

Foresee Crime in Society Using Advanced ML Techniques

SUPREETHA S¹, SNEHA SANJAY GHATTADAHALLI², SINDHU R³, SHREYA P NAIK⁴,
KAMLESHWAR KUMAR YADAV⁵

^{1, 2, 3, 4, 5} Computer Science Engineering, Global Academy of Technology, Bengaluru

Abstract— Crime is considered as an immediate infringement and danger to the general public also, ought to be controlled. The exact expectation and anticipating of crime is very crucial because these are the two factors when combined together would lead the enhancement of the safety of many metropolitan cities. Crime prediction and future forecasting are interconnected because you cannot predict the future forecasting trends without accurately predicting crime. When it comes the human brain although capable of many amazing marvels it is not built in such a way where it can process large and complex data and this would lead to many computational challenges along with the hinders of the early and accurate prediction. The current manual methods that are used by the police department leads to a very low success rate and hinders the early accuracy rate of crime prediction. Through multiple research efforts a few of the predictive algorithms have been picked to better predict crime and give the police direct access to the future trends that may occur in the future. All the studies that have conducted previously show that there is a big lacking factor that exists in achieving accurate prediction of crime based on the use of learning models. So therefore, in this study all of the following algorithms have been applied on the given dataset in order to not just accurately predict crime but also effectively do so. The algorithms that are applied are as follows Naïve Bayes, k-nearest neighbours(KNN), random forest and eXtreme Gradient Boosting(XGBoost), multilayer perceptron(MLP), autoregressive integrated moving average(ARIMA),logistic regression, support vector machine (SVM), and time series analysis by long-short term memory(LSTM) model to more readily fit the wrongdoing information. By using all these algorithms there can be accurate results that can generated which can be used by the police

department to have a better eye on the types of crimes ,crime rates and the vigilantes that would be responsible.

Indexed Terms-- ARIMA , LSTM , Forecast ,Predict ,Crime.

I. INTRODUCTION

Crime is an event that occurs around the world regardless of if the country is developed or under developed[1]. These criminal activities have a very negative effect on

the economy as well as hindering the well-being of people which in turn leads to various forms of social issues. The quality of life would also be effected if criminal activities occur on a daily basis. The occurring crimes also come with its own cost factor which would be put on both the public and the private sectors once again leading to negative impact of the economic well-being. Crime will also directly affect the public safety specifically when considering of the fact of people shifting to new locations and travelling. And each criminal activity would have its own associated consequence. Overall crimes take place due to many different and distinct factors such as critical situations, human behavior and poverty. And cities that are growing and highly populated will see higher crime rates due to multiple factors and different types of social and economical environments. If a city wants to become suitable in a social manner then their aim would be to decrease the rate of crime so that they can have their residents live a peaceful lifestyle and avoid corrupt societies. So taking all these factors into consideration the forecasting and the prediction of crime has gained a lot of popularity in the recent years. Using various machine learning algorithms police can foresee the future crime trends [2]. These algorithms work in a manner where there would a computerized

process that would occur of the analyzing of the previous crime reports and showing police any direct links and ties. This would also help police to see the various hot spots for crime and predict with better and accurate results leading to a safer and secure society and well-being of residents in any given city.

II. LITERATURE SURVEY

Crime prediction has several factors that effect it like the education status, the level of income and the unemployment rate. In this day and age social media also plays a huge role in accurately analyzing and predicting crime. The prediction of crime is done on the basis of focusing on the cause of the crime, once the cause has been found then the pertaining consequences will follow. There are multiple implementation techniques that have been used to acquire a higher accuracy rate. In this literature review section the discussion would be specifically based on the data that is given by the Boston dataset. This dataset will be analyzed to predict the future crime rate along with it trying to reveal the information about the upcoming crime trends. The sole purpose of this section is to bring light to the classification, prediction and the forecasting of the current crime. The results that would be revealed would be solely based on the data that has been collected from the city of Boston. The general expectation rate relies straightforwardly upon the data that is given inside the dataset that is utilized for the forecast reason. Boston is a populated and is a crowded famous city of the U.S. what's more, its dataset is accessible freely at approved stores, relating different characteristics that have been an incredible wellspring of fascination for examiners. With a particular objective to the short, there have been various examinations lately founded on these datasets to foresee exactness and areas of interest wrongdoing locales by applying numerous AI calculations, and sorts of assumption achieved. A portion of the new examinations on the two urban communities are summed up beneath. Fig 2.1 system architecture shows the architecture for the wrongdoing expectation and gauging system will take the crime dataset as input, then it will pre-process the dataset and extract the features and apply the ML algorithms to predict the crimes and apply the ARIMA forecasting model to forecast the crime.

