A COMPARATIVE ANALYSIS OF MACHINE LEARNING AND DEEP LEARNING STRATAGY TO ESTIMATE CRIME PATTERN PREDICTION AND PROGRESSION

R.Brindha[1] , Dr.M.Thillaikarasi[2], Dr.R.Jebakumar[3]
[1]brindha.anbuselvan@gmail.com, [2]selva.thillai69@gmail.com, [3]jebakumr@srmist.edu.in

[1]Research Scholar,Dept of CSE,Annamalai University,Chidambaram,Tamilnadu.

[2]Assistant Professor,Dept of CSE,Annamalai University,Chidambaram,Tamilnadu.

[3]Associate Professor,Dept of CTech,SRMIST,Kattankulathur,Tamilnadu.

Abstract—
Crime will continue to be a big concern to all groups and peoples around the globe, and the complexity of technology and processes that are being exploited to allow for incredibly intricate criminal activities will continue to be a significant threat as well. In order to identify and analyse crime patterns and trends, a rigorous process known as criminal analysis is used. As the usage of digital technologies grows, crime data analysts may be able to aid law enforcement officers in speeding up the investigation process. It's because of this that data mining has become a significant weapon in the fight against crime. Governments and corporations throughout the globe are using it to unearth hidden information from enormous volumes of data, and it is regulated by both. Unstructured information may be mined to uncover previously unknown, useful information, making the system more efficient. Police departments utilise analytical and predictive methods to detect offenders in predictive policing. It has been shown to be really helpful in this area.. To keep up with the ever-increasing volume of crime data, the system will have to automate the evaluation of vast amounts of data that are now stored in warehouses.Additionally, criminals are becoming more good with technology, thus it is vital to deploy new technology in order to keep the police one step ahead of them. Comparisons between deep learning and machine learning for crime prediction are made in this research. This article explains how to extract and classify features for crime pattern analysis. It is possible to begin by preprocessing the input data, and then utilising Deep component level set analysis to identify the most important properties. Gaussian naive bayes classifier and deep learning based multilayer perceptron may be utilised to classify the criminal patterns in this instance. Kaggle's standard crime data has been used in the tests.. Deep learning-based prediction model outperforms machine learning-based model, according to the results.

Index Terms—Crime, Deep component level set analysis, gausian naïve bayes, deep neural network multilayer perceptron, kaggle database
DOINumber: 10.48047/nq.2022.20.7.NQ33580 NeuroQuantology 2022;20(7):4820-4831

I. INTRODUCTION

Crime is a global problem that affects individuals in rich and developing countries alike. People's quality of life and well-being may be negatively impacted by criminal acts, which can lead to social and societal issues. Both the public and the private sectors may bear the financial burden of illegal activity. For those who travel or migrate to a new place, public safety is an important consideration. Indeed, a wide range of consequences may be associated with certain offences. Numerous variables contribute to the occurrence of criminal action, including the nature of human nature and behaviour, significant life events, and overall poverty. Unemployment and gender disparity, high population density, child labour, and illiteracy all contribute to a rise in violent crime. Cities that are growing and becoming more crowded have higher crime rates in both commercial structures and municipal housing zones. Reducing crime is critical for a community's long-term social and economic well-being because it allows its residents to have more active and meaningful lives. As a result, the analysis of crime reports and statistics is critical to ensuring the long-term viability of development. An rise in criminal prediction may be attributed to the use of computational techniques by law enforcement agencies. It is necessary to enhance the predictive algorithms used by police to locate criminals. Crime statistics from various geographic regions have been used to forecast crime types, crime rates, and crime hotspots . In addition, the information is being used to construct a variety of prototype programmes to detect crime hotspots, such as residential and commercial sites. According to the research findings, machine learning and deep learning approaches have been utilised to forecast crimes and maintain a secure environment. Nave Bayes, random forest, DNN, CNN SVM, decision tree and regression are among recent examples of effective machine learning and deep learning algorithms in crime prediction and analysis. In order to prevent crime, it is vital to be able to accurately predict criminal conduct. Historical patterns can be used to reliably determine crime rates, kinds of crimes, and hot locations, however there are computational obstacles as well as development potentials. Only a few studies have carefully examined various machine learning methodologies, although machine and deep learning methods are now the most often used. Learning algorithms are capable of handling non-linear rational data, such as crime prediction, according to studies. Faster processing of high-dimensional data allows for the extraction of data's properties. There are enormous datasets for multiple states, however the literature does not offer relative accuracy for crime prediction from large datasets despite substantial research attempts. Current research suggests that several models may improve the accuracy of violent act prediction and forecasting, particularly in high-crime regions. Using machine learning and deep learning classifiers for classification, we want to investigate crime prediction in the Indian criminal dataset.

(1) In comparison to current literature findings, the prediction accuracy has improved,

(2) An innovative approach is being used to estimate crime rates and high-intensity crime locations for the next years, allowing for more accurate forecasts .

Using data on expected overall crime rates, the police may spend more resources and execute preventative actions against criminals in the identified alleged crime hotspots, which would be beneficial. In comparison to previous results, this study's findings show that additional analysis based on a variety of machine and learning algorithms has improved the efficiency of accurate criminal behaviour prediction. In terms of performance measures and crime prediction accuracy, deep learning has been shown to outperform machine learning. A visual depiction of crime categories and statistics is also provided via exploratory data analysis. According to the paper's structure, the following are the sections: Section 2 discusses an overview of the research on crime prediction. Section 3 provides the problem statement. Section 4 shows how the proposed crime prediction method works in practise. Section 5 provides a comparative analysis of the kaggle Indian crime dataset. Finally, in Section 6, the research is summarised.

