Crime Type and Occurrence Prediction Using Machine Learning

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 analyze 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 period and location where it has happened. Some machine learning algorithms such as Naïve Bayes is implied in this work in order to classify among various crime patterns and the accuracy achieved was comparatively high when compared to pre composed works.

**EXISTING SYSTEM**

In pre-work, the dataset obtained from the open source are first pre-processed to remove the duplicated values and features. Decision tree has been used in the factor of finding crime patterns and also extracting the features from large amount of data is inclusive. It provides a primary structure for further classification process. The classified crime patterns are feature extracted using Deep Neural network. Based on the prediction, the performance is calculated for both trained and test values. The crime prediction helps in forecasting the future happening of any type of criminal activities and help the officials to resolve them at the earliest.

Disadvantages

1. The pre-existing works account for low accuracy since the classifier uses a

categorical values which produces a biased outcome for the nominal attributes with

greater value.

2. The classification techniques does not suited for regions with inappropriate data and real valued attributes.

3. The value of the classifier must be tuned and hence there is a need of assigning an optimal value.

**PROPOSED SYSTEM**

* The data obtained is first pre-processed using machine learning technique filter and wrapper in order to remove irrelevant and repeated data values. It also reduces the dimensionality thus the data has been cleaned. The data is then further undergoes a splitting process. It is classified into test and trained data set. The model is trained by dataset both training and testing .It is then followed by mapping. The crime type, year, month, time, date, place are mapped to an integer for ensuring classification easier.
* The independent effect between the attributes are analysed initially by using Naïve Bayes. Bernouille Naïve Bayes is used for classifying the independent features extracted. The crime features are labelled that allows to analyse the occurrence of crime at a particular time and location. Finally, the crime which occur the most along with spatial and temporal information is gained. The performance of the prediction model is find out by calculating accuracy rate. The language used in designing the prediction model is python and run on the Colab – an online compiler for data analysis and machine learning models.

**Advantages**

1. The proposed algorithm is well suited for the crime pattern detection since most of the featured attributes depends on the time and location.

2. It also overcomes the problem of analyzing independent effect of the attributes.

3. The initialization of optimal value is not required since it accounts for real valued, nominal value and also concern the region with insufficient information.

4. The accuracy has been relatively high when compared to other machine learning

prediction model.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).