International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Performance Analysis for Crime Prediction and Detection Using Machine Learning Algorithms

R. Ganesan[1], Dr. Suban Ravichandran[2]

Submitted: 06/12/2023 **Revised**: 17/01/2024 **Accepted**: 27/01/2024

Abstract: Crime prediction system is essential to identifying and analyzing patterns and trends in crime. In recent era, the main aim of government is to reduce the crime events in all countries. Nowadays a greater number of crimes are perpetrated which affects the typical human's life. Hence, crime prediction is a vital task. Various factors such as criminal behavior, age, place etc. are used to predict the crime pattern. The preeminent objective of this work is to identify the crime groups by using different years of dataset. Machine learning techniques are used to analysis and discover the crime patterns. In this work, a comparative analysis of various techniques like the Decision Tree, Gaussian Naïve Bayes, Linear Support Vector Classifier, Logistic Regression, and Stochastic Gradient Descent are used to suggest the crime pattern mainly based on time and place. Also, various existing feature selection techniques like Linear Regression, Ridge Regression and Polynomial Regression are used to select the salient features from the dataset. The performance of each technique is validated with various performance metrics such as accuracy, precision, recall and f1 score. The Experimental results show that the Linear Regression with Stochastic Gradient Descent model performing better than other techniques.

Keywords: *Crime prediction, Machine Learning, Decision Tree, Linear Regression, detection*

1. Introduction

In recent times, crime rates are increased day by day in various ways such as robbery, drugs, murder, etc. Further, crime activities are varying from zone to zone. Hence, it is an essential to solve the crime activities very fast manner. Nowadays, getting and analysing the crime data are critical. The crime data can be identified by various factors such as location of occurrence, crime detected time, also predicting their future relationship is an essential in crime preventing system [1]. In this research, time and place are considered as main aspects in identifying the crime pattern. Machine Learning (ML) techniques offers to extract the information from the collected datasets also find the relationship between the crime, place, and time. Many researchers stated that identifying the crime pattern is very critical and time-consuming task [2]. It can be resolved by the ML techniques.

In modern era, ML techniques plays a crucial role in every field, in crime prediction also it acts as a major part. With the help of ML, can able to generate the crime committed details and then used to identify the criminals faster than mankind. There various types of ML are there such as supervised, unsupervised, and reinforcement learning method. The supervised learning method subdivided into two types classification and regression. Classification techniques are helps to identify the class from the given data and regression help to find the relationship between crime time and place from the given data. The unsupervised learning method is subdivided into clustering method. Clustering methods helps to group the similar class from the given data. This work utilizes various ML techniques like Decision Tree (DT), Gaussian Naïve Bayes (NB), Linear Support Vector Classifier (SVC), Logistic Regression and Stochastic Gradient Descent (SGD) for predicting the crime pattern based on time and place of the crime committed. As well as, feature selection techniques also play a crucial role in ML. Feature Selection techniques help to find the key features to improve the crime prediction rate. In this work, existing feature selection techniques such as, Linear Regression, Ridge Regression and Polynomial Regression are used to identify the salient features. The main objective of this research is to predict the crime based on time and place of crime committed using the communities and crime dataset.

To find the best classification and feature selection techniques for identifying the crime pattern which helps to government for future detection.

2. Related Work

A number of studies which is related to crime detection analysis is discussed in this section. Jianming et. al [3] proposed a novel big data method to analyse the crime rate. In their work, combination of Gradient Boost Decision Tree (GBDT), Recursive Feature Elimination (RFE) and

1Research Scholar, Department of Information Technology, Faculty of Engineering and Technology, Annamalai Univeristy, Tamilnadu, India. ganeshitlect@gmail.com
2Associate Professor, Department of Information Technology, Faculty of Engineering and Technology, Annamalai Univeristy, Tamilnadu, India rsuban82@gmail.com

Geographical Information System (GIS) technique is used to extract the features.

Das et. al [4] proposed a novel cluster technique to predict the crime pattern. In their work, the proposed technique extracting relations from the newspaper using criminological data. Li

Ding et. al [5] proposed a PerpSearch method that used to identify the crime using crime committed location, types and the personal characteristics. In their work, the data were collected from the Alabama Criminal Justice Information Centre. Tushar et. al [6] predicted the crime with the help of various crime and crime committed location.

