, Crime Trends, Deep Learning, Feature Engineering, Law Enforcement, Public Safety, Resource Allocation, Crime Forecasting.**

I. INTRODUCTION

Crime has always been a problem for every society across the globe, which threatens public safety and economic stability as well as societal well-being. Law enforcement organisations can efficiently allocate resources, make well-informed choices, and even create proactive crime prevention initiatives with the use of crime prediction and analysis. As data science and artificial intelligence (AI) continue to grow rapidly, predictive analytics has become a powerful tool to analyze complex crime patterns, to identify potential trends, and to forecast future occurrences of crimes.

As large scale, high dimensional crime data becomes available, this opens the door to accurate crime prediction using state of the art machine learning (ML) and deep learning (DL) techniques. However, this data is inherently complex with temporal dependencies, spatial heterogeneity, and seasonalities, and thus reliable predictions require robust methodologies. While ARIMA is an effective model for time series forecasting, they are inadequate in modelling the nonlinear patterns and the complex interdependencies. Conversely, deep learning models like recurrent neural networks (RNNs) and its variants, gated recurrent units (GRUs) are good at learning complex temporal patterns, but they may not be good at interpretability and long term seasonal trends. In this study, we introduce a hybrid strategy that combines the advantages of both approaches by utilising a statistical model called SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors) and a deep learning model called Bidirectional GRU. The Bidirectional GRU captures temporal patterns and nonlinear relationships in the data and the SARIMAX models seasonality and the influence of external variables using exogenous variables. We integrate these models to increase the predictive accuracy and to have a more complete picture of the trends of crime.

To assess the efficacy of the proposed hybrid model, we used a rich dataset of district wise crime data for India over a 12 year period (2001–2012). It contains temporal and spatial information, and diverse crime categories (murder, theft, kidnapping, and other Indian Penal Code (IPC) offenses). Feature scaling, label encoding and cyclical encoding of temporal features were part of the data preprocessing pipeline to make the model learn more complex patterns. To demonstrate the advantages of the hybrid architecture, we further expand our methodology by conducting a comparison study of standalone GRU and SARIMAX models. Additionally, the performance metrics of mean absolute error (MAE), mean squared error (MSE), and R^2 score are used to quantitatively measure the model accuracy. The results demonstrate the promise of the hybrid model to produce reliable predictions alongside interpretable insights for practical use.

This research combines statistical rigor with the adaptability of deep learning and makes a contribution to the burgeoning field of crime analytics by providing a new methodology that can be applied to different datasets and regions. The results are intended to assist policymakers, law enforcement agencies and urban planners with improving public safety and allocating resources for crime prevention.

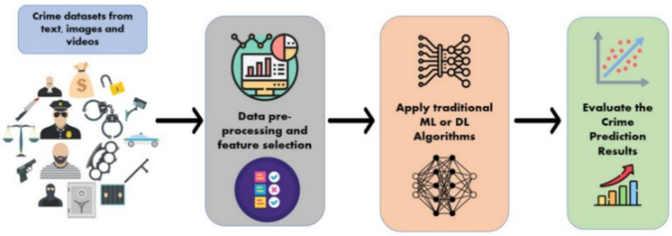

Fig 1: Crime Prediction

II. LITERATURE REVIEW

In recent years, crime prediction using the machine learning (ML) and deep learning (DL) techniques has attracted much attention. Different algorithms to predict crime events have been studied in various ways to help law enforcement agencies to optimize the resource allocation and prevent the crime. We summarize key studies in the field, and highlight their limitations. GRU, CNN and Autoencoders were combined by Selvakumari et al. [1] to classify the crime. The model increased accuracy, however, it was not scalable for large datasets and did not incorporate spatiotemporal dependencies or external factors, thereby limiting the ability to forecast. Decision trees and neural networks are some of the ML and DL models reviewed by Mandalapu et al. [2] for crime prediction. The review is comprehensive but does not introduce a new approach nor does it address challenges in real-time crime forecasting or handling high dimensional data sets.

LSTM networks were used by Safat et al. [3] for crime time series forecasting but they could not capture the spatial patterns or the seasonal fluctuations in crime data; both of which are very important for accurate predictions. Moreover, the model was limited in dealing with factors outside the model. A crime prediction model using deep neural networks (DNNs) introduced by Chun et al. [4] was effective at modeling non-linear relationships. The study didn't address imbalanced datasets or how external variables like weather or socio economic conditions affect crime patterns. Mandalapu et al. [5] showed the significance of feature engineering and model selection, however, they did not investigate hybrid models which combine the power of deep learning with statistical models such as ARIMA to capture temporal and seasonal crime patterns.

