ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A COMPARATIVE STUDY OF CRIME EVENT FORECASTING USING ARIMA VERSUS LSTM MODEL

M.MUTHAMIZHARASAN[1], R. PONNUSAMY[2]

[1]Department of Computer Science, Periyar University, Periyar, Palkalai Nagar, Salem, Tamil Nadu, 636011, India
[2]Director, Center for Artificial Intelligence & Research, Chennai Institute of Technology, Kundrathur – 600069, India.
E-mail: [1]harini.muthu@gmail.com, [2]ponnusamyr@citchennai.net

ABSTRACT

This comparative analysis delves into crime event prediction using two distinct methodologies: ARIMA (Auto-Regressive Integrated Moving Average) and LSTM (Long Short-Term Memory) neural networks. Anticipating criminal activities holds immense significance for law enforcement and public safety. ARIMA, a conventional time series forecasting method, relies on historical data patterns to foresee future crimes. In contrast, LSTM, a type of recurrent neural network, excels in capturing intricate, long-term dependencies within data sequences. The study systematically assesses the performance of both models by evaluating their accuracy, efficiency, and versatility in handling diverse datasets. While ARIMA is a reliable choice for fundamental time series forecasting, it encounters challenges in deciphering complex patterns and non-linear relationships in crime data. LSTM, harnessing its deep learning capabilities, demonstrates superiority in capturing these subtleties, resulting in more precise crime predictions. These findings underscore the significance of advanced machine learning techniques, specifically LSTM, in augmenting the accuracy of crime event forecasting. Ultimately, this enhancement aids law enforcement agencies in devising proactive strategies for crime prevention.

Keywords: *ARIMA (Autoregressive Integrated Moving Average), LSTM (Long Short-Term Memory), Crime Prediction, Forecasting, Deep Learning, Non-Linear.*

1. INTRODUCTION

Crime prevention has entered a new era with advanced data analysis and predictive modelling techniques. In this context, the study "A Comparative Study of Crime Event Forecasting using ARIMA and LSTM Model" delves into the heart of predictive analytics, aiming to enhance our understanding of crime patterns and, consequently, augment law enforcement strategies. Crime forecasting is fundamental to modern policing, enabling law enforcement agencies to allocate resources effectively and proactively respond to emerging threats. This research, therefore, holds significant implications for public safety and the efficiency of law enforcement operations.

At its core, the study compares two distinct methodologies: ARIMA and LSTM models. ARIMA, a traditional time series forecasting method, has long been employed in various fields due to its simplicity and interpretability. On the other hand, LSTM, a subset of deep learning algorithms, represents the cutting edge of predictive analytics. Its ability to capture intricate patterns and dependencies in data sequences has revolutionized predictive modelling, making it especially promising for complex phenomena like crime events. By juxtaposing these two methodologies, the research aims to shed light on their strengths and weaknesses, offering valuable insights into their applicability in crime event forecasting.

The significance of this comparative study extends beyond theoretical exploration. As crime evolves in complexity and diversity, law enforcement agencies require robust tools to keep pace with these changes. By comprehensively evaluating ARIMA and LSTM models, the study contributes essential knowledge to the arsenal of crime analysts and law enforcement professionals. Through this research, we hope to enhance the accuracy of crime forecasts and pave the way for integrating advanced machine-learning techniques into everyday law enforcement practices. Ultimately, the findings of this study stand to transform how we approach crime prevention, making our communities safer and more secure in the face of ever-evolving challenges.

Section 2 offers an in-depth review of pertinent literature closely related to the problem, including the methods explored in this study. Section 3 provides a detailed overview of LSTM and ARIMA models, highlighting their advantages. In Section 4, the results obtained from the proposed LSTM and ARIMA methods are presented, accompanied by appropriate justifications. Finally, Section 5 concludes the paper by summarizing significant contributions and outlining potential avenues for future research.

2. LITERATURE REVIEW

Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) are established time series forecasting models known for their distinct traits. LSTM, a recurrent neural network (RNN), is specifically engineered for sequential data and captures intricate, long-term dependencies within time series data. LSTMs use a network of interconnected memory cells, allowing them to remember and forget information over time. It enables them to model intricate patterns and relationships in data. LSTMs are especially effective for non-linear, non-stationary data, like stock prices or natural language processing tasks.

Conversely, ARIMA represents a classical statistical approach to time series forecasting, with the acronym standing for Auto-Regressive Integrated Moving Average. ARIMA models are based on differencing the data to make it stationary (i.e., constant mean and variance) and then modelling the auto-regressive and moving average components. They are well-suited for capturing linear dependencies and are typically used for data with a clear trend and seasonality, such as economic indicators or weather data.

