IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657
Volume 10, Issue 4, 2022

Crime Rate Analysis Utilizing Unsupervised Learning by K-Means Machine Learning

Sathish,Bumpali kavyasree, Konda Rajeshwari, Kotagiri Varsha

Article Info

Received: 03-07-2022 Revised: 17 -08-2022 Accepted: 12-09-2022

ABSTRACT;

Wrongdoing is a disturbing part of our general public, and its counteraction is an essential assignment. Wrongdoing investigation is an efficient method of recognizing and looking at examples and patterns in wrongdoing. It is of most extreme significance to examine reasons, think about various factors and decide the relationship among different violations happening and find the best appropriate techniques to control wrongdoing. The essential goal of this venture is to recognize different wrongdoings utilizing grouping procedures dependent on the events and routineness. Information digging is utilized for examination, examination and check designs in violations. In this venture, a bunching approach is utilized to break down the wrongdoing information; the put away information is grouped utilizing the K-Means calculation. After the bunching, we can foresee a wrongdoing dependent on its recorded data utilizing characterization. This proposed framework can demonstrate wrongdoing head which have a high likelihood of crime percentage.

Key Words: Crime rate, data mining, k-means, machine learning.

INTRODUCTION

Crimes are the huge danger to the mankind. There are numerous wrongdoings that happen
normal timespan. Maybe it is expanding and spreading at a quick and huge rate. Crimes occur
from little town, town to enormous urban areas. Crimes are of various kind – burglary, murder,

serious issues to the police office as there are gigantic measure of crime information that exist. There is a need of innovation through which the case settling could be quicker.
Crime is increasing considerably day by day. Crime is among the main issues which is growing
continuously in intensity and complexity[1]. Crime patterns are changing constantly because of
which it is difficult to explain behaviors in crime patterns[2]. Crime is classified into

assault, attack, battery, bogus nment,hijacking, manslaughter. violations are expanding there is a need to settle the cases ina lot quicker way. The crime exercises have been expanded at a quicker rate and it is the obligation of police division t control and diminish the crime exercises. Crime forecast and criminal distinguishing proof

various types
like kidnapping, theft murder, rape etc. The law enforcement agencies collects the crime data
information with the help of information technologies(IT). But occurrence of any crime is
naturally unpredictable and from previous searches it was found that various factors like poverty, employment affects the crime rate [3]. It is neither uniform nor random[4].

1Professor, Department of ECE, Malla Reddy Engineering College for Women, Hyderabad, Telangana, India.2UG Scholar, Department ofECE, Malla Reddy Engineering College for Women, Hyderabad,Telangana, India

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 10, Issue 4, 2022

With rapid

increase in crime number, analysis of crime is also required. Crime analysis basically consists of
procedures and methods that aims at reducing crime risk. It is a practical approach to identify
and analyze crime patterns. But, major challenge for law enforcement agencies is to analyze
escalating number of crime data efficiently and accurately. So it becomes a difficult challenge for
crime analysts to analyze such
voluminous crime data without any computational support.

Related Work

Many researches have been done which address this problem of reducing crime and many crime- predictions algorithms has been proposed. The prediction accuracy depends upon on type of data used, type of attributes selected for prediction. In[5], mobile network activity was used to obtain human behavioral data which was used to predict the crime hotspot in London with an accuracy of about 70% when predicting that whether a specific area in London city will be a hotspot for crime or not.

IN[6], data collected from various websites, newsletter was used for prediction and classification of crime using Naive Bayes algorithm and decision trees and found that former performed better.

In[7] , a thorough study of various crime prediction method like Support Vector Machine(SVM), Artificial neural networks(ANN) was done and concluded that there does not exist particular method which can solve different crime datasets problems.

IN[8] , various supervised learning techniques, unsupervised learning technique[9] on the crime records were done which address the

connections between crime and crime pattern for the purpose of knowledge discovery which will help in increasing predictive accuracy of crime. In [10], different approach for predicting like Data mining technique, Deep learning technique, Crime cast technique, Sentimental analysis technique were discussed and it was found that every method have some cons and pros. Every method gives better result for a particular instance. Clustering approaches were used for detection of crime and classification method were used for the prediction of crime, [11]. The K-Means clustering was implemented and their performance is evaluated on the basis of accuracy. On comparing the performance of different clustering algorithm DBSCAN gave result with highest accuracy and KNN classification algorithm is used for crime prediction. Hence, this system helps law enforcement agencies for accurate and improved crime analysis.

