# Categorical Crime Rate Analysis and Prediction

submitted in partial fulfillment of the requirement
for the award of the Degree of

**Bachelor of Technology**
in
**Computer Engineering**

by

**Jainam Jain**
**Jay Mehta**
**Vaidehi Vatsaraj**

under the guidance of

**Dr. Anant V. Nimkar**


**Department of Computer Engineering**
Bharatiya Vidya Bhavan's
Sardar Patel Institute of Technology
(Autonomous Institute Affiliated to University of Mumbai)
Munshi Nagar, Andheri-West, Mumbai-400058
University of Mumbai
October 2020

# Certificate

This is to certify that the Project entitled "Categorical Crime Rate Analysis and Prediction" has been completed to our satisfaction by Mr. Jainam Jain, Mr. Jay Mehta and Ms. Vaidehi Vatsaraj under the guidance of Dr. Anant V. Nimkar for the award of Degree of Bachelor of Technology in Computer Engineering from University of Mumbai.

**Certified by**

**Dr. Anant V. Nimkar**
**Project Guide**

**Dr. Anant V. Nimkar**
**Head of Department**

**Dr. Bhalchandra Chaudhari**
**Principal**


**Department of Computer Engineering**
Bharatiya Vidya Bhavan's
Sardar Patel Institute of Technology
(Autonomous Institute Affiliated to University of Mumbai)
Munshi Nagar, Andheri(W), Mumbai-400058
University of Mumbai
April 2019

# Project Approval Certificate

This is to certify that the Project entitled "Categorical Crime Rate Analysis and Prediction" by Mr. Jainam Jain, Mr. Jay Mehta and Ms. Vaidehi Vatsaraj is found to be satisfactory and is approved for the award of Degree of Bachelor of Technology in Computer Engineering from University of Mumbai.

**External Examiner** **Internal Examiner**

**(signature)** **(signature)**

**Name:** **Name:**

**Date:** **Date:**

**Seal of the Institute**

## Statement by the Candidates

We wish to state that the work embodied in this thesis titled "Categorical Crime Rate Analysis and Prediction" forms our own contribution to the work carried out under the guidance of Dr. Anant V. Nimkar at the Sardar Patel Institute of Technology. We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission.

**Name and Signature:**
**1. Jainam Jain**
**2. Jay Mehta**
**3. Vaidehi Vatsaraj**

# Acknowledgments

It is a pleasure to acknowledge the help and support that has gone to in making this thesis. We express our sincere gratitude to our guide Dr. Anant V. Nimkar for his invaluable guidance. Without his encouragement, this work would not be a reality.

With the freedom he provided, we really enjoyed working under his guidance. We thank our examiner, Dr. Sudhir Dhage for his word of advice. We would like to mention special thanks to the open source community to provide us with the necessary tools and knowledge base required to complete this project.

We thank staff of Computer Engineering Department for giving us all the facilities required to carry out this research work.

We would like to thank all our family members and well wishers for their constant encouragement for all these years, without which we could not have completed this work.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| VAR | Vector Autoregression |
| LSTM | Long Short-Term Memory |
| AIC | Akaike information criterion |
| ADF | Augmented Dickey Fuller |

**Abstract**

Crime rates have been on the rise ever since the population began rising exponentially. The only way of preventing the crimes is to identify a possibility and eliminate it. Our aim is to predict categorical crime rate using its relation with various factors like GDP, Unemployment, Tourism etc. This factors will give better insight to our analysis and will train our model in a much better way. Additionally, there are a myriad of categories of crimes and all of them needed to be treated differently as it has varied consequences. In this project, we aim to predict the categorical crime rates using our own combinational model methodology and analyze their relation with multiple factors.

# Chapter 1

# Introduction

Crimes are a social hindrance and cost our society in several ways. It is a menace that the world faces today. Rise in cases of ghastly murders, sensational robberies, rapes, thefts and kidnappings each day is very daunting. Although government implemented numerous ways to track and decrease the crime rate, the outcomes are not fulfilling. We need more effective ways of resource allocation, planning and awareness creation to curb the increasing crimes.

Additionally, There are multiple categories of crimes and all of them need to be treated differently owing to their varied consequences. Although, all types of crimes have different consequences depending on their nature and severity, they all cause physical, emotional and psychological harm. Crimes are also categorized based on their seriousness into misdemeanor, or a felony[1] . Moreover, crimes occurring at a particular geo-location can also be consequences of socio-economic factors such as literacy rate, population, economic status, urbanization, poverty index, political upheaval [2] . A detailed study of these factors and their impact on the crime rates and incidences will be helpful in devised crime prevention strategies.

