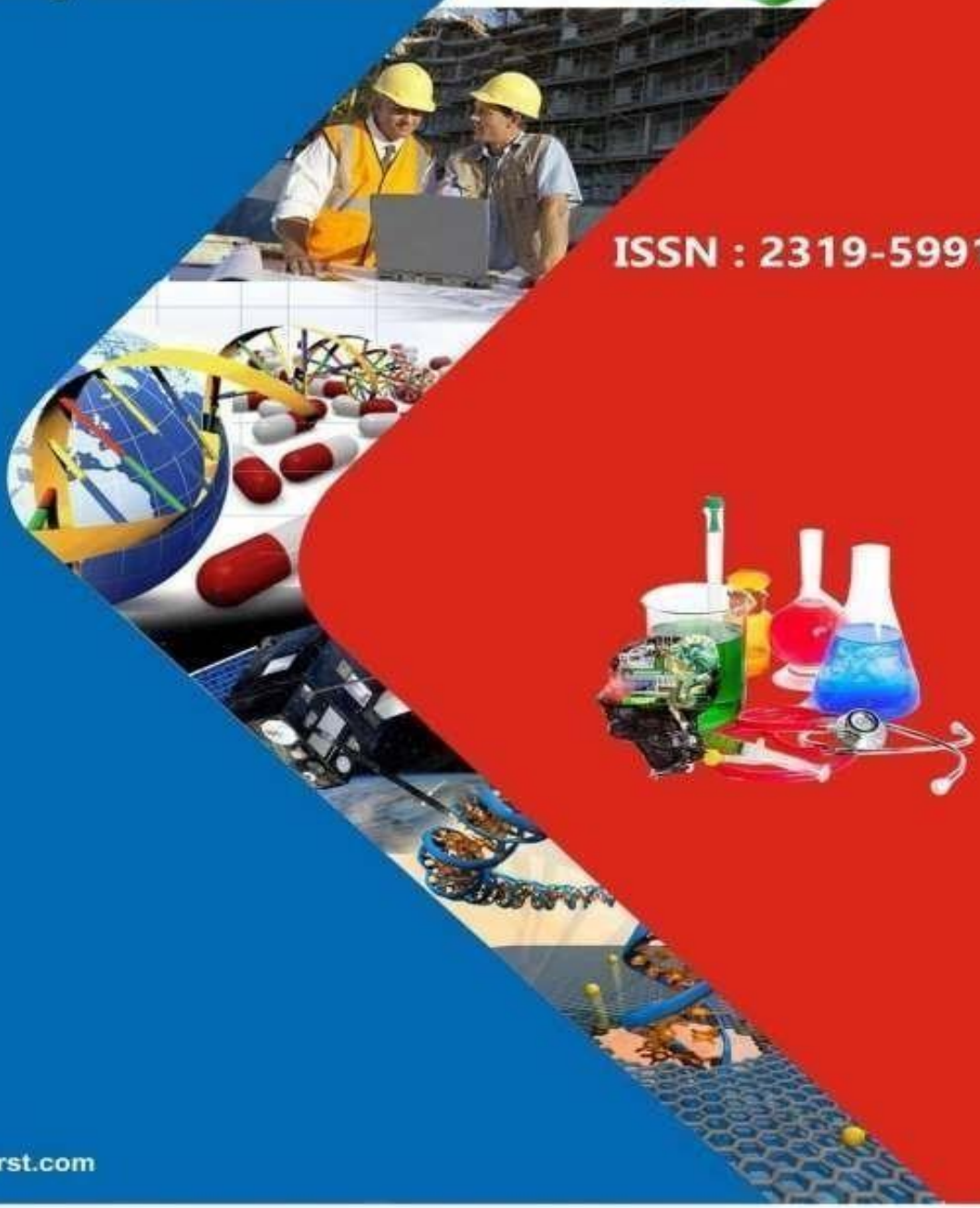
International Journal of Engineering Research and Science & Technology

IJERST

ISSN : 2319-5991



www.ijerst.com

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

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

Crime Type and Occurrence Prediction Using Machine Learning

[1] **Mr. K. LAKSHMAN REDDY,** [2] **VELAGALA NISHANTH REDDY**

[1](Associate Professor), Dept of MCA, S.V.K.P & Dr K.S. Raju Arts & Science College(A) Penugonda, W.G.District,Andhra Pradesh, klreddyin@gmail.com