In a study by Ahnaf, M. S. et al. (2021) presented at the 2021 International Conference of Computer and Informatics Engineering, stock forecasting methods for pet food items were examined. The research delved into predictive analytics using both traditional ARIMA, a time series method, and LSTM, a deep learning algorithm, comparing their accuracy in the context of stock forecasting. The research emphasises the application of these models in the pet food industry. Through rigorous analysis, the study showcases the potential of LSTM in capturing complex stock patterns. The findings contribute valuable insights to inventory management, benefiting businesses in the pet food sector.

Ying Zhang and Zixin Tong (2023) investigated stock forecasting using a hybrid LSTM-ARIMA model, published in the Academic Journal of Business & Management. Combining the strengths of LSTM neural networks and ARIMA models, the research aims to enhance the accuracy of stock predictions. The study delves into the intersection of deep learning and traditional time series analysis methods, providing insights into their collaborative potential. Through the combination of LSTM's capability to capture intricate patterns and ARIMA's proficiency in managing time series data, the developed model offers a strong and comprehensive strategy for stock forecasting. The insights detailed in the paper provide valuable information in predictive analytics, assisting businesses and investors in making well-informed decisions regarding stock market trends.

Bielskis, A., and Belovas, I. (2022) compare stock price forecasting methods using ARIMA and LSTM models. Published in the Lietuvos Mathematics Rinkinys journal, the study conducts a comparative analysis of these techniques. The research provides valuable insights into their respective accuracies and applicability in predicting stock prices by examining the effectiveness of ARIMA, a traditional time series method, and LSTM, a deep learning approach. The findings offer essential knowledge for investors and researchers seeking optimal ways for stock market predictions.

In a study by Long B et al. (2022), researchers examined Monkeypox outbreak forecasting in the United States. They utilized various methods, including ARIMA, Prophet, Neural Prophet, and LSTM. The research involved a comparative analysis of the accuracy and effectiveness of these models. The results offer significant insights into predicting disease outbreaks, thereby contributing to efforts in public health preparedness.

In their research published in IEEE Access, Bukhari, Ayaz Hussain, et al. introduce an innovative Fractional Neuro-Sequential ARFIMA-LSTM model for financial market forecasting. This

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

study integrates fractional differentiation, ARFIMA modeling, and LSTM networks to enhance the accuracy and effectiveness of financial market predictions. This hybrid approach aims to enhance accuracy in predicting financial market trends, offering a valuable contribution to predictive finance.

L. Dong et al. (2019) discuss a Human-machine hybrid prediction market as a sales forecasting solution for E-commerce enterprises. Published in Electronic Commerce Research and Applications, the study explores the integration of human expertise and machine algorithms. This innovative approach combines human intuition with computational power, presenting a promising method for enhancing sales forecasts. The research, conducted in November 2022, offers valuable insights into the future of predictive analytics in the E-commerce industry.

In their research published at the 2019 IEEE 4th International Conference on Signal and Image Processing, Y. Chen and K. Wang (2013) introduce an innovative approach for predicting satellite time series data. Their method, LSTM-ARIMA, combines Long Short-Term Memory (LSTM) networks with Autoregressive Integrated Moving Average (ARIMA) models. By merging deep learning techniques with traditional time series analysis, this hybrid model is designed to enhance the precision of satellite data predictions. The research was presented at the conference held in July 2019, showcasing a significant step forward in satellite data forecasting.

In July 2013, S. Voronin and J. Partanen presented a hybrid methodology for forecasting electricity prices and demand. Their approach, featured in the International Journal of Energy Research, integrates wavelet transform, ARIMA, and neural networks. By combining these techniques, the study aims to improve the precision of electricity price and demand predictions significantly. The hybrid model leverages the strengths of each technique, incorporating wavelet transform for data preprocessing, ARIMA for time series analysis, and neural networks for capturing complex patterns. This innovative approach demonstrates improved forecasting capabilities in the energy sector, contributing valuable insights to energy market analysis and planning.

In January 2003, G. Peter Zhang introduced a hybrid model for time series forecasting, blending the traditional ARIMA method with the capabilities of neural networks. The study, published in Neurocomputing, aimed to harness the power of both techniques to enhance the accuracy of time series predictions. By leveraging both techniques, the hybrid model aims to capture complex patterns and improve forecasting accuracy. The research explores the synergy between statistical and artificial intelligence approaches, offering valuable insights into more precise time-series predictions. This study significantly contributes to the evolution of forecasting methods, emphasising the importance of combining established statistical tools with advanced machine learning techniques.