II. Related Works

[1] here machine learning methods are used to analyse the rate of cybercrime at the state level in a nation, which helps in the classification of cybercrime. By combining Indian data analytic methodologies with security analytics, we are able to analyse and identify transgressions utilising both structured and unstructured data that is fully integrated.With a 99 percent accuracy rate, the testing analysis reports serve as the work's most important strength. [2] Using a variety of machine learning methods, including neural networks, India's unlawful actions are being monitored. An investigation of the impact of unemployment and GDP on robbery, theft, and burglary has been conducted in a number of areas in India. The Granger causality approach has also been used to examine the correlation between crime rates and economic variables.There is a strong correlation between

joblessness and crime rates, which may be reduced by increasing the amount of job opportunities accessible, according to a recent experimental investigation. [3] investigates the use of machine learning to predict criminal conduct. For this research, two different ways of analysing crime data over the last 15 years in Vancouver were used. In order to anticipate crime in Vancouver, methods such as K-nearest-neighbor and boosted decision trees are used. The accuracy of crime prediction models in Vancouver varies from 39 percent to 44 percent. Predicting the kind and number of crimes that will occur in a given location is the major objective of this research. A model has been built using training data sets that were cleaned and converted as part of the data preparation and transformation procedure. Data visualisation software may be used to analyse a data collection and all of its properties.. Recognized and documented are a number of components. [4] in which the discovery of risk indicators and the development of prediction tools is helping to ensure the safety of society.. Numerous statistical and clustering approaches will be used to conduct this investigation. [5] Using neural networks and a Hybrid Deep Learning algorithm, authorities will be able to swiftly detect and analyse criminal activities because of their system's effectiveness, which will lessen the burden on the authorities. As a result of this technology being used in smart city infrastructure, crime detection will become more effective and flexible. [6] Data from social media sites is being utilised to forecast the most common sorts of social media crimes (such as cyber stalking, cyber bullying, cyber hacking, cyber harassment, and cyber scam).Data pre-processing, model creation, and predictive analysis are all components of the system. Building a model to predict crime is essential. Naive Bayes (MNB), KNN, and SVM are some of the approaches used to categorise data into different crime categories (SVM). Many machine learning techniques may be used to assist find the best value for n, as well as how accurate the N-Gram model is at different levels. We can determine the best n value and the accuracy of each system level, including Unigram accuracy, with the use of additional N-Gram language models and machine learning techniques. For these tasks, machine learning models may be improved with the use of spatial-temporal and urban features, according to [7]. In a competition on the website kaggle.com, one of these models came in first place in the category of crime categorization. Machine-learning algorithms are used in a unique way in [8] to improve patrol planning by identifying areas and times that are more likely to be riskier than the short-term average. Our AaaS design can improve existing public safety systems and platforms. The term "algorithm-as-a-service" refers to this style of architecture (AaaS). ROTA, a public safety platform built by Natal as part of the state's smart cities plan, is utilised by the agency. According to the results of an experiment, geographical granularity affects the performance of estimators. "Machine learning and computer vision" should be part of the mix, according to [9]. Our investigation was sparked after reading their work, which detailed the results of certain strategies in action. As a result of these statistical comparisons, crime detection and prevention has radically altered. Machine learning and computer vision may be used by law enforcement or other authorities to detect, prevent, and solve crimes in a more fast and accurate way. [10] simulated datasets of large metropolitan areas and various types of crimes were created, and then algorithms were used to analyse the data. Web server wamp was used to create this application using HTML and CSS as well as PHP. Algorithms that can predict the likelihood of crimes are the focus of the proposed research project. For the purposes of making predictions about criminal activity in specific locations at specific times, we also used decision trees and naive Bayesian classifiers. Faster R-CNN (Region Convolutional Neural Network) systems that recognise objects in an interior setting are demonstrated in [11]. [12] Terrorist groups' involvement in organised crime is the focus of their research, which they hope to add to the current body of knowledge. Terrorist groups that have been studied in criminal justice and terrorism literature are first examined for environmental and organisational factors. Second, it demonstrates the usefulness of inductive research techniques for investigating patterns in the criminal conduct of terrorist groups in order to develop theories about them and classify new terrorist organisations that may emerge in the future as possible threats. In [13], researchers show how a machine learning method, Random Forests, may produce more accurate long-term forecasts of crime at micro locations than other popular approaches. [14] to better fit the crime data, this study used a variety of machine learning algorithms, including logistic regression, support vector machine (SVM), Naive Bayes, k-nearest neighbours (KNN), decision tree, multilayer perceptron (MLP), random forest, and eXtreme Gradient Boosting

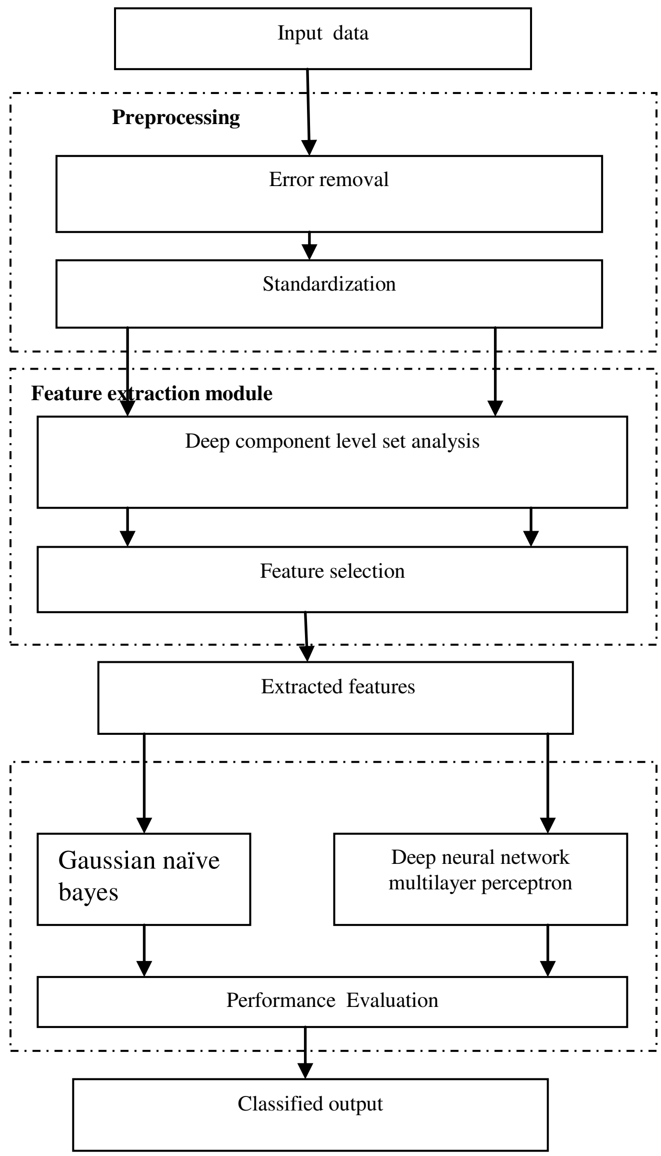
(XGBoost), in addition to time series analysis with long-short term memory (LSTM) and the autoregressive integrated moving average (ARIMA) model. The LSTM scored well on both data sets in terms of RMSE and MAE. Exploratory data analysis predicts more than 35 distinct kinds of crimes, indicating that Chicago's crime rate is reducing year after year while Los Angeles' crime rate is slightly growing. The overall rate of crime in Chicago is likely to level down in the next few years after years of steady rising. According to the ARIMA model, the total number of crimes in Los Angeles has decreased dramatically...

