

CRIME RATE PREDICTION AND ANALYSIS

S.Kothai¹, K.Latha², Dr.K.ReginaBose³

¹Department of Computer Science Engineering , Prince Shri Venkateshwara Engineering college

²Department of Computer Science Engineering , Prince Shri Venkateshwara Engineering college

³Department of Computer Science Engineering , Prince Shri Venkateshwara Engineering college

Abstract - Crime is one of the biggest and dominating problem in our society. There are huge number of crimes committed frequently. Here the dataset consists of the date and the crime rate that has taken place in the corresponding years. We use prophet algorithm to predict the percentage of the crime rate in the future by using the previous data information.

1.Introduction:

Crime is a violation of humanity, often punishable by law. Criminology is a study of crime, interdisciplinary science that investigates and investigates crime and criminal performance data. Criminal activity is now high and the police department is responsible for controlling and reducing criminal activity. There has been tremendous increase in machine learning algorithms that have made crime prediction feasible based on past data. The aim of this project is to perform analysis and prediction of crimes in states using machine learning models. In this project various machine learning models to predict crimes. Various visualization techniques and plots are used which can help law enforcement agencies to detect and predict crimes with higher accuracy. This will indirectly help reduce the rates of crimes and can help to improve securities in such required areas.

2.OBJECTIVES

Aim of the project:

The aim of this project is to analyze and predict the crime rate that helps the investigator or police to solve the crime problems from the huge amount of information that is stored in database.

Data visualization is also used to represent the output in the forms of images and charts.

Scope of the project:

The scope of the project is to prevent crimes or control the crime activities which may occur in the future.

3.PLATFORM

JUPYTER NOTEBOOK

Jupyter Lab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science,

ANACONDA

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda distribution that allows you to launch applications and easily manage conda packages, environments, and channels without using command-line commands.

COMMA SEPARATED VALUE - DATASET

CSV (Comma Separated Values) is a simple file format used to store tabular data, such as a spreadsheet or database. A CSV file stores tabular data (numbers and text) in plain text. Each line of the file is a data record. Each record consists of one or more fields, separated by commas. The use of the comma as a field separator is the source of the name for this file format.

4.LITERATURE REVIEW

Shiju Sathyadevan, Devan M.S, proposed that Day by day the crime rate is increasing considerably. Crime cannot be predicted since it is neither systematic nor random. Also the modern technologies and hi-tech methods help criminals in achieving their misdeeds. According to Crime Records Bureau crimes like burglary, arson etc have been decreased while crimes like murder, sex abuse, gang rape etc have been increased. Even though we cannot predict who all may be the victims of crime but can predict the place that has probability for its occurrence. The predicted results cannot be assured of 100% accuracy but the results shows that our application helps in reducing crime rate to a certain extent by providing security in crime sensitive areas. So for building such a powerful crime analytics

tool we have to collect crime records and evaluate it .It is only within the last few decades that the technology made spatial data mining a practical solution for wide audiences of Law enforcement officials which is affordable and available. Since the availability of criminal data or records is limited we are collecting crime data from various sources like web sites, news sites, blogs, social media, RSS feeds etc. This huge data is used as a record for creating a crime record database. So the main challenge in front of us is developing a better, efficient crime pattern detection tool to identify crime patterns effectively [1].

PAC Duijn, V Kashirin, proposed, policymakers and law enforcement agencies across the globe struggle to find effective strategies to control criminal networks. The effectiveness of disruption strategies is known to depend on both network topology and network resilience. However, as these criminal networks operate in secrecy, data-driven knowledge concerning the effectiveness of different criminal network disruption strategies is very limited. By combining computational modeling and social network analysis with unique criminal network intelligence data from [2].

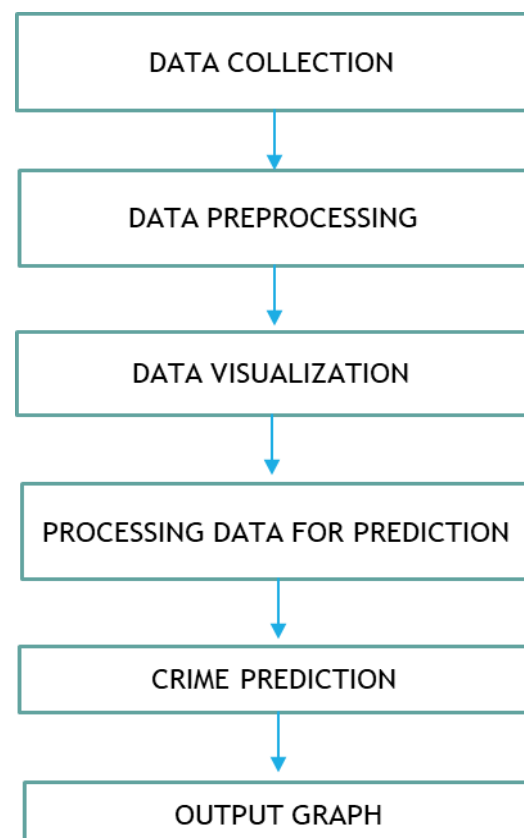
C.P. Chaithanya, N. Manohar, Ajay Bazil Issac, describes Text detection is the method of locating areas in a picture wherever, text is present. Text detection and classification in natural pictures is very important for several computer vision applications like optical character recognition, distinguish between human and machine inputs and spam removal. Currently the challenge in text identifying is to detect the text in natural pictures due to many factors like, low- quality image, unclear words, typical font, image having a lot of color stroke than the background color, blurred pictures due to some natural problems like rain, sunny, snow, etc. The main aim of this work is to identify and classify the text in natural pictures. Here system detects the text and finds the connected regions, chain them together in their relative position. Uses a text classification engine to filter chains with low classification confidence scores [3].

