CRIME PREDICTION SYSTEM

[1]Dachawar Priyanka Shankarrao

[1]Student
[1]Department of Computer Engineering,

[1]M. B. E. Society's College Of Engineering, Ambajogai, India.

Abstract : In this era of recent times, crime has become an evident way of making people and society under trouble. An increasing crime factor leads to an imbalance in the constituency of a country. In order to analyse and have a response ahead this type of criminal activities, it is necessary to understand the crime patterns. This study imposes one such crime pattern analysis by using crime data obtained from Kaggle open source which in turn used for the prediction of most recently occurring crimes. The major aspect of this project is to estimate which type of crime contributes the most along with time pe- riod and location where it has happened. Some machine learning algorithms such as Naïve Bayes is implied in this work in order to clas- sify among various crime patterns and the accuracy achieved was comparatively high when compared to precomposed works.

Key- words: Crime, Analyse, Crime patterns, Kaggle, Estimate, Naïve Bayes, Accuracy.

I. INTRODUCTION

Crime Prediction is a systematic approach for finding the crime patterns and trends. This paper gives different technologies that can be used for building Crime Pre- diction System. By building Crime Prediction System, it speeds up the process of solving crimes and reduces the rate of crime. We have different techniques which are dependent on the data that are previously reported and recorded and time and location. Crime Prediction system uses recorded data and analyses the data using several analysing techniques and later can predict the patterns and trends of crime using any of the below mentioned approaches.

II. NEED OF THE STUDY.

To develop an application which will serve a way to register and track criminals remotely with the help of criminal data and also it will help to maintaining of properdataset of crime and analyzing this data to help in predicting and solving crimes in future.

EXISTING SYSTEM

As the crime rate and criminals are increasing day by day managing, finding and tracking these criminals is a major issue for police personnel. There are applica- tion which will help police department to store the records and data about a crimes and criminal but these applications won't help in finding those criminals. Criminal and crimes details were mainly managed using records books or stored as software records in the database. Previously when a criminal is found guilty the picture of the criminal is being taken and stored in records but these pictures serve no purpose. The existing methods will only help in managing criminal records and those methods will not finding criminals from any location.

Disadvantages

It is not possible to detect criminal from any location
Existing methods only provide data storage and security for data but not live track- ing.There was no application which will find criminals from CCTV footage.
Criminal details were stored manually in a record and it requires lot of work

PROPOSED SYSTEM

The dataset is obtained from the Kaggle repository. This is the domain for the various research-oriented dataset. The dataset contains homicide entries collected from the FBI's supplementary Homicide Report. The dataset consists of 638454 rows and 17 columns and the column metadata. From the dataset, the significant features like State, Year, Month, Crime Type, Crime Solved, Victim Gender, Victim Age, Victim Race, Victim Count and Weapon are chosen as the input features for the system. The features Perpetrator Age,

Perpetrator Sex and Relationship of the perpetrator with the victim are chosen as the target variable to be predicted by the system. We have used SVM algorithm for the prediction .

Advantages

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 find criminals history.
This application can find crimal details with high accuracy
The application is fast, robust, reasonably simple and accurate with a relatively simple and easy to understand algorithms and technique

III. RESEARCH METHODOLOGY

Anaconda Navigator
Anaconda Navigator is a desktop graphical user interface (GUI) included in Ana- conda distribution that allows users to launch applications and manage conda pack- ages, environments and channels without using command-line commands. Naviga- tor can search for packages on Anaconda Cloud or in a local Anaconda Repository, install them in an environment, run the packages and update them. It is available for Windows, macOS and Linux.
Thus, the main difference between Python and Anaconda is that the former is a pro- gramming language and the latter is software to install and manage Python and other programming languages (such as R). In this article, we'll discuss how to use Ana- conda to manage and install packages as well as when to use pip or conda .
Anaconda Navigator makes creation and management of multiple virtual environ- ments a simple affair. In the Environments section you'll see the environment list. When you installed Anaconda, the base (root) environment was created by default. As you can see in the above image, a list of all installed packages is shown.