In [12] , a comparison of classification algorithms, Na¨ıve Bayes and decision tree was performed with an data mining software, WEKA. The datasets for this study was obtained from USCensus 1990. In [13], the pattern of road accidents in Ethiopia were studied after taking into consideration various factors like the driver, car,

road conditions etc.Different classification algorithms used were K- Nearest Neighbour, Decision tree and Naive Bayes on a dataset containing around 18000 datapoints. The prediction accuracy for all three methods was between 79% to 81%.

EXPERIMENTAL SET UP AND METHODS

An unsupervised machine learning model used for making clusters of crime type as crime head labels further a supervised machine learning model is trained by using crime head data .The model is trained to predict the probability that a new crime head that should be reported and we can avoid crimes to be occurred.

PROPOSED SYSTEM

There are many machine learning algorithms available to users that can be implemented on datasets. However, there are two major types of learning algorithms: supervised learning and unsupervised learning algorithms. Supervised learning algorithms work by inferring information or "the right answer" from labeled training data. The algorithms are given a particular attribute or set of attributes to predict. Data preprocessing process includes methods to remove any null values or infinite values which may affect the accuracy of the system. The main

steps include Formatting, cleaning and

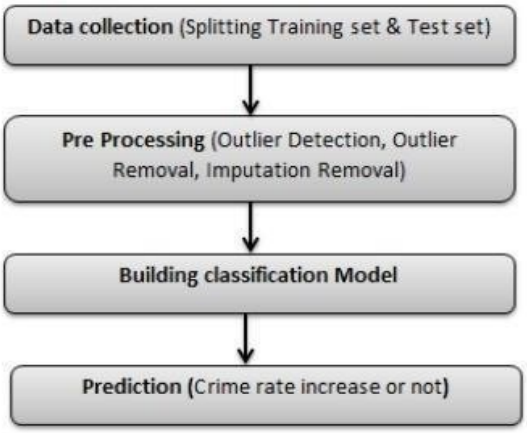
International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 10, Issue 4, 2022

sampling. Cleaning process is used for removal or fixing of some missing data there may be data that are incomplete.

Crimes Prediction ways:

- To utilize the resources identify the hotspots of crimes and allocate vigilante resources such as policeman, police cars, weapons etc. reschedule patrols according to the vulnerability of a place.
- Through that avoid crimes Ensure better civilization through avoiding happening crimes such as murder, rapes, thefts, drug, smugglings etc.

SYSTEM ARCHITECTURE



System Requirements:

Software:

Python

Anaconda navigator

Hardware :

Windows 7,8,10(64 bit)

RAM 3GB

Raspberry pi 3 B+

Raspberry Pi

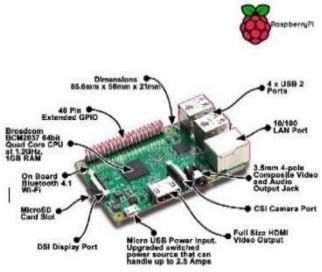
Raspberry Pi is a small single-board Computer developed in UK by Raspberry Pi foundation to promote the teaching of computer science in schools and in developing countries.

Original model become far more popular than anticipated sealing outside of its target market, for uses such as robots.

Processor

The processor at the heart of the Raspberry Pi is a Broadcom BCM28XX.

This is the Broadcom System on Chip (SOC) chip use in the Raspberry Pi. The processor from first to third generations include:



Raspberry Pi 1: Broadcom BCM2835 SOC with 700MHz CPU speed, L2 cache of 128kb with ARM

Compatibility AR1176JZF-S (ARMv6) 32-bit RISC ARM.

Raspberry Pi 2: Broadcom BCM 2836 SOC with 900MHz CPU speed, L2 cache of 256kb with 32-bit quad-core ARM cortex-A7 (ARMv7).

Raspberry Pi 3: Broadcom BCM2837 SOC with 1.2GHz 64-bit quad-core –A53 with 512 kb shared L2 cache (64-bit instruction set ARMv8).

Raspberry Pi 3

Raspberry Pi 3In this project we are using Latest version of Raspberry i.e. Raspberry Pi 3The processor at the heart of the Raspberry Pi 3 is a Broadcom BCM2837, and the later models of the Raspberry Pi 2. The underlying architecture of the BCM2837 is identical to the

BCM2836. The Only significant difference is the replacement of the ARMv7 quad core cluster with quad-core ARM Cortex A53 (ARMv8) cluster.

The ARM cores run at 1.2GHz, making the device about 50% faster than the Raspberry Pi 2 with a on board memory of 1GB RAM. The videocoreIV runs at 400MHz.

METHODOLOGY

1.