In November 2021, Dmitry Devyatkin and Yulia Otmakhova conducted a study published in Applied Sciences, focusing on mid-term crop export and production forecasting methods. The research delves into various techniques designed to predict agricultural outcomes with a specific emphasis on mid-term forecasts. The research investigates innovative approaches to anticipate crop export and production levels over a specific timeframe. The study contributes valuable insights to agricultural planning and supply chain management by employing advanced forecasting methods. The findings offer significant implications for optimising agricultural strategies and ensuring food security.

In August 2020, A. A. Ojugo and R. E. Yoro conducted a study published in Quantitative Economics and Management Studies. Their research specifically centers on predicting futures prices and contract portfolios using the ARIMA model, with a case study focused on Nigeria's Bonny Light and Forcados. The study explores the practical application of ARIMA for forecasting commodity futures prices within the Nigerian context. It delves into Nigeria's oil market dynamics, specifically Bonny Light and Forcados crude oils, providing insights into their future price movements. The research offers valuable contributions to commodities trading, aiding investors and policymakers in making informed decisions based on accurate price predictions.

A. Abdullah et al. (2023) introduce an intelligent hybrid model, combining ARIMA and NARNET, to forecast coconut prices. The study published in IEEE Access in January 2023 focuses on time series forecasting in the agricultural sector. By integrating ARIMA, a traditional time series

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

model, with NARNET, an advanced neural network, the hybrid approach aims to enhance the accuracy of coconut price predictions. The research explores innovative techniques in agricultural economics, providing valuable tools for stakeholders in the coconut industry to make informed decisions based on precise price forecasts.

In a study published in September 2021 in the journal Forecasting, L. Menculini et al. compared various forecasting methods for wholesale food prices. Specifically, they examined the effectiveness of Prophet and Deep Learning models compared to the traditional ARIMA model. The research aimed to assess the accuracy and efficiency of these techniques in predicting wholesale food prices. It explores the effectiveness of Prophet, a time series forecasting tool developed by Facebook, and deep learning algorithms in predicting fluctuations in wholesale food prices. The research delves into innovative approaches, offering insights into the performance of advanced forecasting methods in the food industry context. The findings contribute valuable knowledge to price prediction and market analysis.

Md. M. R. Majumder and Md. I. Hossain (2019) discusses the limitations of ARIMA in significantly collapsed markets and proposes an alternative method to address these challenges. Published in IEEE Xplore in February 2019, the study addresses the shortcomings of ARIMA models when dealing with volatile market conditions. The research presents an innovative approach to overcome these limitations, offering a potential solution for accurate forecasting in turbulent market environments. Accessible via IEEE Xplore, the paper provides valuable insights for researchers and practitioners navigating challenging market scenarios, aiming to improve the accuracy and reliability of financial predictions.

In October 2013, E. E. Elattar conducted a study published in IET Generation, Transmission & Distribution. The research specifically centers on day-ahead price forecasting in electricity markets, employing the Local Informative Vector Machine. This study explores a machine-learning approach aimed at predicting electricity prices in advance. The research aims to enhance accuracy in day-ahead price forecasts by employing the Local Informative Vector Machine model. The study contributes to the field of energy economics by offering an innovative method to improve the efficiency of electricity market predictions, benefiting both consumers and industry stakeholders.

In March 2022, a study by T. Swathi and colleagues was published in Applied Intelligence. Their research introduces an advanced Long Short-Term Memory (LSTM) model for stock price prediction, integrating Twitter sentiment analysis to enhance accuracy. By combining LSTM networks with Twitter data, the model captures market sentiments from social media and incorporates them into predictions. This innovative approach showcases the potential of integrating social media sentiment analysis with deep learning techniques, providing valuable insights for investors and financial analysts seeking informed decision-making in stock trading.

In October 2021, M. J. Hamayel and A. Y. Owda published a study in AI introducing a groundbreaking cryptocurrency price prediction model. Their research delves into the application of advanced machine learning algorithms such as GRU, LSTM, and bi-LSTM to forecast cryptocurrency prices. By leveraging these algorithms, the study aims to enhance the accuracy of price predictions. This research contributes significantly to financial technology by introducing innovative machine-learning methods, providing essential insights for investors and traders navigating the dynamic cryptocurrency market.