Most of the research in this domain cumulatively applies the data mining techniques and does not treat the crimes based on its specifications. Moreover, most of the research focuses on identifying the suspect, drawing relation between the crimes(clustering) or factors considering criminal behaviour. Hence, we wish to predict the crime rate across various categories by only considering the past values but also the impact of relevant factors on them. Artificial Intelligence has presented with efficient ways of achieving this task through multiple algorithms , procedures. Utilizing them, we intend to carry out detailed analysis on the crime rates and its relationship with other socio-economic factors and exploit it to make the predictions.

## 1.1 Motivation

Crimes are a social hindrance and cost our society in several ways. There are a myriad of categories of crimes and all of them needed to be treated differently as it has varied consequences. Most of the research in this domain cumulatively applies the data mining techniques and it does not treat the crimes based on its need. Moreover, most of the research focuses on identifying the suspect, drawing relation between the

crimes(clustering) or factors considering criminal behaviour Hence, more research into effect of more factors such as recidivism of a particular criminal can help the current system a lot. More work concerning patterns specifically to India need to be considered to be precise.

## 1.2 Objectives

The Objectives that we have identified for our research are as follows:

- To study and compare the existing work done on crime rate prediction and its dependence on relevant attributes.
- To predict crime rate using these factors and the past time series across four categories: location, victim, type of crime, act.
- To evaluate the system using a suitable metrics.

## 1.3 Problem Statement

We have chosen the problem statement of Crime rate Prediction, because estimation of possible number of crimes helps the governing authority plan and strategize crime prevention actions, allocate required force, suggest/ enforce surplus curfew/restrictions. Our aim is to predict the categorical crime rate wherein the intended categories are state of crimes, type of crimes, type of victims involved, act-wise. We also aim to identify the impact of factors such as unemployment, recidivism, GDP, population and literacy, tourism. The relation established with these factors will further assist us in identifying the major contributors to crime and will facilitate crime rate prediction.

## 1.4 Contributions

We propose the model which will predict crime rates across various categories based on past year's data. It will establish a relationship with factors such as unemployment, GDP, literacy, tourism and utilize it to predict the crime rates. Using this relationship, our model will understand the interdependence and exploit it to predict the crime rates.

## 1.5 Layout of the Report

A brief chapter by chapter overview is presented here.
Chapter 2: A literature review of different statistical methods for crime rate prediction is presented.
Chapter 3: Block Diagram, design, methodology and metrics for Evaluation will be described in this chapter.
Chapter 4: Models used in our prediction, Implementation Flowchart, Data Preprocessing and Testing.

Chapter 5: Complete method of evaluation and results through survey are discussed.
Chapter 6: Conclusion of the project is discussed in this chapter.
Chapter 7: Discussion on future course of research work
Chapter 8: Research Publication.

# Chapter 2

# Literature Survey

Thorough understanding of various research papers and available work helped us in forming a socially relevant problem statement which has not been dealt with already and getting insights into accurate technology stack in developing solution. statistical models: 1) Weighted Moving Average 2) Functional Coefficient Regression 3) Arithmetic-Geometric Progression, were used to predict crime rates on Indian dataset [1]. Crime cases obtained from NCRB were clustered using K-Means clustering, Agglomerative clustering and DBSCAN algorithms and KNN algorithm was used to predict crime in future [2]. Data fetched from NCRB on violent crimes was used to predict future crime rates using 48, SMO Naïve byes bagging and, the Random Forest and SVM. These models were then stacked to form the final model [3]. Crime events were linked, grouped and visualized using KNN algorithm The probability of future occurrences was predicted using Naive Bayes algorithm [4] .

Data was collected from NCRB and was used to predict crime rate using Decision Tree, KNN, Naive Bayes, Random Forest Considered good attributes like age, laws, etc [5]. It used SmartStep Dataset London on logistic regression, support vector machines, neural networks, decision trees, and different implementations, to predict hotspots[6]. Applied k-means clustering technique to identify different Patterns in crime events [7]. Created correlation between Crimes using fuzzy association rules[8]. Applied Apriori algorithm, K-means Clustering, Naive Bayes Classification and Correlation Regression on Indian Dataset to find patterns in crime events[9]. Implemented K-means Semi-Supervised Learning on Chicago Dataset [10]. Analysed crime patterns using PolyAnalyst Link analysis[11]. Applied Apriori algorithm on time, place and category for crime rate detection [12]. In this paper [14], their system predicted regions which have high probability for crime occurrence and can visualize crime prone areas. In this paper [15], the dataset consists of the date and the crime rate that has taken place in the corresponding years. In this project the crime rate was only based on the robbery. They used linear regression algorithm to predict the percentage of the crime rate in the future by using the previous data information. The date was given as an input to the algorithm and the output was the percentage of the crime rate in that particular year.

