Prediction And Monitoring Of Crime Rate Using K- Means Algorithm In Machine Learning

Dibyahash Bordoloi Head of the Department, Department of Computer Science & Engineering, Graphic Era Hill University, Dehradun, Uttarakhand India, 248002 dibyahashbordoloi@geu.ac.in

Durgaprasad Gangodkar, Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 dgangodkar@yahoo.com

Abstract- A crime is nothing but the illegal or any malpractice done by someone can by punished by the government. This is done when an individual does a cross violation of law, offence against morality or any criminal activities. The types of crime includes felonies, misdemeanors, inchoate offenses and strict liability offenses. The state and central government decides the sort of conduct. The crime rate is increasing due to several illegal activities. By proper monitoring and identification, the illegal activities can be avoided. These activities may mislead the upcoming generations to follow and get into violence. These activities are monitored and predicted by using machine learning through K means algorithm. This is a technique used for clustering of data that are used in the unsupervised machine learning. Thus the increasing crime rate can be predicted and can be monitored easily.

Keywords: Malpractice, law, Government, K-means alogorithm, unsupervised machine learning.

I.Introduction

The crimes are increasing in number due to one's selfishness. They are punished by the government based on Indian Penal Code (IPC). By predicting the crime before it occurs helps the investigation agency to avoid several consequences in future. These crimes may disturb the peaceful environment and tend others to do the same. The capability of predicting the crime at the particular period of time and location leads to provide necessary information to the law enforcement [1]. The analysis of crime data is followed through categorization of crime, analysis of crime, dissemination of the crime and evaluation of the information. Thus the increasing crime rate is predicted through the machine learning involving image processing techniques [2]. The forecasting of crime is helpful to predict the crime at the earlier stage by using different artificial intelligence techniques by identifying the pattern and reports using machine learning tools. The machine learning is a type of artificial intelligence (AI) that allows the software application to become more accurate and reliable by predicting the accurate values with lesser computational time. Here the data are feed and it produces the expected correct values as output parameter [3]. These prediction of crime can be done through predicting the crime, method of predicting the offenders, method of predicting and analyzing the identities and detecting and monitoring the crime victims [4]. The prediction and monitoring of the crime rate helps to analyze the

types of crimes done at the particular period of time. These machine learning can also help to identify the hotspot regions where the crime occurs frequently [5].

II.Proposed system

The crime done by the individual and the victim face a dominating situation in the society. Some crimes are done due to the situation and some are caused by the selfishness of an individual that hurts the other individual. The proposed system aims to provide a machine learning model to predict and analyze the crime and victims at the particular period of time. These crimes are recorded and stored for future references. They are stored in the datasets. These dataset are composed with date and time of the crime for future prediction and references. The prediction of crime using human source leads to time consuming. To obtain accurate output results of the crime the machine learning is proposed. This system enables by predicting the crime rate by accurate detection and identify the cause with relevant to the previous datasets. Thus the proposed system is highly reliable, convenient in usage and gives accurate results. This is done through the training and testing with the dataset. The crime prediction is done using Multi Linear Regression (MLR). The main advantages are to keep the historical record of the crime and it is user friendly and have lesser computational time.

III.Methodology

The machine learning is a kind of artificial intelligence that leads to perform based on the human intelligence. The types of machine learning involves supervised machine learning, unsupervised machine learning, semi-supervised learning and reinforcement learning [6]. The multi linear regression involves the process of predicting the graph between the independent variables and dependent variables. The exact crime patterns gives on the basic crime analysis. The frequency of the crime can be predicted by using the multi linear regression based on the existing data and the crime recognition.
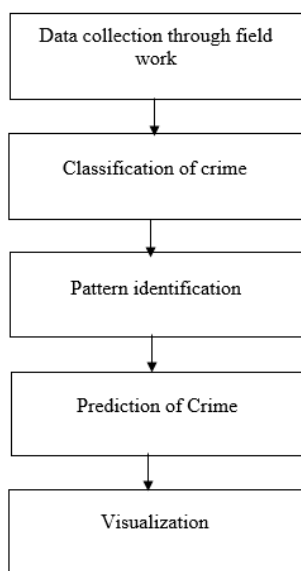

Fig 1 : Stages in crime analysis

The figure 1 shows the stages in crime analysis. The main aim of the project is to detect the crime rate based on the information obtained. The various stages involved the prediction of crime includes data collection, classification, pattern identification, prediction of crime rate and visualization to obtain the accurate results with lesser computational cost.