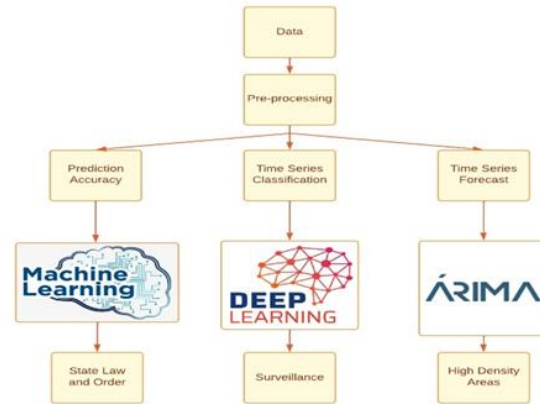


Fig 2.1 : System Architecture

III. PREDICTION AND FORECASTING

Approaches to crime prediction and forecasting has changed emphatically as of late starting from the presentation of business programming bundles. crime prediction alludes to the precision of revealed violations before, while forecasting towards the future crime patterns. Notwithstanding, a speedy outline of crimes has been accomplished by examination specialists through the accessible programming bundles, while for profound examination, just learning approaches may guarantee the ideal arrangement. Accordingly, unique machine learning procedures can be utilized to anticipate wrongdoing designs and along these lines might aid further fundamental activities in light of authentic information. Consequently, This research is split into two sections: 1) prognosis and 2) crime prediction. There are eight distinctive AI calculations are executed to accomplish exceptionally precise prediction in Boston dataset. To obtain the crime prediction accuracy We employed machine learning methods including Decision Trees and Random Forests, XGBoost, Logistic Regression, Naive Bayes, MLP, SVM, and KNN are all examples of machine learning techniques. Point by point data about these AI calculations models engineering is given in the valuable data and the prediction results similarly pick out regions Given a high crime frequency, all sorts of crime, and a recent crime rate In addition, the ARIMA experimental model for data set investigations was utilised to anticipate future crime patterns and investigation. crime estimating in view of time series information was likewise carried out in a later piece of this review.

A period series investigation includes gauging in view of a grouping of occasions or information guides that shapes a series with deference toward time[4]. Research bunches all over the planet have as of late utilized various methodologies, including unaided models, for example, the The The bilinear system, the autoregressive conditional heteroskedastic (ARCH) system, the threshold autoregressive (tar) system and the deep learning model are all examples of such models technique were all used distinguish future patterns. ongoing wrongdoing gauging is dependably basic, particularly in obscure conditions; when and where the following wrongdoing will happen stays hard to foresee precisely. Along these lines, we utilized an ARIMA model for future determining and determined the RMSE to total the extents of the blunders also, wrongdoing forecasts. The subtleties of the ARIMA model are talked about in the SI. The anticipating results show future crime patterns by featuring the crime problem areas, top five crime and generally crime percentages until 2024.light of authentic information. Consequently, This research is split into two sections: 1) prognosis and 2) crime prediction. There are eight distinct AI computations done. accomplish exceptionally precise prediction in Boston dataset. To obtain the crime prediction accuracy, we used machine learning algorithms such as Random forest, Decision tree, Logistic regression, MLP, Nave Bayes, SVM, XGBoost, and KNN are all examples of machine learning techniques. Point by point data about these AI calculations models engineering is given in the valuable data and the prediction results similarly pick out regions The bilinear model, the threshold autoregressive (tar) model, the autoregressive conditional inhomogeneous variance (ARCH) model, and deep learning are all examples of statistical method patterns and investigation. crime estimating in view of time series information was likewise carried out in a later piece of this review. A period series investigation includes gauging in view of a grouping of occasions or information guides that shapes a series with deference toward time[5]. Research bunches all over the planet have as of late utilized various methodologies, including unaided models, for example, the The Bilinear model, threshold autoregressive (tar) model, autoregressive conditional heteroskedastic (ARCH) model, and deep learning technique were all used distinguish future patterns. ongoing wrongdoing

gauging is dependably basic. particularly in obscure conditions. when and where the following wrongdoing will happen stays hard to foresee precisely. Along these lines, we utilized an ARIMA framework for future judgments and RMSE decisions total the extents of the blunders also, wrongdoing forecasts. The ARIMA model's complexity is discussed in SI. The forecast findings emphasise the violent area, the top five crimes, and frequent offences to demonstrate future crime tendencies percentages till 2024.