Spyder
Spyder, the Scientific Python Development Environment, is a free integrated de- velopment environment (IDE) that is included with Anaconda. It includes editing, interactive testing, debugging, and introspection features.
At the Anaconda Prompt (terminal on Linux or macOS), type spyder and press Enter. Spyder should start up just like it did when you launched it from Anaconda Naviga- tor. Close Spyder the same way you did in the previous exercise. At the Anaconda Prompt (terminal on Linux or macOS), type jupyter-notebook and press Enter.
On other platforms, open Anaconda Navigator, scroll to Spyder under Home , and click Launch . If Spyder does not launch via this method or you prefer to use the command line, open Anaconda Prompt (Windows) or your terminal (other plat- forms), type conda activate base then spyder

Python
Python is an interpreted, object-oriented, high-level programming language with dy- namic semantics developed by Guido van Rossum. It was originally released in 1991. Designed to be easy as well as fun, the name "Python" is a nod to the British comedy group Monty Python.
Python was developed by Guido van Rossum in the late eighties and early nineties at the National Research Institute for Mathematics and Computer Science in the Netherlands.
Python is derived from many other languages, including ABC, Modula-3, C, C++, Algol-68, SmallTalk, and Unix shell and other scripting languages.
Python is copyrighted. Like Perl, Python source code is now available under the GNU General Public License (GPL).
Python is now maintained by a core development team at the institute, al- though Guido van Rossum still holds a vital role in directing its progress.

DB SQLite
SQLite is a software library that implements a self-contained, serverless, zero-configuration, transactional SQL database engine. SQLite is the most widely deployed SQL database engine in the world. The source code for SQLite is in the public domain. This tutorial
will give you a quick start with SQLite and make you comfortable with SQLite pro- gramming. SQLite is an in-process library that implements a self-contained, server- less, zero-configuration, transactional SQL database engine. It is a database, which is zero-configured, which means like other databases you do not need to configure it in your system.
SQLite engine is not a standalone process like other databases, you can link it statically or dynamically as per your requirement with your application. SQLite accesses its storage files directly. SQLite does not require a separate server process or system to operate (serverless).
SQLite comes with zero-configuration, which means no setup or administra- tion needed.
A complete SQLite database is stored in a single cross-platform disk file.
SQLite is very small and light weight, less than 400KiB fully configured or less than 250KiB with optional features omitted.
SQLite is self-contained, which means no external dependencies.
SQLite transactions are fully ACID-compliant, allowing safe access from mul- tiple processes or threads.
SQLite supports most of the query language features found in SQL92 (SQL2) standard.
SQLite is written in ANSI-C and provides simple and easy-to-use API.

SQLite is available on UNIX (Linux, Mac OS-X, Android, iOS) and Windows (Win32, WinCE, WinRT).

IV.IMPLEMENTATION

MODULES

Dataset - Pre-processing - Feature Extraction - Classification

MODULES DESCRIPTION

Dataset – Provide dataset (This means that the data collected should be made uniform and understandable for a machine that doesn't see data the same way as humans do.)

Pre-processing –A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

Feature Extraction – Feature Extraction aims to reduce the number of features in a dataset by creating new features from the existing ones (and then discarding the original features). These new reduced set of features should then be able to summarize most of the information contained in the original set of features.

Classification - The Classification algorithm is a Supervised Learning technique that is used to identify the category of new observations on the basis of training data. In Classification, a program learns from the given dataset or observations and then classifies new observation into a number of classes or groups.

V. RESULTS AND DISCUSSION

A systematic approach to identifying crime is crime analysis and prediction. This system can predict and visualise crime-prone areas by predicting regions with a high probability of crime occurrence. We can extract previously un- known, useful information from unstructured data using the concept of data mining

REFERENCES

[1] Shiju Sathyadevan M.S, Surya Gangadharan: Crime Analysis and Pre- diction Using Data Mining,in NetworksSoft Computing(ICNSC),(2014) First International Conference.

[2] .H. Benjamin Fredrick David1,A. Suruliandi: Survey on crime analysis and $pre_x0002_dictionusingdataminingtechniques.Departmento f$ $ComputerScienceandE$ $PredictingCrimeUsingTimeandLocationData$(2019).

[3[Peng Chen,Justin Kurland, Modus Operandi: Time,Place,A Simple Apriori $Al_x0002_gorithmExperi$ ob $jectiveoptimizationalgorithmbasedonAICc, "Frontierso f ComputerScience, vol.12, no.6, pp.12$

[4] N. 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.

[5] 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.

[6] N. Shama, "A Machine Learning Approach to Predict Crime Using Time and Loca- tion Data," pp. 1–52, 2017.

[7] Y. Abouelnaga, "San Francisco Crime Classification," 2016, [Online]. Available: http://arxiv.org/abs/1607.03626.