Crime Prediction System Using Machine Learning

Aaditi Chougule
Dept. of Computer Engineering JSCOE
Savitribai Phule Pune University
Pune, INDIA

Sakshi Kashid
Dept. of Computer Engineering JSCOE
Savitribai Phule Pune University
Pune, INDIA

Pranjal Patil
Dept. of Computer Engineering JSCOE
Savitribai Phule Pune University
Pune, INDIA

Urvesh Thakkar
Dept. of Computer Engineering JSCOE
Savitribai Phule Pune University
Pune, INDIA

Prof. A. B. Ghandat
Dept. of Computer Engineering JSCOE
Savitribai Phule Pune University
Pune, INDIA

Abstract—**One of our society's most pervasive and concerning aspects is crime. Numerous crimes are perpetrated daily, which has negatively impacted the lives of regular citizens.restless. Therefore, stopping the crime from happening is a crucial duty.Artificial intelligence has recently demonstrated its value in practically every industry, and crime prediction is one of them. However, since this data can be consulted in the future, it is necessary to have a proper record of the crime that has occurred. The ability to foresee potential crimes can aid law enforcement organisations in stopping them before they happen.From a strategic standpoint, the capacity to foresee any crime based on time, place, and other factors can assist in giving law police vital information.However, because crimes are rising at an alarming rate, it is difficult to predict them accurately. Therefore, it is crucial to identify potential crimes now in order to prevent them from happening later. Many researchers have recently experimented with predicting crimes using various machine learning techniques and specific inputs. SVM algorithm is used to predict crime. The major goal is to demonstrate the value and potency of machine learning in forecasting violent crimes occuring in a certain area so that authorities may utilise it to lower crime rates in society.**

Index Terms—**Support Vector Machine, Machine Learning, Crime Prediction, SVM**

I. INTRODUCTION

The crime rate is significantly rising every day. One of the major problems that is always intensifying and becoming more complex is crime[1]. It is challenging to describe behaviours in crime patterns since crime patterns are continually changing[2]. Crime can be divided into other categories, such as rape, murder, theft, and kidnapping. With the aid of information technologies, law enforcement organisations gather information about crime (IT). However, the occurrence of any crime is inherently unpredictable, and past research has shown that a number of factors, such as poverty and unemployment, have an impact on the crime rate [3]. It is neither regular nor haphazard[4]. Analysis of crime is necessary due to the sharp rise of crime.

Crime analysis primarily comprises of techniques and procedures meant to lower the risk of crime.Identification and analysis of criminal patterns can be done practically.

However, a fundamental problem for law enforcement organisations is to accurately and efficiently analyse the growing volume of crime data. So, without any computing support, it becomes a challenging task for crime analysts to analyse such vast amounts of crime data. Traditional crime analysis cannot be used when crime data is high dimensional and complex queries have to be processed, necessitating the employment of a potent system for crime prediction. In order to successfully identify criminal patterns, a tool for crime prediction and analysis was required. This essay provides a few approaches that can be used to forecast which types of crimes are more likely to occur when and where they do.In order to identify features and forecast future trends in crime data based on similarities, classification is helpful. Techniques employed in this investigation is Support Vector Machine (SVM).

i)Problem Statement

In all parts of India, the crime rate is rising. Crime analysis techniques in criminology put a strong emphasis on spotting and forecasting patterns and trends. We employ a data mining approach to predict crime hotspots, which helps law enforcement agencies identify certain age groups that are more vulnerable to criminal activities.

ii) Motivation of the Project

High or rising crime rates cause communities to deteriorate because they negatively affect housing prices, community satisfaction, and the desire to move. Identifying the causes of crimes, predicting crimes, and offering solutions are crucial for crime reduction and prevention.It is unrealistic to conduct a manual analysis due to the size of the data sets and the number of algorithms that must be used on crime data. As a result, a platform that can apply any algorithm needed to perform a descriptive, predictive, and prescriptive study

on a vast volume of crime data is essential.

iii) Goal And Objective
1.To undertake descriptive, predictive, and prescriptive analysis in the proposed system.
2.Create a platform that enables descriptive and predictive data analytics methods to be utilised to examine crime data.
3.Examine the spatial and temporal (time of day, day of the week, and season) correlations in crime data using the suggested platform.
4.To reduce the rate of crime.
5.Examine the connections between crime and census statistics.