1. Dataset

The fundamental data collection is gathered from the field work. The features of the dataset includes date, time, year, victim, crime, situation and type of crime. The data can be collected from various online sources and televisions [7]. This is stored as dataset for future references. Hence there is no template for the data analysis of crime records. They are composed with larger volumes of structured and semi structured data. Here the object oriented programming is used to obtain accurate results.

2. Classification

The classification of the acquired data are more important to predict and identify the crimes with correct information for future references. This classification is done based on the types of crime and their consequences [8]. The judgements for the previous crimes are recorded and used for future analysis. The classification of dataset is an important parameter in the detection and identification of the crime rates.

Id	DISTRICT	STATE-todel	ASSAULT ON WOMEN	DACOITY	DOWRY DEATHS
1	Ariyalur	Tamil nadu	64.0	0.0	1.0
2	Chennai	Tamil nadu	45.0	2.0	16.0
3	Chennai rly	Tamil nadu	2.0	1.0	2.0
4	Chennai city	Tamil nadu	61.0	1.0	18.0
5	Coimbatore	Tamil nadu	39.0	11.0	1.0
6	Coimbatore City	Tamil nadu	35.0	1.0	1.0

Fig 2 : Classification of datasets

The figure 2 shows the classification of datasets based on the information obtained through field works.
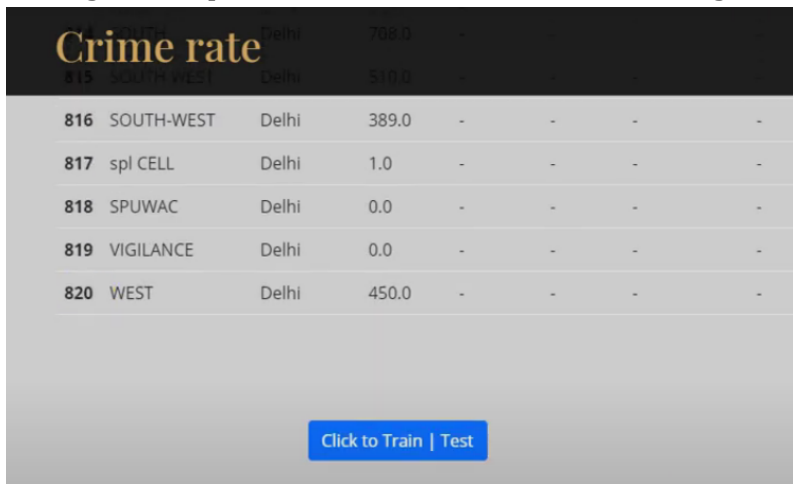
3. Preprocessing

The preprocessing is a technique of converting the obtained raw data into a useful information. The preprocessing is the essential step to obtain the accurate results. To obtain a clean data set, filters are necessary that are obtained by the threshold values.

The crime document is an unstructured form of data and information due to the varied crime and differs from one individual to another [9]. Hence the preprocessing itself helps to obtain an accurate data as output for further image processing techniques. This is proceeded to the pattern identification.

3. Pattern identification

The pattern identification is done by the testing and training process. Here the data sets are trained and tested based on the human intelligence through the machine learning techniques. The training and testing are done by the comparison from the original data to the previous data sets [10-12]. These identification through pattern gives the complete details about the crimes and the victims respectively. The figure 3 represents the outcome after the testing and training of the datasets.
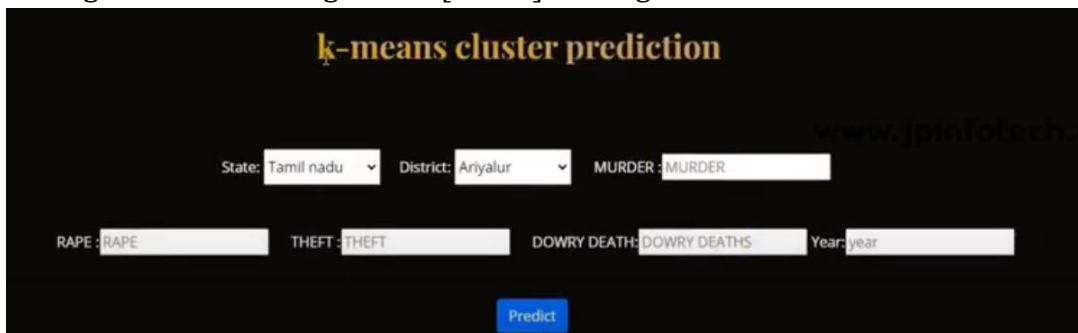