[2]PG, scholar , S.V.K.P & Dr K.S. Raju Arts & Science College(A) Penugonda, W.G.District,Andhra Pradesh, vnishanthreddy48@gmail.com

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

1.INTRODUCTION

Crime has become a major thread imposed which is considered to grow relatively high in intensity. An action stated is said to be a crime, when it violates the rule, against the government laws and it is highly offensive. The crime pattern analysis requires a study in the different aspects of criminology and also in indicating patterns. The Government has to spend a lot of time and work to imply technology to govern some of these criminal activities. Hence, use of machine learning techniques and its records is required to predict the crime type and patterns. It imposes the uses of existing crime data and predicts the crime type and its occurrence bases on the location and time. Researchers

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

undergone many studies that helps in analysing the crime patterns along with their relations in a specific location. Some of the hotspots analysed has become easier way of classifying the crime patterns. This leads to assist the officials to resolve them faster. This approach uses a dataset obtained from Kaggle open source based on various factors along with the time and space where it occurs over a certain period of time. We implied a classification algorithm that helps in locating the type of crime and hotspots of the criminal actions that takes place on the certain time and day. In this proposed one to impose a machine learning algorithms to find the matching criminal patterns along with the assist of its category with the given temporal and spatial data.

2.LITERATURE SURVEY

Crime are of different type that occurs at different locations around the various geographical location. Many research scholars have been suggesting a mechanism to analyse the relationship between crime and social variables that includes unemployed individuals, earning amount, level of education and so on. Suhong Kim and Param Joshi [1] proposed two different machine learning models which is used for prediction, K nearest neighbour algorithm (KNN) Crime are of different type that occurs at different locations around the various geographical location. Many research scholars have been suggesting a mechanism to analyse the relationship between crime and social variables that includes unemployed individuals, earning amount, level of education and so on. Suhong Kim and Param Joshi [1] proposed two different machine learning models which is used for prediction, K nearest neighbour algorithm (KNN) and decision tree approach. The accuracy obtained ranges between 39 to 44 percent when predicting crime patterns and finding the crime type. Benjamin Fredrick David. H [2] imposed a data mining technique that involves evaluating and inspect large pre-existing datasets in accordance to deliver more information. The extraction of new patterns is cross checked with predefined datasets available. Shraddha S. Kavathekar [3] used association rule mining in predicting crimes. Some Machine learning algorithms including

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

Deep Neural Network (DNN) and Artificial Neural Network (ANN) have been implied. A deep neural network works more accurately using the feature level dataset. Using DNN, entirely connected convolution layers has been used in building the prediction model, mainly for multilabelled data classification. It was implemented using Tenserflow that is an API mainly designed for Deep learning technique with the dropout layers. These findings suggest that when there is more count of missing values, there is a need for pre-processing because crimes do not occur in the same manner but focuses on some particular areas. Artificial Neural Network [ANN] is based on the prognosis by trend analysis in solving problems. It comprises of enormous amount of processing constituent that works altogether in building a model. Chandy and Abraham [4] proposed a random forest classifier in extracting the features for data processing using cloud computing. The extracted features are request number, user identification, expiry time, time of arrival nd memory requirement. After feature extraction, the prediction of work load is done by using the trained data that has been perceived from the learning stage that allows to learn the details of the extracted features from user's request. Rohit Patil, Muzamil Kacchi, Pranali Gavali and Komal Pimparia [5] suggests an Apriori algorithm for frequent patterns and the result obtained from K-means is used. Due to increase in crime rate over these recent years, system has to handle an enormous amount of data which requires more time to analyse them manually. Hence, advance machine learning approaches like K means clustering has been used. A literature survey on Spatial and Temporal Hotspot prediction of crime [6] proposed a study to categorize and evaluate the location and time of the crime hotspot detection techniques by performing (SLR) Systematic Literature Review. Fuzhan Nasiri, Zakikhani, Kimiya and Tarek Zayed [7] suggested a failure prediction model that helps in detecting the corrosion in the pipelines of gas transmission. Most of the prediction model depend absolutely on the experimental tests data or involving

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

some of the limited historical data records. This helps in ignoring the corrosion from various geographical circumstances. Nikhli Dubey and Setu K. Chaturvedi [8] imposed pertinent analysis of data mining approaches for the detection of the impeding future crime. A Computational mechanism to classify the crime using machine learning techniques [9] proposed a malleable computational implementation tool to analyse the crime rate in a country helps in classifying cybercrimes. Hyeon-Woo Kang and Hang-Bong Kang [10] suggested a fusion method based on Deep Neural Network in predicting the criminal activities from the feature level data with sufficient parameters.