II. RELATED WORK

Prediction of crime type and occurence using a machine learning algorithm. [1]. One of the most complicated systems ever created by humans is the cyberspace; although widely used, the resources of cyber-technology are only dimly grasped by the majority of users. Cyber attacks in the past typically followed a random attack pattern to trick unwary targets. More data has shown that information about cyber attacks is disseminated around people and hacker forums in the virtual environment. This study suggests leveraging open source intelligence to find writings about cyber dangers on the surface web and deep web hacker forums. Our methodology can offer law enforcement organisations and cybersecurity professionals trustworthy data that can be used to create control and containment tactics for cyberattacks based on technical indicators of threats from the surface web.

Fuzzy C-Means Algorithm for Crime Analysis and Prediction [2] Crime analysis is a methodical technique for pinpointing crime hotspots.The identified crime regions, which are mostly based on crime type, are useful in lowering the crime rate. Finding high-crime regions may be done quite simply, and the crime rate can then be examined. As computer systems proliferate, crime data analysts can assist criminal nvestigators in their analysis of crimes. Using preprocessing and clustering, extract the crime hotspots from a structured data. We are concentrating mostly on crime variables from prior years when analysing the causes of crimes, including crime details of persons involved and other aspects. This approach primarily focuses on the areas where crimes will occur.

Utilizing the K-Nearest Neighboring Algorithm for Crime Prediction [3]. It is not unusual for people in a developing nation like India to hear of crimes occurring frequently. We must always be conscious of our surroundings because cities are rapidly becoming more urbanised. We will attempt to monitor crime rates using the KNN prediction algorithm in order to prevent the unfortunate. It will make a stab at predicting the type of crime, as well as when, where, and

how it might happen. This information will show the patterns of
crime in a certain location, which could be useful for criminal investigations. Additionally, it will show us the areas with the highest crime rates.

Empirical analysis for forecasting and predicting crime using deep learning and machine learning techniques [4]. The threat of crime and violation is intended to be controlled. Computationally, accurate crime predictions and future trend analysis can help to improve metropolis safety. Big data makes it difficult for people to quickly and accurately detect and forecast crime because of their limited capacity for processing complex information. Numerous computational opportunities and problems are presented by the precise calculation of the crime rate, types, and hotspots from historical trends. Despite significant research efforts, a better predictive algorithm is still required to lead police patrols in the direction of criminal activity. Lack of prior research prevents learning models-based crime predictions and prediction from being accurate.

III. STEPS OF IMPLEMENTATION

Since predictive modelling offers a method that can both develop a model and generate predictions, it was utilised to make predictions. This approach entails a variety of machine learning algorithms that can examine characteristics in the training data in order to make predictions. It is divided into two main categories: classification of patterns and regression. In order to create predictions regarding continuous variables, regression models are based on an investigation of the relationship between trends and variables. As an outcome of the prediction, classification's task is to assign a specific class label to each data value.
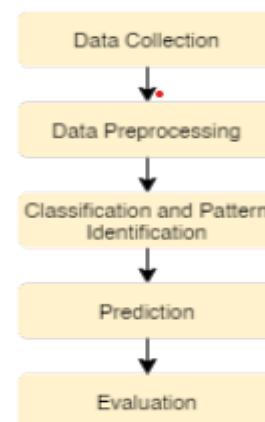
A. Gathering data and pre-processing:
During the data collection process, knowledge is gathered from a number of sources and used to build machine learning models. The information must be maintained in a way that is appropriate for the problem.

In essence, data pre-processing comprises methods to remove infinite or null values from data that could harm the performance of the model. This stage involves transforming the data collection into a format that machine learning models can use.

B. Model Selection:

Machine that supports vectors (SVM) The Support Vector Machine is effective for problems involving regression, time prediction series, and categorization. Support vector machines perform similarly to recurrent neural networks. Consequently, SVM has been used to predict crime hotspots [16] as well as conditions like diabetes and pre-diabetes. because it can produce nonlinear relation prototypes coherently. For time series expectation, it works well.

C.   Classification:

A classification problem involves using predictive modelling to forecast the class label for a certain input data example.An approach for supervised machine learning called SVM can be applied to classification or regression issues. Data are transformed using a method known as the kernel trick, and based on these modifications, it determines the best output boundary.

D.   Prediction

:The term "prediction" refers to the output of an algorithm that has been trained on historical data and applied to current data when estimating the likelihood of a particular outcome.

E.   Evaluation:Evaluating a machine learning model involves using a variety of evaluation techniques to understand its performance as well as its advantages and disadvantages. Early in the research process, a model's efficacy must be assessed; model evaluation also aids in model monitoring.

IV. PROPOSED IMPLEMENTATION

Dataset: For a machine that doesn't see data the same way that people do, the data collected should be made standard and intelligible.