III. PROBLEM STATEMENT

The majority of functions in the present crime management system are performed manually, such as sending complaints, taking measures against crimes, viewing status, and so on. It is necessary to increase the number of people working in this system to search down criminal records. System security is not provided by the current system in place. The current method is time-consuming and not particularly user-friendly, and it should be replaced. As a result, under the current system, anybody who want to file a complaint against a crime must do so via the police. Retrieving historical criminal records takes a significant amount of time. Because all labour is done on paper under the existing system, it is very difficult to maintain the security of crime report data. Another significant role is the identification of errors in previously entered data as well as the cross-verification of data. These are completed manually, and the process would be time-consuming. As a result, the present system places an increased burden on the authorised person, but under our proposed system, the user may register on our website and submit reports of crime and complaints against a certain city or individual.

IV. PROPOSED WORK

A. When constructing a model that specifies and separates data classes or ideas, the goal is to use this model to forecast the class of objects whose class label has not yet been created. Artificial neural networks, decision trees, naive Bayesian classification, support vector machines (SVM), and voting techniques like as k-nearest neighbours and weighted voting may all be used to build classification models. In order to build the models, we examine a large amount of training data (i.e., data objects whose class label is unknown). A dataset may be used to generate a collection of models that can be used to predict an unknown class label. The data was retrieved from a https://www.kaggle.com/rajanand/crime-inindia/discussion/37689according to 40+ variables, the 2001 state-level crime data is categorised.



4823

Figure 1 Schematic representation of the suggested methodology

a. Pre-processing

The data must be cleaned in order to eliminate any unnecessary letters or words. Once the data has been cleaned up, formal feature extraction approaches may be used. The following are the stages involved in data preparation:

- Text extraction from the body.
- All characters are lowercased, and nonalphanumeric characters are removed.
- Text processing-related white spaces are eliminated.

Data cleansing should be done in a consistent manner. First, the data body (non-text filtering) is filtered, then paragraphs, phrases and words are inspected at the paragraph, phrase and word levels (text normalization). Unfortunately, the study was hampered by noisy data. The first stage in creating a high-quality dataset is data cleansing. The words from the topic, body, and signature components of each data set are designated by U, as shown in Equation (1)

$$U = \{V1, V2, V3\} \quad (1)$$

As shown in Equation (2), V1 is a subset of E that contains a collection of terms from an data's subject line.

$$V1 = \{r : r \in E\} \quad (2)$$

As shown in Equation (3) and (4), V2 and V3 are also subsets of E, with V2 and V3 comprising a set of words in the body and signature parts.

$$V2 = \{s : s \in E\} \quad (3)$$
$$V3 = \{t : t \in E\} \quad (4)$$

To separate these white spaces, special symbols, semicolons, commas, periods, and quotations Hat matrix standardiation are frequently employed.

$$P_s = \{b_1, b_2, \dots, b_q\} * P_t (5)$$
$$|P_t| = m$$
$$P_s \subset P_t$$

Where P_s represents the Hat matrix function

The output of hat matrix standardiation was themultiple words and extraneous words were removed from each data . Additionally, the words in each data were changed to lowercase, which aids in the stop-word reduction process.Let w1 be a collection of unwanted words in the subject line of an data. In the subject line of the $V1'$, there was useful information. As seen in Equation (6), the undesirable words from the subject line were deleted.

$$V1' = P_s - w1 \quad (6)$$

Let w2 be a group of undesired words i $V2'$ contained useful information. As shown in Equation (7), the undesirable words from the body portion were deleted .

$$V2' = V2 - w2P_s \quad (7)$$

After removing all of the unnecessary terms, the processed email (E') had just the most important and valuable words, as shown in Eq (8). E' was added to the processed dataset as a processed email.

$$E' = V1' \cup V2'^{\wedge} P_s \quad (8)$$

Finally the standardized data can be obtained by using the equation ,

$$\prod_{E'} = \mathbb{E}_s\left[\sum_{t=s+1}^{\infty} \delta^{t-s} \mathbb{I}_t\left(\min\{p_{A_t}, p_{B_t}\} - \mu\right)\right] = \frac{\delta\lambda}{1-\delta} \mathbb{E}[\min\{p_A, p_B\} - \mu]$$
(9)

b. Feature extraction

The feature extraction model in this paper is the deep component level set analysis model. As stated in Eqns (10) and (11), the DCLS model characterizes crime features as a sequence of semantically connected terms and their distributions on the opinion:

$$|\nabla\phi| = 1 \quad (10)$$

$$\text{Where } \phi \begin{cases} > 0a \in \Omega/\partial\Omega \\ = 0a \in \partial\Omega \\ < 0a \in P^2/\Omega \end{cases}$$

$$\phi_2(a,b) = p\phi_1 \begin{bmatrix} \frac{(a-x)cos\theta + (b-y)sin\theta}{p}, \\ \frac{-(a-x)sin\theta + (b-y)cos\theta}{p} \end{bmatrix}$$
(11)

The multinomial features can be defined by using the following equations,

$$\min_{Z(q)} \int p(|\nabla I(Z(q))| Z'(q)| dq$$
(12)

Where $p(|\nabla I|) = \frac{1}{1+|\nabla I|^2}$

$$p(X_i|q, \alpha) \sim Multinomial(\alpha)$$
(13)