All of the above papers, predicts the overall rate, but no categories are considered for better assistance in the crime prevention analysis. So, our main focus is to predict categorical crime rate using its relation with various factors like GDP,

Unemployment, Tourism, etc.

## 2.1 Research Gaps Identified

1. Predicts the overall rate, but no categories are considered for better assistance in the crime prevention analysis. [1]
2. Focuses mainly on clustering and does not consider contributing factors. [2]
3. Considers only one category and does not provide in-depth analysis and prediction. [3]
4. Mainly focuses on visualizing the data: Does not provide robust mechanism for predicting future data. [4]
5. Does not Predicts purpose and crime type using factors like GDP, tourism,etc. [5]
6. Does not cater to crime rates and focus only on locations. [6]

# Chapter 3

# Design

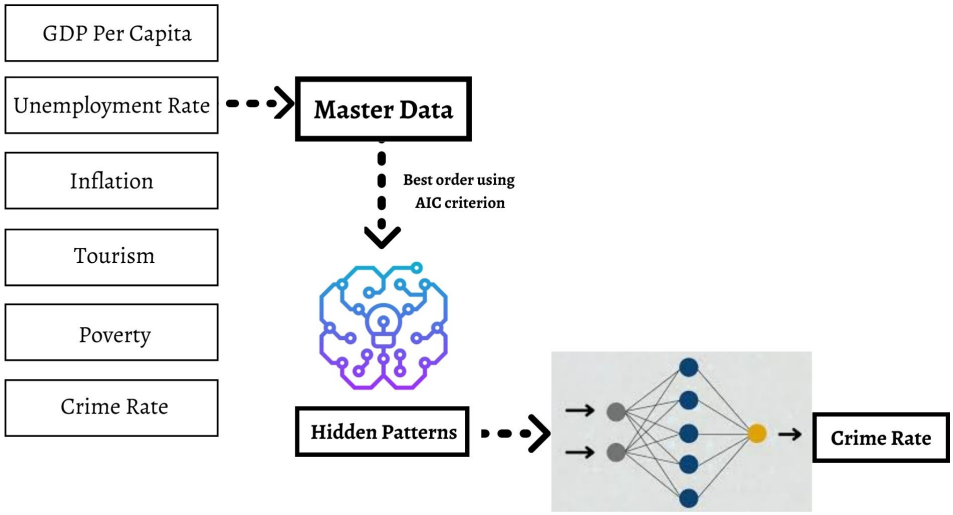


Figure 3.1: Methodology Block Diagram

As shown in figure 3.1, The objectives are satisfied by creating a combinational model to extract the hidden patterns between data points of a time series and that of multiple attributes and to utilize them to predict the future values. This model is trained on the dataset created by collating time series of multiple factors gathered from various resources.

## 3.1 Dataset

The dataset considered for this study is obtained from US Department of Justice, Office of Justice Programs and Uniform Crime Reporting Program websites and consists of data from 1960-2021. Out of multiple attributes considered, following attributes showed major correlation with the crime rates.

Negative correlation signifies an inverse relationship whereas positive correlation signifies a direct relationship. (Not proportional).

Table 3.1: Correlation Table.

| Index No | Attribute Name | Correlation |
|---|---|---|
| 1 | GDP Per Capita | -0.95 |
| 2 | Unemployment Rate | 0.48 |
| 3 | Tourism | -0.94 |
| 4 | Recidivism | 0.54 |
| 5 | Literacy | -0.89 |

## 3.2 Prediction Basis

For demonstration purposes, we consider following categories:

- Crime type: Many types of crimes such as Burglary, Murder, Rape happen daily and all of them need to be handled and prevented in different ways. Hence, we intend to predict the crime rates of those crimes separately. For demonstration and testing purpose, we consider the crime type as 'Murder'. This unlawful killing of human is a common crime and leads to a direct loss of human. The dataset for this crime type was readily available on the websites mentioned in the above section.
- Crime act: Although ample acts are in place to ensure safety of citizens, prohibited offenses are not eliminated. Maintaining the data of these incidences and predicting the future data not only allows the officials to raise awareness about them but also make modifications in order to make it more stringent. For this study, we have considered The National Stolen Property Act of 1934 to execute time series analysis. This act forbids certain incidences relating to stolen property and forgery.
- Victim involved: Crimes can also be categorized based on the victims affected in the act. It is observed, that most of the crimes affect a particular group of citizens. For example, The most common victim involved in the rape cases are women. Hence, protection and training is provided to women to curb the cases of rape. In such scenarios, predicting the crimes happening against a particular category of citizens is helpful. We have considered victim category as the 'Family and Children'.
- Crimes in a state: The crime rates are predicted across the state to have an overview of its preparedness. For the purpose of this study, we consider California, a state in the Western United States.

