Special Issue - 2020

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCAIT - 2020 Conference Proceedings

Forecasting India's Crime Rate

M Sushma Reddy1, Mythri K P1,
Supriya S1, Muppala Manasa1
1Department of CSE,
Cambridge Institute of Technology,
Bengaluru, India

Vimala Devi J2,
2Department of CSE,
Cambridge Institute of Technology,
Bengaluru, India

Abstract - **Because of advancement in technologies like data analytics and criminal sciences, crime forecasting is made possible. Crime forecasting plays a major role in making decisions and planning strategies to reduce the number of crimes. Our main aim is to compare the forecasting models which are "time series model-ARIMA (Auto-Regressive Integrated Moving Average)" and "regression model-Linear Regression". The data we are using for this research has crime count value from 2001-2013. After the execution, time series model's accuracy rate is around 96% and the other is about 60% from which we can conclude that former is better than the latter.**

Keywords - Crime Forecasting, ARIMA, Linear Regression.

I. INTRODUCTION

Recent reports show that India has high growth in its population than other countries. Population in India is about 1.36 billion. As the report says that India has high illiteracy rate, high population density, low job opportunities, these have become one of the reasons for high crime rate in India. Due to different problems in every state, different states have different crime rate. From 2019 statistics of crime rate in India, Delhi has high crime count per one lakh population compared to others whereas Tamil Nadu holds the second position in high crime rate. So this has been a major worry for the government in India. NCRB report shows that recently urban areas are facing high crime rates [1]. The statistics also show that India's crime rate is increasing consistently for the past 10 years.

So in today's world security has been a major concerns and issue. It has been the top priority of every system to protest the individuals and their rights. By analyzing the crimes of previous years and forecasting for future years, we can take preventive measures to reduce crime rate.

ARIMA model has been productively employed to forecast economic, marketing, production and others since it was initiated by Box and Jenkins [2]. ARIMA has proved that it is one of the best tools for forecasting crime and provides appropriate discernment into various depending factors of the time series data [15]. Among the data mining tools, Linear Regression is assumed to be one of the best tools for crime forecasting [17]. In this research, we forecast crime using two different models and compare them to find which might be better. The models are Linear Regression model and ARIMA model.

II. RELATED WORK

This section of work defines the techniques which are carried out to forecast crime rates during previous years. Now-a-days, making short-term forecasting crime using time series model is a new field for researching [15].

In this paper they did a comparative study among the models ARIMA, HES and SES which forecast the crime data of China, where they concluded that ARIMA is the best fitting model compared to others [4]. This research included a crime forecasting using BoxJenkin's arima model to predict the currency imitation in Gujarath state [3]. Maintaining the Integrity of the Specifications.

This study presented crime forecasting using fuzzy alpha-cut along with ARIMA model. They concluded that this combination generates more précised forecasting values [6]. This study proposed crime forecasting using time series models like ARIMA and exponential smoothing, where they proved that time series models are best for crime forecasting [15]. This study presented a study where he compared Box-Jenkins method with Regression on sales data, in which he concludes that Regression is better than Box-Jenkin [16]. This paper has a study on crime forecasting using Linear Regression for the crime data of Bangladesh [17].

III. METHODOLOGY

As we talked earlier, we used two methodologies which are discussed below as following,

A. Box-Jenkins Model

Box-Jenkins method's basic idea is described in the Fig 1. Box-Jenkins model is a three stage approach [8]. The stages are as following

- Identification: In this first stage, we use all the data to help select a model that summarizes the data to its best.

- Estimation: In this phase, we use the data to train the model's parameters.

- Diagnostic Checking: In this last stage, fitted model is evaluated using data to check for the parts that may be enhanced to get a best fitted model.

If the data set has following conditions, then that data set is good for forecasting,

- The residuals will be uncorrelated.
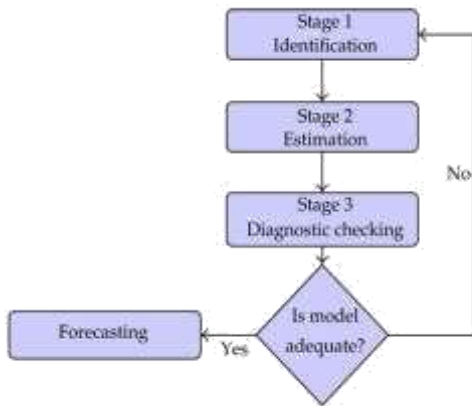
- The residuals will have zero mean.

