

# Risk Prediction of Crime data using Arima and LSTM

<sup>1</sup>Aravindarajan V, <sup>2</sup>Shalini M, <sup>3</sup>Swetha E, <sup>4</sup>Vijaya Lakshmi G M

<sup>1</sup>Assistant Professor, <sup>2,3,4</sup>Students  
Information Technology Department  
K.L.N College of Engineering, Sivagangai, Tamil Nadu, India

**Abstract:** Urbanization makes a ton of social issues. One of these issues intrinsic in all urban areas of the world is crime. Police information bases gather a lot of information that could be broke down all together crime percentages. The examination of crime and prediction of number of crime stays quite possibly the most fascinating issue for researchers. For a non-industrial nation like India, it isn't new that individuals know about violations happening regularly. With the quick urbanization of urban areas, we need to continually know about our environmental elements. To keep away from the sad, this work attempt to notice crime percentages by the crossover forecast technique. Based on deep learning methods such as an auto-regressive integrated moving average (ARIMA) model and a long short term memory (LSTM) model, we established a hybrid Model. This model will predict the crime percentage with more accuracy. Our investigation gives a thorough manual for crime prediction with respect to crime prediction with more accuracy by comparing supervise classification machine learning algorithm.

**Index Terms:** crime prediction, violence, criminal case, risk prediction.

## INTRODUCTION:

A large portion of the current crime analysis strategies just portray the past crime circumstance. In any case, in the crime predictive analysis method, regression analysis, risk area analysis, data mining (classification technique, cluster technique), geographic profiling, hot spot analysis, proximity-repeated modelling. It is feasible to make more precise and compelling crime predictions by quickly and successfully handling a wide variety of data like time and space analysis. Additionally, it is feasible to predict crime and victims, showing the probability that people or gatherings will endure crime. This is a strategy for examining focuses in peril zones or designated by crime type, including violence, robbery, and isn't very different from the methods utilized in criminal and crime prediction. Studies focused on historical criminal event records and related data analysis, concentrates on zeroed in on criminal behavior psychology, and instances of crime prediction system benefits that can notice criminal events and view important insights in light of crime prediction methods. Among these cases, research on data analysis technology and crime prediction system assumes a fundamental part in diminishing the frequency of crime on account of the new extreme improvement of machine learning technology.

## MOTIVATION OF THE PROJECT:

Among many social issues, crime issues have become more-savvy, progressed, expanded, and widened, making it challenging to anticipate and adapt to violations before they happen. The greater part of the examinations on the event of crimes that have been directed up to this point comprise of various investigations focusing on the past, such as the increment or lessening in the crime percentage contrasted with the earlier year, and concentrates, for example, anticipating the event of wrongdoing and it are lacking to plan countermeasures.

## PROBLEM STATEMENT:

Accurate crime prediction is complicated but necessary for the prevention of criminal acts. The accurate estimation of the crime rate, types and hotspots from past patterns creates many computational challenges and opportunities. Crime Prediction based on machine learning is the current mainstream for prediction analysis. However, only a few studies systematically compare different machine learning method.

## Literature Survey:

### 1. Crime Spatiotemporal Prediction with Fused Objective Function in Time Delay Neural Network

Author: A. Ghazvani

Findings:

Urbanization has been speeding up social and economic transformations in urban communities, the smallest social units in a city. However, urbanization brings challenges to urban management and security. Therefore, a system of risk prediction of crimes may be essential to crime prevention and control in urban communities and its system improvement. To tackle crime-related problems in urban communities, this paper proposes a model of daily crime prediction by combining Long Short-Term Memory Network (LSTM) and Spatial-Temporal Graph Convolutional Network (ST-GCN) to automatically and effectively detect the high-risk areas

in a city. Topological maps of urban communities carry the dataset in the model, which mainly includes two modules - spatial-temporal features extraction module and temporal feature extraction module - to extract the factors of theft crimes collectively. We have performed the experimental evaluation of the existing crime data from Chicago, America. The results show that the integrated model demonstrates positive performance in predicting the number of crimes within the sliding time range.