IV. DATA PREPROCESSING

For this project we have used dataset of the city Boston. We have obtained dataset from the Kaggle site which is in csv format. our goal is to achieve good crime predictive accuracy. Boston's dataset contains various crimes that was committed in the city[6]. Boston, formally the Boston is the capital of Massachusetts, the most populous city in the state, and the 24th most populated city in the United States. As a result, the situation causes police personnel to change their tactics, resulting in a downward trend in recent years. With 22 variables and over 7 million cases, the newly released dataset includes extensive data on time, area (i.e., latitude and longitude), and crime kinds. From 1999 to 2018, a visual representation of the percentage and metrics of violent crime in Boston, Massachusetts. The FBI's Uniform Crime Reporting (UCR) Program divides severe crime into four categories: murder/murder and nonnegligent homicide, assault, burglary, and disturbed attack. Violations are classified as severe In the UCR Program, these transgressions are defined as those involving power or the threat of authority. The crime rate in Boston MA in 2018 was 622.45 per 100,000 people, a 6.99 percent decrease over 2017. The 2017 crime rate in Boston MA was 669.2 per 100,000 people, a 5.4 percent decrease from 2016. The 2016 crime rate in Boston MA was 707.4 per 100,000 people, a 0.09 percent increase over 2015. The crime rate in Boston, MA in 2015 was 706.79 per 100,000 people, a 2.6 percent decrease from 2014. The 2019 crime percentage in Boston, MA is 258 (City-Data.com wrongdoing record), which is tantamount to the U.S. normal. It was higher than in 78.3% U.S. urban areas. The 2019 Boston crime percentage fell by 6% contrasted with 2018. The quantity of murders remained at 42 a diminishing of 14 contrasted with

2018. Over the most recent 5 years Boston has seen decline of vicious wrongdoing and decline of vandalism. in 2019 violent crime rate was : 329.1 in 2018 violent crime rate was : 343.3 in 2017 violent crime rate was : 368.4 in 2016 violent crime rate was : 385.2 in 2015 violent crime rate was : 380.5 in 2014 violent crime rate was : 396.4 in 2013 violent crime rate was : 421.9 in 2012 violent crime rate was : 450.8 in 2011 violent crime rate was : 459.6 In 2010 violent crime rate was : 488.6. In 2009 violent crime rate was : 529.5. In 2008 violent crime rate was : 587.2. In 2007 violent crime rate was: 616.3. In 2006 violent crime rate was: 714.2. In 2005 violent crime rate was: 701.7. In 2004 violent crime rate was: 501.6. In 2003 violent crime rate was: 505.6. In 2002 violent crime rate was : 496.8. In property crime rate in 2018: 157.4. In property crime rate in 2017: 168.4. In property crime rate in 2016: 166.5. In property crime rate in 2015: 188.2. In property crime rate in 2013: 235.8. In property crime rate in 2012: 250.3. In property crime rate in 2011: 271.1. In property crime rate in 2010: 275.3. In property crime rate in 2009: 273.8. In property crime rate in 2008: 312.0. In property crime rate in 2007: 362.0. In property crime rate in 2006 is 405.5. In property crime rate in 2005 is 424.9. In property crime rate in 2004 is 356.3. In property crime rate in 2003 is 360.4. The amassed crude information from online archives ordinarily contains immaterial data and blunders.

Fig 4.1 depicts an graphical representation of the Top ten crimes that have been observed in the city of Boston seen in the Boston Dataset. The general information will probably have clamor, irregularities, anomalies, fumbles, and missing characteristics or, all the more essentially, information is conflicting to begin technique. Subsequently, the choice of significant information is important to dispose of inconsistencies against the anomalies, commotion, missing qualities, and different inconsistencies, and consequently change over the impractical information into conceivable is reasonable to achieve data dealing with. Furthermore, assortment for the more mind-publishing content to a blog structure is constantly required keeping in view the ongoing creating pace of information in business, industry applications, science, and examination organization. The information pre-processing solidifies information arranging, exacerbated by blend, cleaning, systematization, and change of information; information decline tasks, subsequently lessening the diverse plan of the information, seeing or ousting unessential and loud parts from the information through component affirmation, event decision, or discretization casings, and consequently finally helps to create genuinely significant information to make precise wrongdoing expectations. Hence, the bootstrap arbitrary testing technique an over highlight choice strategy, which is likewise normal since it is the most un-one-sided strategy to create appraisals of populace boundaries; explicitly when the dataset is large. At first the datasets were inspected from various sources and to take normal credits.