Azeez and Aravindhar [6] proposed a hybrid solution based on using deep learning in conjunction with ARIMA for crime prediction. Although it had promising results, it was prone to overfit and was not tested on various crime categories and large datasets. Computer vision and machine learning were used by Shah et al. [7] for crime forecasting with surveillance data. The model proved promising for real time crime prediction, however, it was not able to address the issue of data quality and scalability, and was not integrated with temporal and spatial data for accurate predictions.

For instance, Kang and Kang [8] employed multi modal data such as social media and crime records for crime prediction. Combining data sources was effective, but privacy and data quality concerns were not, and scalability in real world applications was untested. Real time crime forecasting was conducted by Wang et al. [9] using deep learning. Nevertheless, their model was never fully evaluated against different crime categories, and the ability of the model to adapt to new crime patterns was never completely tested. While Wang et al. [10] introduced a ternarization framework for real time crime forecasting, their model's reliance on fixed class thresholds may not be appropriate for all crime types, especially those with skewed distributions. The study did not also consider hybrid approaches that combine statistical models with deep learning for better accuracy.

These studies show the promise of ML and DL for crime prediction, but also indicate several challenges that are not addressed, such as dealing with large datasets, capturing spatiotemporal dependencies, and using external factors. To fill these gaps, this research integrates deep learning (GRU) with statistical models (SARIMAX) to make more accurate and interpretable crime predictions.

III. METHODOLOGY

The crime trends prediction and analysis in this research was done through a hybrid model based on Gated Recurrent Units (GRU), Convolutional Neural Networks (CNN), and SARIMAX. The methodology included dataset preprocessing, model development and evaluation.

1. Data processing

The dataset contained crime data across Indian states and districts from 2001-2012, including multiple crime categories.
Missing values were addressed using mode imputation.
Categorical features like STATE/UT and DISTRICT were label-encoded, and YEAR was transformed into cyclical features (YEAR_sin, YEAR_cos) to capture temporal trends.
Features were normalized using MinMaxScaler to ensure uniform scaling.

2. Model Development

2.1 GRU-Based Deep Learning Model

The GRU model, designed to handle sequential dependencies, was constructed with:
A Bidirectional GRU layer to capture long-term dependencies.
A dense output layer predicting values for each crime category.

2.2 SARIMAX Time-Series Model

SARIMAX was used for time-series forecasting, leveraging historical crime data and exogenous features like YEAR_sin and YEAR_cos. The ARIMA and seasonal orders were tuned to model crime trends and seasonality effectively.

3. Hybrid Model

Proceedings of the 8th International Conference on Trends in Electronics and Informatics (ICOEI-2025)
IEEE Xplore Part Number: CFP25J32-ART; ISBN: 979-8-3315-4460-7

The final predictions were derived by combining GRU and SARIMAX outputs using a weighted average formula:
Hybrid Prediction=0.7×GRU Prediction+0.3×SARIMAX Prediction
This approach integrated the strengths of both models, balancing non-linear pattern detection and time-series forecasting.

4. Evaluation and Visualization
The model's performance was evaluated using metrics like MAE, MSE, and R-Squared (R^2).
Visualizations, such as pie charts, depicted the distribution of predicted crime types for specific regions and years.