3. 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

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

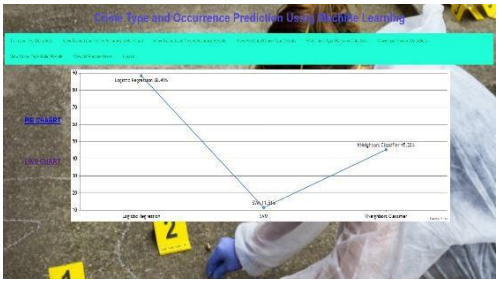
cleaned. The data is then further undergoes a splitting process. It is classified into test and trained data set. The model is trained by dataset both training and testing .It is then followed by mapping. The crime type, year, month, time, date, place are mapped to an integer for ensuring classification easier.

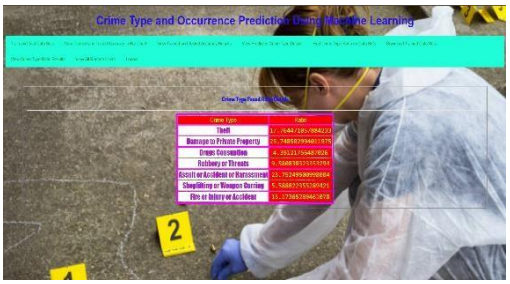
- The independent effect between the attributes are analysed initially by using Naïve Bayes. Bernouille Naïve Bayes is used for classifying the independent features extracted. The crime features are labelled that allows to analyse the occurrence of crime at a particular time and location. Finally, the 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.

4. OUTPUTSCREENS

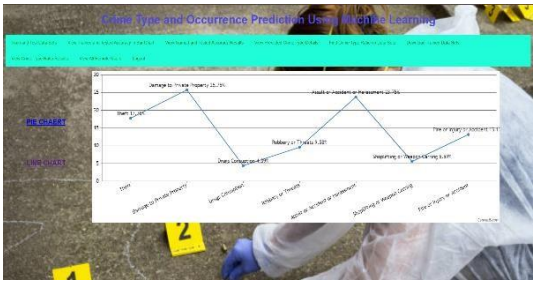




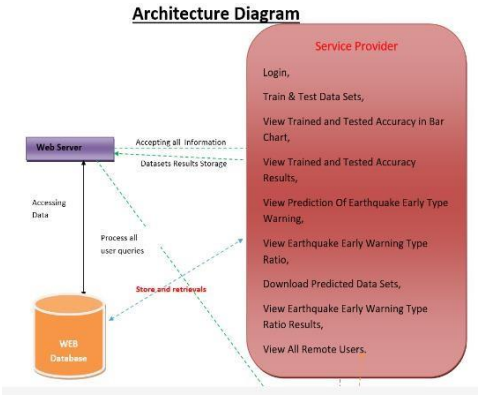
International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024



SYSTEM ARCHITECTURE



- **Web Database**
- **Web Server**
- **Remote User**
- **Service provider**

- **<u>Web Server :</u>**

➢ The web server acts as the central point of interaction for users and service providers.

➢ It hosts the application responsible for handling user requests, processing data, and providing predictions. Technologies like Flask (Python), Node.js (JavaScript), or Django (Python) can be used to develop the web server.

- **<u>Web Database</u>**

➢ The database stores historical crime data required for training machine learning models and for making predictions.

➢ Depending on the scale and complexity of the data, you might choose a relational database (e.g., PostgreSQL, MySQL) or a NoSQL database (e.g., MongoDB) for storage.

- Ensure proper indexing and normalization of data to optimize query performance.

- **<u>Remote User Interface</u>**:

➢ Remote users interact with the system through a user-friendly web interface.

➢ The interface allows users to input parameters such as location, time, and other relevant features for prediction.

➢ Results from predictions are displayed to the users in a meaningful format.

➢ The interface should be responsive and intuitive, providing a seamless user experience.

1) **Register & Login:** Once the User register successfully by providing all the required details and then he need to login

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

with Username and Password. The user can perform the following operation

2) **View user profile:** The user can view their profile data.

3) **Predict crime type :** Here, the user will give the data and will predict the ambulance positioning type whether ambulance found or not found.