# Chapter 4

# Implementation

The uncertainty in the time series data makes the prediction of these attributes a very tedious task.[12] Many statistical models such as ARIMA, SARIMA, EARCH have proved to be effective in such predictions and can be compared based on nature of data. However, combination of these models with neural networks have shown exemplary results in many use cases.[13] Hence we put forth a combinational model of a statistical algorithm and a deep learning model to predict the crime rates.

## 4.1 Algorithms Used in Combinational Model

**Vector AutoRegression (VAR)**

Vector AutoRegression is a multivariate forecasting algorithm. Herein, each variable is defined as a linear combination of its past values and past values of other attributes in the given problem. It can be represented as follows: [10]

$$\boldsymbol{Y_{1,t}} = \boldsymbol{\alpha_1} + \boldsymbol{\beta_{11,1}}\boldsymbol{Y_{1,t-1}} + \boldsymbol{\beta_{12,1}}\boldsymbol{Y_{2,t-1}} + \boldsymbol{\epsilon_{1,t}}$$

$$\boldsymbol{Y_{2,t}} = \boldsymbol{\alpha_2} + \boldsymbol{\beta_{21,1}}\boldsymbol{Y_{1,t-1}} + \boldsymbol{\beta_{22,1}}\boldsymbol{Y_{2,t-1}} + \boldsymbol{\epsilon_{2,t}}$$

When multiple time series are considered for prediction of one variable, all the variables need to be treated symmetrically. VAR algorithm is structured to include an equation explaining the evolution of a parameter based on the lags of all the parameters in the model. Therefore, it is highly efficient in capturing the evolution and inter-dependencies between the variables. [11]

**Long Short-Term Memory (LSTM)**

LSTM are a type of Recurrent Neural Network capable of retaining information for a long period of time. In addition to the standard units maintained by RNN, LSTM includes a 'memory cell' for storage and a set of gates for control. This model is capable of learning important parts of the sequence and forget the less important parts.

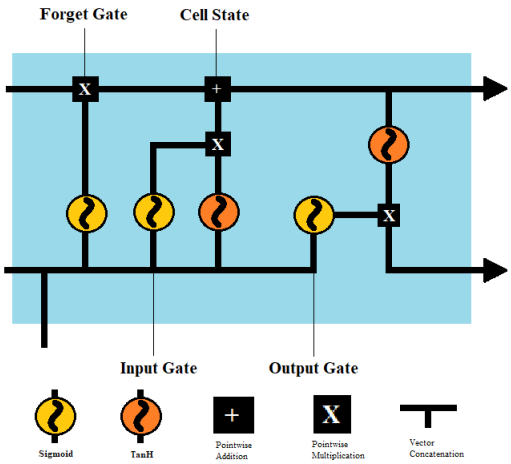


Figure 4.1: LSTM Block Diagram

## 4.2 Implementation steps

### 4.2.1 Data Preprocessing

Obtaining the data from official, authentic sources provided an advantage in terms of its quality. Thus, we had to perform very few pre-processing actions on the dataset.

- Checking for NA values: No NA values were found in the dataset.
- Conversion of data type: Datset presented the numbers in string format(eg. 16,67,299) and were converted into integer data types.
- Maximum Absolute Scaling: This step is applied on the time series of 'Tourism' and 'Population', as the numbers were huge. We used sin and cos functions to keep them within a range.

### 4.2.2 Data Testing

#### Testing Causation using Granger's Causality Test

Vector AutoRegression is based on the idea that each of the time series has significant impact on the target value. Thus, to test the relationship between considered attributes on the target variable i.e. 'Crime Rate', we executed Granger's Causality Test. This test identifies causal relationship for all possible combinations of variables and stores the p-value for the relations. We found that the p-value for relationship between attributes and crime rate are less than 0.05(significance level), concluding that the attributes cause the 'Crime Rate' (Considered attributes have impact on the target variable).

#### Cointegration Test

Group of time series is said to be cointegrated if the linear combination of them is stationary.[7] We performed Johanson's Cointegration Test using Statsmodel package in python. The 'True significance' obtained for all attributes assured us of the long run statistically significant relationship between the attributes.