Where α is the polynomial index

Additionally, a Dirichlet random variable may be used to model the crime value distribution, as shown in Eqn. (14):

$$E(z_1, z_2, \phi) = \int_{\Omega} \{(u - z_1)^2 H(\phi) + (u - z_2)^2 (1 - H(\phi)) + \mu |\nabla H(\phi)|\} dadb \quad (14)$$

u and z are factors that determine the crime characteristics function, which is $H(a)$:

$$H(a) = \begin{cases} 1, & a \geq 0 \\ 0, & a < 0 \end{cases} \quad (15)$$

The characteristics variables θ_{jn}, δ_{mn}a can be solved using Gibbs Sampling:

$$\begin{cases} \theta_{jn} = \frac{feature\ count_j^n + \sigma_n}{\sum_{n=1}^{N} specialized\ count_j^n + \sigma_n} \\ \delta_{mn} = \frac{feature\ count_m^n + \gamma_n}{\sum_{m=1}^{M} specialized\ count_m^n + \gamma_n} \end{cases}$$
(16)

Eqns. (10, 11) may be used to determine the specialized parameter characteristics a^* and b^*, which are specified by its specialization.

$$a^* = \frac{(a-x)cos\theta + (b-y)sin\theta}{s}$$
(17)

$$b^* = \frac{-(a-x)sin\theta + (b-y)cos\theta}{s}$$
(18)

As a consequence, opinion on the crime network is more similar, is less similar. It may use this feature to compute similarity and construct a similarity model based on it.

$$d_{Euclid}(a^*, b^*) = \left[\sum_j (x_j - y_j)^2\right] = \left[2 - 2\sum_j x_j y_j\right]$$
(19)

It is based on the notion of cluster analysis that the crime features of same weightage are aligned in an order. In viewpoint d, we utilize Equation . (20) to obtain the log frequency weight of term t.

$$v_{(a^*,b^*)} = \begin{cases} 1 + log_{10} pf_p, b & if \quad pf_p, b > 0 \\ 0 & otherwise \end{cases}$$
(20)

When few characteristics are almost identical, cosine similarity returns one, and when the characteristics are extremely distinct, it returns zero. We must normalize the similarity between them (Eqn. 21).

$$\cos(x, y) = \frac{\sum_{j=1}^{n} x_j y_j}{\sqrt{\sum_{j=1}^{n} x_j{}^2} \sqrt{\sum_{j=1}^{n} y_j{}^2}}$$
(21)

The cosine similarity between two normalised vectors, x and y, is just the dot product of their cosine similarity (Eqn. 22).

$$Cos\phi = \begin{cases} 1 & -\{(u - z_1)^2 - (u - z_2)^2 + 2\lambda(H(\phi) - H(\psi))\} \geq 0; \\ - & -\{(u - z_1)^2 - (u - z_2)^2 + 2\lambda(H(\phi) - H(\psi))\} < 0. \end{cases}$$
(22)

Where U=Ξ_{is}, Z=Ξ_{ts}, ϕ=$\hat{\beta}$, H=$\tilde{X}_{ij}$

The new variables are aligned for convenience,
Where,

$$\Xi_{is} = \sum_j (Y_{ij} - \langle Y_{i.} \rangle)(\Omega_{sj} - \langle \Omega_{s.} \rangle)$$

$$\Xi_{ts} = \sum_j (\Omega_{tj} - \langle \Omega_{t.} \rangle)(\Omega_{sj} - \langle \Omega_{s.} \rangle)$$

$$\hat{\beta} \equiv \hat{\Xi} \times_s \Sigma^{-1},$$

$$\tilde{X}_{ij} = X_{ij} \times exp - \sum_{s=1} \hat{\beta}_{is}(\Omega_{sj} - \langle \Omega_{s.} \rangle)$$

$$\tilde{X}_{ij} = X_{ij} \times \left(\frac{(\prod_{k=1}^{M} Z_{1K})^{\frac{1}{M}}}{Z_{1j}} \right)^{\frac{r_{i1}}{\sigma_i^2}}$$

The overall feature function can be represented as,
$r_{i1} = \sum_{j=1}^{M} (\log X_{ij} - \langle \log X_i \rangle)(\log Z_{1j} - \langle \log z_{1.} \rangle), \sigma_i^2$
(23)

$$Where\ \sigma_i^2 = \sum_{j=1}^{M} (\log Z_{1j} - \langle \log z_{1.} \rangle)^2.$$

From the above function the selected features can be extracted.

C.classification

It's common practise to utilise these algorithms to forecast future outcomes from historical data. To classify anything, you need to know what you're looking for. The class label may be predicted using this approach if there are enough training examples. Numerous data classification methods exist, such as the use of artificial neural networks (ANNs) and support vector machines (SVMs). An unknown class may be predicted from the data by building a collection of models. Training and test sets are created by classification to divide the dataset into two half (independent set). Machine learning algorithms must first be tested on a set of practise data before being applied to new data. Gaussian naive bayes was used as a starting point for this analysis. Using a decision tree as a model, an item's goal value (expressed in the branches) may be predicted (represented in the leaves). It is possible to use classification trees when there are few possible values for the target variable. In this case, leaves denote class labels, and branches indicate the qualities that lead to those class labels being combined. A data visualisation was shown in this manner for the first time:

$$x_k = \frac{\sum_{i=1}^{m} c_{ik} a_{ij}}{\sum_{i=1}^{m} c_{ik}}$$

(24)

Regression trees (usually real numbers) are used as the goal variable in the proposed method. A variation of Quinlan's ID3 technique, the C4.5 approach, is used to build Gaussian naive bayes The nodes of the tree are selected based on entropy.

$$E_{ik} = \begin{cases} 1 \quad if\ \|a_i - x_k\|^2 = \min_{1\le k\le z}\|a_i - x_k\|^2 \\ \quad 0 \quad , otherwise \end{cases}$$

(25)

The qualities with the highest entropy were chosen because they increase the degree of uncertainty in the result. Splitting the source set into subsets according to an attribute value test may be used to train a tree. Recursive partitioning is the technique of recursively dividing each derived subset.

$$(c, X, \propto, \lambda) = \sum_{i=1}^{m}\sum_{k=1}^{c} c_{ik}\|a_i - x_k\|^2 - \beta m \sum_{k=1}^{z} \alpha_k In\, \alpha_k - \gamma \sum_{i=1}^{m}\sum_{k=1}^{z} c_{ik}\, In\, \alpha_k - \lambda(\sum_{k=1}^{z} \alpha_k - 1) \qquad (26)$$

