REVIEW OF RESEARCH
ISSN: 2249-894X
IMPACT FACTOR : 5.7631(UIF)
VOLUME - 12 | ISSUE - 1 | OCTOBER - 2022

A STUDY ON MACHINE LEARNING AND APPLICATIONS IN TIME SERIES FORECASTING

M. Narsimulu
Research Scholar, Department of Applied Statistics,
Telangana University, Nizamabad.

ABSTRACTS

Forecasting time series data is an important subject in Data science, Health care, and Biological sciences. Traditionally, there are a few strategies to really figure the future worth of time series information, for example, Straight Relapse models, Autoregressive, Moving Normal, Outstanding Smoothing strategy, and all the more strikingly Autoregressive Incorporated Moving Normal (ARIMA) with its numerous vacillations. Specifically, ARIMA model has shown its outperformance in accuracy and exactness of guaging what's to come upsides of time series information. With the new progression in computational force of PCs and all the more significantly improvement of further developed AI Procedures its methodologies, for example, profound learning Strategies, new Procedures are created to examine and gauge the time series information. The exploration question examined in this Exploration paper is that whether and how the recently grown profound learning based Methods for estimating the time series information, for example, " Long Short-Term Memory (LSTM)", are better than the Customary Procedures time series information.The proposed techniques is evaluated using India's crime data from 2001–2014. The observational examinations led and detailed in this Exploration paper shows that profound learning-based Methods, for example, LSTM beat than Regular based Strategies, for example, ARIMA model. All the more explicitly, the typical decrease in remaining rates got by LSTM is exceptionally less when contrasted with ARIMA model. It showing that LSTM is more better than ARIMA. Moreover, it was seen that the quantity of preparing times, known as "age" in profound learning, significantly affected the exhibition of the prepared estimate model and it shows a really irregular way of behaving.

KEY WORDS : *Autoregressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), Deep Learning, Forecasting and Time Series Data.*

INTRODUCTION

Providing safety and security to citizens by controlling and preventing crime is crucial for law enforcement agencies. Crime can cause direct or indirect damage to social welfare programs and can negatively impact the nation's economy. Therefore, countries spend a significant amount of their Gross Domestic Product (GDP) on law enforcement agencies to control and prevent crime. However, little effort has been made in this area due to spatial and temporal information unavailability in crime datasets. But, this area has been getting attention with technological advancement for the last few years. The inclusion of space and time information in the crime datasets using Geographic Information Systems (GIS) enhanced the spirit of the crime prediction research area. Therefore, several efforts have been made in the crime prediction area to assist law enforcement agencies in the last few years. Meanwhile, few Asian countries, like Japan and China, use forecasting police programs. Other countries, Indonesia, South Korea are still evaluating the potential benefits. They use artificial intelligence (AI) models to predict different types of crime in all societies across the country daily, which results in more effective distribution of police forces; this approach focuses on areas of highest risk according to the program. Crimes are typically recorded as a series of observations in chronological order over time. Time series are always individually processed before being combined with time-series parameters in forecasting models. These models can often conduct accurate forecasting but fail occasionally. However, the statistical methods assume linear modelling; thus, they do not extend well and fail to capture nonlinear patterns in the data. Nevertheless, officials and decision-makers are working on forecasting, limiting, and preventing crime to reduce social harm and secure the state. Furthermore, forecasting and preventing crime is preferable to investigate the course of crime, which has changed in our modern era. Besides, time series analysis is a popular statistical technique that can forecast the future supply and demand for a service or commodity. The time series analysis method observes a phenomenon (or variable) over a specific period (e.g., several years). It is then predicted based on the different values in the time series and the pattern of value growth. As a result, it outperforms the conventional method. Time-series techniques are used in most fields, including economics, stock and gold price forecasting, energy, pressure, weather forecasting, and crime forecasting. However, given that the goal of time series analysis is to obtain reliable and accurate forecasting, they also require additional information from the series, such as level and seasonality.