## **2. Predictability comparison of three kinds of robbery crime events using LSTM**

Author: Y. Mei

Findings:

Prediction of crime occurrences is of great significance for underlying criminal risk reduction and public security improvement. In all kinds of crime events, robbery is a more serious and noticeable one that threatens our everyday life. Different kinds of robbery events have different characteristics, and therefore their predictability varies. Previous studies mainly focus on forecast of all kinds of crime events at a fixed spatiotemporal scale, however, special predictability comparison between different robbery events at multiple scales is insufficient. In this paper, we use LSTM to capture long-short space-time dependencies of robbery events, and explore the predictability differences of three kinds of robbery crimes at different spatial and temporal scales. Daily crime occurrence data of Atlanta, US, from 2009 to 2017, and crime data of Baltimore, US, from 2011 to 2016, are respectively used in experiments, and the correlation coefficient between the predicted value and the true value is used to evaluate the model performance. Three robbery events, including commercial robbery, pedestrian robbery, as well as residence robbery, are compared. Experimental results show that the predictability of these three kinds of robbery events differs from each other and varies across different cities. In Atlanta, the commercial robbery is more predictable than that of the residential robbery and pedestrian robbery, with a higher correlation coefficient up to about 0.75. While in Baltimore, the predictability of residential robbery and commercial robbery achieves up to about 0.9, much higher than that of pedestrian robbery. In addition, experiments demonstrate different degree of predictability changes of three robbery events as the spatiotemporal scale changes, reflecting the different geographical distribution characteristics of the three robbery events.

## **3. Short-term load forecasting for Spanish insular electric system**

Author: JE. Caro

Findings:

In any electric power system, the Transmission System Operator (TSO) requires the use of short-term load forecasting algorithms. These predictions are essential for appropriate planning of the energy resources and optimal coordination for the generation agents. This study focuses on the development of a prediction model to be applied to the ten main Spanish islands: seven insular systems in the Canary Islands, and three systems in the Balearic Islands. An exhaustive analysis is presented concerning both the estimation results and the forecasting accuracy, benchmarked against an alternative prediction software and a set of modified models. The developed models are currently being used by the Spanish TSO (Red Eléctrica de España, REE) to make hourly one-day-ahead forecasts of the electricity demand of insular systems.

## **4. Forecasting based on an ensemble autoregressive moving average - adaptive neuro - fuzzy inference system - neural network - genetic algorithm framework**

Author: JF. Prado

Findings:

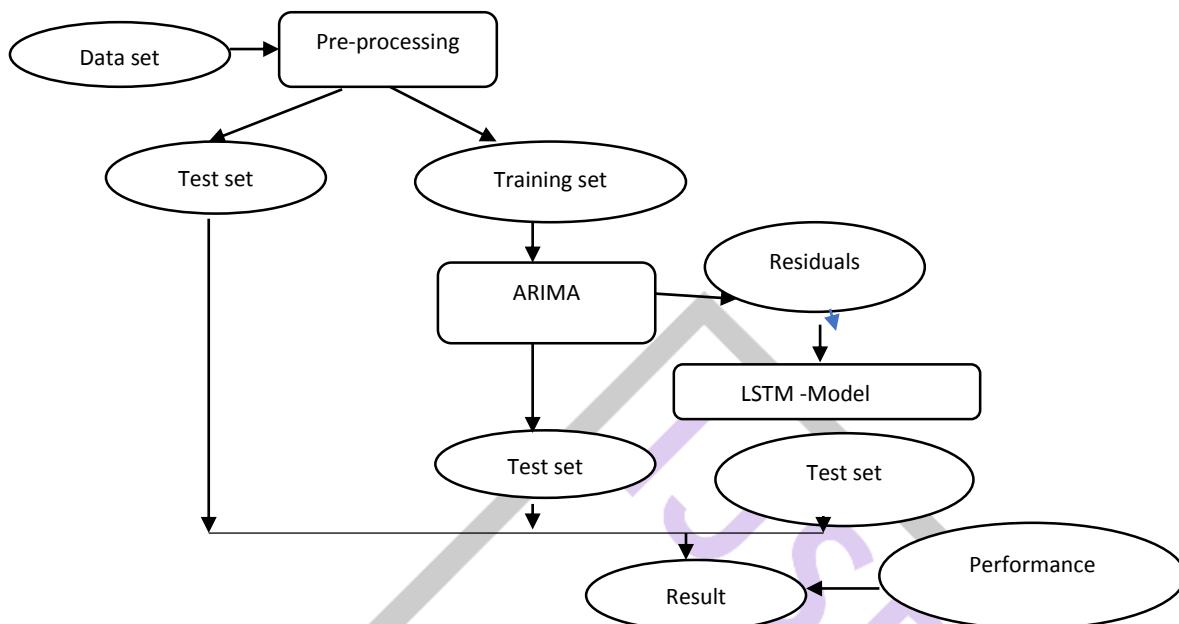
This paper proposes a novel ensemble methodology comprising an auto regressive integrated moving average, artificial neural network, fuzzy inference system model, adaptive neuro fuzzy inference system, support vector regression, extreme machine learning, and genetic algorithm to forecast aggregated, long-term energy demand. After comparing the framework with several benchmark methods by the loss functions mean squared error and mean absolute percentage error, and applying a model confidence set this work suggests that the proposed method improves forecasting accuracy over previous approaches. The proposed approach resulted in a mean squared error decrease of 22.3% and mean absolute percentage error by 33.1% with respect to the best artificial intelligence and econometric models in a sample study. Post-processing optimization of the forecasting ensemble in this methodology improves prediction accuracy. The approach developed herein provides an addition to the field for how hybridized models and augmented forecasting accuracy can be improved. Continued improvements to forecasting techniques are extremely important especially in areas where there are upper bound constraints on supply and lower bound on minimum operation levels.