Birks et. al [7] proposed a novel cluster technique for identifying the crime pattern. For this analysis, the robbery data is used which was collected from the area of United Kingdom. Felson et. al [8] identifies the crime rate at the time of COVID – 19. For this analysis, Detroit, Michigan robbery data is used. For detecting the crime, three basic crime patterns in four combination of graph theory were implemented by Wang and Zhang [9]. The prediction is based on the crime happened place and time.

Aarthi et. al [10] used K means cluster technique to detect the crime pattern. The system proposed to extract the crime data also grouping the crime pattern. Chen and Kurland [11] proposed a technique to identify the crime pattern with the help of three attributes namely Time, Place and Modus Operandi. In their work, Apriori algorithm is used for detecting the crime. Zero Crime system is implemented by Mugdha Sharma [12] to identify the suspicious crime action. In this work, DT method is used for crime detection process.

System for Identifying the Influential Members of a Criminal Organization (SIIMCO) tool was proposed by Taha and Yoo [13]. It can identify the most bigwig of a criminal grouping also finds the current leader of that group. For validating, the SIIMCO system, two real time datasets such as Enron email corpus and Nodobo mobile phone records are used. Benjamin and Suruliandi [14] illustrate the various ML techniques which has been used in the crime prediction system. The review is based on Text, Crime pattern, Geo – location, Prisoner and Communication method. In their paper, various classification and clustering techniques are analysed by qualitatively and quantitatively

In ancient day, the crimes are in the level of burglary. Historically, solving the crimes has been right of the criminal justice and law enforcement specialists [14]. Criminology refers to predict the crime and criminal traits. The police departments, detective agencies, and crime branches are helps to find the real characteristics of criminal. But nowadays, the crime levels are increased in the way of robbery, assault, sexual assault, homicide, etc. Hence, the crime department faces more difficulties in order to predict the criminal. The crime activities are varies based on various factors such as place, time, anger, drug consumption, etc. Predicting these relationships are the most important feature in criminology [1]. Consequently, it needs a detailed dataset for identifying these relationships. Justified by these facts, ML techniques are used in this work to predict the crime pattern which is mainly based on time and place.

From the literature survey it is observed that various ML plays a crucial role in crime pattern prediction. This work gives overview of several ML techniques to detect the crime and grouping them based on time and place of crime committed. Classification techniques are used to predict the crime whereas regression techniques are applied to relationships between the crime committed place and time for further process. This work uses ML techniques such as DT, Gaussian NB, Linear SVC, Logistic Regression and SGD for crime detection process. DT requires less attempt for preparing the data during pre-processing but it needs more time to train the system. Though Gaussian NB mandatory to have a large number of data samples for achieving a good result, it does not need to spend much time for training the samples. However linear SVC is not convenient for large datasets, it relatively more efficient. Logistic Regression is very fast at classifying unknown records also it is not apt if the number of observations is lesser than the number of attributes. SGD applied for larger datasets; it can converge faster as it causes updates to the parameters more frequently. Due to frequent updates, the step taken towards the minima are very noisy. Similarly, various feature selection techniques are applied to find the important features from the given crime dataset. In this work, Linear Regression, Ridge Regression and Polynomial Regression techniques are used for finding salient features. Each techniques have own way for prediction process. Hence, it is needed to find the best ML technique for crime detection process. To the best of knowledge, as far as there is no research carried to compare the performance of above-mentioned techniques for crime prediction. Motivated by this fact, a comparative analysis of various ML techniques is conducted in this work and also the performance is evaluated. Based on that results, more efficient technique is identified for crime prediction process using time and place.