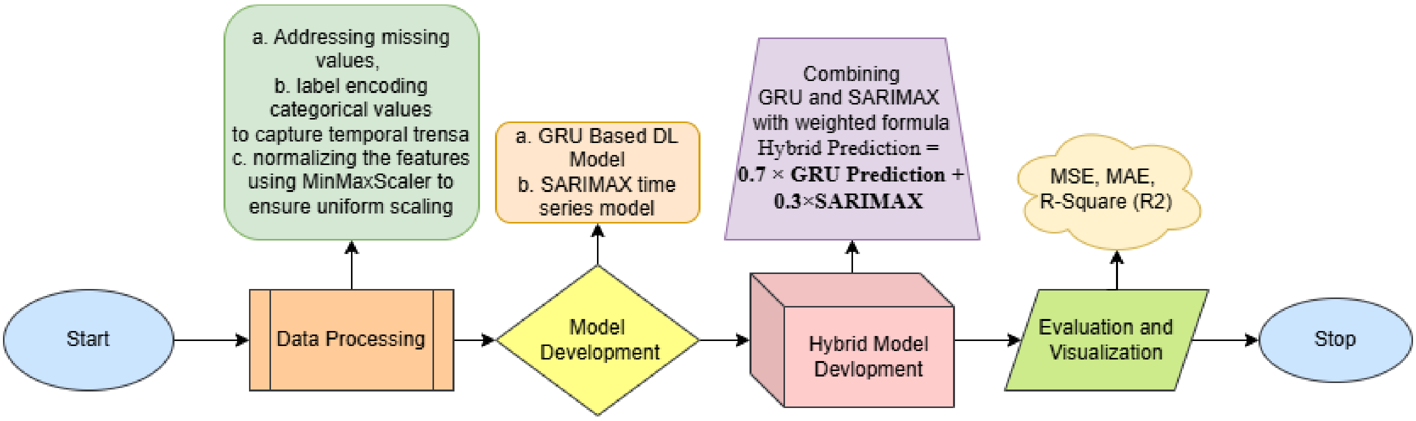

Fig 2: Methodology Diagram

A. Dataset Description

Dataset Overview
This study utilizes a dataset obtained from **Kaggle**, comprising crime statistics reported across various districts within states and union territories (UTs) in India. The dataset contains **9,017 rows** and **35 columns**, with each row representing a unique combination of a district and year. The data provides insights into diverse categories of crimes under the Indian Penal Code (IPC), enabling a comprehensive analysis of crime trends, distributions, and correlations. It can be observed from the following reference [21].

Dataset Attributes
The dataset includes the following key attributes:

1. **Location and Temporal Information:**
   STATE/UT: Name of the state or union territory.
   DISTRICT: Name of the district within the state/UT.
   YEAR: Year corresponding to the crime data.
2. **Crime Categories**: The crime statistics are categorized into multiple types, summarized below:
   Example:
   MURDER: Total number of murder cases reported.
   ATTEMPT TO MURDER: instances of attempted homicide.
   CULPABLE HOMICIDE NOT AMOUNTING TO MURDER: Homicide cases
   KIDNAPPING & ABDUCTION: Total cases of kidnapping and abduction.
3. **Aggregate Metrics:**
   TOTAL IPC CRIMES: Aggregated total of all IPC crime cases for a given district and year.

Dataset Characteristics

- Number of Rows: 9,017
- Number of Columns: 35
- Granularity: Each row corresponds to a unique district-year combination.
- Numerical Columns: All crime-related columns contain non-negative integers representing the count of reported cases.

Visualization:

1. **Top 10 States with Highest Total IPC Crimes**: A horizontal bar plot showcasing the states with the highest crime rates.

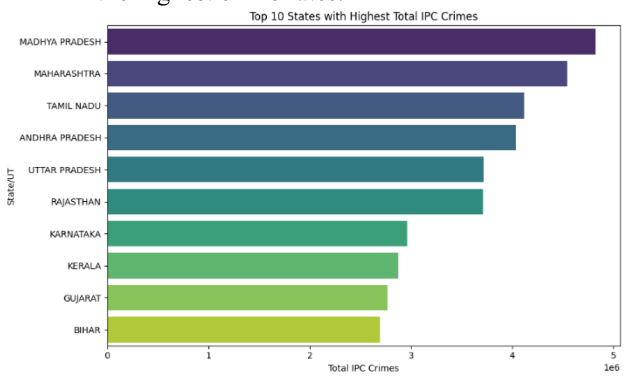

Fig 3: Bar chart of Top 10 States with Highest Total IPC Crimes

2. **Yearly Trend of Total IPC Crimes**: A line plot illustrating the trend of IPC crimes over the years.

