WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Crime Forecasting and it's Models

Sanjana Patel, Zoya Naaz

Dept. CSE

ABSTRACT---

in this paper we will be specifically discussing about different types of AI models that are used for crime forecasting depending on the data and the provided time, also learn how differently models will behave depending on their type ,technology and algorithms that they use to study a certain amount of data. Crime forecasting plays a very important role in the analysis of criminal data that can be used to train an AI to predict future crimes ,that can help the law forces to prevent crimes before they happen.

Keywords—machine learning, crime forecasting, AI models, real time forecasting

Introduction

Crime forecasting is the technique through which we can predict any time of crime before it happens , and for that we can use different types of forecasting models. There are different types of crime like robbery, murders and others. Crime forecasting helps the law agencies through which we can prevent crimes from happening .

There are different types of models like linear regression, auto regressive models , ARIMA, random forest etc.

Problem Statement—

Checking different types of models by applying data and analysing the crime rate.

Crime rate predictions depend on different models and algorithms , there are different types of crimes and crime rate difference depending on the data and timeline of the criminal data.

To correctly check and analyse the crime the models that are already present among them ARIMA model and linear regression are some of the models that have given nearly accurate precision with accord to timeline of the data and it's type.

Machine learning is one of the trendy technology that is being for different types of new innovations and is an integral part of artificial intelligence and in with regard to crime forecasting it is plays a very important as it can adapt to change to quickly with the changing timeline and fluctuating data that changes every now and then, therefore the models of machine learning that adapt to this technology most efficiently are considered to be the most accurate result providing models for forecasting.

AI helps the law forces to analyse the data which is present in real time like cctvs, footprints, blood samples and DNA analysis and through machine learning algorithms it learns from it's past data and predicts the crimes according to their previous hotspots where they are most likely to happen again.