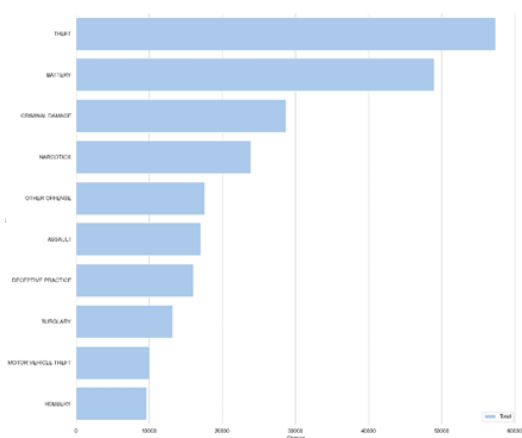


Fig 4.1 Top Ten Crimes



Fig 4.2 Activity Diagram

Fig 4.2 shows activity diagram which is graphical representation A process overview that includes step-by-step activities and actions that enable selection interaction and concurrency. Activity diagrams in the Unified Modeling Language are meant to depict data flows that overlap both computational and organisational processes, as well as associated activities. Although activity diagrams generally depict the overall flow of control, you may also include features that depict the movement of data between one or more datastores. Altogether, there are 9 normal credits in both datasets, and information cleaning was guaranteed by eliminating every missing worth. For execution, Python (rendition 3.6) structure was utilized with various Most libraries for information changes, such as B.imblearn and sklearn. The last contributions evaluated in this investigation were identified as ID dates, wrongdoing essential so rt, depiction of the wrongdoing, area, year, postal division and police locale. As a result, theThe data is separated into a test set (30%) and a training set (70%) (30 percent). 70% of the time. There were 106550 occurrences of Boston after the preprocessing step [7]. The primary performance measures in this study are accuracy, precision, recall, and f1 score. Fig 4.2 shows Activity diagram which are a A graphical depiction of

a sequential activity and action process that allows for selection, iteration, and concurrency. Activity diagrams in the Unified Modeling Language are meant to depict data flows that overlap both computational and organisational processes, as well as associated activities. The activity diagram generally depicts the overall flow of control, but it may also include components depicting data flow across activities via one or more data storage.

V. RESULTS

A. ANTICIPATING EFFICIENCY

The particular research adapted various boundaries directed towards survey exhibition coming from different theorems, and that finer rebound a genuine data file operation. Eight distinctive principles were adopted towards the Bostan dataset into interrogating the elaborated guessing preciseness of skilled models Evidently, these estimations have not been executed together for Boston dataset Thus, the fundamental motivation to pick this city is populace thickness, which announced greater unlawful act percentages in the former times alongside enormous information. Majorily carried out calculations have various techniques to refine that information which includes supervised or the superintended, unsupervised or unrestrictedly learning approaches. carried out calculations have various techniques directed towards sharpening the information that-a-way includes supervised or the superintended, unsupervised or unrestrictedly and reinforcement or beefing up training path. In addition, Random Forest and XGBoost were also carried out in order to encourage all

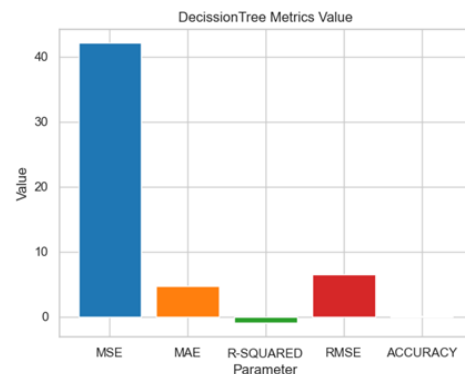


Fig 5.1 Decision Metrics Value

the overall approaches of group learning. Fig 5.1 shows a picture of Decision the design the huge choice, while SVM and Naïve Bayes are utilized in order to achieve improvised grouping as well as KNN considering extreme regression in order to oversee subordinate factors Logistic relapse is finished nearby MLP whichever recommends Multi layer among the perceptron. In conderation of this large number of numerical articulations help to look for moved along precision to the best of their capacity, with other execution assessments, for example, accuracy, review and F1-score The precision appraises the portion of occasions which may be really grouped to accomplish the best edge for wrongdoing expectation[8].