Fig 3: Training and testing

4. Prediction of crime

The prediction of crime is done through the K-means algorithm. Here the data sets are arranged in a sequential manner and the required information are identified using the key terms in the systems. It gives the complete details by identifying from the stored datasets. The accurate results are obtained through the k-means algorithm [13-15]. The figure 4 shows the k-means cluster prediction algorithm.

Fig 4: Cluster prediction

IV.Simulation results

To develop the online platform for the detection and recognition of crime and victim, an application is implemented using visual studio code. This system holds the complete programmed dataset to give the accurate results as output. These datasets are identified and gathered, preprocessed, classified based on the events occurred, pattern identification, prediction of crime and by visualization.
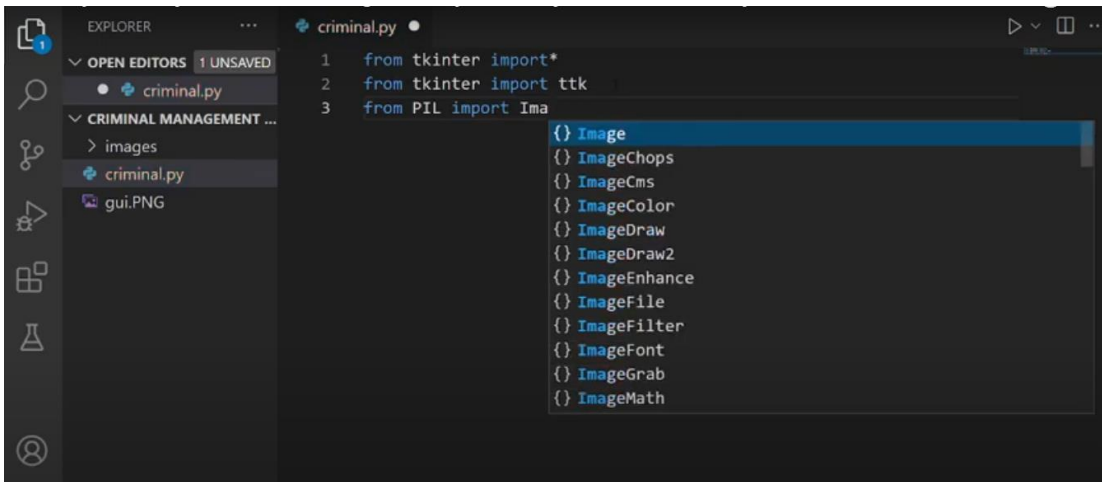
Fig 5 : Visual Studio Code

The figure 5 represents the visual studio code for the detection and identification of crime and victims using python. After the completion, this is enabled as national crime agency which gives the complete detailed information through the online source. This software performs the task of monitoring the spot areas frequency to avoid several consequences. It holds the details and problems faced by the victims and the court judgements that are allotted for the crime with regarding to the Indian Penal Code (IPC). The judgements and proceedings of the both High court and the Supreme court are listed with date and the day when it is allotted. It includes the complete advanced details of the crime with the accurate date and time as shown in figure 6
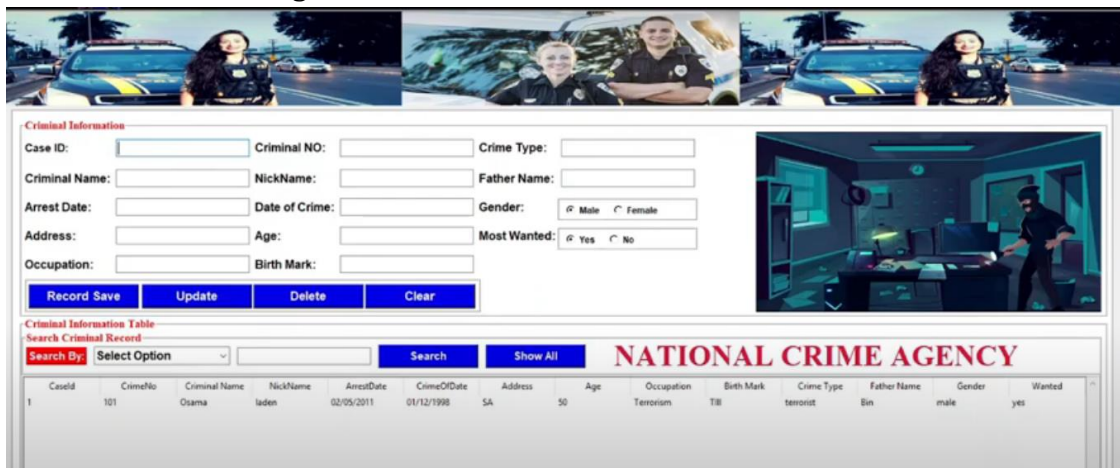