Where $c, X, \propto, \lambda$ represents the subset of the nodes

Recursion ends when the target variable in a node's subset has the same value across all nodes, or when splitting no longer improves the accuracy of the predictions. Classification values may also be represented as follows,

$$\alpha_k{}^{(t+1)} = \sum_{i=1}^{m} c_{ik}/m + (\beta/\gamma)\, \alpha_k{}^{(t)}\left(In\, \alpha_k{}^{(t)} - \sum_{s=1}^{z} \alpha_s{}^{(t)} In\ \alpha_s{}^{(t)}\right)$$

(27)

Deep learning will be used in the second stage of machine learning deployment. We'll focus at deep neural network multilayer peceptrons since various CNN methodologies employ diverse network designs, which is well-known. A CNN model has three layers: pooling, convolution, and the final layer of connections. Under the convolution layer, notions like weight sharing and local perception are employed to classify the input data. The continuous function of an area is therefore reflected by the local perceptions of that region. From the point of view of law enforcement, it is conceivable to make a distinction. By sharing the weights, CNNs are able to minimise the amount of time it takes to process data. So, let's pretend that dot I represents a graph generated from data on previously committed criminal acts in order to illustrate the investigation into public crime. Equation 28 provides the solution.

$$dot_i(a,b) = z^{(t)} - \left| filter_{j,i} \left\{ \alpha_k^{(t+1)} | \alpha_k^{(t+1)} < 1/m\,, k = 1, \dots, z^{(t)} \right\} off_i \right| (28)$$

4826

Moreover, a,b depicts the number of total crime factors in the list;$z^{(t)}$ illustrates the feature rate of jth formed graph using features obtained; $filter_{j,i}$ depicts the weight required for neural network; off_i depicts the offset; $delta(a)$ represents the activation function.

The pooling technique is necessary for reducing the number of features in order to obtain the opinion and increase the efficiency of training. An additional convolutional layer is then added to the final analysis result that also includes the activation function as a subject matter.

$$\propto_k^*\ {}^{\wedge\prime} = {}^{\propto_k^*}\Big/_{dot_i(a,b)\sum_{s=1}^{z^{(t+1)}} \propto_k^*}$$

(29)

Accuracy is estimated using activation function sigmoidal error rate. As a result, the estimated error function is as follows:.

$$E1 = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (30)$$

Where p_j illustrates the input signal and $E1$ depicts the significant rate of output error in terms of present layer and which is expressed using equation (31)

$P_j = \delta(E_k - 1), E = \delta(b_k - 1) * er_{j-1}$(31)

Where $E_k - 1$ is the outcome signal obtained in the final layer, and er_{j-1} is the outcome of last layer. Since the sigmoid function's monotonicity from the absolute rate, the equation is satisfied.

$\Sigma_B = \sum_{i=1}^{K} P(\omega_i)(m - m_i)(m - m_i)^t$
(32)

Which signifies that the function utilised does not cause any weakening of error in output. Moreover, for a CNN with K layers, the output error in the upper limit is expressed as follows.

$\Sigma_B < P_j \sum_{k-1}^{K} \frac{{er_{j-1}}^2}{b_{k=1}^2}$(33)

To total up, if the error formed from the activation function satisfies and expressed in the given equation (33), the crime was coorectly predicted.

Algorithm 1Crime prediction framework

Input : sequence of n data
Output: crime classifiaction

Number of iterations is denoted by the framework

Initialize$U = \{V1, V2, V3\}$ and $P_s = \{b_1, b_2, \ldots, b_q\} * P_t$

For P=1,2,…n do

1. $\prod_{E'} = \mathbb{E}_s\left[\sum_{t=s+1}^{\infty} \delta^{t-s} \mathbb{I}_t(\min\{p_{A_t}, p_{B_t}\} - \mu)\right] = \frac{\delta\lambda}{1-\delta} \mathbb{E}[\min\{p_A, p_B\} - \mu]$
2. Feature ext=semantic connection
3. $\phi_2(a,b) = p\phi_1 \begin{bmatrix} \frac{(a-x)cos\theta + (b-y)sin\theta}{p}, \\ \frac{-(a-x)sin\theta + (b-y)cos\theta}{p} \end{bmatrix} h_t : \varepsilon_t = \frac{1}{2}\sum_{i=1}^{n} D_t \left(1 - h_t(a_i, b_i) + \sum_{b \neq b_i} q_t\,(i,b) h_t(a_i, b)\right), h_t : A \times B \to [0,1]$
4. Cos sim* operative function
5. ${\alpha_k}^{(t+1)} = \sum_{i=1}^{m} c_{ik}/m + (\beta/\gamma)\, {\alpha_k}^{(t)} \left(In\, {\alpha_k}^{(t)} - \sum_{s=1}^{z} {\alpha_s}^{(t)} In\ {\alpha_s}^{(t)}\right)$
6. $Update\ function$
7. $Output:\ \Sigma_B < P_j \sum_{k-1}^{K} \frac{{er_{j-1}}^2}{b_{k=1}^2}$

V. PERFORMANCE ANALYSIS

In this part of the proposed research, an efficiency of a suggestd technique based on deep and machine learning methods developed was analyzed under python environment.