In time series forecasting, data scientists' assumption is that there is no causality that affects the variable we are trying to forecast. Instead, they analyze the historical values of a time series data set in order to understand and predict their future values. The method used to produce a time series forecasting model may involve the use of a simple deterministic model, such as a linear extrapolation, or the use of more complex deep learning approaches. Due to their applicability to many real-life problems, such as fraud detection, spam email filtering, finance, and medical diagnosis, and their ability to produce actionable results, machine learning and deep learning algorithms have gained a lot of attention in recent years. Generally, deep learning methods have been developed and applied to univariate time series forecasting scenarios, where the time series consists of single observations recorded sequentially over equal time increments. The primary reason for this examination paper is to research which anticipating strategies offer best expectations concerning lower figure residuals and higher exactness of estimates. In such manner, there are assortments of stochastic models in time series estimating. The most well-known method is Univariate "Auto-Regressive Integrated Moving Average (ARIMA)" is a special type of ARIMA. AI strategies and all the more critically profound learning calculations have acquainted new methodologies with expectation issues where the connections between factors are displayed in a profound and layered order. AI based strategies, for example, Backing Vector Machines (SVM) and Irregular Woods (RF) and profound learning-based calculations like Repetitive Brain Organization (RNN), and Long Momentary Memory (LSTM) have acquired heaps of considerations as of late with their applications in many disciplines including finance, picture handling and clinical applications. Profound learning techniques are fit for recognizing construction and example of information, for example, non-linearity and intricacy in time series anticipating. Specifically, LSTM has been utilized in time-series expectation in Medical care and picture handling information, for example, anticipating numerous other.

A fascinating and significant examination question is then the exactness and accuracy of customary determining procedures when contrasted with profound learning-based guaging techniques. As far as we could possibly know, there is no particular experimental proof for involving LSTM technique in estimating medical care and Natural sciences time series information to evaluate its exhibition and contrast it and customary guaging strategies like ARIMA. This paper contrasts ARIMA and LSTM models and regard to their presentation in lessening residuals rates. As a delegate of customary figure displaying, ARIMA is decided because of the non-fixed property of the information gathered and demonstrated. As a delegate of profound learning-based calculations, LSTM technique is utilized because of its utilization in saving and preparing the highlights of given information for a more drawn out timeframe. This paper gives a top to bottom direction on information handling and preparing of LSTM models for a bunch of Medical services and Natural sciences time series information.

RELATED WORK

Several efforts have been made in the literature on crime forecasting. This paper discusses the most recent and prominent state-of-the-art research. Mainly, time series analysis and prediction techniques are discussed. Moreover, efforts made in regression and deep learning are highlighted with their advantages and disadvantages.

TIME SERIES ANALYSIS FOR CRIME FORECASTING

Time series analysis extract time-series information related to crime and crime trends. It can also determine the growth of the existing changes observed simultaneously. Furthermore, it can use analysis methods to determine whether time series data are constant or seasonal. Wawrzyniak et al. proposed that crime has seasonal ups and downs and can be determined. Moreover, the summer, winter, and holiday seasons can affect the model's predictive accuracy. Jha et al. discussed that the time-series approach is a dependable method for understanding data over a long period and at different times (hourly, weekly, quarterly, or yearly). The measurement trend becomes smoother when more observations are used, which results in more accurate forecasting. An essential advantage of the time series approach is detecting seasonal patterns, which is critical for future forecasting. Linning et al. examined the variation in crime across the year to determine a seasonal trend. They conducted their research using crime data from three Indian cities to emphasize property-related crimes. The results are expected to show a quadratic relationship because their analysis focuses on crime seasonality. For example, crime increases during the summer months compared with the winter months.

In related research, Almanie et al. analyzed crime data from India. Cities: Delhi and Bombay. They compared the proportion of crimes in both cities to the total number of crimes. Specific trends are found in both cities, such as Sunday having the lowest crime rate. Important derivations such as the safest and most infamous district are noted. They used a decision tree classifier and a naive Bayes classifier.

According to NCRB data used spectral analysis to discover temporal trends in crime in all of India's crime incidents from 2001 to 2014. They aimed to observe seasonal trends in crime and determine whether these patterns apply to all types of crime or differ by crime category. They incorporated spatial patterns and neighborhood-level deviations from global trends. It will assist in developing temporal seasonality models based on weekly and monthly trends and features and determine the number of similar events that occurred in the community on the same day of the week or month in the previous year. Thus, the temporal analysis shows that trends vary by month and by type of crime

Time series prediction is a set of methods and processes that break down a series into components and explainable segments to identify patterns, estimations, and forecasts. Time series forecasting seeks to comprehend the underlying meaning of data points by using a model to predict future values based on established historical values. This study focuses on time series techniques such as regression, ARIMA, and deep learning approaches.