Fig 6: Criminal Management system software

V.Conclusion

The major objective of the system is to reduce the crime occurring in the day to day life to obtain a peacefully environments. Due to several situations the crimes are happening. To reduce the crime rate and to identify the crime and victim, the online detection and analysis of crime are done through the machine learning algorithm. The need for using machine learning are they function based as human intelligence. Hence by using the k-means algorithm in the machine learning, the crime rates are detected and identified with higher precision followed with lesser computational time.

References

[1] Sivanagaleela, B., and S. Rajesh. "Crime analysis and prediction using fuzzy c-means algorithm." 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2019.

[2] Catlett, Charlie, et al. "A data-driven approach for spatio-temporal crime predictions in smart cities." 2018 IEEE International Conference on Smart Computing (SMARTCOMP). IEEE, 2018.

[3] Kim, Suhong, et al. "Crime analysis through machine learning." 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, 2018.

[4] Dash, Saroj Kumar, Ilya Safro, and Ravisutha Sakrepatna Srinivasamurthy. "Spatio-temporal prediction of crimes using network analytic approach." 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018.

[5] Yi, Fei, et al. "An integrated model for crime prediction using temporal and spatial factors." 2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018.

[6] Yadav, Romika, and Savita Kumari Sheoran. "Crime prediction using auto regression techniques for time series data." 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE). IEEE, 2018.

[7] Agarwal, Shubham, Lavish Yadav, and Manish K. Thakur. "Crime Prediction Based on Statistical Models." 2018 Eleventh International Conference on Contemporary Computing (IC3). IEEE, 2018.

[8] Feng, Mingchen, et al. "Big data analytics and mining for effective visualization and trends forecasting of crime data." IEEE Access 7 (2019): 106111-106123.

[9] Elluri, Lavanya, Varun Mandalapu, and Nirmalya Roy. "Developing machine learning based predictive models for smart policing." 2019 IEEE International Conference on Smart Computing (SMARTCOMP). IEEE, 2019.

[10] Wang, Shihuai, and Kunxiaojia Yuan. "Spatiotemporal analysis and prediction of crime events in atlanta using deep learning." 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC). IEEE, 2019.

[11] Li, Zhe, et al. "Spatio-temporal pattern analysis and prediction for urban crime." 2018 Sixth International Conference on Advanced Cloud and Big Data (CBD). IEEE, 2018.

[12] Ghankutkar, Surabhi, et al. "Modelling machine learning for analysing crime news." 2019 International Conference on Advances in Computing, Communication and Control (ICAC3). IEEE, 2019.

[13] Yadav, Romika, and Savita Kumari Sheoran. "Modified ARIMA model for improving certainty in spatio-temporal crime event prediction." 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE). IEEE, 2018.

[14] Biswas, Al Amin, and Sarnali Basak. "Forecasting the trends and patterns of crime in Bangladesh using machine learning model." 2019 2nd international conference on intelligent communication and computational techniques (ICCT). IEEE, 2019.

[15] Yadhunath, Ramshankar, et al. "Identification of Criminal Activity Hotspots using Machine Learning to Aid in Effective Utilization of Police Patrolling in Cities with High Crime Rates." 2019 4th

International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS). Vol. 4. IEEE, 2019.