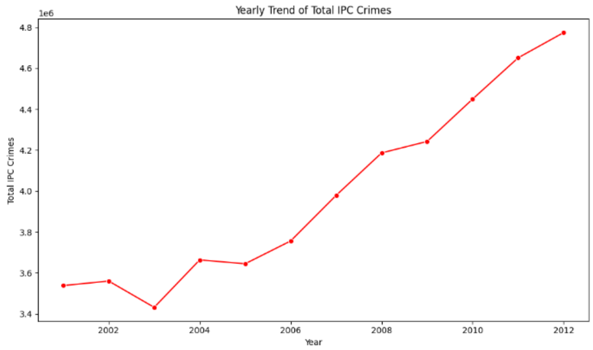

Fig 4: Line Plot of Yearly Trend of Total IPC Crimes

C. Understanding Algorithm Development:

The proposed hybrid model combines **Gated Recurrent Units (GRU)** and **SARIMAX** to address both non-linear and linear temporal patterns in crime forecasting. Below is the mathematical justification for each algorithm and their integration.

1. Gated Recurrent Units (GRU)

A recurrent neural network optimised for sequential data is called a GRU. In order to capture long-term relationships without fading gradients, its architecture uses gating methods to regulate the information flow. The core equation of GRU are:

Hidden State:

ht=zt∗ht−1+(1−zt)∗h~t [1]

This is done by combining previous hidden state and candidate state together with update gate. The GRU's output at each timestep t, yt, is given by: yt=Wy·ht+byy [1]

For this project:

Inputs xt: Encoded features (YEARsin, YEARcos, STATE/UT, DISTRICT

Output yt: Predicted crime counts for each category.

2. Seasonal ARIMAX (SARIMAX)

SARIMAX extends ARIMA by incorporating seasonal components and exogenous variables. It models the crime data as a combination of:

- Seasonal Components: SARIMA(P,D,Q,m): $\Phi P(B^{m})\nabla_{m}^{D} yt=\Theta Q(B^{m})\epsilon t$
- Exogenous Variables: The model incorporates external predictors Xt: yt=β0+β1Xt,1+…+βkXt,k+ϵt
- SARIMAX predicts crime data as: Yt = ARIMA component + Exogenous Variables

In this research: Exogenous variables: YEAR sin, YEAR cos, YEAR scaled. Seasonal differencing handles cyclic crime trends over years.

3. Hybrid Model Integration

The hybrid model combines the predictions from GRU (y_{GRU}) and SARIMAX ($y_{SARIMAX}$) using weighted averaging to leverage the strengths of both models:

$$\mathbf{y_{Hybrid}=w1\cdot y_{GRU}+w2\cdot y_{SARIMAX}}$$

Where: w1+w2=1

In this project: w1=0.7 (GRU) and w2=0.3 (SARIMAX).