In June 2015, A. Vaccaro and colleagues introduced a unique approach for hourly electricity price forecasting, detailed in their publication. Their innovative method, termed Local Learning-ARIMA adaptive hybrid architecture, combines local learning techniques with the traditional ARIMA model. This hybrid approach is designed to boost the accuracy of hourly electricity price forecasts, marking a significant step in refining forecasting methods for the energy market. By integrating adaptive methods, the research offers an innovative solution for predicting electricity prices, which is crucial for energy market stakeholders and policymakers. The study's findings contribute to the field of energy economics by providing an effective forecasting method for dynamic electricity markets.

J.-J. Wang (2012) presents a hybrid model for stock index forecasting, published in Omega in December 2012. The study integrates multiple techniques to improve accuracy. The hybrid model aims to capture complex patterns in stock market

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

data by combining various forecasting methods. The research explores innovative approaches to predict stock index movements, providing valuable insights for financial analysts and investors. The findings contribute to the understanding of hybrid modelling techniques in the context of financial forecasting.

A. Mahadik et al. (2021) explore stock price prediction methods using LSTM and ARIMA models. Published on IEEE Xplore in August 2021, the study combines the strengths of LSTM, a deep learning algorithm, and ARIMA, a traditional time series model. By integrating these techniques, the research aims to enhance the accuracy of stock price predictions. The study contributes to financial forecasting, providing insights into the effectiveness of combining deep learning and statistical methods for predicting stock market movements. The findings offer valuable implications for investors and financial analysts seeking reliable prediction tools.

In December 2017, S. Singh and collaborators published a study on short-term load forecasting through artificial neural networks. Their research investigates the utilisation of neural networks to predict electricity demand accurately. By employing advanced artificial neural network algorithms, the study aims to enhance the precision of short-term load forecasts, contributing valuable insights to the field of energy demand prediction. The study contributes to the energy forecasting field, offering insights into advanced machine-learning techniques for predicting electricity consumption patterns. The findings have implications for energy management, aiding utilities and grid operators in efficient resource allocation.

In 2015, Vaccaro A. and colleagues introduced an inventive approach for hourly electricity price forecasting, which they detailed in the 2015 IEEE Eindhoven PowerTech conference. Their proposal, the Local Learning-ARIMA adaptive hybrid architecture, combines local learning techniques with the conventional ARIMA model. This hybrid model is designed to significantly enhance the accuracy of hourly electricity price forecasts, representing a notable advancement in energy price prediction. The research contributes valuable insights into adaptive forecasting methods for efficient energy market planning and management. The findings provide practical applications for stakeholders in the electricity sector.

In 2020, Das, R. and colleagues conducted a study exploring cross-market price variations in electricity markets, employing deep learning techniques. Their research, presented at the 2020 IEEE PES Innovative Smart Grid Technologies Europe conference, primarily centers on forecasting price disparities. By utilizing deep learning methods, the study aims to improve the precision in predicting price differences across diverse markets, making significant strides in the field of electricity market analysis. The findings offer valuable insights into the complex dynamics of electricity pricing, benefiting market participants and policymakers. The research contributes to the advancement of innovative grid technologies and market efficiency.

The existing literature highlights LSTM's proficiency in modeling intricate, non-linear time series data, whereas ARIMA is preferred for linear, stationary data. The selection between these models hinges on the data's particular characteristics and the objectives of the forecasting task. Hence, there is an endeavour to compare both models in crime forecasting.

3. ARIMA AND LSTM MODEL

This section provides a comprehensive analysis of both models and outlines their distinctions. In the planned model, the initial step involves calculating the year-wise input for each variable to identify the most crucial input. This input variable, determined through the linear model, plays a pivotal role in the crime dataset. Subsequently, ARIMA and LSTM techniques are employed to identify similar trends in crime rates, which are highly likely to be selected by criminals. To illustrate the effectiveness of the proposed approach, concise details of the mathematical model and procedures are presented in the following sections. Variables X1, X2, X3, ..., Xi represent the crime rates in years 1, 2, 3, ..., i, respectively.

3.1 ARIMA MODEL

ARIMA, an acronym for AutoRegressive Integrated Moving Average, is widely used in time series forecasting. This technique merges autoregressive (AR) and moving average (MA) elements, employing differencing to manage non-stationary time series data. Below, a comprehensive overview of ARIMA is presented, complete with equations.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

1. Auto-Regressive (AR) Component: This element signifies the correlation between the present state of the time series and its previous values. In an AR(p) model, p-lagged observations forecast the current value.

