Crime Analysis & Prediction

Using Data Mining

Abstract:

The governing authority or the state, to maintain social order and stability of the society may impose an organized or formalized system. This system consists of rules & regulations which are created after lots of discussions and debates. The administering agencies then codify these rules into laws, assign police citizens to ensure that the citizens comply with those laws. The violation against these laws or the system is considered as a crime. The rate of crime in India is increasing drastically over the last few years. That is why there is a need for more advanced preventive measures that are to be implemented, for that proper analysis and prediction on the previous year's data can help develop the laws required. This paper essentially concentrates on two main things, which are Data Visualization and Data Prediction. The study has been done on a raw data set, first to deep dive into the problem by analyzing it using various data visualization techniques. Afterward, implementing powerful machine learning and data mining algorithms to extract the information out and to further find out the hidden relationships and unique crime patterns among the data.

Introductions:

Crime is an intentional action violating the criminal code imposed by the governing or administering authority, for which an individual or a group of individuals can get punished. Therefore, every crime violates the law, but not every violation of the law is considered a crime. Like breaches of contract and other civil law will not lie among the category of crime, it will lie under the category of “offenses” or in “infractions”. In India, the crimes are so rampant that in about an hour, a total of 187 cognizable IPC (Indian Penal Code) crimes and 443 SLL (Special and Local Laws) crimes get committed. There’s an annual increase of 1.6% in the registration of cases (50,74,635 cases) and the increase in crime rate per 100,000 population has increased from 383.5 in 2018 to 385.5 in 2019 [5]. More than one-fifth of all registered crime cases (10,50,945) were classified as violent crimes (e.g. – murder, kidnapping, assault, death by negligence, etc.) [5]. These figures can be reduced if preventive measures are introduced after proper analysis and prediction of crime data. The conventional process of analysis includes the study of crime reports and then discovering unique patterns, series, trends, and inclinations.

In this paper, the raw crime data sets used are confirmed and verified by the NCRB (National Crime Records Bureau), which proves its authenticity and assurance [7][8][9]. The data sets used are from a period of 2001 – 2019 with various kinds of parameters like based on place of occurrence, State/UT – wise, type of crime, etc. For the Data Visualization part of the paper, Tableau has been used as the visualization tool and for the Data Prediction part, Python programming language is used along with some libraries or modules like Sci-Kit [10][11].

Related Work:

3. Methodology:

For optimum and organized analysis of crime in India, various visualization techniques and machine learning algorithms have been implemented. Classification of the analysis has been done below in three sub-parts.

* 1. Data Collection & Preprocessing

Multiple data sets have been utilized for Data Mining, and the data sets used are reliable, real, and verified by the NCRB (National Crime Records Bureau of India) [12][13]. The data sets are acquired by the official website of the Indian Government [7][8][9]. The study has been done on various parameters like based on the type of crimes, the place of occurrence of crime, and state-wise as well. In this phase, the history of crimes from the year 2001 – 2019 has been considered. In the preprocessing phase, removal of inconsistent data (such as missing values, redundant information, etc.), joining two or more data sets constructively, and transformation of data as required for the visualization and prediction of crime has been done.

* 1. Data Visualization
     1. Module 1: Heat Map representation of different types of crime

1. Personal Crimes
2. Property Crimes
3. Statutory Crimes
4. Inchoate Crimes
   * 1. Module 2: Visualization of Crime State-Wise from 2001 - 2019
     2. Module 3: Visualization of Crime based on the place of occurrence
     3. Module 4: Crime against different kinds of people
   1. Crime Prediction

Conclusion:

Future Scope:

References:

[1] <https://ieeexplore.ieee.org/abstract/document/4053200/>

[2] [https://ieeexplore.ieee.org/abstract/document/6906719/](https://ieeexplore.ieee.org/abstract/document/6906719/%20)

[3] [https://link.springer.com/content/pdf/10.1007/s00146-014-0539-6.pdf](https://link.springer.com/content/pdf/10.1007/s00146-014-0539-6.pdf%20)

[4] [https://iopscience.iop.org/article/10.1088/1742-6596/1000/1/012046/meta](https://iopscience.iop.org/article/10.1088/1742-6596/1000/1/012046/meta%20)

[5] <https://en.wikipedia.org/wiki/Crime_in_India>

[6] <https://en.wikipedia.org/wiki/Data_analysis>

[7] <https://data.gov.in/catalog/crime-india-2018>

[8] <https://data.gov.in/catalog/crime-india-2015>

[9] <https://data.world/rajanand/crime-in-india>

[10] <https://www.tableau.com/>

[11] <https://scikit-learn.org/stable/>

[12] https://en.wikipedia.org/wiki/National\_Crime\_Records\_Bureau

[13] https://ncrb.gov.in/