DATA COLLECTION

Data used in this paper is a set of crime head and year wise total cases occurred records. This step is concerned with selecting the

ISSN 2347–3657
Volume 10, Issue 4, 2022

subset of all available data that you will be working with. ML problems start with data preferably, lots of data (examples or observations) for which you already know the target answer. Data for which you already know the target answer is called labelled data.

2. DATA PRE-PROCESSING

Data Preprocessing is a data mining technique used to transform the raw data into useful and efficient format. The data here goes through 2 stages

1. **Data Cleaning**: It is very important for data to be error free and free of unwanted data. So, the data is cleansed before performing the next steps. Cleansing of data includes checking for missing values, duplicate records and invalid formatting and removing them. 2. Data Transformation: Data Transformation is transformation of the datasets mathematically; data is transformed into appropriate forms suitable for data mining process. This allows us to understand the data more keenly by arranging the 100‟s of records in an orderly way. Transformation includes Normalization, Standardization, Attribute Selection.

3**. Exploratory** data analysis(EDA)

Exploratory data analysis(EDA) is an approach to understand the datasets more keenly by the means of visual elements like scatter plots, bar plots, etc. This allows us to identify the trends in the data more accurately and to perform analysis accordingly.

EVALUATION MODEL

Model Evaluation is an integral part of the model development process. It helps to find the best model that represents our data and how well the chosen model will work in the future. Evaluating model performance with the data used for training is not acceptable in data science because it can easily generate overoptimistic and over fitted models. Performance of each classification model is estimated base on its averaged. The result will be in the visualized form. Representation of classified data in the form of graphs.

Accuracy is defined as the percentage of correct predictions for the test data. It can be calculated easily by dividing the number of correct predictions by the number of total predictions.

Training

For training, it is simple. We only need to fit our x_train (input) and y_train (output/label) data.

Testing

To evaluate the model, we need to predict the sentiment using our x_test data and comparing the predictions with y_test (expected output) data.

CONCLUSION

Crime forecast is one the latest things in the general public. Crime forecast expects to decrease Crime events. It does this by foreseeing which kind of crime might happen in future. Here, examination of crime and expectation are performed with the assistance of different methodologies. From the outcomes got we saw that the preparation season of SVM is exceptionally high subsequently it ought to be stayed away from for this dataset. Anyway which model will work best is absolutely reliant upon the dataset that is being utilized.

In this framework, we will order and bunch to work on the precision of area and example based violations. This product predicts as often as possible happening crime, particularly for specific state, and events.

REFERENCES

[1] Andrey Bogomolov, Bruno Lepri, Jacopo Staiano, Nuria Oliver, Fabio Pianesi, Alex Pentland.”Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data”,in ACM International Conference on Multimodal Interaction (ICMI 2014).

[2] ShijuSathyadevan,Devan M. S.,Surya S Gangadharan, First,”Crime Analysis and Prediction Using Data Mining” International Conference on

Networks Soft Computing (ICNSC), 2014.

[3] Sunil Yadav, Meet Timbadia, Ajit Yadav, Rohit Vishwakarma and NikhileshYadav,”Crime pattern detection,analysis and prediction,International Conference on Electronics, Communication and Aerospace

ISSN 2347–3657
Volume 10, Issue 4, 2022

Technology(ICECA), 2017.
[4] [4] AmanpreetSingh,NarinaThakur,AakankshaSharma,"A review of supervised machine learnin algorithms",3rd International Conference on Computing for Sustainable Global Development,2016.
[5] Bin Li,YajuanGuo,YiWu,JinmingChen,YuboYuan,XiaoyiZhang,"An unsupervised learning algorithm for the classification of the protection device in the fault diagnosis system",in China International Conference on Electricity Distribution (CICED),2014.
[6] Varshitha D N Vidyashree K P,Aishwarya P Janya T S,K R Dhananjay Gupta Sahana R,"Paper on Different Approaches for Crime Prediction system",International Journal of AID 2010), pp. 14-19, 2010
Engineering Research Technology (IJERT), ISSN: 2278- 0181, 2017

[7] R. Iqbal, M. A. A. Murad, A. Mustapha, P. H. ShariatPanahy, and N. Khanahmadliravi, "An experimental study of classification algorithms for crime prediction," Indian J. of Sci. and Technol., vol. 6, no. 3, pp. 4219-4225, Mar. 2013.
[8] Malathi. A,Dr. S. Santhosh Baboo,"An Enhanced Algorithm to Predict a Future Crime using Data Mining",International Journal of Computer Applications (0975 – 8887) Volume 21– No.1, May 2011.
[9] T. Beshah and S. Hill, "Mining road traffic accident data to improve safety: role of road-related factors on accident severity in Ethiopia," Proc. of Artificial Intell. for Develop. (