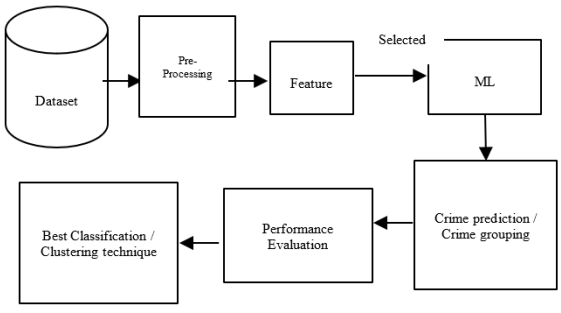

Fig 1. BLOCK DIAGRAM

Fig 1 illustrates the entire process of crime prediction process. The process starts with the pre-processing procedure. In this step, the given input data is verified whether it has any duplicate data or missing values. Further it given into the feature extraction procedure. In that, the key features from the pre-processed data are selected. Consequently, the extracted features are fed into the either ML techniques to identify the crime or crime pattern. Also, the performances of each classifier are validated to find the appropriate technique for crime detection process.

Organization of the paper

The remaining part of this paper is organized as follows: Section 2 describes the methodologies which is used in the crime prediction process; the performances of each classifier is illustrates in section 3 finally the work concludes in section 4.

3. Methodology

Machine Learning Techniques

ML techniques are used to identify the crime pattern with the help of training samples from the given data. It is break into two categories: supervised and unsupervised. Supervised learning has known label dataset further it is divided into two sub parts such as training and testing. The known training dataset are well trained with the prediction technique to find the target outcome from the unknown testing dataset. The supervised learning is divided into classification and regression [15]. The following subsection describes various supervised learning techniques and feature selection techniques that were used in this work.

3.1 Supervised Learning Techniques

It is the one type of ML technique. It includes classification and regression methods. Classification techniques are used to predict the target class using the independent attributes. Regression techniques are used to find the exact value of target attribute with the help of independent attributes.

Decision Tree

DT is a supervised learning method, also it is a tree structure model. It consists of one root or parental node, several branches and leaf nodes. Each intermediate node i.e., branches have one root node. It further divided into several branches until it reaches the leaf node which is target class [16]. The splitting involves that the number of one target are placed one side and remaining target classes are other side. The splitting will continue until it reaches the accurate target class. For splitting it uses Gini impurity or information gain function. The pseudo code for the DT is illustrated in [17].

$$Information\ Gain\ (T,X) = Entropy\ (T) - Entropy(T,X) \quad (1)$$

where, T implies to the current state and X is the selected feature.

$$Gini\ Index = 1 - \sum (p)\text{^}2 = 1 - [(p+)\text{^}2 + (p-)\text{^}2] \quad (2)$$

where, p+ implies the probability of yes and p- implies the probability of no.

Gaussian Naïve Bayes

Gaussian NB [18] is a type NB supervised learning method. It mainly based on the Bayes theorem. It predicts the target class based on the probability rule. The NB algorithm conducts differently for discrete and numeric features. When considering the likelihood probability $P(X|C)$, that means the number of times the value was observed divided by the total number of observations. The class which has the greatest probability is selected as target class. Based on the data distribution, the NB classifier is break into several types. In this work, Gaussian NB is used to predict the crime. When the predictors assume a continuous value, it concludes that sample values are from the gaussian distribution. The pseudo-code for the Gaussian NB is given in [19]. The formula that used in this technique to calculate the probability is

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}\right) \quad (3)$$

Linear Support Vector Classifier

Support Vector Machine (SVM) is a supervised learning method also it is used for both classification and regression problem. When it is applied for classification, it is denoted as SVC and it is used in regression problem it known as Support Vector Regression (SVR). The SVM classifies the target class based on the hyperplane [20]. The hyperplane also decision boundaries, it helps to separates the target class. To minimize the wrong prediction, the SVM creates the hyperplane in a repetitive aspect. The main objective of the SVM is to find the maximum marginal hyperplane to predict the target class. It can be done in two ways: Initially, it creates the hyperplane iteratively and eventually, it selects the hyperplane which selects the target class more accurately. The pseudo-code for the SVM classifier is defined in [21].