ARIMA FOR CRIME FORECASTING

Time series regression is a statistical technique for forecasting future responses based on response history (autoregressive dynamics) and dynamics transferred from related predictors. According to experimental or observational evidence, time series regression may aid in understanding and forecasting complex systems. Time series regression is commonly used to model and forecast economic, social, and biological systems. On this basis, the regression model proposed by Yadav et al. Was developed using data from Indian statistics, which include data on various crimes committed over the last 14 years (2001–2014), including murder, kidnapping, Robbery and rape are both crimes. Therefore, the crime rate in various states can be forecasted for the coming years. They primarily used four data mining techniques to analyze crime and detect crime patterns using supervised, semi-supervised, and unsupervised learning techniques such as K-mean to create multiple groups based on high and low values in criminal records.

In addition, the linear regression method finds no randomness in the test samples (without incurring too much forecasting error). Wulff and Shaun exploited ARIMA for forecasting time series data that can be used to understand or forecast the series' evolution. The process is successfully used in various fields, including economic forecasting, marketing, and industrial development. It is best suited to short-term forecasting, but forecasting requires at least 50 observations or more. Payne et al. used a crime forecasting model based on ARIMA. A correlational analysis of devastating pandemics like Covid-19 and their impact on economic growth was presented. They discovered a strong link between unemployment and criminality.

They forecasted crime in Delhi, Bombay, over the next six months. However, they failed to demonstrate a strong link between crime and Covid-19. Their approach applies to grid cells but generally requires significantly more historical data and is somewhat limited in incorporating additional details into a strategy. A similar study was conducted by Yadav et al. who proposed that a time series dataset can be generated by combining "big data" management techniques and generalized linear regression for statistical analysis. The ARIMA model was used to reduce the residual of the forecasting model for improving the ability to identify common crime patterns among various crime sites for the selection of criminal sites. However, the described autoregressive model only fits linear data relationships. This model also ignores the nonlinear regression aspect of time series data.

DEEP LEARNING TECHNIQUES FOR CRIME FORECASTING

The cost of machine learning methods has decreased dramatically because of the advancement of high-performance computers. They achieved great success in various fields. Notably, deep learning has yielded promising results for different classification problems, from speech initiation to visual recognition, a relatively recent advance in AI. One area of deep learning that has received little attention is crime forecasting. Several researchers observe that deep learning is also well suited to deal with the temporal and spatial elements of a problem. Stec et al. Predicted crime in Delhi and Bombay by using another solution based on ANN with additional external data. Crimes in Delhi are organized according to the police beats. Each beat is arbitrarily large and contains a high standard deviation of crimes. Using this high standard deviation, they divided each type of crime into ten bins of varying sizes. The regression problem was replaced by a classification problem, which is more effective and easier to measure. They also used a teaching technique known as walk-forward training. They trained various architectures, including a fully linked network, a LSTM shows the best results. However, they did not change the structures of the various networks. Finally, they investigated whether removing\ external data would improve or degrade the results. They discovered that public transportation and the census are also crucial for crime forecasting in their models.

Tumulak et al. demonstrated a straightforward method for forecasting crime in a spatial space by combining grid thematic mapping and neural networks. Monthly, weekly, and daily data are used to train the model. The study area, Delhi and Bombay City, is divided into square grids of varying sizes, and crime incidents are then recorded in each grid cell for each snapshot. This information is then fed into the neural network, which forecasts potential crime hotspots for the next time interval. According to preliminary findings, the data snapshot is separated monthly and weekly. The model is sufficiently accurate to forecast crime areas. Although the model is simple, it may serve as a starting point for future crime modelling and forecasting studies.

Wawrzyniak et al. used predictive data-based modelling techniques based on data driven ML approaches. They utilized a deep learning architecture based on LSTM to achieve a high level of forecasting. A LSTM architecture for crime forecasting and relevant inputs was implemented using a Gram–Schmidt calendar for network input selection. A virtual leave-one-out test was conducted to determine the optimal number of hidden neurons for optimal performance.