Special Issue - 2020

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCAIT - 2020 Conference Proceedings

Fig. 1. Box Jenkins Method

"Auto Regression Integrated Moving Average (ARIMA) model" is applied when the data is nonstationary, so the initial step of differencing is applied to eliminate non-stationarity. In this we used Seasonal ARIMA model, which is denoted by the ARIMA $(p, d, q)(P, D, Q)m$

p is number of time lags of autoregressive

model d is degree of differencing

q is an order of moving average term m

is number of periods in each season

If the data series is stationary then d=o. While it will be non stationary for d>0, then we have to do differencing to convert it into stationary. I value is defined by order of difference, if it is 0 then time series data is stationary.

B. Linear Regression

According this regression method, dependent variable is always associated with some of the independent variables. So this method uses this association and forecast the values. This regression approach is shown in the following Fig 2,
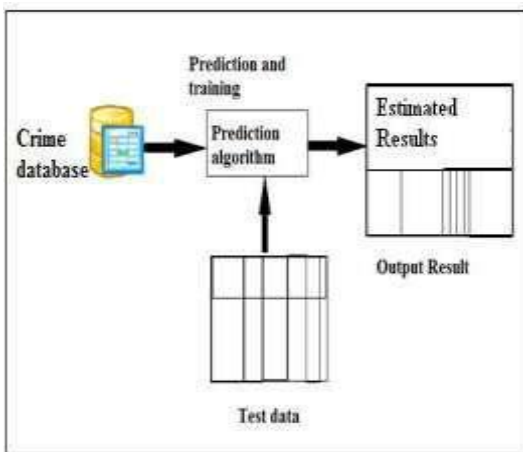

Fig. 2. Regression model

This regression approaches linearly to model the association between the independent and dependent variables. This model gives the enough information about how the "output variable" is affected by the "input variable". "Variable Y (target variable)" is predicted by a linear function of another "variable X (input variable)", given m training examples of the form "$(x1,y1), (x2,y2), …, (xm,ym)$, where xi ε X and yi ε Y".

IV. RESULTS

This section shows the outcomes that are obtained from the two methodologies as we discussed before. They are as the following,
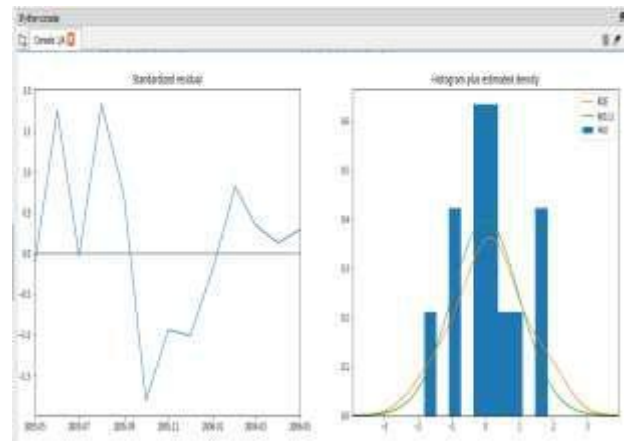
A. ARIMA Model

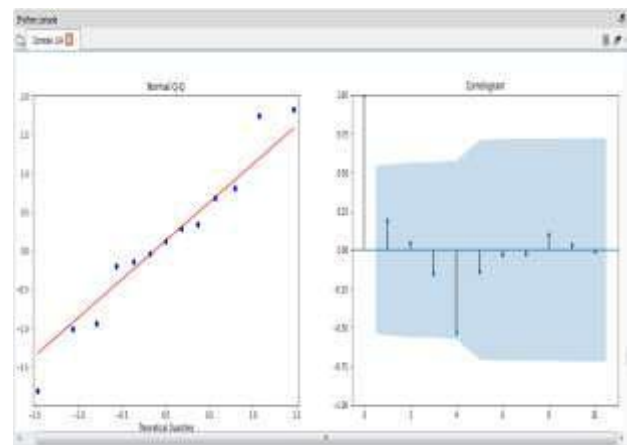Fig. 3. ARIMA Model's Diagnostic series

Fig. 4. ARIMA Model's Diagnostic series

Special Issue - 2020

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
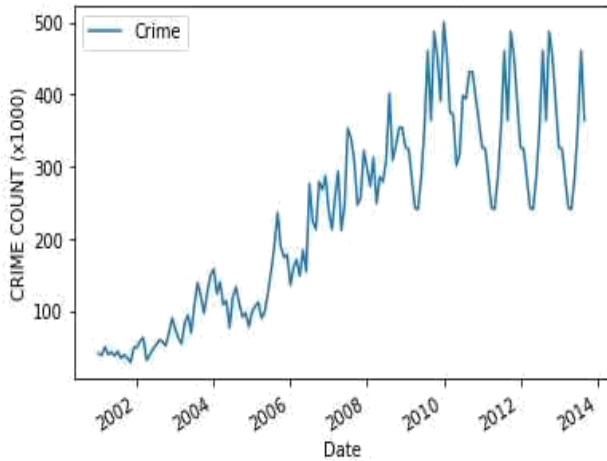**NCAIT - 2020 Conference Proceedings**

