

Crime Rate Analysis Classification using Machine Learning

¹G Geetha Devi, ²T. Akhila Datta, ³Nidigonda. Hari Priya

¹Assistant Professor, Department of Information Technology, Bhoj Reddy Engineering College for Women, Hyderabad, India

² & ³Students, Department of Information Technology, Bhoj Reddy Engineering College for Women, Hyderabad, India

ABSTRACT

Article Info

Volume 9, Issue 5

Page Number : 621-625

Publication Issue

September-October-2022

Article History

Accepted : 10 Oct 2022

Published : 30 Oct 2022

Crime is obvious that the rate of crimes were increasing day by day in all societies in world, but we personally do believe that there are a lot which can be done by both the governments and the individuals to reduce the crimes in communities. Crime analysis is a well-organized way of detecting and examining patterns and trends in crime. We should give utmost importance to study the reasons behind the crimes, so that we can prevent various crimes occurring and we can be able to find suitable solutions to prevent them. When people cannot find work, they have all the free time in the world. They think of crimes as a shortcut to obtaining and processing the riches of life, without any hardwork. To my mind, the overwhelming majority of people tend to participate in activities assisting the government to keep the society a safe place for their own families and the others and for all age groups. Our main aim of this project is to distinguish various crimes using clustering techniques based on the occurrences and regularity. In this project, the crime data is classified using the Support Vector Machine, Decision Tree, Random Forest Algorithm. This proposed system can indicate the areas which has more probability of occurring crimes so that we can easily identify the crimes based on the previous history and we can take measures to prevent the occurrences of crimes.

Keywords: Clustering Techniques, Criminal Intent, Data Mining

I. INTRODUCTION

Crime is increasing considerably day by day. Crime is among the main issues which are growing continuously in intensity and complexity. Crime patterns are changing constantly because of which it is difficult to explain behaviors in crime patterns. Crime is classified into various types like kidnapping, theft murder, rape etc. The law enforcement agencies collect the crime data information with the help of information technologies. 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 . It is neither uniform nor random. 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. The crime

activities have been increased at a faster rate and it is the responsibility of police department to control and reduce the crime activities. Crime prediction and criminal identification are the major problems to the police department as there are tremendous amount of crime data that exist. There is a need of technology through which the case solving could be faster

II. RELATED WORK

Many researches have been done which address this problem of reducing crime and many crime-predictions algorithms has been proposed. Below are some researches:

Marchant, R., Haan, S., Clancey, G., Cripps : Applying machine learning to criminology-semi parametric spatial demographic Bayesian regressions. A comparison of classification algorithms, Naive Bayes and decision tree was performed with an data mining software, WEKA. The datasets for this study was obtained from USCensus. The pattern of road accidents 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%.

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 Engineering Research Technology (IJERT), ISSN: 2278-0181, 2017. 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.

Sunil Yadav, Meet Timbadia, Ajit Yadav, Rohit Vishwakarma and Nikhilesh Yadav, "Crime pattern detection, analysis and prediction, International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2017. 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.

III. PROPOSED SYSTEM

This project has undergone the following process: Initially, the data set is prepared manually. After identifying the relationships and visualizing the data, we create a regression model for forecasting the percapita. For this model, we have used Multi Linear regression model. Other models such as the Linear Regression and Logistic Regression models were also tested, but the Multi Linear regression produced the minimal error while training the model. This regression model predicts the percapita of Crime rate that is going to be happen in future by taking different parameters.

Logistic Regression: In statistics, the logistic model (or logit model) is used to model the probability of a certain class or event existing such as pass/fail, win/lose, alive/dead or healthy/sick. Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, logistic regression (or logit regression) is estimating the parameters of a logistic model (a form of binary regression). Mathematically, a binary logistic model has a dependent variable with two possible values, such as pass/fail which is represented by an indicator variable, where the two values are labeled "0" and "1".

KNN: In pattern recognition, the k-nearest neighbors algorithm (k-NN) is a nonparametric method used for classification and regression. In both cases, the input

consists of the k closest training examples in the feature space. The output depends on whether k -NN is used for classification or regression. In k -NN classification, the output is a class membership. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor. In k -NN regression, the output is the property value for the object. This value is the average of the values of k nearest neighbors. Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MLR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable. So far, we have seen the concept of simple linear regression where a single predictor variable X was used to model the response variable Y . In many applications, there is more than one factor that influences the response. Multiple regression models thus describe how a single response variable Y depends linearly on a number of predictor variables. In essence, multiple regression is the extension of ordinary least-squares (OLS) regression that involves more than one explanatory variable.