Logistic Regression

Logistic regression also a supervised learning method. It is an efficient ML technique, which is used to

find the likelihood and classify the continuous and discrete datasets. In this work, it is used when the target class is categorical [22]. It does not need any linear relationship between the known variables and target variables. It separates the one target class against all other target classes with linear boundary by considering the remaining target classes are one class. The pseudo-code for the logistic

regression is illustrated in [23]. The general logistic regression equation can be written as

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n \quad (4)$$

$$log\left[\frac{-y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n \quad (5)$$

Stochastic Gradient Descent

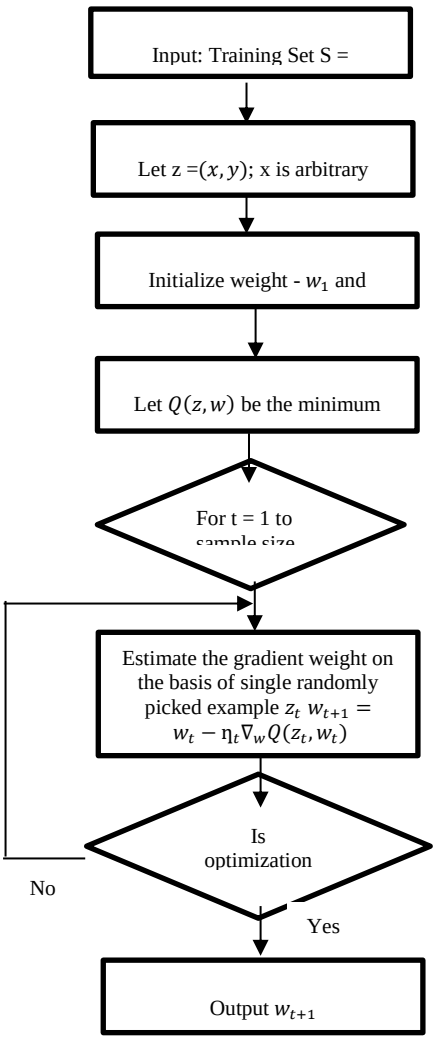

Fig. 2 Flow Diagram for SGD

Gradient Descent is a ML technique that used to optimize the predictive model cost function. Various Gradient Descent techniques are used for predictive model. In this work SGD [24] classifier is used in crime prediction process. The working function of SGD classifier is illustrated in Figure 2. SGD classifiers trains one instance at a time also updates each training samples parameter at a time. The SGD classifier reduces the cost function of the prediction and hence it does not require to recognize the previously visited instances. The pseudo – code for the SGD classifier is given in [25].

3.2 Feature Selection Techniques

Feature selection techniques are used to identify the salient attributes from the given dataset also removes the irrelevant attributes from the dataset. It mainly divided into three types such as filter, wrapper and embedded method. Filter method selects the important feature with the help of correlation with target attribute whereas Wrapper method removes the irrelevant features with help of classifiers. Embedded method combines the merit of both filter and wrapper method.

Linear Regression

Linear regression technique is used to remove irrelevant features from the crime dataset. It is chiefly used in predictive analysis. The two main concepts of linear regression [26] are identifying the target values also find the more efficient attributes which was mainly influenced in target value. The pseudo – code for the linear regression is given in [27]. It is defined as

$$y = b_0 + b_1x + \in \quad (6)$$

where, y denotes target class or dependent attribute, x is a predictor or independent attributes, b_0 illustrates the intercept of the line, b_1 denotes linear regression coefficient and $\in$ is a random error.

Ridge Regression

Ridge regression technique, also known as L2 regularization. It is an embedded type feature selection method. It used to find the important of all features further it helps to reduce the coefficient of features [28]. Compare to other methods, it implicitly revokes the effect of less suitable features. It is widely used in large dataset also it reduces the error in prediction process. The pseudo – code for the ridge regression is given in [29]. It is denoted as

$$y = x_i^T\beta + e_i \quad (7)$$

where, y is a target class, x represents independent attributes, β is a regression coefficient and e illustrate the residual error.

Polynomial Regression

Polynomial regression [30] technique also used to find the important features from the dataset. Not like linear regression, polynomial regression does not require any linear relationship between the input attributes and target value. It works based on the Gauss Markov theorem to minimize the variance of the coefficient for attributes. The pseudo – code for the polynomial regression is given in [31]. It can be expressed as

$$y = \theta_0 + \theta_1x + \theta_2x^2 + \cdots + \theta_nx^n \quad (8)$$

where, y represents target class, $\theta_1, \theta_2, \ldots, \theta_n$ are the coefficient, x illustrates the independent attributes, n is the polynomial degree.

4. Experimental Results and Discussions

This section illustrates the dataset used, performance metrics which were used to analyse the performance of each technique and discussed the results of each technique.