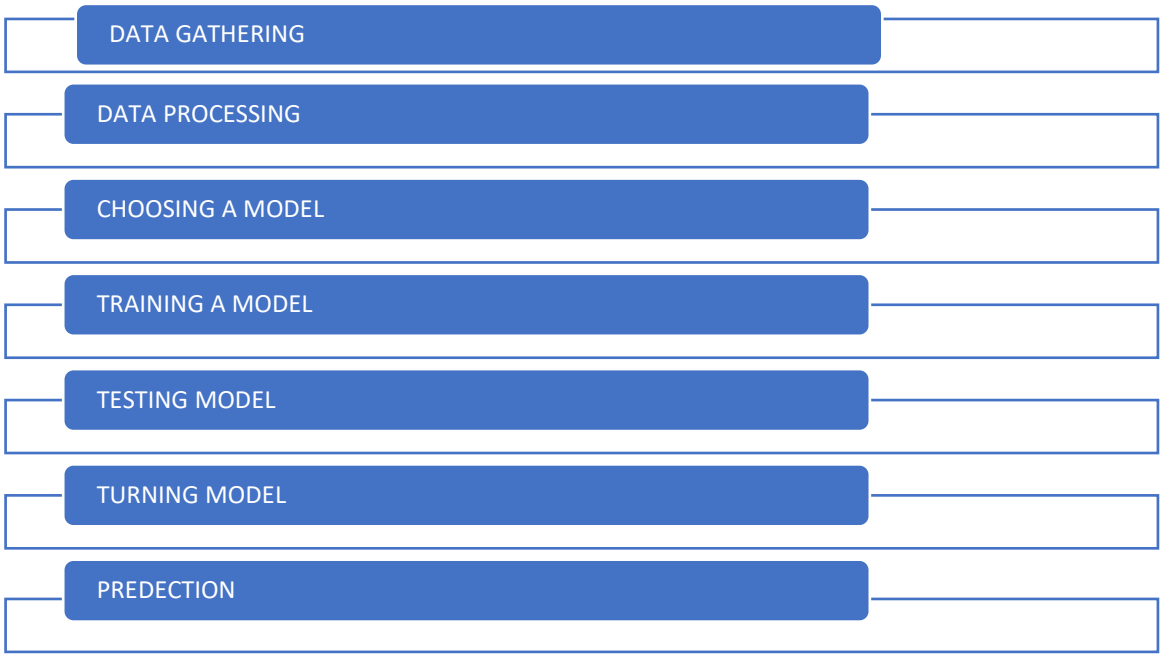
Reference	Crime model	User data description
[2]	• Artificial Neural Network (ANN) • Nonlinear Autoregressive with Exogenous Input	Crime record data from Selangor and Kuala Lumpur, Malaysia in 2013.
[8]	• Particle swarm optimization (PSO) • Support Vector Regression (SVR) • ARIMA	Historical USA crime rate dataset with external economic data from 1960 to 2009.
[10]	• Exponential Smoothing • ARIMA	Weekly burglary counts over a nearly five year period for the City of Pittsburgh.
[11]	• Interpolation method • Kernel density estimation (KDE)	Crime dataset in Arlington, Texas from 2007 to 2008.
[12]	• Artificial Neural Network (ANN) • Support vector machine (SVM)	Crime dataset comprising aggregated counts of crime and crime related events categorized by police department in USA.
[15]	• Autoregressive (AR)	Crime dataset in Chicago, USA from 2001 to 2014.
[16]	• ARIMA	50 weeks of property crime data in China.
[17]	• Linear regression • Adaptive regression • Decision stump algorithms	Crime and communities dataset in UCI machine learning repository.
[18]	• Random walk • Brown's simple exponential smoothing • Holt's two-parameter linear exponential smoothing	Crime data collected from computer aided system and offense report at Pittsburgh from 1991 to 1998.
[19]	• Space-time Autoregressive	USA regional and state rates of violent and property crimes.
[20]	• Eigenvector Spatial Filtering • Generalized Linear Mixed Model	Vehicle burglary crime data in Plano, Texas from 2004 to 2009.
[21]	• Linear Regression	Complaint dataset from Internet Crime Complaint Centre (IC3) from 2000 to 2005.
[22]	• Artificial Neural Network (ANN) • Probabilistic model	Crime related and spatial dataset in Pakistan.
[23]	• Fuzzy Logic • Self-organizing Map (SOM)	Monthly crime data of 20 county police bureaus in Taiwan from 2003 to 2004.
[24]	• Genetic programming	Crime and communities dataset in UCI machine learning repository.
[25]	• Random forest • Support vector machine (SVM) • Correlation analysis	Crime dataset in India from 2001 to 2012.

Limitations of crime forecasting

The word crime has multiple types depending on the situation in which the victim was, so basically anything that causes harm and hinders the peace of an individual in physical or mantel manner is defined as crime. Crime forecasting is applied to crimes that can be tracked in real time with the help of technology and human resources but the crimes that are not put into serious categories like murders or visual assaults they cannot be tracked or predicted using crime forecasting such as domestic violence or sexual assaults because in many cases either people do not report them or there is no pervious record present for the machine to learn regarding a certain person or area, that is where crime forecasting fails and the cases can be only solved by the help of human experience .

As we know that crime forecasting uses higher level of technology so it cannot be understood by everyone as not everyone in the law forces may not be able to control all the tools that are present in technology and also it will be very hard to apply these types of technologies in the rural areas of any country regardless of it's ranking , there are still areas where people do not have proper knowledge about computers and it also costs a lot to put these types of technologies in common use because of it's compact operations and it also there is no technology which can give 100 % result.

Methodology used for forecasting



- Collect data relevant to your target of analysis- before starting any kind of working processes we need data to that has to be processed.
- Organize data into a single dataset- when we collect information on any kind of topic that is most probably always in a distorted form
- Clean your data to avoid a misleading model.
- Create new, useful variables to understand your records.
- Choose a methodology/algorithm.
- Build the model.
- Train that model according to the type of data'
- Test and turn the model.
- Then apply the model and get the result.