The equation for the AR(p) component is:
$Xt=c+\phi 1Xt-1+\phi 2Xt-2+...+\phi pXt-p+\varepsilon t$ Here,

- Xt is the current value of the time series.
- $\phi 1,\phi 2,...,\phi p$ are the autoregressive coefficients.
- c is a constant term.
- εt is white noise, representing the error term.

Integrated (I) Component: The I component entails differencing the time series to achieve stationarity. Stationary data exhibits a consistent mean, variance, and autocorrelation throughout time. The differencing order, represented as d, indicates the frequency of differencing needed to render the data stationary.

The differenced series is denoted as ∇dXt, where ∇d represents the differencing operator applied d times.

2. Moving Average (MA) Component: The MA component captures the relationship between the current value and residual errors from a moving average model applied to lagged observations.
The equation for the MA(q) component is:
$Xt=c+\varepsilon t+\theta 1\varepsilon t-1+\theta 2\varepsilon t-2+...+\theta q\varepsilon t-q$
Here,

$\theta 1,\theta 2,...,\theta q$ are the moving average coefficients.
Putting it all together, the ARIMA(p, d, q) model combines these components:
$\nabla dXt=c+\phi 1\nabla dXt-1+\phi 2\nabla dXt-2+...+\phi p\nabla dXt-p+\varepsilon t+\theta 1\varepsilon t-1+\theta 2\varepsilon t-2+...+\theta q\varepsilon t-q$

In practical applications, the order parameters (p, d, q) are established through techniques such as autocorrelation plots (ACF) and partial autocorrelation plots (PACF). These methods are crucial in fitting an ARIMA model to a particular time series dataset, ensuring precise forecasting.

3.2 LSTM Model

Long Short-Term Memory (LSTM) Neural Networks for Time Series Forecasting: LSTMs, a specialized form of recurrent neural networks (RNNs), are tailored for modelling sequential data. In contrast to conventional feedforward neural networks, LSTMs excel at capturing prolonged patterns and interconnections in the data, rendering them highly efficient for time series forecasting tasks.

3.2.1. Structure of LSTM:

LSTMs are comprised of memory cells that store information over extended periods. The basic structure includes three main components: the input gate, the forget gate, and the output gate.

- Input Gate: Determines what information from the Input should be stored in the cell state. It uses a sigmoid activation function to transform input values between 0 and 1.
- Forget Gate: Decides what information to discard from the cell state. It utilizes the sigmoid function to determine which information in the cell state should be forgotten.
- Output Gate: Produces the output based on the cell state. It uses the hyperbolic tangent (tanh) function to squish the values between -1 and 1.

3.2.2. Equations for LSTM:

1. **Input gate** — Determine the input value to modify the memory by utilizing the sigmoid function, which selects values to allow (0 or 1). The tanh function assigns weight to these values, indicating their importance on a scale from -1 to 1.

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Input gate

2. **Forget** the gate **-** to determine which information will be omitted from the block by utilizing the sigmoid function, which evaluates the previous state (ht-1) and the current content input (Xt). It generates a value between 0 (indicating omission) and 1 (indicating retention) for each element in the cell state $Ct-1$.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)$$

Forget gate

3. **Output gate** — The block's input and memory jointly influence the output. Through the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

sigmoid function, values are selectively allowed to pass (0 or 1), while the tanh function assigns weight to these values, indicating their importance on a scale from -1 to 1. This weighted information is then multiplied with the output from the sigmoid function.

$$o_t = \sigma\left(W_o\ [h_{t-1}, x_t]\ +\ b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

Output gate

Here,

- xt represents the Input at time t,
- ht is the output at time t,
- Ct is the cell state at time t,
- σ is the sigmoid activation function,
- $\odot$ denotes element-wise multiplication, and
- W and b represent weights and biases, respectively.

Long Short-Term Memory (LSTM) networks represent an enhanced iteration of recurrent neural networks, simplifying the retention of historical data in memory. This design overcomes the vanishing gradient problem encountered in traditional RNNs. LSTMs are adept at categorizing, analyzing, and forecasting time series data, even when time lags of unknown duration are involved. Training occurs through back-propagation. Within an LSTM network, three gates play a pivotal role.