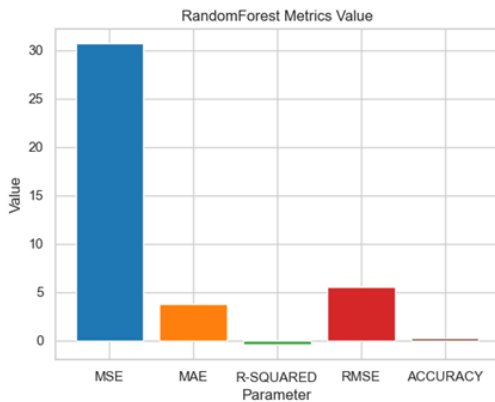


Fig 5.2 RandomForest Metrics Value

Random Forest which is an very important algorithm which is being used in this project is an Supervised Machine Learning theorem that is being utilized generally in Classification and Regression issues. This assembles Decision trees on serveral examplar and hold their greater chunk choice in appreciation of characterization and habitual if there should arise an occurrence of relapse. One of the primary highlights of the Random Forest calculation is that it can manage the enlightening file containing relentless variables as because of backslide and outright factors as by virtue of arrangement. It provided improved results upon order errors. The XGBoost calculation performs better compared to different calculations and is definitely or more precision on the Boston dataset, as many quantities of inventive calculations execute behind XGBoost. Naïve Bayes, MLP, and SVM estimations likewise accomplish unrivaled openness on the Boston informational index with near most extreme. The

decision tree which is a majorly known algorithm gets a very good and descent accuracy on the Boston data. The MLP and KNN algorithms which is very essential algorithms need have provided calculations which has reached the highest accuracy on the Boston data. The calculated relapse model picks the certifiable association between components to accomplish ideal outcomes. here, it portrays steady execution with 90% exactness on the Boston data and keeping in mind that the Naïve Bayes calculation accomplishes practically similar outcomes on the Boston. the closeness in this way genuinely relies on how as frequently as conceivable the violations happened already, and anticipating. rare crimes in the population involved can lead to low precision Regardless, SVM calculation accomplishes an ordinary outcomes[9]. classification quality is typically assessed on the exhibition of goal capacities like accuracy, review and F1-score. The review presents the applicable examples that are recovered aside classier, though the accuracy is the rate of accurately divided tests. The two limits at the same time in like manner, upgrade the two goals with a contrary relationship, but the F1-score would basically mean as weighted ordinary of review and precision. XGBoost shows further developed results for accuracy, review, F1- score than different representation.

B. TIME SERIES ANALYSIS THROUGH LSTM

LSTM a rich variety of RNN engineering, which is a methodology which could be added on to show successive information. Plan of LSTM makes it a powerful response for the fight to come the vanishing incline issue of RNNs. It uses memory prepared for tending to the long circumstances in progressive data. Former to or Before LSTM execution the records had been pre processed to lessen noise after which converted into desk bound records. Time collection facts are generally in non-desk bound shape and must be converted into desk bound shape for less difficult coping with and higher classification. Subsequently, the Dickey-Fuller Exam is directed to examine for fixed information in an accepted manner furthermore, to additionally assess fitting blunder scores. outcomes give inside and out direction from information handling and preparing the LSTM model for a few bunch of time-series information. For time-series information, various kinds of mistakes are ordinarily estimated, like the scale-subordinate blunder and rate

blunder. Thus, two known scale-subordinate mistake measures were utilized, specifically, alongside the number of epoch and batch size. Accurately, the square base of the normal of the squared contrasts between the expected and real discernments. Subsequently, it is the typical across the test of the outright contrasts between the ordinary and veritable pieces of information where each of the specific differentiations have unclear cost. The display estimations of LSTM shows the execution of relating model in testing data instead preparation information or the readiness data. There isn't any proof schooling the community with the identical dataset extra than as soon as could enhance the accuracy of the prediction. Sometimes, the exhibition even deteriorates, showing that the prepared models are overfitting. In any case, obviously setting the amount of ages to 1 delivers reasonable assumption model. exhibition of LSTM is by all accounts satisfactory for time series investigation. Recently, the average majority of mean bad behavior types show a diving design, which proposes a further reduction in the greater part of the evaluations in the general time broadens. In any case, it isn't appropriate when the generally up pattern is connected to other unlawful act and offenses. The time series characterization is possibly an immediate marker, however it can't be investigated as estimation of unique prices, byte fairly as an information driven model

C. PRECURSORY DATA ANALYSIS

This segment talks about the itemized intermittent bits of knowledge of the Boston dataset statistics. The time span wrongdoing count alludes back to the wide assortment of wrongdoing episodes, simultaneously as focused energy wrongdoing locales are the new spot wrongdoing regions with respect to the area degree area. Fig 5.3 shows the daily arrests in Boston city similarly Fig 5.4 depicts weekly arrests and Fig 5.5 depicts monthly arrests in the city of Boston. The consequences are acquired with the aid of using the use of the investigating module available in Python, wherein the unlawful act price is the crime depend standardize with aid of using population for time. The review distinguishes 8 unique wrongdoing types for Boston dataset. shows the yearly patterns for Boston, showing a huge lessening in the crime percentage. Past concentrates on Boston likewise propose that environmental factors, for example, cruel atmospheric

conditions or the colder time of year season might diminish wrongdoing and may incline toward individuals and inhabitants.