Fig. 5. Forecasting results using ARIMA

B. Regression Method

The below figure Figure 5 shows forecasted crime rate values of future years which is implemented using Linear Regression.
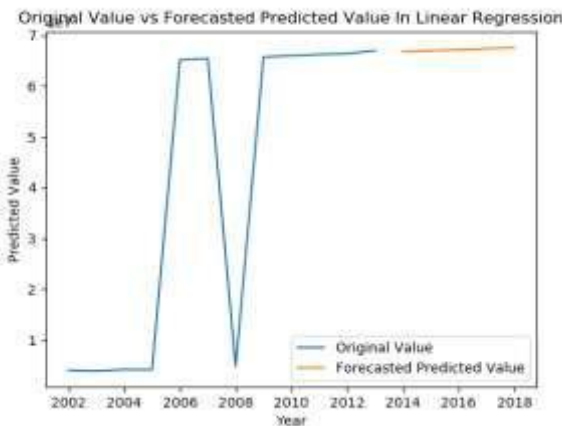

Fig. 6. Forecast results using Linear Regression

V. CONCLUSION

As we know that there is high crime rate in India, so to take preventive measures crime forecasting helps the officials. In this we did a comparision between two models i.e., Box-Jenkin and regression which showed the accuracy values of about 96% and 60% respectively. According to our research, we can conclude that arima is better than linear regression. In future we will try to create optimized models for any specific crime.

REFERENCES

[1] Web portal of NCRB http://ncrb.gov.in/

[2] Box, George EP, and David A. Pierce, "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," Journal of the American statistical Association 65.332, 1970, pp. 1509-1526.

[3] Shrivastav, and Anand Kumar, "Applicability of Box Jenkins ARIMA model in crime forecasting: A case study of counterfeiting in Gujarat state," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) 1.4, 2012, pp. 494.

[4] Chen, Peng, Hongyong Yuan, and Xueming Shu, "Forecasting crime using the arima model," Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08, Fifth International Conference on Vol. 5, IEEE, 2008

[5] Rattner, and Arye, "Social indicators and crime rate forecasting," Social Indicators Research 22.1 ,1990, pp. 83-95

[6] Noor Maizura Mohamad Noor, Astari Retnowardhani, Mohd Lazim Abd and Md Yazid Mohd Sanam, "Crime forecasting using ARIMA model and fuzzy alpha-cut," Journal of Applied Sciences 13.1, 2013,pp. 167-172.

[7] Groff, Elizabeth R., and Nancy G. La Vigne, "Forecasting the future of predictive crime mapping," Crime Prevention Studies 13, 2002, pp. 29-58.

[8] Williams, Billy, Priya Durvasula, and Donald Brown, "Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models," Transportation Research Record: Journal of the Transportation Research Board 1644, 1998, pp. 132-141.

[9] Gorr, Wilpen, Andreas Olligschlaeger, and Yvonne Thompso,. "Shortterm forecasting of crime," International Journal of Forecasting 19.4, 2003, pp. 579-594.

[10] Flaxman, and Seth, "A General Approach to Prediction and Forecasting Crime Rates with Gaussian Processes," Heinz College Technical Report, 2014.

[11] Gorr, Wilpen, Andreas Olligschlaeger, and Yvonne Thompson, "Assessment of crime forecasting accuracy for deployment of police," International Journal of Forecasting, 2000, pp. 743-754

[12] Razana Alwee, Siti Mariyam Shamsuddin and Roselina Sallehuddinl, "Hybrid support vector regression and autoregressive integrated moving average models improved by particle swarm optimization for property crime rates forecasting with economic indicators," The Scientific World Journal ,2013.

[13] Mutangi, Kudakwashe, "Time Series Analysis of Road Traffic Accidents in Zimbabwe," International Journal of Statistics and Applications 5.4 , 2015, pp. 141-149.

[14] Manish Kumar, Athulya S, Mary Minu MB, Vidya Vinodini M D, Aiswaria Lakshmi K G, Anjana S, and Manojkumar TK, "Forecasting of Annual Crime Rate in India: A case study," IEEE, 2018.

[15] S Nanda, "Forecasting: Does the Box-Jenkins Method Work Better than Regression?" sage journals,1988.

[16] Md. Abdul Awal, Jakaria Rabbi, Sk Imran Hossain, and M M M Hashem, "Using Linear Regression to forecast future trends in crime of Bangladesh," IEEE, 2016.

[17] Priyanka Gehra, Dr Rajan Vohra, "Predicting Future Trends in City Crime Using Linear Regression," IJCSMS, 2014.