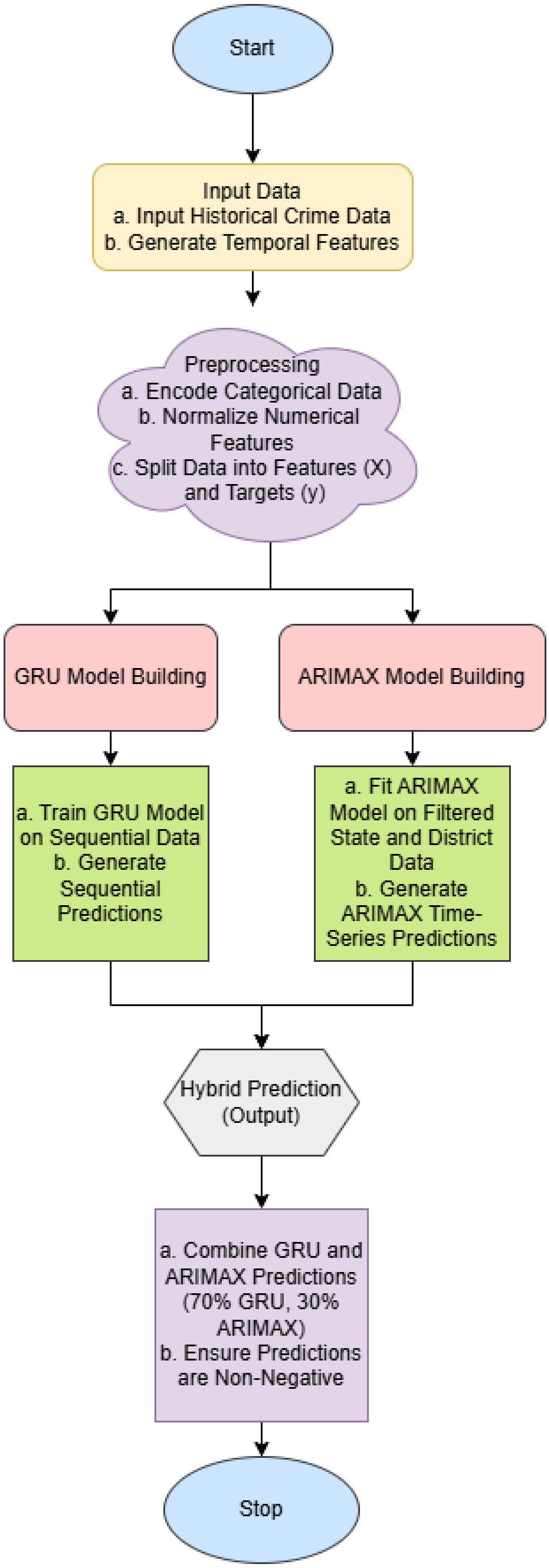

Fig 5: Algorithm Flowchart

IV. ARCHITECHTURURAL DESCRIPTION

The suggested architecture blends machine learning algorithms, real-time data processing capabilities, and visualization libraries into one cohesive system of predicting and analyzing crime patterns. The sequence is initiated by an extensive data preprocessing layer where missing values are managed through mode imputation in order to preserve the integrity of datasets. Categorical features like STATE/UT and DISTRICT are encoded with label encoding, while numerical features are normalized using Min-Max Scaling to guarantee the uniformity of model inputs. In order to represent temporal patterns, the YEAR feature is converted into cyclical factors with sine and cosine functions and represented as YEAR_sin and YEAR_cos.

After preprocessing, the prediction layer is implemented using a two-model scheme. A Bidirectional GRU model is applied to extract complicated spatiotemporal patterns from crime data by utilizing the model's capacity to process both forward and backward sequences for improved contextual accuracy. Concurrently, an ARIMAX model conducts trend analysis with the inclusion of exogenous variables, offering strong statistical forecasting. The two models are then combined in a hybrid prediction system that blends their outputs using a weighted averaging method, where 70% is assigned to the GRU model and 30% to the ARIMAX model. This process improves predictive precision by capitalizing on the advantages of both deep learning and statistical approaches.

The hybrid model output is piped into a visualization and insights layer with interpretability and user interaction in mind. Crime distribution is depicted in pie charts by category, state, and district, whereas heatmaps are used to visualize spatial patterns over regions. Users can dynamically input parameters and observe predictions and visualizations in real time through an interactive web interface built with Flask. The data management and backend layer takes care of storage in CSV files, providing quick and organized access to historic crime data. The Flask app also takes care of user input, model inference, and result rendering via the interface. Lastly, system performance is stringently tested against standard benchmarks, such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and the R^2 score, allowing for a numerical measurement of model accuracy and stability. Collectively, these elements create a scalable and actionable architecture for real-time prediction and analysis of crime trends.

V. EXPERIMENTATION AND RESULTS

The proposed crime prediction system was evaluated based on its ability to predict crime trends and provide actionable insights using historical data. The results were analyzed from both quantitative and qualitative perspectives, focusing on prediction accuracy, visualization effectiveness, and system usability.

1. Prediction Accuracy

The hybrid framework combining the Bidirectional GRU model and ARIMAX demonstrated strong predictive performance across various crime categories and geographical locations. The GRU model was trained for 50 epochs, achieving a final loss of **0.0021**, which reflects its ability to minimize prediction errors during training. The evaluation metrics for the system are summarized in Table

Table 1: Performance Metrics for the Prediction Models

Model	MAE	MSE	R^2 Score
Hybrid Model	0.24155	0.0978	0.1294

The hybrid model demonstrated superior performance over individual models, achieving significantly lower MAE and MSE values, indicating its effectiveness in minimizing prediction errors and enhancing forecasting accuracy.

ARIMAX Model Representation

The ARIMAX model is utilized to incorporate both time-series dependencies and external influencing factors:

$$CrimeRate_t = \alpha + \sum_p \beta_p CrimeRate_{t-p} + \sum_q \gamma_q Error_{t-q} + \sum_m \delta_m ExternalFactor_{t-m} + \epsilon_t$$

Fig 6: ARIMAX Model Equation Representations.