**Stationarity Test**

The VAR model requires the time series to be stationary. Hence, we checked the dataset for stationarity using Augmented Dickey-Fuller Test (ADF Test). This test uses 'Unit Root Test' and formulates a hypothesis which based on the value of p-value is either accepted/ rejected. In the case of our dataset, not all series were found to be stationary at the first attempt. Hence, we made the series stationary by finding the discrete difference.

### 4.2.3 Splitting the Data

Data was split in 4:1 ratio for training and testing purposes respectively.

### 4.2.4 Finding the Best VAR order

To select the best order for VAR model, we consecutively applied increasing orders. The order giving the least value for AIC was selected for the final model. (The AIC criterion asymptotically overestimates the order with positive probability.) [8]

### 4.2.5 Fit the VAR model

After finding the best lag order using AIC criterion, we fit the VAR model using the best order. To implement VAR algorithm, statsmodel package from python was used.

### 4.2.6 Structure of LSTM

Following LSTM model is trained on fitted values obtained from the VAR model. After fitting the VAR model on the data, fitted data and raw data are divided into several windows of certain length. This length depends on the values of 'Look Back' and 'Look ahead'. As an initial step, we defined an 'Adam Optimiser' and allocated learning rate to it. This learning rate is later tuned using hyper parameter tuning. After this step, a 3 layered computational model is defined by using an LSTM, RepeatVector, LSTM and Timedustributed. The RepeatVector inserted in between two LSTM layers in this structure, acts a bridge between the first encoder model and the next decoder model. It adds the inputs a certain number of times for the next layer. Eventually a dense layer is applied to every sample input using TimeDistributed. The input paremeters for each of the layers are fine tuned using hyper parameter tuning as explaied in the next step.

### 4.2.7 Hyperparameter Tuning

After fitting the final model, we optimized the model using hyperparameter tuning. We applied KerasGridSearch and KerasRandomSearch on the final model in order to find the optimum values for the parameters.

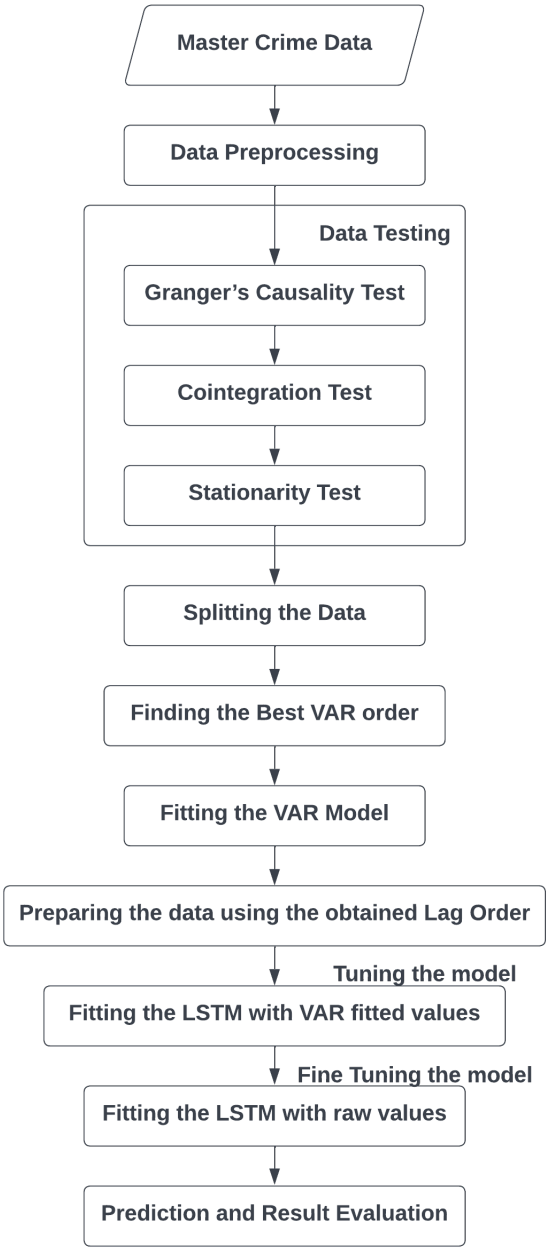


Figure 4.2: Implementation Flowchart

# Chapter 5

# Results and Discussion

The model proposed was trained on the crime data from 1960-2011 and the accuracy is tested on the remaining data from 2012-2021 (i.e. span of 10 years). Finally, we calculate the Root Mean Squared Error to quantitatively display the result. The RMSE was found to be promising for all the categories and the model proposed by us displayed better results compared to the original LSTM model. Fig - shows the visualization of RMSE(Root Mean Squared Error) given by both the models.