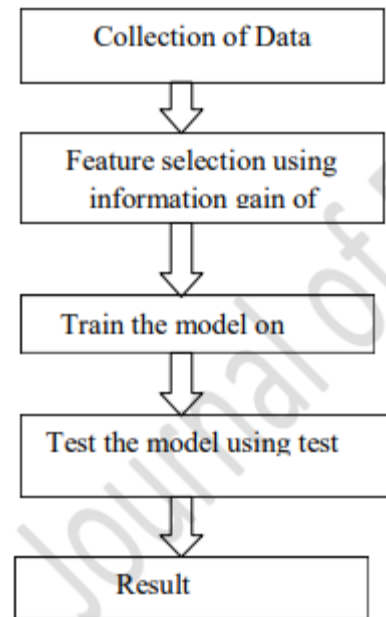


Fig 1 : System Architecture

IV. RESULTS

This module mostly center around examination of wrongdoings in different state. The information sources and techniques used to direct anticipating incorporate different sort wrongdoing measurements, overview of the overall individuals information, writing surveys and factual models that extrapolate wrongdoing patterns into what's to come. Calculations models that portray the way of behaving of see past qualities can be utilized to figure future wrongdoing patterns by projecting a period series examination of wrongdoing patterns into what's to come. Any prescient model undertaking to show a connection between specific indicator and a reliant variable. To guarantee the more noteworthy precision those models should distinguish and foresee the extension and nature of various variables that will influence wrongdoing and exploitation later on. in this part summarizes the paper and make aware about the future crime based on algorithms and crime data set we find out the crime rate in various section like age based, male vs female, area based and monthly crime rates. The data sources and methods used to guide forecasting include various type crime statistics, survey of the general people data, literature reviews and statistical models that extrapolate crime trends

into the future. Algorithms models that describe the behavior of observe past values can be used to forecast future crime trends by projecting a time series analysis of crime trends into the future. Any predictive model endeavor to show a relationship between certain predictor and a dependent variable. To ensure the greater accuracy those models must identify and predict the scope and nature of a number of factors that will influence crime and victimization in the future. This research paper about future crime rate predictions are much more specific and precise. The accuracy table of different algorithm accuracy in this part summarizes the paper and make aware about the future crime based on algorithms and crime data set we find out the crime rate in various section like age based, male vs female, area based and monthly crime rates. The data sources and methods used to guide forecasting include various type crime statistics, survey of the general people data, literature reviews and statistical models that extrapolate crime trends into the future. Algorithms models that describe the behavior of observe past values can be used to forecast future crime trends by projecting a time series analysis of crime trends into the future. Any predictive model endeavor to show a relationship between certain predictor and a dependent variable. To ensure the greater accuracy those models must identify and predict the scope and nature of a number of factors that will influence crime and victimization in the future. This research paper about future crime rate predictions are much more specific and precise. The accuracy table of different algorithm accuracy.

V. CONCLUSION

With the assistance of AI innovation, it has become simple to figure out connection and examples among different information's. The work in this task chiefly rotates around foreseeing the kind of wrongdoing which might occur assuming that we know the area of where it has happened. Utilizing the idea of AI we have assembled a model utilizing preparing informational index that have gone through information cleaning and information change. The

model predicts the sort of wrongdoing with exactness of 0.789. Information representation helps in investigation of informational index. The charts incorporate bar, pie, line and disperse diagrams each having its own qualities. We produced chart and found intriguing insights that aided in understanding Indias violations datasets that can help in catching the variables that can help in guarding society.

VI. REFERENCES

- [1]. K-means Clustering - <https://nevonprojects.com/crime-rate-prediction-using-k-means/>
- [2]. Mehmet Sait, and Mustafa Gok. "Criminal prediction using Naive Bayes theory." Springer 28.9 (2016).
- [3]. Ahishakiye, E., Anisha, C., Dhanashree., and , I.,2017. Crime Prediction Using Decision Tree(J48)Classification Algorithm. International Journal Computer and Information Technology, 6(03) 2017.
- [4]. Anisha., C.Dhanashree, and A.Arпита ,. Application for analysis and prediction of crime data using data mining. International journal of advanced computational engineering and networking, 2017.
- [5]. Sivaranjani, S., S. Sivakumari, and M. Aasha. "Crime prediction and forecasting in Tamilnadu using clustering approaches." Emerging Technological Trends (ICETT), International Conference on. IEEE, 2016.
- [6]. Ahishakiye, A., Taremw, and D., Omulo, E.O., 2017. A Performance Analysis of Business Intelligence Techniques on Crime Prediction. International Journal Computer and Information Technology, 6(03) 2017.
- [7]. Mahmud, Nafiz, et al. "Crimecast: A crime prediction and strategy direction service." 19th International Conference on. IEEE, 2016.
- [8]. Babakura A, Sulaiman MN, Yusuf MA. Improved method of classification algorithms

for crime prediction. In Biometrics and Security Technologies (ISBAST), IEEE 2014.

- [9]. Almanie T, Mirza R, Lor E. Crime prediction based on crime types and using spatial and temporal criminal hotspots. International journal of data mining and knowledge management (2015).
- [10]. Kiani, Rasoul, Siamak Mahdavi, and Amin Keshavarzi. "Analysis and prediction of crimes by clustering and classification." international journal of advanced research in artificial intelligence (2015).

Cite this article as :

G Geetha Devi, T. Akhila Datta, N. Hari Priya, "Crime Rate Analysis Classification using Machine Learning", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 9 Issue 5, pp. 621-625, September-October 2022.

Journal URL : <https://ijsrst.com/IJSRST2182942>