### **Proposed System:**

This study plans to recognize the best model for anticipating the crook rate by consolidating Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory Neural Network (LSTM) models with exogenous meteorological factors. The

information of everyday crime rate were gathered as the preparation set, and the testing set. The ARIMA and LSTM models consolidated and uncombined with exogenous meteorological factors were taken on to fit the crime rate occurrence by utilizing the information of the preparation set. The gauging exhibitions of the four fitted models were confirmed by utilizing the information of the test set.

### System Architecture:



### Advantage:

- More accurate
- The ARIMA model filters linear tendencies in the data and passes on the residual value to the LSTM model
- For both linear and non-linear data.

### Limitation:

- Less accuracy
- Only for linear data

### System Requirements:

- **Hardware Tools**
  1. A PC with windows OS
- **Software Tools**
  1. Python IDE 3.7.6
  2. Software Package
  - Numpy
  - Matplotlib
  - Pyramid-arima

### CONCLUSION:

In this project, two deep neural network models ARIMA and LSTM were utilized to anticipate crime rate. Regardless of their intricacy in modelling, the results were empowering. As per the results, the ARIMA and LSTM model prediction were superior to the traditional model as far as the pattern and variety of crime rate. This proposed model lies in the fact that the ARIMA and LSTM model can create more exact results than the past model. Additionally, the training time of the ARIMA and LSTM model is less with a multiple LSTM layer. The proposed network is demonstrated to be a modern technique in modelling, examining, and anticipating criminal rates.

**REFERENCES:**

**Base paper:** <https://ieeexplore.ieee.org/document/9276416>

Additional Requirement:

- [1] Khalid, S., Khan, S. A., & Ifzal, S. Q. (2021). A Fuzzy Logic-Based Framework for Mapping Crime Data on Established Sociological Hypothesis for Societal Disorder Identification and Prevention. IEEE Access, 9, 80197–80207. doi:10.1109/access.2021.3083542
- [2] D. M. Raza and D. B. Victor, "Data mining and Region Prediction Based on Crime Using Random Forest," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 980-987, doi: 10.1109/ICAIS50930.2021.9395989.
- [3] S. Ghankutkar, N. Sarkar, P. Gajbhiye, S. Yadav, D. Kalbande and N. Bakereywala, "Modelling Machine Learning For Analysing Crime News," 2019 International Conference on Advances in Computing, Communication and Control (ICAC3), 2019, pp. 1-5, doi: 10.1109/ICAC347590.2019.9036769.
- [4] N. Baloian et al., "Crime prediction using patterns and context," 2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2017, pp. 2-9, doi: 10.1109/CSCWD.2017.8066662.
- [5] S. Tarlekar, R. Bhosle, E. D'souza and S. Sheikh, "Geographical Crime Rate Prediction System," 2021 IEEE India Council International Subsections Conference (INDISCON), 2021, pp. 1-6, doi: 10.1109/INDISCON53343.2021.9582218.