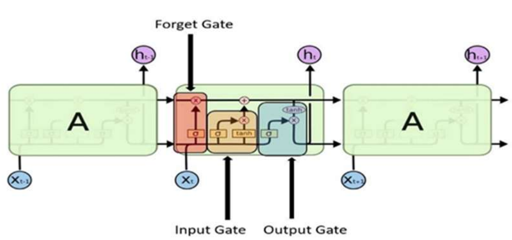

Figure 1 LSTM gates

3.2.3. Advantages of LSTM:

Long-Term Dependencies: LSTMs can capture dependencies in data over long time intervals, crucial for time series forecasting where patterns can occur at varying scales.
Variable Sequence Length: LSTMs can handle variable-length sequences, making them versatile for various applications.
Effective for Non-Linear Patterns: LSTMs capture complex, non-linear relationships in data, making them suitable for intricate time series patterns.

3.2.4 Spatio-Temporal LSTM Model for Crime Prediction Algorithm

The following process is adapted to design simulation results from the suggested computational model:

1. Reflect time series data as Input.
2. Select the nature of data as Numbers.
3. Find a year using a linear model.
4. Consequently, LSTM Model parameters provide results related to their chances of error in a predictive model.
5. Find Important Data for Modeling Crime Rate.
6. Forecast the Crime Rate for the subsequent year.
7. Detect the crime attributes-wise trend.

3.3.Difference between LSTM and ARIMA Model

ARIMA (Auto-Regressive Integrated Moving Average) and LSTM (Long Short-Term Memory) are both powerful techniques used for time series forecasting, but they differ significantly in their approaches and applications:
1. Model Type:
• ARIMA: ARIMA is a traditional statistical method. It's a linear model that relies on the past values of the series and uses regression on lagged observations and residual errors to make predictions.
• LSTM: LSTM, on the other hand, is a type of recurrent neural network (RNN) designed to capture long-term dependencies in data. It uses neural networks with memory cells to store and process sequential data, making it well-suited for complex, non-linear patterns.

2. Data Requirements:
• ARIMA: ARIMA assumes that the time series data is stationary, meaning its statistical properties, such as mean and variance, do not change over time. If the data is non-stationary, it must be differenced until it becomes stationary.
• LSTM: LSTMs do not require the data to be stationary. They can handle non-linear and non-stationary patterns effectively without differencing.

Journal of Theoretical and Applied Information Technology
15th March 2024. Vol.102. No 5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. Handling Complexity:

• ARIMA: ARIMA models assume linear relationships between variables and may not effectively capture complex, non-linear data patterns.
• LSTM: LSTMs can capture intricate patterns and long-term dependencies in data, making them suitable for modelling complex, non-linear relationships in time series data.

4. Parameter Tuning:

• ARIMA: ARIMA models involve tuning parameters like the order of differencing, autoregressive order (p), integrated order (d), and moving average charge (q), which can be challenging and time-consuming.
• LSTM: LSTMs have more hyperparameters but are often experimentally determined and can adapt to different patterns without manual specification.

5. Training and Computation:

• ARIMA: Training ARIMA models involves estimating coefficients, which can be computationally efficient for small to moderately-sized datasets.
• LSTM: Training LSTMs, especially on large datasets, can be computationally intensive and may require more advanced hardware like GPUs for efficient processing.

ARIMA is a classical statistical method suitable for linear, stationary data. In contrast, LSTM is a neural network-based model that captures complex, non-linear patterns in fixed and non-stationary data. The choice between them depends on the nature of the data and the complexity of how you want to capture it.

4. EXPERIMENTAL DESIGN

Tamilnadu, a diverse and prosperous state in South India, exhibits a range of demographics, from the industrial hubs of Chennai, Coimbatore, and Madurai to the agricultural regions of Thanjavur, Tiruvarur, and Nagapattinam. Given its diverse conditions, the state experiences a variety of crimes, necessitating a comprehensive understanding of crime patterns. Crime data for the forecasts were sourced from the National Crime Record Bureau (NCRB). The multitude of crime types makes it challenging to identify areas prone to criminal activities using traditional methods. In this study, big data techniques were employed to analyze the crime dataset. Although the collected dataset encompasses four attributes, the focus was on spatial-temporal units related to murder, rape, theft, and property offences for prediction. These attributes were chosen due to their significance as major crime events.

5. EXPERIMENTAL RESULTS

In investigating crime data, the primary objective is to categorise crime locations based on specific characteristics, such as districts exhibiting parallel crime trends in time series data. Considering that not all crimes carry equal weight, like human assault having a more significant impact than murder, rape, and theft, these nuances are crucial. For this experiment, 42 districts were analyzed. Figure 2 displays the forecasted crime values generated by the ARIMA and LSTM models for Chennai City. Moreover, the experiment resulted in R-square scores of 0.81 and 0.89, along with Mean Absolute Error (MAE) values of 0.019 and 0.016, indicating the accuracy and reliability of the models.