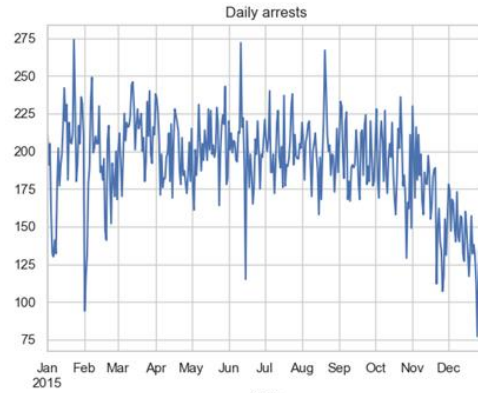


Fig 5.3 Daily Arrests

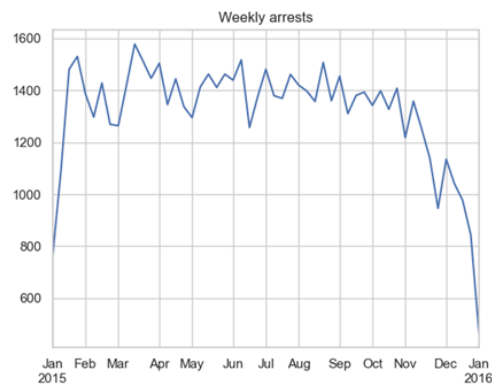


Fig 5.4 Weekly Arrests

In the year of 2019 the crime percentage or the crime amount in Boston, MA was 258 which is nearly equivalent to United States of America. normal. It was Greater upon in 78.3% United States of America. urban areas. In the year of 2019 Boston wrongdoing cost lowered down upto 6% appeared differently in relation.

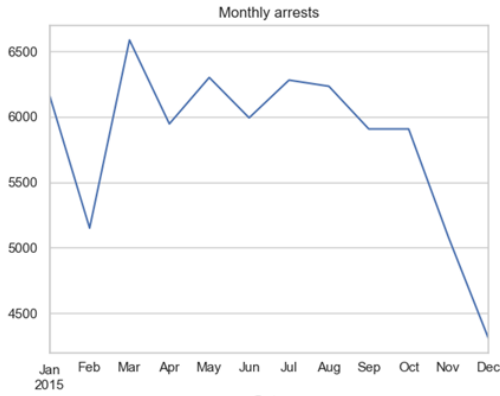


Fig 5.5 Monthly Arrests

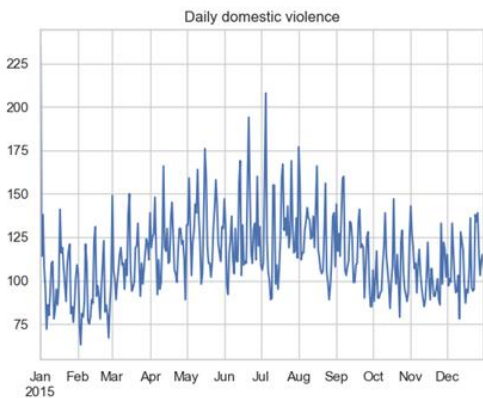


Fig 5.6 Daily Domestic Violence

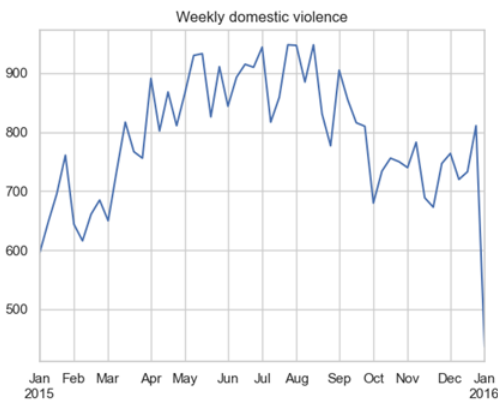


Fig 5.7 Weekly Domestic Violence

2018. The quantity of murders remained at 42 - a decline of 14 opposed upon 2018. In the latest 5 years Boston saw massive decline on fierce wrongdoing also decline of vandalism. There were 16 bad behavior regions or the areas in Boston with the most important wrongdoing rates

with extra wide encounters. Moreover, up coming crime frequency regions were likewise concentrated by utilizing an ARIMA model, It is examined in following articulation.