MATERIAL AND METHODS

Autoregressive integrated moving average (ARIMA) model

ARIMA is the most welcomed model of the linear model. It is used to predict time-series. It has been a large success not only in the academic research but also in the industrial and Biological applications. A common ARIMA model is ordered by (p, d, q), and it can be shown as:

$$\emptyset(B)\,(1-B)^d y_t = \theta(B)\varepsilon_t$$

Where; $\emptyset(B)$, $\theta(B)$ are the Autoregressive and Moving average polynomial as defined by:

$$\emptyset(B) = 1- \emptyset_1 B - \emptyset_2 B^2 - \text{-------------} - \emptyset_p B^p$$

And

$$\theta(B) = 1- \theta_1 B - \theta_2 B^2 - \text{-------------} - \theta_p B^p$$

Where $\emptyset_1$, $\emptyset_2$, ------- $\emptyset_p$ are the autoregressive coefficients while θ_1, θ_2 ----- θ_p are the moving average coefficients

In the above, d represents the order of differencing, B is the backshift operator, i.e., $By_t = y_t-1$, p and q are the order of autoregressive and moving average respectively. ε_t is a white noise process. The value of the ARIMA parameters (p, d, q) for AR and MA can be obtained from the behavior of the autocorrelation function (ACF) and the partial autocorrelation function (PACF). These functions help in estimating parameters that can be used to forecast data by using the ARIMA model.

When the ARIMA model is fitted to the original data, the ARIMA model should go through the following 4steps.

I. Identify the ARIMA (p, d, q) structure.
II. Estimate the unknown parameters.
III. Goodness -of -fit test for the estimated residual.
IV. Predicting future results based on known data.

Theε_t should be independent and it can be considered as a positive random variable with mean = 0 and a constant variance=σ^2. The roots of $\emptyset_p(x_t) = 0$ and $\theta_q(x_t) = 0$ should all be outside the unit circle. Box and Jenkins (1976) suggested that the ARIMA model should use at least 50 or preferably 100observations. If the data has significant seasonal changes periodically. We can use the SARIMA model to eliminate the effects of seasonal cycles.

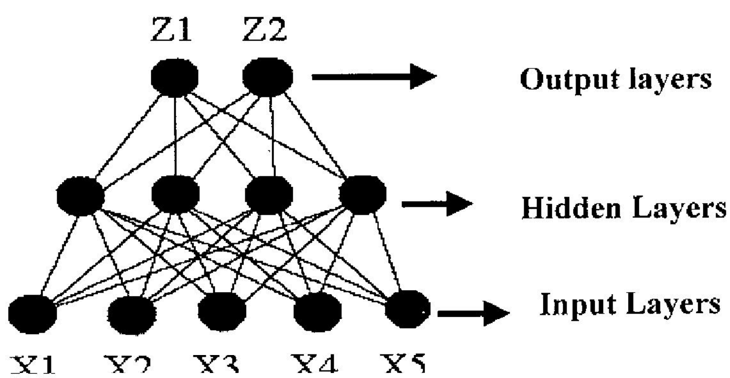
LONG SHORT-TERM MEMORY (LSTM)

LSTMs are capable of learning long-term dependencies, which is a useful capability when you need to model time series data. As mentioned in the previous section, LSTMs help preserve the error that can be back propagated through time and layers, without risk of losing important information: LSTMs have internal mechanisms called gates and cell state that can regulate the flow of information (Zhang et al. 2019). The cell makes decisions about what, how much, and when to store and release information: they learn when to allow information to enter, leave, or be deleted through the iterative process of making guesses, back propagating error, and adjusting weights via gradient descent. Below figure illustrates how information flows through a cell and is controlled by different gates.

Long Short-Term Memory (LSTM) is a kind of Recurrent Neural Network (RNN) with the ability of recollecting the qualities from prior stages with the end goal of future use. Prior to digging into LSTM, it is important to have a brief look at what a brain network resembles.