the data mining is data analyzing techniques that used to analyze crime data previously stored from various sources to find patterns

and trends in crimes. In additional, it can be applied to increase efficiency in solving the crimes faster and also can be applied to automatically notify the crimes. However, there are many data mining techniques. In order to increase efficiency of crime detection, it is necessary to select the data mining techniques suitably[4].

The application is developed as a Windows application by using TKinter-Python for crime prediction. Machine learning concepts and implementation are used here for performing crime analysis and prediction, which aid the ease of understanding the data in multiple ways and further predict it with good accuracy. The algorithms used behind this work are K-means and decision tree algorithm. The proposed work aims at analyzing the data mining concepts for clustering and classifying the crime prediction. The results show that the classification method outperforms in terms of detection and accuracy. Experimented results are evaluated for error calculation, and they are further analyzed in this study[5].

5.PROPOSED METHODOLOGY



1. DATA PREPROCESSING:

Firstly import the Data Set and explore it. Then discover the unnecessary and missing values. Date are not in DateTime data-type convert this first. Dropping unnecessary columns in a Dataset. Renaming columns to a more recognizable set of labels. Skipping unnecessary rows in a CSV file.

2.DATA VISUALIZATION:

Use Date Column as an Index for analysis based on Date. Then select the required columns for visualization. Completely remove the dates we are not certain about and replace them with NumPy's NaN. Now design the aesthetic graph using Seaborn. Display the required graph using Matplotlib. The Visualization made for Analysis:

1.Top 15 Primary Types of Crime

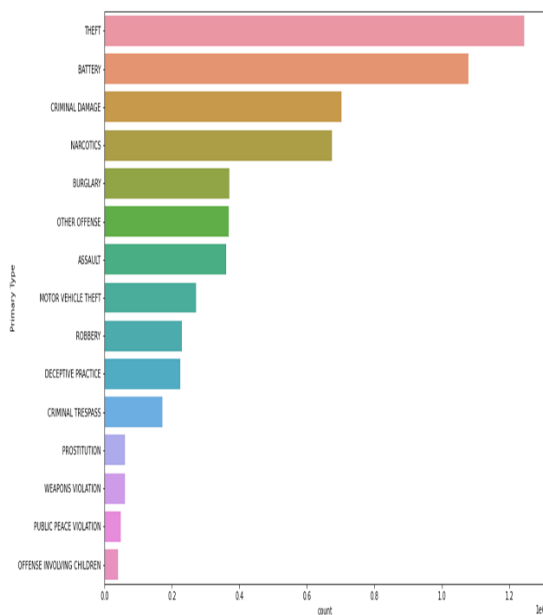


Fig.1

2.Top 15 Locations where crimes occurred

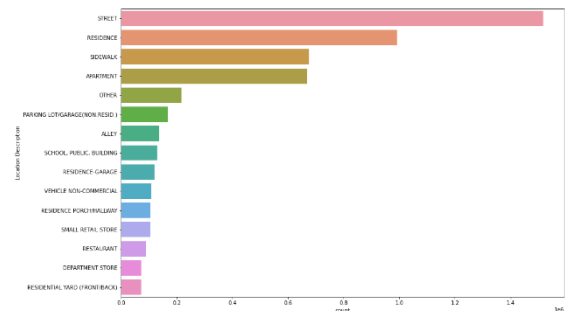


Fig.2

3.Crimes per year

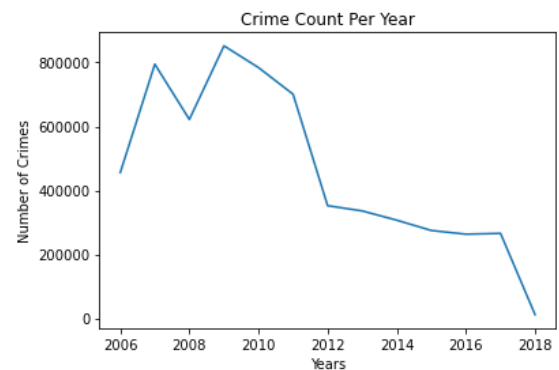


Fig.3

4.Crimes per Month

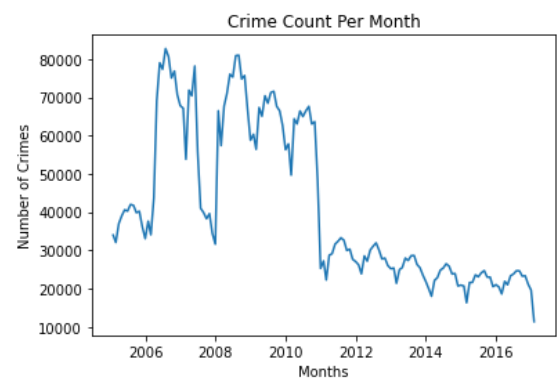


Fig.4

5. Visualize Crimes per Quarter

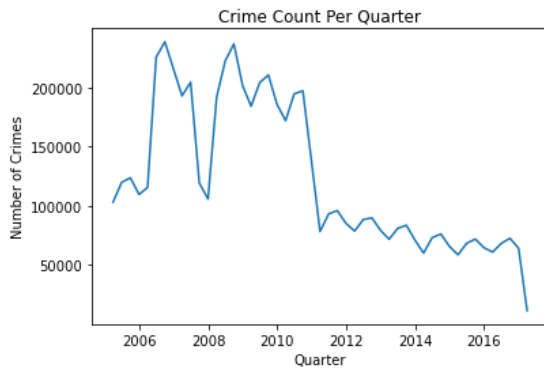


Fig.4

3. CRIMERATE PREDICTION:

Algorithm used :Prophet

Data processing is done for prediction. Make new dataframe with Date and it's count Monthly bases to feed the data to Prophet. Change Column names as the prophet takes ['ds','y'].The y column must be numeric, and represents the measurement we wish to forecast. Predictions are then made on a dataframe with a column ds containing the dates. A suitable dataset that extends into the future a specified number of days using the helper method `model.make_future_dataframe()`. Plot the forecast by calling the `model.plot` method and passing in your forecast dataframe. To see the forecast components use the `model.plot_components` method. The interactive forecast is made using `plot_plotly` method.

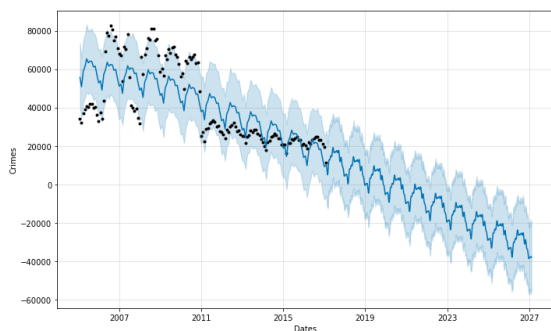


Fig.5

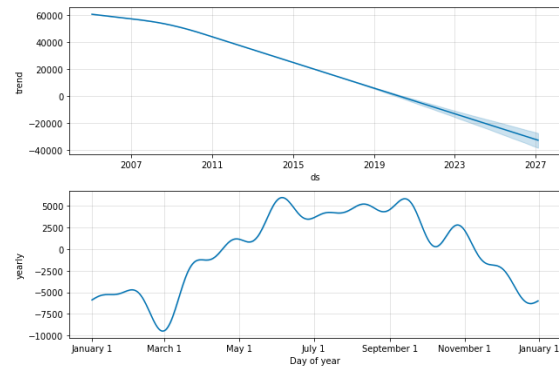


Fig.6

FORECAST GRAPH:



Fig.7

6.CONCLUSION

This project helps to analysis and predict the crime rate. From the encouraging results, we believe that crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis. Visual and intuitive criminal and intelligence investigation techniques can be developed for crime pattern.

7.REFERENCES

- [1] Machine Learning in Production: Developing and Optimizing Data Science Workflows and Applications (Addison-Wesley Data & Analytics)
- [2] The Complete Reference - Martin C. Brown
- [3] Ashish Sharma, Dinesh Bhuriya, Upendra Singh. "Survey of Stock Market Prediction Using Machine Learning Approach", ICECA 2017.

[4] Ubon Thansatapornwatana, "A Survey of Data Mining Techniques for-Analyzing Crime Patterns", *Second Asian Conference on Defense Technology ACDT IEEE*, pp. 123-128, Jan 2016.

[5] Crime Rate Analysis Based on Kmeans Clustering and Decision Tree Algorithm-
Jogendra Kumar, M. Sravani, Muvva Akhil.