Fig 5.8 Monthly Domestic Violence

Fig 5.6 shows the daily domestic violence, Fig 5.7 shows weekly domestic violence and Fig 5.8 shows monthly domestic violence . The unlawful act variety and its assessed forces were considerably held greater significant in order to decide the expected opportunity of wrong-doings events. murder, assaults , burglaries , attacks , robberies , robbery , auto burglaries fire related crime were the primary violations saw in Boston . Fig 5.4 shows a picture of the Daily arrests in the Boston city.

D. FORECASTING WITH AN ARIMA MODEL

Time series determining exhibits its significance in building a successful model, particularly in the field of applied sciences. The plan of LSTM makes it a convincing solution for the fight to come the vanishing tendency issue of RNNs. It utilizes memory equipped for addressing the drawn-out conditions in successive information. LSTM guarantees further developed learning for time series by catching the construction of successive information all the more normally what's more, even performs progressive handling for complex worldly undertakings. Time series classification errands are unique from customary classification and relapse prescient displaying issues and have been viewed as trying as far as information digging throughout the previous twenty lifetimes.

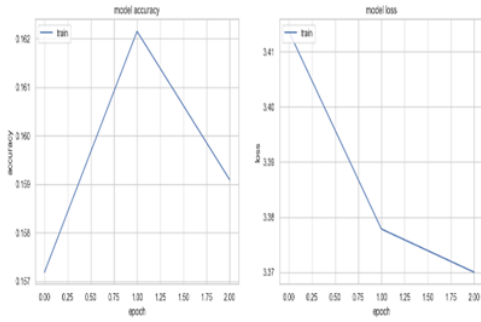


Fig 5.9 Model Accuracy and Model Loss

Fig 5.9 shows a graphical representation of the comparisons between the Model accuracy and the Model loss. From electronic prosperity records to organize insurance, for all intents and purposes generally certified world applications request time-series data for game plan[10].A point by point. Portrayal of LSTM is given in the SI. Preceding LSTM execution, the information were preprocessed to diminish commotion and afterward changed into fixed information. Time series information are as a rule in non-fixed structure and must be changed into fixed structure for more straightforward taking care of and better grouping. Hence, the Dickey-Fuller test is directed to check for fixed information in a standard manner what's more, to additionally assess the proper mistake scores.

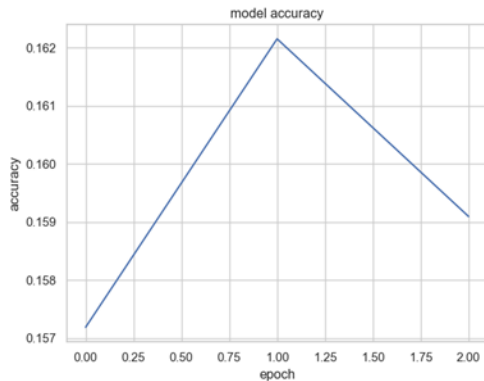


Fig 5.10 Model Accuracy

Fig 5.10 depicts a picture of Model Accuracy in which a graph has been plotted for epoch with respect to Accuracy of the model. The results give inside and out direction from information handling and preparing of the LSTM model for a bunch of time-series information. For time-series information, various

kinds of mistakes are typically estimated, like the scale-subordinate mistake and rate mistake. There is no proof preparation the organization with the same dataset at least a few times would work on the exactness of the expectation. Every so often, the show even disintegrates, exhibiting that the pre-arranged models are overfitting. In any case, evidently setting the quantity of ages to 1 produces a sensible expectation model. The different frequencies integrate the ordinary, step by step, month to month, quarterly, and yearly results. The mean wrongdoing thickness region for Chicago has a serious variety diagram basically in everyday and week after week information, though the month to month and quarterly information have moderate variety plans. The general interaction includes fostering a capacity that ascertains and presents the moving normal of the occasions in the neighborhood of the occasions. As of late, most of mean wrongdoing types show a descending plans, which recommends a further decrease in most of figures in the general time stretches. Be that as it may, it is not relevant when the generally up diagram is connected to other criminal offenses. Fig 5.11 depicts a picture of ARIMA Metrics Value The time series order is potentially a prompt pointer, but it can't be treated as an evaluation of specific values, but it is rather as an information driven model.