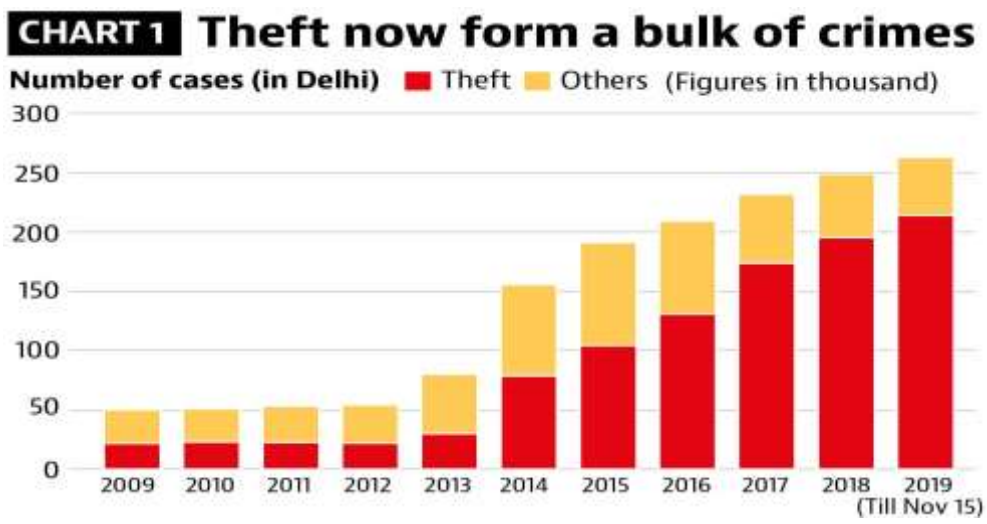
The basic steps that are involved mainly are Data gathering, Data processing, choosing a model, training that model, testing the model then we make evaluations through the predictions. These steps help us to make a proper predictions and make them as accurate as possible therefore it is very important for all of these steps to be executed properly.

Criminology Literature

The data that is kept and maintained for the later use and study of crime is known as it's literature therefore all kinds of data that are used for the study of a crime is known as criminology literature .

All types of stats regarding a case or the technologies that were used with all the main clues of a criminal case are covered in criminology literature. These types of data is mainly defined as the time series data. To check and navigate these types of data we use different types of AI models that are statistical also known as the hybrid models . statistical models follow algorithms like linear regression, exponential smoothing and autoregressive integrated moving average(ARIMA).

Some examples of statistical models are as follows:



(theft cases in delhi 2019)

Models used for forecasting

In this section mainly we will know about the models that are being used in crime forecasting for predictions giving it a different types of models that are being used.

There are mainly two types of models methodologies that are used that are :

1.Traditional Model methodology

The traditional model in the forecasting function mainly focus on the way to work manually on any data in form of statistical form on manual paper which cannot support the change in type of data and their different factors and changing behaviour of data.

2.Artificial intelligence model methodology

There is one more type of model that is used in forecasting and is much reliable and accurate than the traditional model which is the AI model.

There are different types of models present in the AI part. All of them have different accuracy levels , working factors and give different types of predictions depending on the data given to them. Few of these models will be explained in this review paper detailing on their working behavior and limitations.

We will mainly discuss the models that follow the Artificial intelligence models.

1. Artificial Neural Network(ANN): it is used to learn hidden patterns in a criminal dataset. The softwares that are used in this are kereas and tensorflow.
2. Linear Regression : it is a technology based on the concept of supervised learning used for forecasting . it is one of the most used algorithms to preidict the statistical data. Uses softwares like MS Excel,spss,origin and many more.
3. Auto regressive model: it uses previous data to predict the output . it is a statistical model that mainly uses python libraries as it's software and mainly relies on them for the representation of the data. The libraries used are NuMPY, pandas, Matpoltlib .
4. Decision stump algorithm: they are mainly used as building blocks for more complex models, such as decision tress. The working duration of the model depends on the amount of data it is being given to process. The softwares used for this are WEKA, MATLAB, Orange.
5. Fuzzy Logic: it is a type of algorithm that considers all the possibilities of a truth or situation rather than just true or false. It values the data into different degrees and then designates them a region. The software used for are panadas, skfuzzy and numpy.
6. Correlational analysis: it is a statistical method that is used to discover if there is a relationship between any two datasets. The software used for this is CorrLAB.
7. Random Forest Regression: it is used to perform regression and classification with the help of aggregation and bootstrap (bagging). The software used for this is Scikit-learn.
8. Auto regressive integrated moving average(ARIMA): it works at the collection of mean absolute error of deviating a model at a very complex dataset. The software used for this is Minitab.