A	B	C	D	E	F	G	H	I	J	K	L	M	N
crime_type	crime_date	crime_year	crime_reason	suspect_name	victim_name	victim_age	crime_location	suspect_age	suspect_sex	suspect_origin	suspect_facial_feature	suspect_weapon_type	victim_nationality
Murder	22June.	2010	Criminal Grou	Mudassar	Ankur	Medium	Delhi	Young	Female	Rajasthan	Limp	Metal Rod	India
Sexual Abu	17June.	2009	Government (	James	Surya	Young	Delhi	Medium	Male	Haryana	Tattoo (Bicep)	Knife	India
Murder	06June.	2012	Political Grou	Irfan	sunny	Young	Delhi	Young	Male	Maharashtra	Cut (Lip)	Gun	India
Sexual Abu	31May.	2011	Military Offici	Bunty	Rishabh	Medium	Delhi	Medium	Male	Rajasthan	Tattoo (Bicep)	Unarmed	Sri Lanka
Assault	28May.	2009	Government (	Kamla	Sahil	Young	Delhi	Young	Female	Haryana	Cut (Lip)	Chain	India
Sexual Abu	27May.	2012	Political Grou	Raju	Akash	Elder	Delhi	Medium	Male	Bihar	Scar (Head)	Others	India
Assault	27May.	2011	Political Grou	Ali Shaaban	Akshat	Elder	Delhi	Young	Male	Maharashtra	Limp	Others	India
Rioting	21May.	2010	Unknown Fire	Mahad Salad A	Omkar	Elder	Mumbai	Medium	Male	Ghaziabad	Cut (Lip)	Gun	Iraq
Assault	09May.	2009	Political Grou	Justin	Raunak	Young	Delhi	Elder	Male	Uttar Pradesh	Scar (Head)	Gun	Russia
Sexual Abu	07May.	2008	Political Grou	Ali Ahmed Abc	Tauseen	Elder	Kolkata	Elder	Male	Rajasthan	Tattoo (Bicep)	Others	India
Assault	07May.	2007	Political Grou	Rajesh Mishra	Manoj	Medium	Delhi	Medium	Male	Rajasthan	Scar (Head)	Unarmed	India
Robbery	24April.	2006	Criminal Grou	Abukar Hassan	Parth	Elder	Kolkata	Medium	Male	Haryana	Scar (Head)	Others	India
Rioting	20April.	2005	Government (	Lee	Manish	Young	Kolkata	Young	Male	Rajasthan	Cut (Lip)	Metal Rod	Brazil
Mugging	20April.	2004	Political Grou	Rémi Ochlik	Shyam	Medium	Delhi	Elder	Male	Uttar Pradesh	None	Unarmed	India
Robbery	19April.	2011	Military Offici	Marie Colvin	Nikhil	Medium	Mumbai	Medium	Male	Bihar	Cut (Lip)	Chain	India
Assault	26March.	2002	Military Offici	Rami al-Sayed	Ashwin	Elder	Pune	Young	Male	Ghaziabad	Limp	Others	India
Sexual Abu	22March.	2001	Military Offici	Mario Randolf	Palash	Young	Mumbai	Elder	Male	Bihar	Tattoo (Bicep)	Unarmed	India
TSNS	19March.	2000	Military Offici	Mazhar Tayyar	Prathamesh	TSNS	Bangalore	TSNS	Male	Uttar Pradesh	TSNS	TSNS	UAE
Robbery	18March.	2009	Military Offici	Hassan Osman	Nithin	Elder	Delhi	Medium	Male	Rajasthan	Tattoo (Bicep)	Gun	India
Mugging	18March.	2012	Political Grou	Nicholas	Nishant	Young	Mumbai	Medium	Male	Haryana	Limp	Unarmed	Iraq
Rioting	07March.	2010	Political Grou	Ahmed Ismail	Ajith	Elder	Mumbai	Young	Male	Uttar Pradesh	Cut (Lip)	Others	India
Sexual Abu	21Feb.	2011	Criminal Grou	Anas al-Tarsha	Siddharth	Medium	Kolkata	Young	Male	Uttar Pradesh	None	Chain	India
Mugging	14Feb.	2005	Government (	Nicole	Rajeev	Elder	Delhi	Medium	Male	Haryana	Cut (Lip)	Knife	India
Rioting	11Feb.	2010	Criminal Grou	Rita	Karan	Elder	Pune	Elder	Female	Haryana	Scar (Head)	Knife	India
Assault	18Jan.	2009	Local Residen	Stacy	Judy	Young	Kolkata	Medium	Male	Maharashtra	Tattoo (Bicep)	Knife	Nicaragua

Figure 2 Sample dataset

Figure 2 represents the sample dataset which are retrieved from Kaggle. 4827

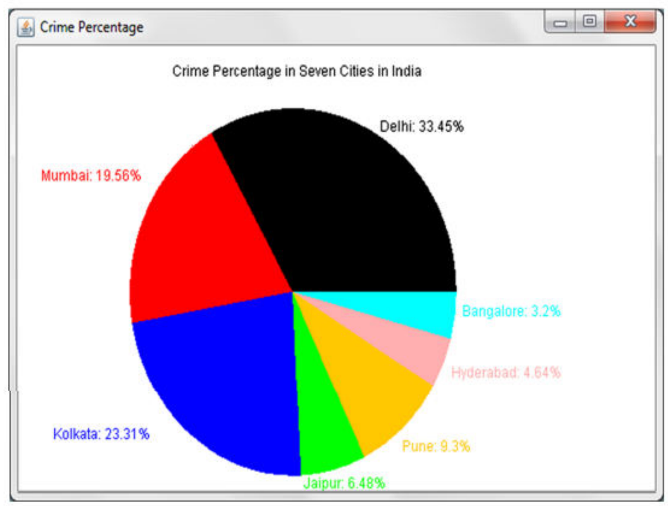


Figure 3 Crime rates in % of top seven Indian cities

We implement a conventional Gaussian naïve bayes and deep neural network multilayer perceptron for crime identification. Both the classifier can predict the top 7 cities that have high range of crime in India as depicted in figure 3.

Specifications like accuracy, precision, F1 score, and recall are used to verify the recommended technique's behaviour. For this assessment, there are four concepts to take into account: It's important to know the difference between a true positive, true negative, a false positive, and a false negative (FN). The algorithm's definition of "TP" refers to the data values that are correctly identified as positive. The data numbers that the system correctly classifies as negative are referred to as TN. Figures that are recognised as positive but not the precise ones constitute FP. It is defined as the numbers that are identified as negative but not the precise ones.

- TP means crime is present
- TN means crime is not present and is not identified.
- FP means crime is not present but is identified.
- FN means crime is present but is not identified.