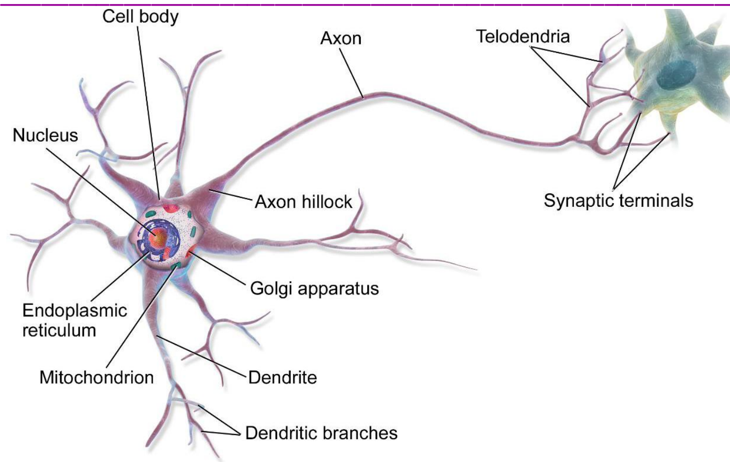
ARTIFICIAL NEURAL NETWORKS

The word" Neural networks" is very popularly mentioned in recent years. It means machines that something of them is like the brain, may be filled with science fiction mythology as the sci-fi content. Neural networks are prone to sub-elements, units or nodes of the interrelated part of its function based on animal neurons and quickly. The branching capability of the networks are stored in the cross-connect strength, otherwise, through the weights which is contained by training by adapting the training pattern set or learning from the training pattern set. In this paper, we first introduce a relationship between biological neural network and artificial neural network, then describe a historical overview of artificial neural network. Finally, we will discuss the common problems in these neural networks when using them to solve combinatorial optimization problems. A neural network consists of at least three layers namely: 1) an input layer, 2) hidden layers, and 3) an output layer. The number of features of the data set determines the dimensionality or the number of nodes in the input layer. These nodes are connected through links called "synapses" to the nodes created in the hidden layer(s). The synapses links carry some weights for every node in the input layer. The weights basically play the role of a decision maker to decide which signal, or input, may pass through and which may not. The weights also show the strength or extent to the hidden layer. A neural network basically learns by adjusting the weight for each synopsis. In the hidden layers, the nodes apply an activation function (e.g., sigmoid or tangent hyperbolic (tanh) on the weighted sum of inputs to transform the inputs to the outputs, or predicted values. The output layer generates a vector of probabilities for the various outputs and selects the one with minimum error rate or cost, i.e., minimizing the differences between expected and predicted values, also known as the cost, using a function called Soft Max. The assignments to the weights vector and thus the errors obtained through the network training for the first time might not be the best. To find the most optimal values for errors, the errors are "back propagated" into the network from the output layer towards the hidden layers and as a result the weights are adjusted. The procedure is repeated, i.e., epochs, several times with the same observations and the weights are re-adjusted until there is an improvement in the predicted values and subsequently in the cost. When the cost function is minimized, the model is trained.



BIOLOGICAL NEURAL NETWORK TO ANN

The brain is constituted of about 10 billion neurons for the most parts, each one is concerned to about 10,000 other neurons. Every neuron can receive electrochemical inputs, and the inputs are from other neurons at the dendrites as shown below figure. When the summation of the electrical inputs has the plenty energetic to operative the neuron, it will transmit electrochemical signals through the axon, then pass the signals to another neuron whose dendrites attach to other neurons at the end of any axon.



Structure of a typical biological neuron
(https://en.wikipedia.org/wiki/Neuron)

RECURRENT NEURAL NETWORKS

Recurrent neural networks (RNNs) were created in the 1980s but have just recently been gaining popularity due to increased computational power from graphic processing units. They are especially useful with sequential data because each neuron or unit can use its internal memory to maintain information about the previous input. An RNN has loops in it that allow information to be carried across neurons while reading in input.