9. Interpolation,KDE(Kernel Density Estimation): Helps in converting the unstable and unrefined data into more refined data that can also be applied on different location despite of the data type.the software used for this is seaborn.

10. SVM(support vector machine): It classifies the data and set the boundries such as if the new data enters it does not need to realanlyze the data for classification. The software used is scikit-learn.

Conclusion

The conclusion that we came upon after completion of the paper is that the technology is something that always keeps upgrading and keeps getting better and has very less chances of continuation , it is also fast and can do wonders in any field that it is put in . same goes for the technology of forecasting whether it is about weather data to forecast or crime data. Crime forecasting is latest booming technology which makes predictions using data that is already present using different types of models and algorithms. It has started a new revolution in the tech and AI industry. After all the type of advancement still the local and simple access of this technology is not very common in underdeveloped countries in as these techs require very advanced serves. To make these technologies more common we need to develop much more simple and user friendly modules that can operated by a person with basic knowledge of computers.

Future scope

We can create a new type of software based on the algorithms that are already present but with new modules that can be easy for the user to work and enter the data in any filtered or unfiltered form and still get their desired result with high accuracy and fast working. This paper presented the technologies that can be used by law agencies to prevent the crimes in future.

Citations

1) Ch. Mahendra, G. Nani Babu, G. Balu Nitin Chandra , A. Avinash , Y. Aditya "Crime Rate Prediction" from journal of engineering(JES),vol 11.,5.05.2020, **www.jespublication.com**

2) Bao Wangy*, Duo Zhangy, Duanhao Zhangy, P. Jrey Brantinghamz, and Andrea L. Bertozziy Deep Learning for Real Time Crime Forecasting, arXiv:1707.03340v1

3) Ourania Kounadi, Alina Ristea , Adelson Araujo Jr. and Michael Leitner "A systematic review on spatial crime forecasting", https://doi.org/10.1186/s40163-020-00116-7

4) Sunday Adeola ajagbe, Ifedotun Roseline Idowu, John B. Oladosu PhD and Ademola o. adesina PhD "Accuracy of Machine Learning Models for Mortality Rate Prediction in a Crime Dataset",Online:ISSN2645-2960; https://www.researchgate.net /publication/348002862

5) Adeniyi Olatunbosun, Olubayo Oluduro "Crime Forecasting and Planning in Developing Countries: Emerging Issues" DOI:10.3968/j.css.1923669720120801.600,CanadianSocialScience, https://www.researchgate.net/publication/267860904

6) Wilpen Gorr , Richard Harriesb" Introduction to crime forecasting", www.elsevier.com/ locate / ijforecast Alif Ridzuan Khairuddin, Razana Alwee and Habibollah Haron "A review on applied statistical and artificial intelligence techniques in crime forecasting", doi:10.1088/1757-899X/551/1/012030,

7) Leslie W. Kennedy & Marco Dugato "Forecasting Crime and Understanding its Causes. Applying Risk Terrain Modeling Worldwide", European Journal on Criminal Policy and Research (2018) 24:345–350https://doi.org/10.1007/s10610-018-9404-3

8) Neil Shah, Nandish Bhagat and Manan Shah "Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention", Shah et al. Visual Computing for Industry, Biomedicine, and Art (2021) 4:9, https://doi.org/10.1186/s42492-021-00075-z

9) theft data of 2019 (image data taken from Hindustan Times) https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.hindustantimes.com%2Fcities%2Ftheft-most-common-crime-in-delhi-shows-ncrb-data%2Fstory-0qLPCcjvVU6pJ9GbouUeHO.html&psig=AOvVaw0wIC7vpF-ArWvYslcO47di&ust=1690187555946000&source=images&cd=vfe&opi=89978449&ved=2ahUKEwjrvqO4taSAAxUQbmwGHWsTD6cQjRx6BAgAEAw