We also contrast the suggestedtechnique with a few of the existing techniques concerning these parameters.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Negative + False\ Positive + True\ Negative}$$

(34)

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

(35)

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

(36)

$$F_1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (37)$$

$$TPR = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

(38)

$$FPR = \frac{False\ Postive}{False\ Positive + True\ Negative}$$

(39)

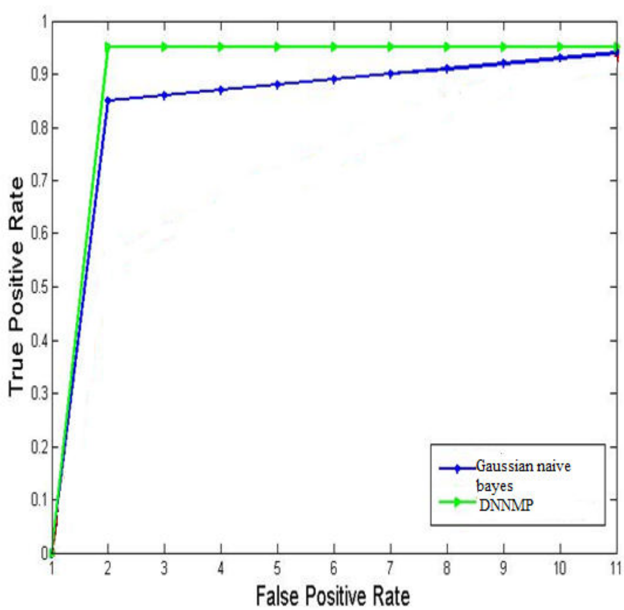
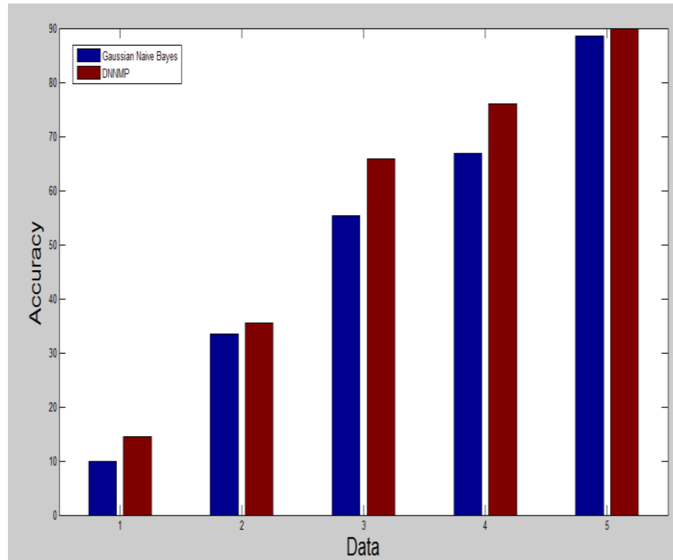


Figure 4 False positive rate Vs. True positive rate

Figure 4 shows that among two of the classifiers, DNNMP has the greatest AUC OF 98% and GNB acquires slighlt variation over DNNMP which is 96% .



4828

Figure 5 Number of images Vs. Accuracy

Figure 5 shows the proposed classification method, showing a maximum accuracy yield of 90%, which is better than GNB(80.62)

Hence among the two classifier the DNNMP expressed satisfied results than other classifier. To prove the efficiency of the outperformed DNNMP it can be compared with the [15] existing methodology.

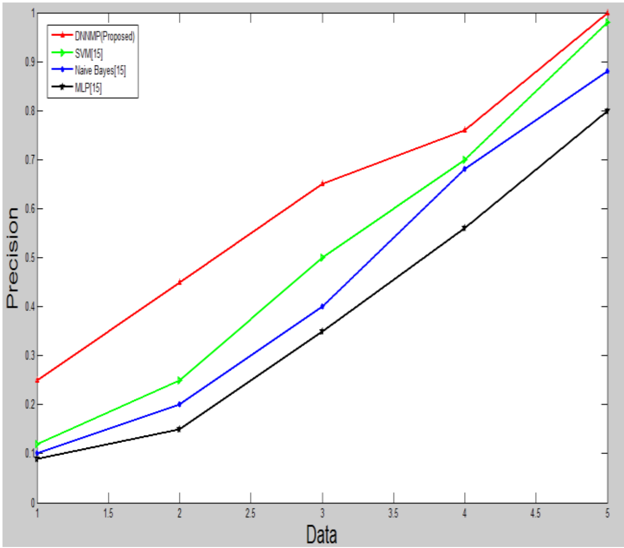


Figure 6 Number of images Vs. Precision

As can be seen in Figure 6, the developed model has a far better accuracy rate (99%) than the one currently in use.

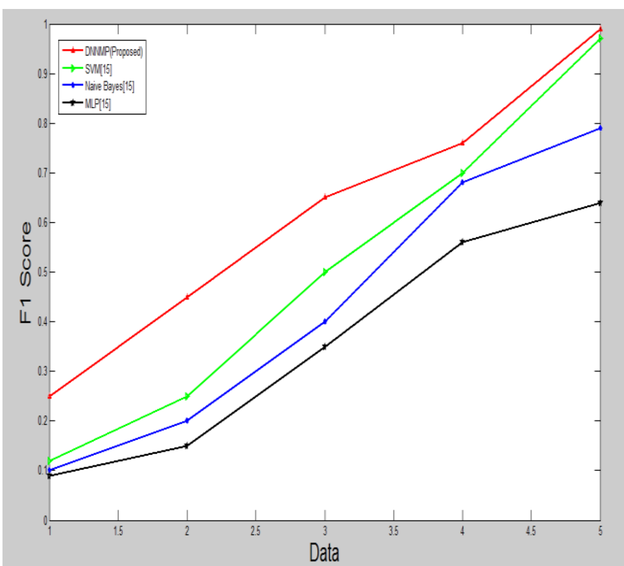


Figure 7 Number of images Vs. F1 score

Figure 7 shows an F1 rating for the solution. The results demonstrate that, according to the suggested strategy, the coefficient of F1 score was 98 percent.