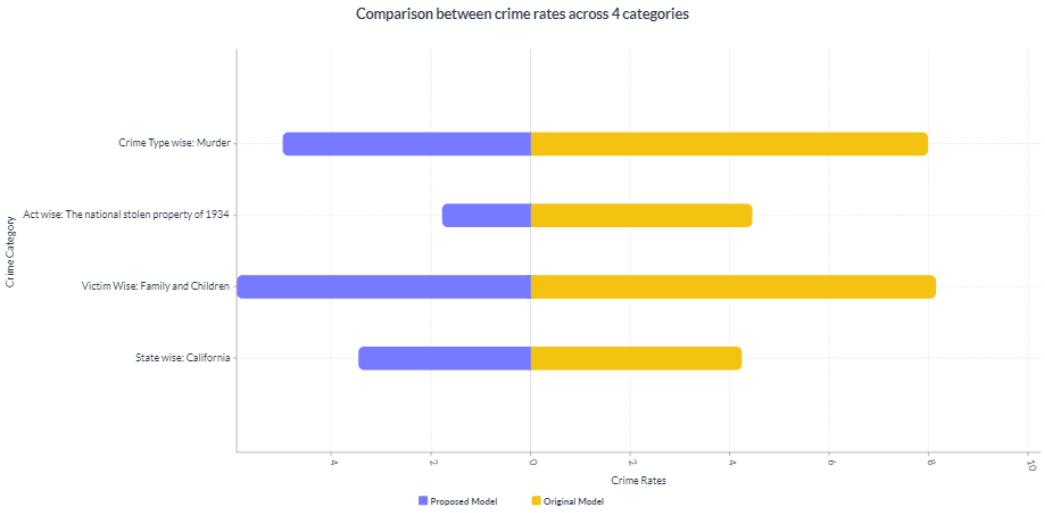


Figure 5.1: RMSE across different categories

Table 5.1: Result Table.

| Crime rate category | RMSE of proposed model | RMSE of original LSTM |
|---|---|---|
| Crime Type wise: Murder | 4.98 | 7.98 |
| Act wise: Act of 1934 | 1.78 | 4.45 |
| Victim Wise: Family and Children | 5.90 | 8.14 |
| State wise: California | 3.46 | 4.24 |

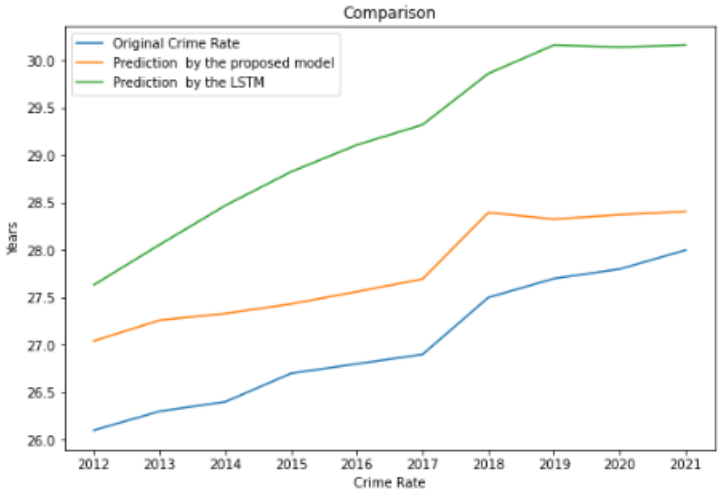


Figure 5.2: Crime rate predictions across the the category 'Crimes against the national stolen property act of 1934'

observe from Figure [5.2], that the crime rates observed against 'The national stolen property of 1934' do not change drastically over the consecutive years and hence the time series is somewhat smooth. Thus, we infer from the graph that VAR was able to deduce the linear relationship between the data points and in turned provided much better results than the original LSTM model.

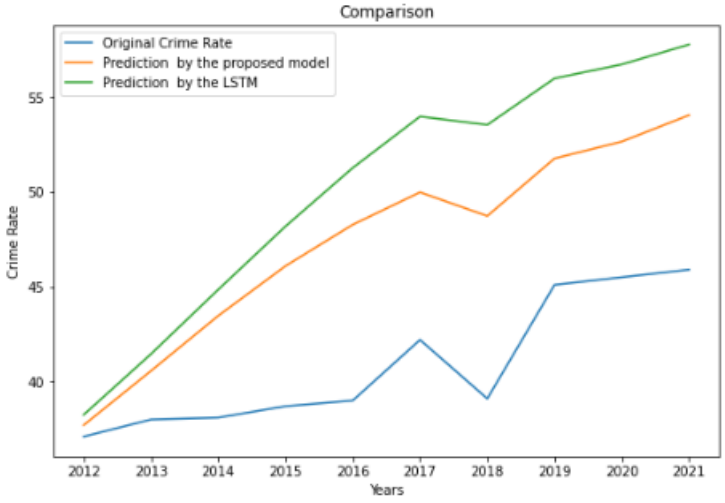


Figure 5.3: Crime rate predictions across the the category 'Victim: Family and Children'