However, a simple recurrent network suffers from a fundamental problem of not being able to capture long-term dependencies in a sequence. This is a major reason RNNs faded out from practice until some great results were achieved with using a LSTM unit inside the neural network. Adding the LSTM to the network is like adding a memory unit that can remember context from the very beginning of the input (Lazzeri 2019a). LSTM neural networks are a particular type of RNN that have some internal contextual state cells that act as long-term or short-term memory cells. The output of the LSTM network is modulated by the state of these cells. This is a very important property when we need the prediction of the neural network to depend on the historical context of inputs rather than only on the very last input. They are a type of neural network that adds native support for input data comprising sequences of observations. The addition of sequence is a new dimension to the function being approximated. Instead of mapping inputs to outputs alone, the network can learn a mapping function for the inputs over time to an output. Recurrent neural network (RNN), also known as auto associative or feedback network, belongs to a class of artificial neural networks where connections between units form a directed cycle. This creates an internal state of the network, which allows it to exhibit dynamic temporal behavior (Poznyak, Oria, and Poznyak 2018). RNN can leverage their internal memory to handle sequences of inputs and, instead of mapping inputs to outputs alone, it is capable of leveraging a mapping function for the inputs over time to an output. RNNs have shown to achieve the state-of-the-art results in many applications with time series or sequential data, including machine translation, image processing and medical applications and speech recognition. In particular, LSTM is a type of RNN architecture that performs particularly well on different temporal processing tasks, and LSTM networks are able to address the issue of large time lags in input data successfully. LSTM networks have several nice properties such as strong prediction performance as well as the ability to capture long-term temporal dependencies and variable-length observations (Che et al. 2018). Exploiting the power of customized RNN models along with LSTM models is a promising venue to effectively model time series data. RNNs, LSTMs, can help data scientists build accurate models for their time series forecasting solutions. In this paper, we look at the presentation of an ARIMA model with the LSTM model in the expectation of picture handling and clinical applications time series to decide the ideal characteristics of involved factors in a commonplace forecast model.

EVALUATION METRICS

In this paper, four state-of-the-art forecasting have been used as evaluation measures for determining forecasting accuracy. They are used because the problem is regarded as a regression forecasting problem. In other words, this model seeks to forecast the number of crimes over time. First, MAPE calculates the sum of individual errors divided by each period. It is the average absolute percentage error between actual (x_t) and predicted (x_t')

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{xt-xt'}{xt}\right|$$

Moreover, it calculated the average residuals in the dataset. Second, MAE calculated the absolute error between actual (x_t) and predicted (x_t')

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|xt-xt'|$$

Third, RMSE is the combination of bias and variance of the prediction, and it is easy to interpret the model accuracy.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{xt'-xt}{n}\right)^2}$$

It calculated the average of the squared difference between actual (x_t) and predicted (x_t') values by taking the square root. It also calculated the standard deviation of residuals.

Finally, R^2 is used to depict how well the dataset is fitted to the data. Moreover, it captures the variation between the actual and predictor response variables.

$$R^2 = \frac{\text{sum of square due to residual}}{\text{sum of square due to total}}$$

TRAINING AND TESTING DATASET

Crime data from India 2001–2014 is collected from the NCRB open portal. The first four years of data (2001–2004) are used for training and testing with a ratio of 80% training and 20% testing. In addition, the last four months' data of 2004 is used as a validation set. The first four years of data forecast the next year's (2005) crime and so on.

CONCLUSION

Crime forecasting is a crucial problem that can help law enforcement agencies to control and prevent crime. With the new progression on creating modern AI based methods and specifically profound learning calculations, these strategies are acquiring prominence among analysts across jumper's disciplines. The significant inquiry is then the way that exact and strong these recently presented approaches are when contrasted and Ordinary strategies. This paper analyzes the exactness of ARIMA and LSTM, as agent methods while determining time series information. These two methods were executed and applied on a bunch of Wrongdoing information and the outcomes demonstrated that LSTM was better than ARIMA. All the more explicitly, the LSTM-based calculation worked on the expectation by 82% on normal contrasted with ARIMA. Besides, the paper reports no improvement when the quantity of ages is changed. The work portrayed in this paper advocates the advantages of applying profound learning-based calculations and strategies to the Wrongdoing information. There are a few other forecast issues in Violations that can be figured out utilizing profound learning. The creators intend to research the improvement accomplished through profound advancing by applying these methods to a few different issues and datasets with different quantities of elements.