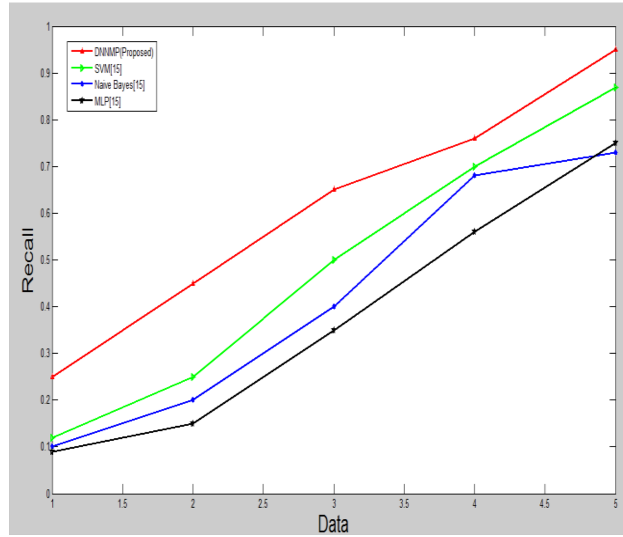


Figure 8 Number of images Vs. Recall

A maximum recall yield of 96 percent is shown in Figure 8, which is superior than traditional approaches..

The overall crime prediction error rate can be illustrated using RMSE as follows,

The RMSE is defined as:

$$RMSE = \sum_{i=0}^{a-1} \sum_{j=0}^{b-1} \sum_{s=0}^{2} ||g\,(i,j,s) - f\,(i,j,s)\,||^2 \quad (40)$$

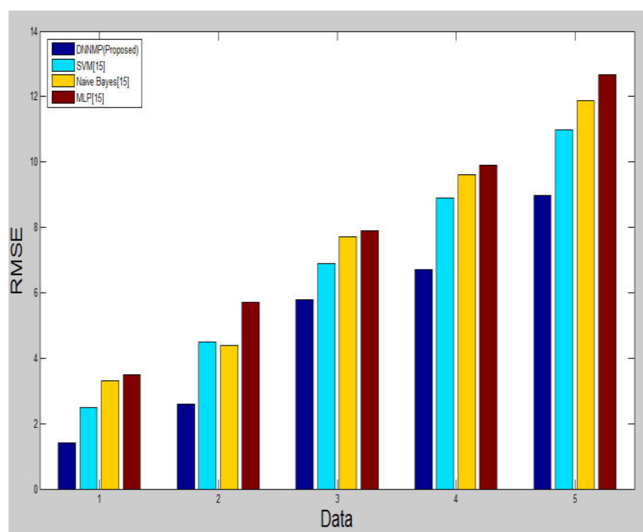


Figure 9 Data Vs. RMSE

Over crime prediction the suggested DNNMP has low amount of error rate which is an average of 8.2% that is very low when compared to other existingmechanisms.

Thus from the results obtained it was revealed that the suggested DNNMP express satisfied results when compared to other exsiting mechanisms.

VI. CONCLUSION

As a result of issues including poverty, migration, unemployment, a lack of education and a culture of corruption, crime in India is on the rise. It is difficult and time-consuming for law enforcement authorities to conduct a manual search of the database of offenders or use the services of a computer data analyst to do so. Computerized methods for crime prediction are being proposed to help fight crimes and identify perpetrators. To test DNNMP's ability to forecast crime data, the system was able to accurately predict an unknown category of crime data with a 90% accuracy rate, which is acceptable enough that the system may be depended on for crime prediction. Our suggested data mining technology may be used by law enforcement authorities to speed up the investigative process. By analysing and sifting through a large amount of data, the Suggested can speed up the process of solving crimes. As a result, DNNMP may assist Indian law enforcement authorities in their efforts to protect people.

REFERENCES

[1] R. Ch, T. R. Gadekallu, M. H. Abidi, and A. Al-Ahmari, "Computational system to classify cyber crime offenses using machine learning," *Sustainability,* vol. 12, p. 4087, 2020.

[2] M. Mittal, L. M. Goyal, J. K. Sethi, and D. J. Hemanth, "Monitoring the impact of economic crisis on crime in India using machine learning," *Computational Economics,* vol. 53, pp. 1467-1485, 2019.

[3] S. Kim, P. Joshi, P. S. Kalsi, and P. Taheri, "Crime analysis through machine learning," in *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2018, pp. 415-420.

[4] S. R. Bandekar and C. Vijayalakshmi, "Design and analysis of machine learning algorithms for the reduction of crime rates in India," *Procedia Computer Science,* vol. 172, pp. 122-127, 2020.

[5] S. Chackravarthy, S. Schmitt, and L. Yang, "Intelligent crime anomaly detection in smart cities using deep learning," in *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*, 2018, pp. 399-404.

[6] Z. Abbass, Z. Ali, M. Ali, B. Akbar, and A. Saleem, "A framework to predict social crime through twitter tweets by using machine learning," in *2020 IEEE 14th International Conference on Semantic Computing (ICSC)*, 2020, pp. 363-368.

[7] J. Borges, D. Ziehr, M. Beigl, N. Cacho, A. Martins, S. Sudrich*, et al.*, "Feature engineering for crime hotspot detection," in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 2017, pp. 1-8.

[8] A. Araújo, N. Cacho, L. Bezerra, C. Vieira, and J. Borges, "Towards a crime hotspot detection framework for patrol planning," in *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 2018, pp. 1256-1263.

[9] N. Shah, N. Bhagat, and M. Shah, "Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention," *Visual Computing for Industry, Biomedicine, and Art,* vol. 4, pp. 1-14, 2021.

[10] S. Ramasubbareddy, T. Aditya Sai Srinivas, K. Govinda, and S. Manivannan, "Crime prediction system," in *Innovations in Computer Science and Engineering*, ed: Springer, 2020, pp. 127-134.

[11] S. Saikia, E. Fidalgo, E. Alegre, and L. Fernández-Robles, "Object detection for crime scene evidence analysis using deep learning," in *International Conference on Image Analysis and Processing*, 2017, pp. 14-24.

[12] J. Semmelbeck and C. Besaw, "Exploring the determinants of crime-terror cooperation using machine learning," *Journal of quantitative criminology,* vol. 36, pp. 527-558, 2020.

[13] A. P. Wheeler and W. Steenbeek, "Mapping the risk terrain for crime using machine learning," *Journal of Quantitative Criminology,* vol. 37, pp. 445-480, 2021.

[14] W. Safat, S. Asghar, and S. A. Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning

Techniques," *IEEE Access,* vol. 9, pp. 70080-70094, 2021.

[15] Safat, W., Asghar, S., & Gillani, S. A.. Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques. *IEEE Access*, *9*, 70080-70094, 2021.