Dataset Description

This work used Communities and Crime [32] dataset that was downloaded from the University of California Irvine (UCI) ML Repository. The dataset includes 1994 records with 128 features. The dataset combines United States (US) communities' records with socio-economic data from the 1990 US Census, law enforcement data from the 1990 US LEMAS survey, and crime data from the 1995 Federal Bureau of Investigation, Uniform Crime Reporting.

Performance Metrics

There are various performance metrics are used to examine each technique performance for crime prediction mainly based on time and place. In this work Accuracy, Error rate, Precision, Recall, Specificity, F1 Score, Area Under the Curve (AUC) and Mean Absolute Error (MAE) metrics are used for performance evaluation.

Accuracy:

Accuracy metric is used to find how exactly the techniques predict the outcome. The range is between 0 to 1. The value nearby 1 indicates the better identification. The formula is defined by

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

where, TP – True Positive; TN – True Negative; FP – False Positive; FN – False Negative

Error Rate:

The defective of predicted target values are known as error. If the outcomes are categorical, the error is denoted as error rate. The formula is defined by Error Rate = 1 – Accuracy (10)

Precision:

Precision metric is used to examine the fractions of correctly classified instances among the ones classified as positive. The range is varied from 0 to 1. The value close to 1 denotes low false positive prediction. The formula is defined by

$$\text{Precision} = \frac{TP}{TP+FP} \quad (11)$$

Recall:

Recall metric is used to examine the fractions of correctly classified instances with all relevant instances. The value is varied from 0 to 1. The value nearby 1 illustrates the less false negative prediction. The formula is defined by

$$\text{Recall} = \frac{TP}{TP+FN} \quad (12)$$

Specificity:

Specificity metric helps to value the ratio of incorrectly classified instances to all relevant instances. The value is ranges from 0 to 1. The value near by 1 demonstrates the less negative prediction whereas 0 denotes the more negative prediction. The formula is derived by

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (13)$$

F1 Score:

F1 Score is used to evaluate the weighted harmonic mean of the precision and recall measurements. The value is varied from 0 to 1. The value close to 1 illustrates the better precision and recall whereas value close to 0 denotes low precision and recall. The formula is defined by

$$\text{F1 Score} = 2 * \frac{PPV.TPR}{PPV+TPR} \quad (14)$$

where, PPV – Positive Precision Value; TPR – True Positive Rate

AUC:

AUC metric is used to calculate the accumulative measure of performance across all classification thresholds. The value is varied from 0 to 1. The value close to 1 illustrates the better performance of classification techniques. The formula is defined by

$$\text{AUC} = \text{TPR vs. FPR} \quad (15)$$

where, FPR - False Positive Rate

MAE:

MAE is the error metric. It is used to evaluate the average of difference between original and predicted values. The range is varied from 0 to ∞. Value close to 0 has better prediction.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y_i}| \quad (16)$$

where, y_i – the sample belongs to the class i.

Results and Discussion

This section evaluates the performance of each existing prediction techniques such as DT, Gaussian NB, Linear SVM, Logistic Regression and SGD to predict the crime detection. Consequently, the performance of each existing feature selection techniques such as Linear Regression, Ridge Regression and Polynomial Regression are examined for selecting the key features from the crime dataset to improve the prediction performance.

Experiment 1: Performance evaluation of existing techniques for crime detection

Table 1 illustrate the performance of each existing techniques for crime prediction based on time and place of crime committed.

Table 1. Performance comparison of prediction techniques for crime detection

Model	Performance Metrics (%)								
	Accuracy	Error Rate	Kappa	Precision	Recall	Specificity	F1 Score	AUC	MAE
DT	75.00	25.00	73.96	80.40	80.90	81.66	80.64	78.63	0.15
Gaussian NB	76.10	23.90	74.00	91.10	69.20	91.65	78.65	78.79	0.14
Linear SVM	79.60	20.40	77.61	84.50	83.40	86.61	83.94	82.00	0.12
Logistic Regression	72.20	27.80	70.98	78.00	77.00	78.55	77.49	76.12	0.16
SGD	**80.00**	**20.00**	**79.18**	**81.00**	**85.00**	**82.96**	**82.95**	**84.21**	0.10

From the table 1, it is evident that the SGD technique predicts the crime more accurately than other existing techniques. Due to updating the parameters more frequently it gives better prediction rate with fast computation.

