Predictive Modeling for Topographical Analysis of Crime Rate

Prof. Vrushali V. Kondhalkar[1], Antara More[2], Bhawana Singh[3], Jaspreet Singh Rahi[4], Sudipti Ranjan[5]

[1]Prof. Vrushali V. Kondhalkar, Dept. Of Computer Engineering, Pune, Maharashtra, India
[2]Antara More, Dept. Of Computer Engineering, Pune, Maharashtra, India
[3]Bhawana Singh, Dept. Of Computer Engineering, Pune, Maharashtra, India
[4]Jaspreet Singh Rahi, Dept. Of Computer Engineering, Pune, Maharashtra, India
[5]Sudipti Ranjan, Dept. Of Computer Engineering, Pune, Maharashtra, India

---***---

Abstract - *Criminal activities are increasing all over the world. It is important to reduce crime as it directly affects the country's economic growth. Therefore, there is an urgent need for security agencies to fight and reduce crime in the community. The proposed system helps us to detect crime and resolve criminal cases quickly based on data collected using machine learning strategies. The system helps to predict the type of crime in a particular area based on crime patterns. In this project, we will be using a machine learning method. Contains important information about crime reporting such as date, type of crime, location of the crime, etc. The data is downloaded from a database called kaggle.com and is pre-processed so that we can extract the most important natural features of crime reporting such as roads or a few places, dates, and times, and areas with a higher crime rate than others. This data is used as an incentive to predict and resolve crime at an instant rate. This project will help us to find a way to improve the crime detection system, the type of crime that will occur in a particular area, and the way to improve investigative efforts of any kind of crime.*

Keywords: Machine learning, Crime prediction

1. INTRODUCTION

Today crime is on the rise worldwide. It affects the quality of life and the development of economic well-being and the dignity of the nation. It directly affects the nation's financial growth by burdening the government with the financial burden due to the need for more police, and criminal justice courts. In terms of public safety, there is a need for more sophisticated ways to improve crime analysis to protect their communities. Accurate forecasts of crime help to reduce crime but remain problematic as crime relies on many complex issues. The basic pattern of crime and its relationship to the region or region helps us to identify and predict crime in a particular area. According to a previous study, it is clear that in every city there are fewer roads or areas with a higher crime rate than others. Crime can be predicted as criminals become more active and active in their comfort zone. When they succeed they try to repeat the crime in the same place. The occurrence of a crime depends on several factors such as criminal intelligence, local security, etc. Usually, Criminals choose the same location and time to try the next crime. While it may not be true in all cases, the chances of recurrence are high, as per the study, and this makes crime predictable. Predicting crime patterns is an important function in developing more effective crime prevention strategies or developing investigative efforts based on the availability of prior data such as case information, location, date, and time. Here, we use machine learning techniques to predict crime and its types in crime hotspots. Machine Learning is a form of practical intelligence that helps us to identify patterns using data analysis. There are three stages:

1) The dataset is extracted from the official site.

2) With the help of a machine learning algorithm, using python as core we can predict the type of crime that will occur in a particular area.

3) The model would be trained for prediction. The training would be done using the training data set which will be validated using the test dataset uploaded using the Kaggle website.

2. MACHINE LEARNING

For analyzing the collected data we will use a Classification algorithm i.e. Random Forest which is a Supervised Machine Learning technique that is used to identify the category of new observations based on training data.

2.1 Algorithm:

2.1.1 Random Forest Algorithm:

It uses a tree-like graph to show the possible results. If you enter a training database with objectives and features in the decision tree, it will create a set of rules. These rules can be used to make predictions. There are two stages in the Random Forest algorithm, one for random forest creation, and the other for predicting from a random forest classifier created in the first phase.

A. Random Forest Creation :

1) Randomly select "K" features from total "m" features where k << m

2) Among the "K" features, calculate the node "d" using the best split point

3) Split the node into daughter nodes using the best split

4) Repeat the a to e steps until the "l" number of nodes has been reached

5) Build a forest by repeating steps a to d for "n" number times to create "n" number of trees

B. Prediction Using Random Forest Classifier :

1) Takes the test features and uses the rules of each randomly created decision tree to predict the outcome and stores the predicted outcome (target)

2) Calculate the votes for each predicted target

3) Consider the high voted predicted target as the final prediction from the random algorithm.

3. LITERATURE REVIEW

3.1 Crime Prediction & Monitoring Framework Based on Spatial Analysis

In this, the authors, Hitesh Kumar, Reddy Toppi Reddy, Bhavna Sardinia, and Ginika Mahajana provided a framework for viewing criminal networks and analyzing them with various machine learning algorithms using various Google Maps and R packages. First, raw data sets are processed and visualized based on need. Machine learning algorithms are used to extract information from these large databases and to detect hidden connections between the data which are also used to report and detect key crime patterns for crime analysts to analyze these criminal networks through various interactions. detection of crime predictions

3.2. Analyzing Crime Through Machine Learning

In this Suhong Kim, Param Joshi, Parminder Singh Kalsi, and Pooya Taheri provide a crime-based model of Vancouver. Vancouver's crime data for the past 15 years is analyzed using two different data processing methods. Guessing machine learning models KNN and an advanced decision tree were used to determine the accuracy of crime forecasts between 39% to 44%.