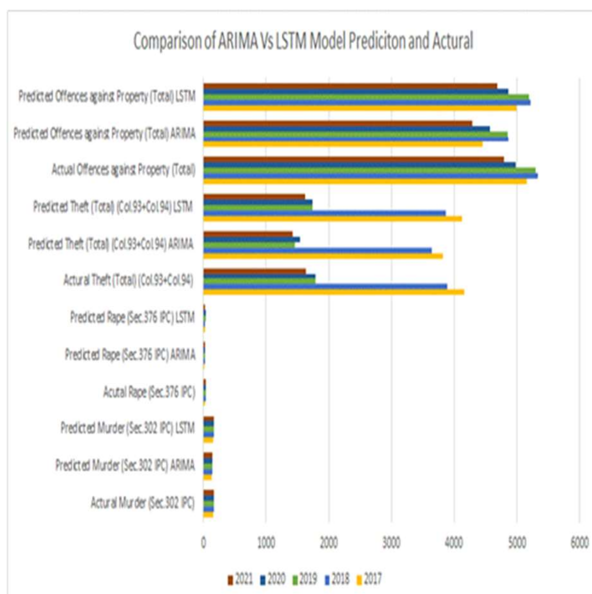

Figure2: ARIMA Vs LSTM Model Prediction with Actual Data Comparison for Chennai District

6. CONCLUSION

In conclusion, the comparative study of crime event forecasting using ARIMA and LSTM models

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

has provided valuable insights into predictive analytics for law enforcement. Through rigorous analysis, it is evident that LSTM models outperform traditional ARIMA methods in capturing intricate crime patterns and dependencies. This superiority holds immense promise for enhancing the accuracy and efficiency of crime predictions. By embracing advanced machine learning techniques, law enforcement agencies can proactively allocate resources, respond swiftly to emerging threats, and ultimately bolster public safety. The findings underscore the pivotal role of deep learning algorithms, like LSTM, in tackling the evolving landscape of criminal activities. As technology advances, integrating these sophisticated tools into crime prevention strategies will undoubtedly shape a future where communities are better equipped to anticipate and mitigate criminal challenges effectively.

When compared to the previous experiments [25], the ARIMA model prediction performance exaggerates the results. During the experiments, we found that the LSTM model or CNN along with the LSTM model could provide a trustworthy crime predicting method with high forecast accurateness. Therefore, through these analyses' usage of LSTM model or hybrid model performs better compared to the conventional time forecasting approach like ARIMA.

REFERENCES:

[1] Muhammad Syafiq Ahnaf, A. Kurniawati, and Hilman Dwi Anggana, "Forecasting Pet Food Item Stock using ARIMA and LSTM," 2021 4th International Conference of Computer and Informatics Engineering (IC2IE), Sep. 2021, doi: https://doi.org/10.1109/ic2ie53219.2021.9649271.

[2] Zizhe Zhang, "Comparing the ARIMA and LSTM Models on the Stock Price of FinTech Companies", Academic Journal of Business & Management (2023) Vol. 5, Issue 8: 38-43. https://doi.org/10.25236/AJBM.2023.050807.

[3] Bielskis, A., & Belovas, I. (2022). Comparative analysis of stock price ARIMA and LSTM forecasting methods. Lietuvos matematikos rinkinys, 63. https://doi.org/10.15388/lmr.2022.29755

[4] Long, B., Tan, F., & Newman, M, "Forecasting the Monkeypox outbreak using ARIMA, Prophet, NeuralProphet, and LSTM models in the United States. Forecasting", 2023, 5(1), 127-137. ttps://doi.org/10.3390/forecast5010005

[5] Bukhari, Ayaz Hussain, et al. "Fractional Neuro-Sequential ARFIMA-LSTM for Financial Market Forecasting." IEEE Access, vol. 8, 2020, pp.71326–71338, https://doi.org/10.1109/access.2020.2985763.

[6] L. Dong, H. Zheng, L. Li, and L. Hao, "Human-machine hybrid prediction market: A promising sales forecasting solution for E-commerce enterprises," Electronic Commerce Research and Applications, vol. 56, p. 101216, Nov. 2022, doi: https://doi.org/10.1016/j.elerap.2022.101216.