- **Service Provider Interface:**
  - Service providers, such as law enforcement agencies or city governments, access the system through a dedicated interface.
  - This interface provides additional functionalities tailored to the needs of service providers, such as data analytics, reporting tools, or integration with existing systems.
  - Authentication and authorization mechanisms ensure that only authorized users can access sensitive functionalities and data.
  - Implement authentication mechanisms (e.g., OAuth, JWT) to verify the identity of users and service providers.
  - Regularly update software components and apply security patches to mitigate potential vulnerabilities.

1) **Login:** Here we can login with Username and Password

2) **Browse IOT Datasets and Train and Test Data Sets :** Here, We browse the dataset and it will train the data set and test the data set

3) **View Trained and Tested Accuracy in Bar chart:** Here, after trained and tested the accuracy of the data the result will be displayed by bar charts.

4) **View Trained and Tested Accuracy Results:** Here, it will check the accuracy of the trained and tested data.

5) **View Prediction of Crime Type:** It will view the prediction of the Crime Whether What Crime Happened.

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

6) **View Ratio Analysis:** It will view the status by the ratio analysis. It display the ratio of CrimeType

7) **Download Predicted Data Sets:** It will automatically download the Predicted datasets. It will perform after predicting the data.

8) **View All Remote Users:** Here, we can see the list of all the remote users who are registered and their status.

MODULES

Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as

Train and Test Data Sets,View Trained and Tested Accuracy in Bar Chart,View Trained and Tested Accuracy Results,View Predicted Crime Type Details,Find Crime Type Ratio on Data Sets,Download Trained Data Sets,View Crime Type Ratio Results,View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like POST CRIME DATA SETS, PREDICT CRIME TYPE, and VIEW YOUR PROFILE.

5. CONCLUSION

In this paper, the difficulty in dealing with the nominal distribution and real valued attributes is overcome by using two classifiers such as Multi nominal NB and Gaussian NB. Much training time is not required and serves to be the best suited for real time predictions. It also overcomes the problem of working with continuous target

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

set of variables where the existing work refused to fit with. Thus the crime that occur the most could be predicted and spotted using Naïve Bayesian Classification. The performance of the algorithm is also calculated by using some standard metrics. The metrics include average precision, recall, F1 score and accuracy are mainly concerned in the algorithm evaluation. The accuracy value could be increased much better b y implementing machine learning algorithms.

6. REFERENCE

[1] Suhong Kim, Param Joshi, Parminder Singh Kalsi,Pooya Taheri, “Crime Analysis Through Machine Learning”, IEEE Transactions on November 2018.

[2] Benjamin Fredrick David. H and A. Suruliandi,“Survey on Crime Analysis and Prediction using Data mining techniques”, ICTACT Journal on Soft Computing on April 2012.

[3] Shruti S.Gosavi and Shraddha S. Kavathekar,“A Survey on Crime Occurrence Detection and prediction Techniques”, International Journal of Management, Technology And Engineering , Volume 8, Issue XII, December 2018.

[4] Chandy, Abraham, "Smart resource usage prediction using cloud computing for massive data processing systems" Journal of Information Technology 1, no. 02 (2019): 108-118.

1. [5] Learning Rohit Patil, Muzamil Kacchi, Pranali Gavali and Komal Pimparia, “Crime Pattern Detection, Analysis & Prediction using Machine”, International Research Journal of Engineering and Technology, (IRJET) e-ISSN: 2395-0056, Volume: 07, Issue: 06, June 2020

[6] Umair Muneer Butt, Sukumar Letchmunan, Fadratul Hafinaz Hassan, Mubashir Ali, Anees Baqir and Hafiz Husnain Raza Sherazi, “Spatio-Temporal Crime Hotspot Detection and Prediction: A Systematic Literature Review”, IEEE Transactions on September 2020.

[7] Nasiri, Zakikhani, Kimiya and Tarek Zayed, "A failure prediction model for corrosion in gas transmission pipelines", Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, (2020).

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 2, 2024

[8] Nikhil Dubey and Setu K. Chaturvedi, "A Survey Paper on Crime Prediction Technique Using Data Mining", Corpus ID: 7997627, Published on 2014.

[9] Rupa Ch, Thippa Reddy Gadekallu, Mustufa Haider Abdi and Abdulrahman Al-Ahmari, "Computational System to Classify Cyber Crime Offenses using Machine Learning", Sustainability Journals, Volume 12, Issue 10, Published on May 2020.

[10] Hyeon-Woo Kang and Hang-Bong Kang, "Prediction of crime occurrence from multimodal data using deep learning", Peerreviewed journal, published on April 2017.