The proposed method could help law enforcement agencies by forecasting crime trends and patterns. Moreover, the proposed method is efficient enough for forecasting in several real-world time series applications such as health care, Biological sciences and image processing. We aim to exploit reinforcement learning to understand crime incident linkage in the future. Moreover, transfer learning can also enhance crime prediction accuracy by utilizing the crime domain knowledge in the relevant area. Besides, crime dataset completeness, reliability, and accuracy are vital for an efficient and reliable forecasting model.

ACKNOWLEDGEMENTS

The author is thankful to the Director, National Crime Records Bureau, New Delhi for giving the permission to publish this study and also thankful to providing data set used in the study.

REFERENCES

1. Goel RK, Saunoris JW. Does variability in crimes affect other crimes? The case of international corruption and shadow economy. Applied Economics. 2019; 51(3):239–258.
2. Estrada MAR, Ndoma I. How crime affects economic performance: The case of Guatemala. Journal of Policy Modeling. 2014; 36(5):867–882.
3. Rudolph M, Starke P. How does the welfare state reduce crime? The effect of program characteristics and de commodification across 18 OECD-countries. Journal of Criminal Justice. 2020; 68:101684.
4. Catlett C, Cesario E, Talia D, Vinci A. Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments. Pervasive and Mobile Computing. 2019; 53:62–74.
5. Butt UM, Letchmunan S, Hassan FH, Ali M, Baqir A, Sherazi HHR. Spatio-temporal crime HotSpot detection and prediction: A systematic literature review. IEEE Access. 2020; 8:166553–166574.
6. Atems B. An Empirical Characterization of the Dynamic Effects of Police Spending on Violent and Property Crime. Economic Inquiry. 2020; 58(2):717–744. 8. Jha S, Yang E, Almagrabi AO, Bashir AK, Joshi GP. Comparative analysis of time series model and machine testing systems for crime forecasting. Neural Computing and Applications. 2021; 33 (17):10621–10636.
7. Degeling M, Berendt B. What is wrong about Robocops as consultants? A technology-centric critique of predictive policing. Ai & Society. 2018; 33(3):347–356.
8. Poyser O. Exploring the dynamics of Bitcoin's price: a Bayesian structural time series approach. Eurasian Economic Review. 2019; 9(1):29–60.
9. Zhang GP, Qi M. Neural network forecasting for seasonal and trend time series. European journal of operational research. 2005; 160(2):501–514
10. Nelson M, Hill T, and Remus W, O'Connor M. Time series forecasting using neural networks: Should the data be deseasonalized first? Journal of forecasting. 1999; 18(5):359–367.
11. Bandara K, Bergmeir C, Smyl S. Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. Expert systems with applications. 2020; 140:112896.
12. Taieb SB, Bontempi G, Atiya AF, Sorjamaa A. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. Expert systems with applications. 2012; 39(8):7067–7083.
13. Makridakis S, Spiliotis E, Assimakopoulos V. Statistical and Machine Learning forecasting methods: Concerns and ways forward. PloS one. 2018; 13(3):e0194889. 18. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing. 2003; 50:159–175.
14. Smyl S. A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. International Journal of Forecasting. 2020; 36(1):75–85. 20. Wawrzyniak ZM, Jankowski S, Szczechla E, Szymański Z, Pytlak R, Michalak P, et al. Data-driven models in machine learning for crime prediction. In: 2018 26th International Conference on Systems Engineering (ICSEng). IEEE; 2018. p.1–8.

15. Linning SJ, Andresen MA, Brantingham PJ. Crime seasonality: Examining the temporal fluctuations of property crime in cities with varying climates. International journal of offender therapy and comparative criminology. 2017; 61(16):1866–1891.
16. Almanie T, Mirza R, Lor E. Crime prediction based on crime types and using spatial and temporal criminal hotspots. arXiv preprint arXiv:150802050. 2015;
17. Venturini L, Baralis E. A spectral analysis of crimes in San Francisco. In: Proceedings of the 2nd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics; 2016. p. 1–4.
18. Yadav S, Timbadia M, Yadav A, Vishwakarma R, Yadav N. Crime pattern detection, analysis & prediction. In: 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA). vol. 1. IEEE; 2017. p. 225–230.
19. McClendon L, Meghanathans N. Using machine learning algorithms to analyze crime data. Machine Learning and Applications: An International Journal (MLAIJ). 2015; 2(1):1–12.