Hybrid Model Prediction

To combine predictions from both models, a weighted averaging approach is applied:

$$FinalPrediction = w_1 \times GRU_{Prediction} + w_2 \times ARIMAX_{Prediction}$$

Fig 7: Hybrid Model Equation Representation

2. Visualization

The system's visualizations provided intuitive representations of the predictions, enabling stakeholders to gain deeper insights into crime trends:

- Pie Charts: Predicted crime distributions for specific states and districts in a given year were represented as pie charts. For example, in State Andhra Pradesh, the system predicted that thefts and assaults accounted for 40% and 25% of the total crimes, respectively, in the year 2012.
- Heatmaps: State-wise crime trends across all categories were visualized using heatmaps. These visualizations highlighted high-crime regions and facilitated comparative analysis. For instance, State Gujarat consistently recorded the highest number of thefts and robberies, as depicted in the heatmap for the 2001–2012 period.

Model's Output:

As can be seen from Fig. 8, the system interface is offered as an index page that easily navigates the user among the prediction choices. The procedure is initiated through entry of crime information at location-specific levels, as can be seen in Fig. 9 where the entry includes state, district, and year to facilitate generation of predictions. The system, thereafter, outputs precise predictions including crime distribution per location (Fig. 10) and by crime (Fig. 11), with all-around visual results. The ability of the system to produce meaningful, data-driven distributions of crime types is also further illustrated in Fig. 12, where pie chart representations provide for better insight into predicted trends.

Fig 8: Index Page

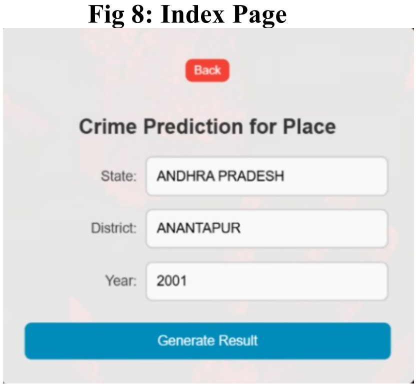

Fig 9: Input for Crime Prediction According place

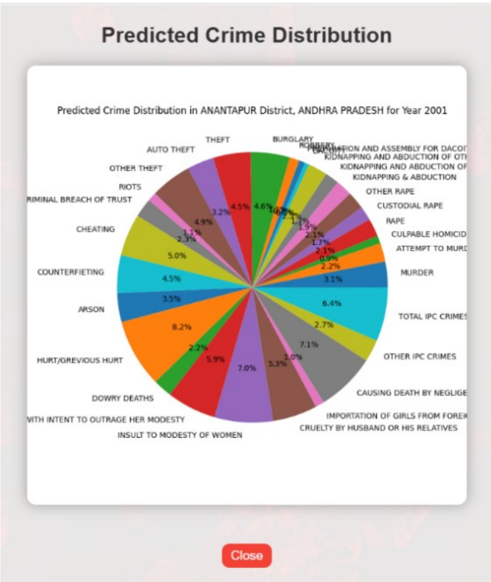

Fig 10: Output for Crime Prediction According place

Fig 11: Input Crime Prediction According crime type

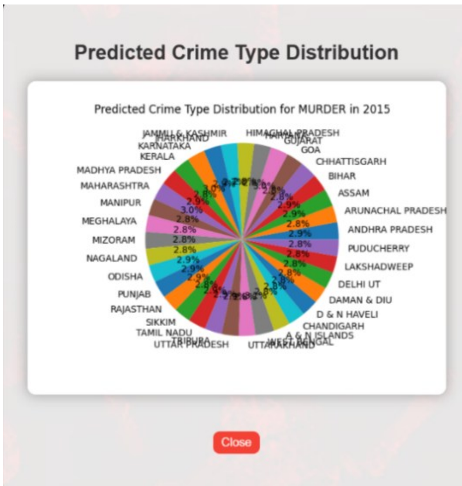

Fig 12: Output Crime Prediction According type

VI. CONCLUSION

The proposed crime prediction system is a major step in using machine learning and statistical methods to solve complex problems of crime analysis and prevention. The hybrid framework combines Bidirectional GRU networks with ARIMAX, utilizing the advantages of both deep learning and traditional time series models to predict crime trends with a high degree of accuracy and reliability. This pattern identification across temporal and spatial dimensions makes the system useful for policymakers, law enforcement agencies, and researchers. Results show that the model is robust and performs well with a final training loss of 0.0021 after 50 epochs and mae of 0.24155995908834355, mse of 0.09782950270830254 for the hybrid framework, where individual models are out performed. Pie charts and heatmaps are great ways to visualize these complex predictions and turn them into actionable insights that can be used to drive data informed decisions. Moreover, the web interface is accessible and usable for various stakeholders.