3.3 Using machine learning algorithms to analyze crime data

In this case, Lawrence McClendon and Natarajan Meghanathan used Linear Regression, Additive Regression, and Decision Stump algorithms using the same set of limitations, communities, and uncommon crime databases to conduct comparative studies between violent crime patterns from this database and statistical data. real estate of the state of Mississippi provided by neighborhoodscout.com

3.4. Decision Tree Algorithm Based System for crime reporting at the University

In this model, Adewale Opeoluwa Ogunde, Gabriel Opeyemi Ogunleye, and Oluwaleke Oreoluwa proposed a program to investigate and detect criminals of any crime committed within Redeem's University. For crime detection at universities, Previous details of both crime and crime were collected from the Student and Development Services (DSSD) Unit. The data was processed in advance to obtain clean and accurate data. The Iterative Dichotomiser 3 (ID3) decision-making algorithm derived from the WEKA mining software was used to analyze and train data. The acquired model was then used to develop a system that demonstrated the hidden relationship between crime-related data, in the form of cutting trees. This result was then used as a basis for information on the development of a crime forecasting system.

3.5. Criminal Prediction Analysis in India using the Hybrid Clustering method

In this process, Dr.J. Kiran, and Kaishveen proposed a crime prediction framework based on the naïve Bayes classifier. The naïve Bayes classifier is compared to the KNN classifier. The proposed techniques are applied to Anaconda and the simulation results show that the naïve Bayes has high accuracy and a short duration of action.

3.6. Summary of Literature Review

Title	Publication and year	Author	Technical details
Crime Prediction & Monitoring Framework Based on Spatial Analysis	International Conference on Computational Intelligence and Data Science (ICCIDS 2018)	Hitesh Kumar Reddy Toppi Reddy, Bhavna Sardinia, Ginika Mahajana	The author provides a framework for visualize criminal networks and diversity analysis machine-learning algorithms using Google Maps once various packages for R.
Crime Analysis Through Machine Learning	IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON 2018)	Suhong Kim; Param Joshi ; Parminder Singh Kalsi; Pooya Taheri	Suhong Kim et al. provide a machine-based crime reporting model in Vancouver. Vancouver's crime data for the past 15 years is analyzed using two different data processing methods. Guessing machine learning models KNN and an advanced decision tree were used to determine the accuracy of crime forecasts between 39% to 44%.
Using Machine Learning algorithms to analyze crime data	An International Journal (MLAIJ) Vol.2, No.1, (March 2015)	Lawrence McClendon and Natarajan Meghanathan*	The author has used Linear Regression, Additive Regression, and Decision Stump algorithms using the same limited set of features, communities, and uncommon crime databases to conduct comparative studies between violent crime patterns from this data and actual crime statistics in the Mississippi status provided by neighbors out.
A Decision Tree Algorithm Based System for Predicting Crime in the University	Machine Learning Research 2017; 2(1): 26-34	Adewale Opeoluwa Ogunde1, *, Gabriel Opeyemi Ogunleye2, Oluwaleke Oreoluwa1	The author has proposed a system to investigate and detect criminals for any crime committed within Redeem's University. For crime detection at universities, Previous details of both crime and crime were collected from the Student and Development Services (DSSD) Unit. The data was processed in advance to obtain clean and accurate data. The Iterative Dichotomiser 3 (ID3) decision-making algorithm derived from the WEKA mining software was used to analyze and train data. The acquired model was then used to develop a system that demonstrated the hidden relationship between crime-related data, in the form of cutting trees. This result was then used as a basis for information on the development of a crime forecasting system.
Prediction Analysis of Crime in India Using a Hybrid Clustering Approach	2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (ISMAC)I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (ISMAC)	J. Kiran; K Kaishveen	J. Kiran; K Kaishveen proposed a crime prediction framework based on the naïve Bayes classifier. The naïve Bayes classifier is compared to the KNN classifier. The proposed techniques are applied to Anaconda and the simulation results show that the naïve Bayes has high accuracy and a short duration of action.

4. DRAWBACKS OF EXISTING SYSTEMS

Current strategies are based on an analysis of crime scenes or theory with integrated crime data. However, it is difficult to quantify the likelihood of a future crime based on an accurate definition of a past crime. Therefore, existing methods are not suitable for adapting to different environments and criminal practices.

5. CONCLUSION

Work on this project is mainly focused on predicting the type of crime and crime that may occur in the future. Using the concept of machine learning we create a model using a set of training data that we have encountered. Predicting crime patterns is an important function in developing more effective crime prevention strategies or developing investigative efforts based on the availability of prior data such as case information, location, date, and time. Here, we use machine learning techniques to predict crime and its types in crime hotspots.

6, REFERENCES

[1] Hitesh Kumar Reddy Toppi Reddy, Bhavna Sardinia, Ginika Mahajana, "Crime Prediction & Monitoring Framework Based on Spatial Analysis ", International Conference on Computational Intelligence and Data Science (ICCIDS 2018).

[2] Suhong Kim; Param Joshi; Parminder Singh Kalsi; Pooya Taheri, "Crime Analysis Through Machine Learning", 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON).

[3] Lawrence McClendon and Natarajan Meghanathan*, "USING MACHINE LEARNING ALGORITHMS TO ANALYZE CRIME DATA", An International Journal (MLAIJ) Vol.2, No.1, March 2015.

[4] Adewale Opeoluwa Ogunde1, *, Gabriel Opeyemi Ogunleye2, Oluwaleke Oreoluwa1, "A Decision Tree Algorithm Based System for Predicting Crime in the University ", Machine Learning Research 2017; 2(1): 26-34.

[5] Dr.J.Kiran, Kaishveen., "Prediction Analysis of Crime in India Using a Hybrid Clustering Approach",2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC), 2018 2nd International Conference.