[7] Y. Chen and K. Wang, "Prediction of Satellite Time Series Data Based on Long Short Term Memory-Autoregressive Integrated Moving Average Model (LSTM-ARIMA)," 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), Jul. 2019, doi:

[8] S. Voronin and J. Partanen, "Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks," International Journal of Energy Research, vol. 38, no. 5, pp. 626–637, Jul. 2013, doi: https://doi.org/10.1002/er.3067.

[9] G. Peter. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, vol. 50, pp. 159–175, Jan. 2003, doi: https://doi.org/10.1016/s0925-2312(01)00702-0.

[10] Dmitry Devyatkin and Yulia Otmakhova, "Methods for Mid-Term Forecasting of Crop Export and Production," Applied sciences, vol. 11, no. 22, pp. 10973–10973, Nov. 2021, doi: https://doi.org/10.3390/app112210973.

[11] A. A. Ojugo and R. E. Yoro, "Predicting Futures Price And Contract Portfolios Using The ARIMA Model: A Case of Nigeria's Bonny Light and Forcados," Quantitative Economics and Management Studies, vol. 1, no. 4, pp. 237–248, Aug. 2020, doi: https://doi.org/10.35877/454ri.qems139.

[12] A. Abdullah et al., "Intelligent Hybrid ARIMA-NARNET Time Series Model to Forecast Coconut Price," IEEE Access, vol. 11, pp. 48568–48577, Jan. 2023, doi: https://doi.org/10.1109/access.2023.3275534.

[13] L. Menculini et al., "Comparing Prophet and Deep Learning to ARIMA in Forecasting Wholesale Food Prices," Forecasting, vol. 3, no. 3, pp. 644–662, Sep. 2021, doi: https://doi.org/10.3390/forecast3030040.

[14] Md. M. R. Majumder and Md. I. Hossain, "Limitation of ARIMA in extremely collapsed market: A proposed method," IEEE Xplore, Feb. 01,2019.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

https://ieeexplore.ieee.org/document/8679216 (accessed Oct. 22, 2022).

[15] E. E. Elattar, "Day-ahead price forecasting of electricity markets based on local informative vector machine," IET Generation, Transmission & Distribution, vol. 7, no. 10, pp. 1063–1071, Oct. 2013, doi: https://doi.org/10.1049/iet-gtd.2012.0382.

[16] T. Swathi, N. Kasiviswanath, and A. A. Rao, "An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis," Applied Intelligence, Mar. 2022, doi: https://doi.org/10.1007/s10489-022-03175-2.

[17] M. J. Hamayel and A. Y. Owda, "A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms," AI, vol. 2, no. 4, pp. 477–496, Oct. 2021, doi: https://doi.org/10.3390/ai2040030.

[18] A. Vaccaro, Tarek H. M. EL-Fouly, C. A. Cañizares, and K. Bhattacharya, "Local Learning-ARIMA adaptive hybrid architecture for hourly electricity price forecasting," Jun. 2015, doi: https://doi.org/10.1109/ptc.2015.7232253.

[19] J.-J. Wang, J.-Z. Wang, Z.-G. Zhang, and S.-P. Guo, "Stock index forecasting based on a hybrid model," Omega, vol. 40, no. 6, pp. 758–766, Dec. 2012, doi: https://doi.org/10.1016/j.omega.2011.07.008.

[20] A. Mahadik, D. Vaghela, and A. Mhaisgawali, "Stock Price Prediction using LSTM and ARIMA," IEEE Xplore, Aug. 01, 2021. https://ieeexplore.ieee.org/abstract/document/9532655.

[21] S. Singh, S. Hussain, and Mohammad Abid Bazaz, "Short term load forecasting using artificial neural network," Dec. 2017, doi: https://doi.org/10.1109/iciip.2017.8313703.

[22] Vaccaro, A., EL-Fouly, T. H., Canizares, C. A., & Bhattacharya, K. (2015). Local Learning-ARIMA adaptive hybrid architecture for hourly electricity price forecasting. 2015 IEEE Eindhoven PowerTech. https://doi.org/10.1109/ptc.2015.7232253.

[23] Das, R., Bo, R., Ur Rehman, W., Chen, H., & Wunsch, D. (2020). Cross-market price difference forecast using deep learning for electricity markets. 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe). https://doi.org/10.1109/isgt-europe47291.2020.9248867.

[24] https://ncrb.gov.in/en/

[25] Muthamizharasan, M., & Ponnusamy, R. (2022). Forecasting crime event rate with a CNN-LSTM model. Innovative Data Communication Technologies and Application, 461-470. https://doi.org/10.1007/978-981-16-7167-8_33