Pre-processing: Real-world data typically includes noise, missing values, and may be in an undesirable format, making it impossible to build machine learning models on it directly. Data pre-processing is necessary to clean the data and prepare it for a machine learning model, which also improves the model's accuracy and effectiveness.

Feature Extraction: By extracting new features from the current ones in a dataset, feature extraction tries to reduce the amount of features in the dataset (and then discarding the original features). The majority of the information in the original collection of features should then be summarised by this new, smaller set of features.

Classification: Based on training data, the classification method is used to categorise fresh observations. It is a type of supervised learning approach. In classification, a programme makes use of the dataset or observations that are provided to learn how to categorise fresh observations into various classes or groups.

V. SYSTEM ARCHITECTURE
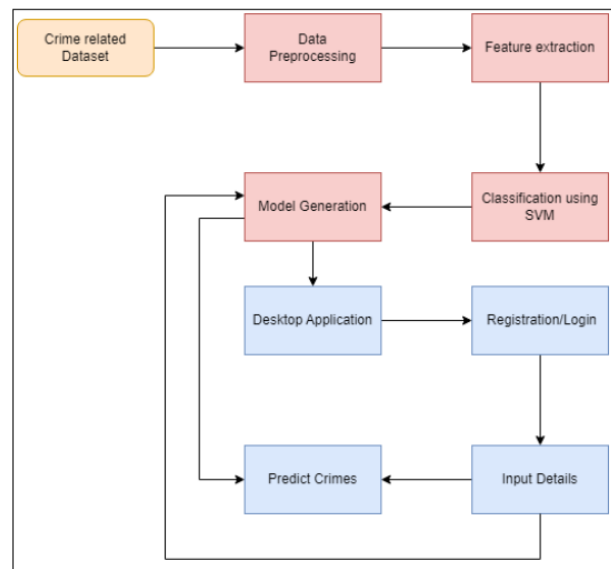
A.   Benefits of proposed system

➢   It helps police personnel to track and find criminals easily.

➢   It requires very less men power and the cost of the operation is very less.

➢   In this application the information cannot be manipulated or lost and data will be safe.

➢   This system can be used by police or investigation department to recognize criminal from their faces.

➢   This application can recognize faces in different lighting conditions with high accuracy

➢   The application is fast, robust, reasonably simple and accurate with a relatively simple and easy to understand algorithms and technique.

VI. ALGORITHM

Support Vector Machine:

One of the most well-liked supervised learning algorithms, Support Vector Machine, or SVM, is used to solve Classification and Regression problems. However, it is largely employed in Machine Learning Classification issues. The SVM algorithm's objective is to establish the best line or decision boundary that can divide n-dimensional space into classes, allowing us to quickly classify fresh data points in the future. A hyperplane is the name given to this optimal decision boundary.

SVM selects the extreme vectors and points that aid in the creation of the hyperplane. The approach is referred described as a "support vector machine" because of these extreme circumstances.

An SVM model maps the instances as points in space with as much room as feasible separating the examples of the different categories.Then, based on the side of the gap on which they fall, new samples are projected into that same area and predicted to belong to a category. Two different SVMs

1. Linear SVM: Data that can be separated into two classes by a single straight line are used for linear SVM. The classifier used is referred to as a Linear SVM classifier, and this type of data is known as linearly separable data.

2. Non-linear SVM: Non-linear SVM is used for non-linearly separated data, which indicates that a dataset is non-linear if it cannot be classified using a straight line, and a non-linear SVM classifier is utilised.

VII.RESULT

GUI Main:

Fig.1 show the GUI page of system . The system provide 3 buttons for user . 1. Login for user to login in the system 2.registration button for new user registration 3.exit button use to exit from the system
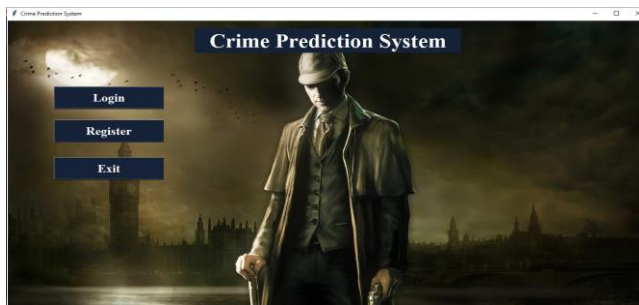

Fig. GUI Main Page

Registration Form:

The registration page is use to register new user.user need to enter complete information required in form.after registration user is able to login into the system.
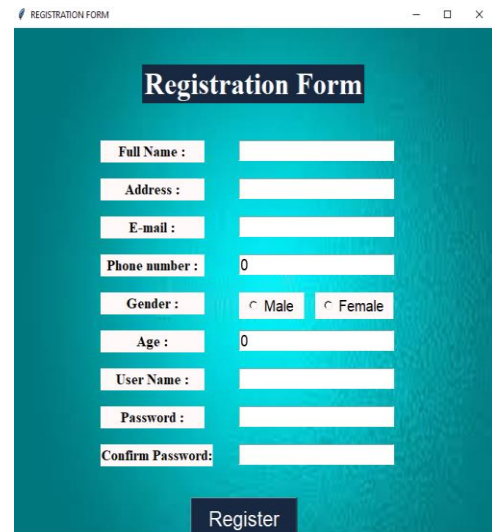

Fig.Registration Form

Login Form:

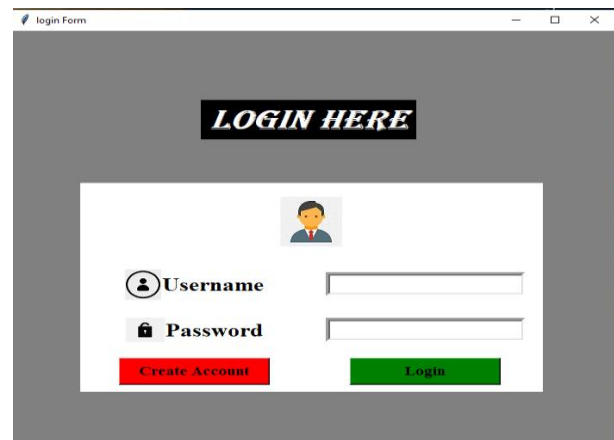After successfully registration user can login using login form.

Fig.Login Form

Welcome Page:

This page show some options like Graph , thef ,murder, kiddnaping, rap.

Fig.Welcome Page

Graph for crime Cases:

When user click on graph option 4 buttons will display.1.Thef,2.murder,3.kiddnaping,4.Rape
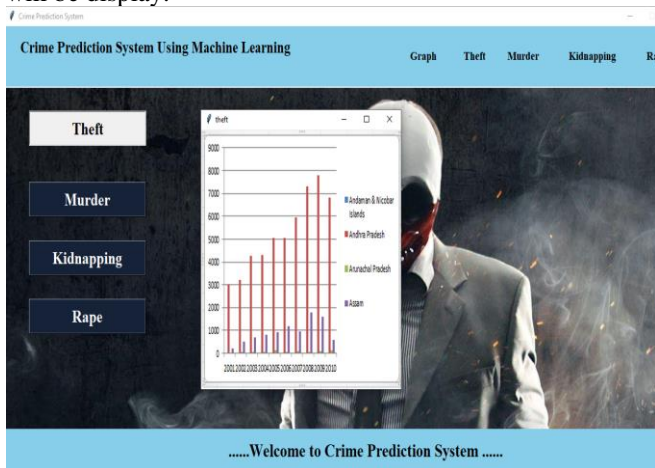When user click on button that particular crime Cases graph will be display.

Fig.Graph For crime Cases

Crime Prediction Case:

Whenever thef ,murder,kiddnaping or rape option will select,3buttons will be display 1.Data processing, 2.Model Traning, 3.Crime Prediction
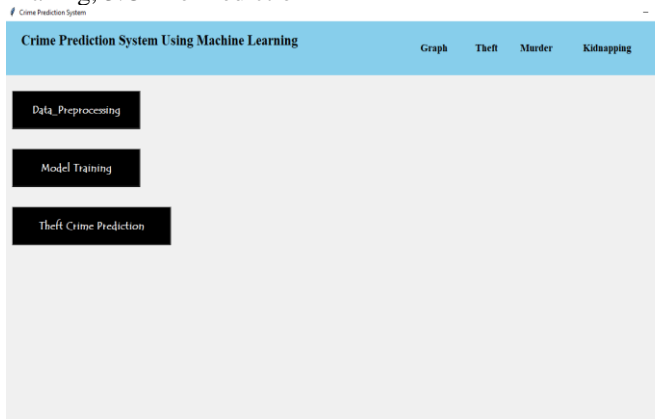

Fig.Crime Prediction Cases

Data Processing and model training:

While click on data preprocessing data will be split and model will be train after click on model training button.
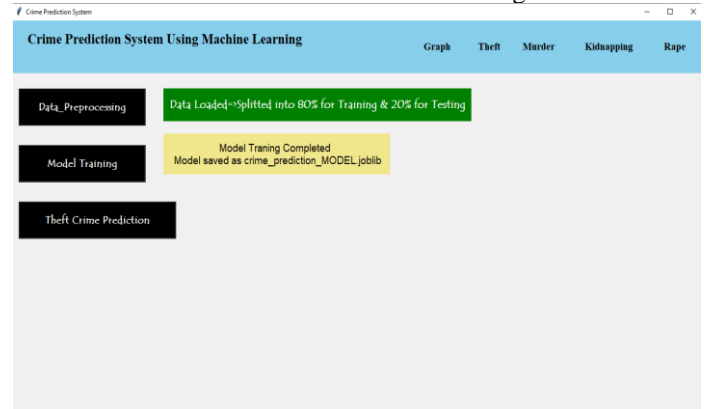

Fig.Data processing and model training

Crime Prediction:

To predict crime ,click on thef crime prediction button. Select area name,year and Sub group name.Submit detail and total crime cases will be display.

Fig.Crime Prediction

VIII. FUTURE SCOPE

1)SVM will be used to compare the information embedded in Twitter and darknet forums after the model has been trained on a bigger deep web data set.

2) Gathering and annotating additional surface and deep web data for public release.

Statement of Scope:

An in-depth knowledge of all criminological concepts is necessary for the analysis of crime patterns, as is the capacity to spot trends. The government will need to invest a large amount of time and energy into implementing technology in order to control some of these illicit activities. To aid in the examination of crime patterns and their links in a particular place, researchers conducted numerous studies. The suggested method classifies temporal and spatial data using

machine learning techniques to identify matching crime patterns.

Application:

1.In Company
2.In Office
3.In Banking
4.All User

IX. CONCLUSION

Crime analysis and prediction are a methodical way to spot crime. By foreseeing places with a high likelihood of crime occurring, this system can forecast and visualise areas that are criminally inclined. Using the idea of data mining, we can find previously undiscovered, meaningful information in unstructured data. This paper has made an effort to use techniques like exploratory data analysis, data analysis, and modelling, which entail methodical steps like the gathering and compilation of primary and secondary data, analysis of pertinent historical and contemporary data using the development of a linear regression model, and building a model with the most correlated features. Some significant findings have been made using these methods, such as the strong association between urbanity and crime rate and population density.

Although a lot of effort is being done in this field, there is no fixed standard for the datasets that are gathered by various organisations and police forces. The factors influencing crime from the used dataset are found in this research article. If the employment ratio had been included in the dataset, it might have improved the model. Additionally, because there are fewer rows in the dataset, the model could not function the same way for upcoming events. A uniform rule for data collection should be utilised in order to better understand crime behaviours. This will assist in developing a common paradigm that will aid all local police departments and intelligence gathering organisations in reducing crime.

REFERENCES

[1] Shiju Sathyadevan M.S, Surya Gangadharan: Crime Analysis and Prediction Using Data Mining,in NetworksSoft Computing(ICNSC),(2014) First International Conference. https://ieeexplore.ieee.org/document/6906719.

[2] H. Benjamin Fredrick David1,A. Suruliandi: Survey on crime analysis and prediction using data mining techniques.Department of Computer Science and Engineering,Manonmaniam Sundaranar University, India

[3] JesiaQuader Yuki, Md.MahfilQuaderSakib, ZaishaZamal, Khan Mohammad Habibullah, Amit Kumar Das: Predicting Crime Using Time and Location Data(2019). https://www.researchgate.net/publication/335854157

[4] Peng Chen,Justin Kurland, Modus Operandi: Time,Place,A Simple Apriori Algorithm Experiment for Crime Pattern Detection(2018).9th International Conference on IISA

[5] M. Hou, Y. Yang, T. Liu, and W. Peng, "Forecasting time series with optimal neural networks using multi-objective optimization algorithm based on AICc," Frontiers of Computer Science, vol. 12, no. 6, pp. 1261–1263, 2018.N.

[6] H. M. Shamsuddin, N. A. Ali, and R. Alwee, "An overview on crime prediction methods," 6th International Student Project Conference ICT, (ICT-ISPC), 2017, vol. 2017-January, pp. 1–5, 2017, doi: 10.1109/ICT-ISPC.2017.8075335.

[7] S. Kim, P. Joshi, P. S. Kalsi and P. Taheri, "Crime Analysis Through Machine Learning," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, 2018, pp. 415-420, doi: 10.1109/IEMCON.2018.8614828.

[8] N. Shama, "A Machine Learning Approach to Predict Crime Using Time and Location Data," pp. 1–52, 2017.