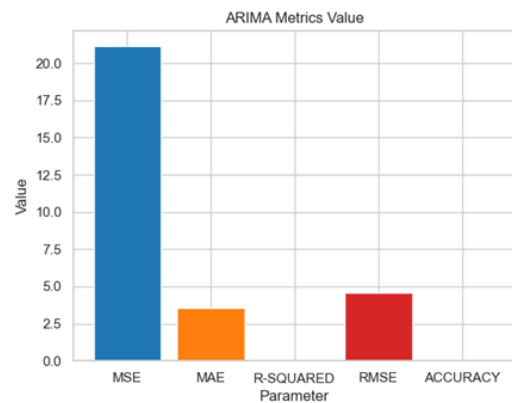


Fig 5.11 ARIMA Metrics Value

VI. CONCLUSION

Crimes cause hovoc in society. Police department works hard to control crime. In order to help the police authorities predictive systems is needed. This helps in improving crime analytics & provides safety and

security to cities by preventing crimes. We have achieved good predictive accuracy by using various machine learning algorithms on Boston dataset. We performed data pre-processing then we divide the dataset into testing set and training set, afterwards we examined the performance parameters. We made use of LSTM algorithm for time series analysis. Exploratory data analysis showed extensive visualizations with respect to crime particulars which includes crime rates in various periods and crime types. We implemented ARIMA model to predict trends with respect to crime rate and hotspots. For future work we will expand this study using imagery data from satellites and various learning techniques can be implemented for visual data of various crime datasets.

VII. FUTURE ENHANCEMENTS

We have executed this crime prediction making use of ARIMA and LSTM machine learning algorithms. In this project we have further utilised a deep learning algorithm for time series analysis, we have analysed crime in Boston city using RMSE and these parameters can be enhanced.

REFERENCES

- [1] G. Mohler, "Marked point process hotspot maps for homicide and gun crime prediction in Chicago," *Int. J. Forecasting*, vol. 30, no. 3, pp. 491–497, Jul. 2014.
- [2] A. Iriberry and G. Leroy, "Natural language processing and e-government: Extracting reusable crime report information," in *Proc. IEEE Int. Conf. Inf. Reuse Integr.*, Las Vegas, IL, USA, Aug. 2007, pp. 221–226.
- [3] V. Pinheiro, V. Furtado, T. Pequeno, and D. Nogueira, "Natural language processing based on semantic inferentialism for extracting crime information from text," in *Proc. IEEE Int. Conf. Intell. Secur. Informat.*, Vancouver, BC, Canada, May 2010, pp. 19–24.
- [4] B. Wang, P. Yin, A. L. Bertozzi, P. J. Brantingham, S. J. Osher, and J. Xin, "Deep learning for real-time crime forecasting and its ternarization," *Chin. Ann. Math., B*, vol. 40, no. 6, pp. 949–966, Nov. 2019.
- [5] S. Chackravathy, S. Schmitt, and L. Yang, "Intelligent crime anomaly detection in smart cities using deep learning," in *Proc. IEEE 4th Int. Conf. Collaboration Internet Comput. (CIC)*, Philadelphia, PA, USA, Oct. 2018, pp. 399–404.
- [6] H.-W. Kang and H.-B. Kang, "Prediction of crime occurrence from multimodal data using deep learning," *PLoS ONE*, vol. 12, no. 4, Apr. 2017, Art. no. e0176244.
- [7] A. Fidow, M. Hassan, M. Imran, X. Cheng, C. Petridis, and C. Sule, "Suggesting a hybrid approach mobile apps with big data analysis to report and prevent crimes," in *Social Media Strategy in Policing (Security Informatics and Law Enforcement)*, B. Akhgar, P. S. Bayeri, and G. Leventakis, Eds. Cham, Switzerland: Springer, 2019, pp. 177–195.
- [8] P. J. Brantingham, M. Valasik, and G. O. Mohler, "Does predictive policing lead to biased arrests? Results from a randomized controlled trial," *Statist. Public Policy*, vol. 5, no. 1, pp. 1–6, Jan. 2018.
- [9] A. Nasridinov and Y.-H. Park, "A study on performance evaluation of machine learning algorithms for crime dataset," *Adv. Sci. Technol. Lett.*, vol. 90, pp. 90–92, Dec. 2014.
- [10] Stec and D. Klabjan, "Forecasting crime with deep learning," 2018, arXiv:1806.01486. [Online].