As per figure [5.3], In case of crime rates against 'victim: Family and Children', the predictions given by the proposed model follow almost same trend as the original series. The huge difference in the predictions of initial years may be due to the low value of look back. Moreover, towards the end of test data (i.e. year 2021), the predicted values goes up by a higher margin than the original date. This may be

attributed to the spike in inflation and poverty and a disproportionate increase in the crime rates against the particular victim category.

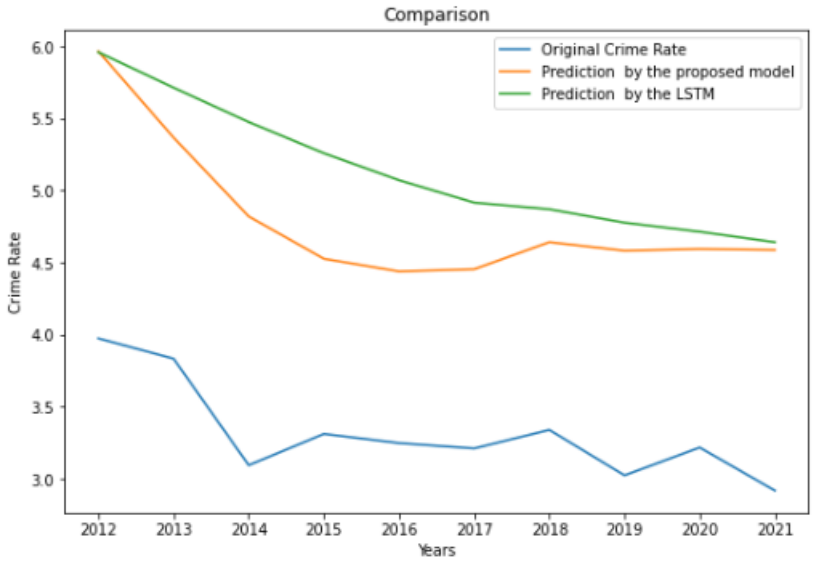


Figure 5.4: Crime rate predictions across the the California state

We can observe from figure [5.4], that California somewhat follows the trend in crime rate over the years. However, the difference between predicted crime rate and original crime rate for the years 2020-21 is too high. We learned that California is one of only 15 states in the USA that saw a dip in violent crime rate even during the covid situation. On the other hand, due to rise in factors such as Unemployment rate, poverty, inflation, the proposed model predicts a higher rate of crimes.

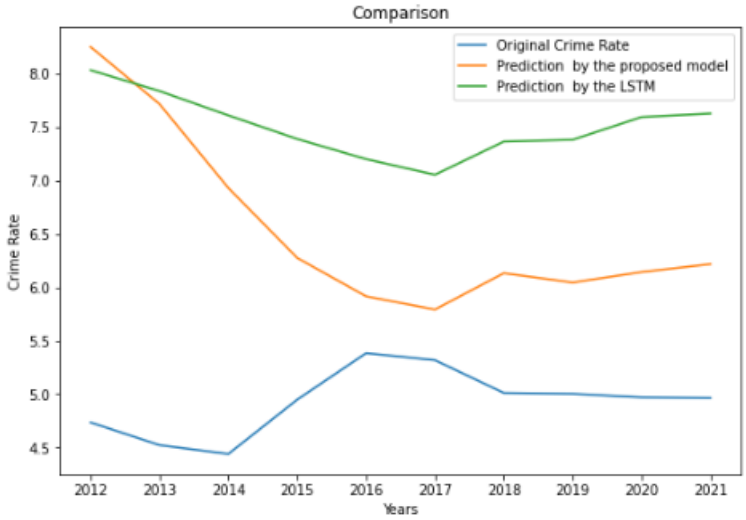


Figure 5.5: Crime rate predictions across the category crime type: 'Murder'

From the figure [5.5], we understand that in this crime category, the initial difference between the predicted values and original values is high. Although, due to feedback

mechanism, the error keeps decreasing and the values move closer to the original crime rate.