Though strong, the system is limited in that it relies on high quality data and is not easily generalizable to other datasets. In future work we will address these challenges by incorporating real time data streams, expand the geographical scope to encompass global crime datasets, and explore more advanced hybridization techniques. Finally, this research also demonstrates how machine learning and statistical models can be combined to create a transformative crime prediction tool. This provides the basis for the development of more complex, scalable, and flexible systems to enable proactive crime prevention strategies and to create safer communities.

REFERENCES

[1] Selvakumari, S. Jeya, and V. Joseph Peter. "Crime Classification Using GRU, CNN And Autoencoder Techniques." Educational Administration: Theory and Practice 30.5 (2024): 2950-2964.
[2] Mandalapu, Varun, et al. "Crime prediction using machine learning and deep learning: A systematic review and future directions." IEEE Access 11 (2023): 60153-60170.
[3] Safat, Wajiha, Sohail Asghar, and Saira Andleeb Gillani. "Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques." IEEE access 9 (2021): 70080-70094.
[4] Chun, Soon Ae, et al. "Crime prediction model using deep neural networks." Proceedings of the 20th Annual International Conference on digital government research. 2019.
[5] Mandalapu, Varun, et al. "Crime prediction using machine learning and deep learning: A systematic review and future directions." IEEE Access 11 (2023): 60153-60170.
[6] Azeez, Jazeem, and D. John Aravindhar. "Hybrid approach to crime prediction using deep learning." 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, 2015.
[7] Shah, Neil, Nandish Bhagat, and Manan Shah. "Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention." Visual Computing for Industry, Biomedicine, and Art 4.1 (2021): 9.
[8] Kang, Hyeon-Woo, and Hang-Bong Kang. "Prediction of crime occurrence from multi-modal data using deep learning." PloS one 12.4 (2017): e0176244.
[9] Wang, Bao, et al. "Deep learning for real-time crime forecasting and its ternarization." Chinese Annals of Mathematics, Series B 40.6 (2019): 949-966.
[10] Wang, Bao, et al. "Deep learning for real time crime forecasting." arXiv preprint arXiv:1707.03340 (2017).
[11] Zhang, Binjie, et al. "Multi-Step Occurrence Prediction of Urban Crimes with Enhanced GRU-ODE-Bayes Model." 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2022.
[12] Kowsrihawat, Kankawin, Peerapon Vateekul, and Prachya Boonkwan. "Predicting judicial decisions of criminal cases from Thai Supreme Court using bi-directional GRU with attention mechanism." 2018 5th Asian Conference on Defense Technology (ACDT). IEEE, 2018.
[13] Wang, Shun, et al. "MRAGNN: Refining urban spatio-temporal prediction of crime occurrence with multi-type crime correlation learning." Expert Systems with Applications 265 (2025): 125940.
[14] Mishra, Prativa, et al. "A Novel Approach to Predicting the Risk of Illegal Activity and Evaluating Law Enforcement Using WideDeep SGRU Model." 2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC). IEEE, 2024.
[15] Mishra, Prativa, et al. "A Novel Approach to Predicting the Risk of Illegal Activity and Evaluating Law Enforcement Using WideDeep SGRU Model." 2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC). IEEE, 2024.
[16] Wu, Jiahui, et al. "Enhancing short-term crime prediction with human mobility flows and deep learning architectures." EPJ data science 11.1 (2022): 53.
[17] Pandey, Ajeet, and Piyush Kumar. "BGRU-MTRA: bilinear GRU networks with multi-path temporal residual attention for suspicious activity recognition." Neural Computing and Applications (2024): 1-28.
[18] Liang, Weichao, et al. "Towards hour-level crime prediction: A neural attentive framework with spatial–temporal-categorical fusion." Neurocomputing 486 (2022): 286-297.
[19] Du, Yingjie, and Ning Ding. "A systematic review of multi-scale spatio-temporal crime prediction methods." ISPRS International Journal of Geo-Information 12.6 (2023): 209.
[20] Utsha, Rittik Basak, et al. "Deep Learning Based Crime Prediction Models: Experiments and Analysis." arXiv preprint arXiv:2407.19324 (2024)
[21] https://www.kaggle.com/datasets/khanmohammadanas/district-wise-crimes-in-india