The relationships between attributes such as crime rate, tourism rate, inflation tend to be complex and the inadequacy in the data points makes it difficult to deduce the relationships. Hence, The main idea behind the combination of VAR and LSTM model is that VAR is able to filter and study patterns on training data of the time series and exploit these relationships for predicting the future values using LSTM. VAR model is able to apprehend intertwined dynamics of a time series. Feedback taking place between multiple variables enhances understanding in their relationships. Recurrent neural networks have short term memory due to repeated use of the recurrent weight matrix. This leads to a loss in information(vanishing gradient problem). However, in an LSTM model, an identity function replaces the recurrent weight matrix and is controlled by series of gates which in turn creates a long term memory function. Moreover, although VAR is known to particularly identify linear relationships between attributes, LSTM is capable of identifying non-linear relationships due to usage of different gates. All these factors contributed to the accuracy of the model.

# Chapter 6

# Conclusions

An estimate of crime rate is crucial for the governing bodies to ensure enough resources to be able to tackle the situation. A few methods are already in place for identification in patterns of the crimes. We performed extensive literature survey to understand and work on existing gaps and proposed our model (an intelligent blend of Artificial Intelligence models for time series analysis). We exploited the characteristics of two models to predict the crime rates and optimized them using hyper-parameter tuning. With our model, we were able to achieve the minimum Root Mean Squared Error of 1.78 even with the data points of only 50 years for training purposes. This research may act as an accession in this problem domain. The crime rates predicted across victim categories may help the authority provide special security and invigilance to the most vulnerable categories. Additionally, the crime rates committed under specific law can assist the judicial authorities in making necessary amendments to the laws. In a nutshell, this research will help the officials take required measures to ensure public safety.

# Chapter 7

# Future Scope

For the purpose of this study, we have focused on predicting crime rates across 4 categories. However, the model can be extrapolated to predict the crime rates across more categories. Along with the usage of Vector Auto-Regressor, an ensemble of statistical models and recurrent neural networks can be created to further improvise the model. Crime hotspots can be created using the predicted crime rates.

# Chapter 8

# Research Publication

<Research Publication of Your Project should be here>

# Bibliography


[1] https://www.legalmatch.com/law-library/article/what-are-the-different-types-of-crimes.html

[2] S. Agarwal, L. Yadav and M. K. Thakur, "Crime Prediction Based on Statistical Models," 2018 Eleventh International Conference on Contemporary Computing (IC3), 2018, pp. 1-3, Noida, India, doi: 10.1109/IC3.2018.8530548.

[3] Peter M Visser, Marcel A Groothuis and Jan F Broenink, 'Multi-disciplinary design support using Hardware-in-the loop simulation ,' *Proceedings of the* $\mathbf{5^t h}$ *Progress Symposium on Embedded Systems*, October 20, 2004, NBC Nieuwegein, NL pp 206-213.

[4] Qing Li, and Caroline Yao: 'Real-time concepts for embedded systems,' *CMP Books.*

[5] M.C. Chandorkar, Deepakraj M, 'Control of Parallel Connected Inverters in Standalone ac Supply Systems,' *IEEE Transactions Industry Applications.* vol. 29, no. 1, pp. 136-143, March 1993.

[6] Ram Kelkar, Ronnie A. Wunderlich and Leonard J. Hitchcock,'Software breadboards for power electronics circuits,' *IEEE Transactions on Power Electronics* vol. 6, no. 2 pp. 170-179, April 1991.

[7] Granger, Clive. (2004). Time Series Analysis, Cointegration, and Applications. American Economic Review. 94. 421-425. 10.1257/0002828041464669.

[8] https://faculty.washington.edu/ezivot/econ584/notes/varModels.pdf

[9] Rainer Kraus and Hans Jurgen Mattausch, 'Status and Trends of power semiconductor device models for circuit simulation,' *IEEE Transactions on Power Electronics* vol. 13, no. 3, pp. 452-466, May 1998.

[10] https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python/

[11] https://towardsdatascience.com/vector-autoregressive-for-forecasting-time-series-a60e6f168c70

[12] Priyamvada and R. Wadhvani, "Review on various models for time series forecasting," 2017 International Conference on Inventive Computing and Informatics (ICICI), 2017, pp. 405-410, doi: 10.1109/ICICI.2017.8365383.

[13] L. Xiong and Y. Lu, "Hybrid ARIMA-BPNN model for time series prediction of the Chinese stock market", 2017 3rd International Conference on Information Management (ICIM), Chengdu, 2017, pp. 93-97.

[14] Sathyadevan, Shiju S., Devan Gangadharan, Surya. (2014). Crime Analysis and Prediction Using Data Mining. 10.1109/CNSC.2014.6906719.

[15] Ch. Mahendra , G. Nani Babu , G. Balu Nitin Chandra , A. Avinash , Y. Aditya, 2020. CRIME RATE PREDICTION. ISSN NO:0377-9254.