Table 2 shows the performance comparison various existing feature selection techniques with various prediction model for the crime detection based on time and place of crime committed.

Experiment 2: Performance evaluation of existing feature selection techniques for crime prediction

Table 2. Performance comparison of various feature selection techniques

Feature Selection	Model	Performance Metrics (%)						
		Accuracy	Error Rate	Kappa	Precision	Recall	F1 Score	AUC
Linear Regression	DT	77.76	22.24	75.86	84.22	84.00	84.10	79.87
	Gaussian NB	79.19	20.81	76.88	93.42	75.12	83.27	81.78
	Linear SVM	81.00	19.00	79.18	87.41	87.61	87.50	82.97
	Logistic Regression	75.43	24.57	73.32	80.04	88.18	83.91	77.17
	SGD	**82.89**	17.11	**80.98**	**84.56**	**88.94**	**86.69**	**85.66**
Ridge Regression	DT	76.90	23.10	74.95	82.10	82.98	82.53	78.97
	Gaussian NB	78.42	21.58	75.66	92.42	73.00	81.57	81.14
	Linear SVM	80.39	19.61	78.05	85.49	85.34	85.41	82.79
	Logistic Regression	73.42	26.58	70.96	79.65	80.75	80.19	75.85
	SGD	81.13	18.87	78.86	82.98	87.32	85.09	83.50
Polynomial Regression	DT	75.79	24.21	74.78	81.00	81.65	81.32	78.18
	Gaussian NB	76.49	23.51	73.45	91.77	70.41	79.68	78.74
	Linear SVM	79.04	20.96	77.43	84.90	84.17	84.53	83.68
	Logistic Regression	72.72	27.28	70.77	79.54	78.54	79.03	75.71
	SGD	80.51	19.49	78.65	81.22	85.87	83.48	82.05

From the table 2, it is observed that all feature selection techniques with SGD prediction model outperforms other existing prediction model. All three existing feature selection techniques commonly selected 10 attributes among all other attributes. The selected features are pctWInvInc, PersPerOccupHous, racePctWhite, PctKids2Par, RentHighQ, MalePctDivorce, NumUnderPov, NumStreet, PctOccupMgmtProf and population. These attributes have absolute value of coefficient, so there are the one which contributes to low crime as well as high crime rate.

Experiment 3: Performance evaluation of existing feature selection techniques with SGD prediction model

From the table 1, it is perceived that the SGD prediction model more accurately predict the crime based on the time and place. The below graphical representation shows the performance comparison of various existing feature selection techniques with SGD model.

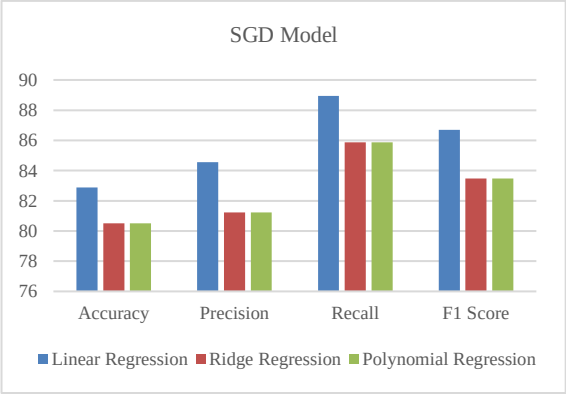

Fig 3 Performance comparison of existing feature selection techniques with SGD model

Figure 3 illustrates that the Linear Regression technique with SGD model outperforms other existing feature selection techniques. Also, the performance of SGD model for crime prediction improves with the Linear Regression technique. The prediction rate SGD with Linear Regression 82.89% is higher than without Linear Regression 80.00%.

5. Conclusion

In this work, a comparative analysis of various prediction techniques and feature selection techniques has been examined to identify the efficient prediction model and features selection techniques for crime prediction. The DT, Gaussian NB, Linear SVM, Logistic Regression and SGD methods are considered for prediction as well as the Linear Regression, Ridge Regression and Polynomial Regression feature selection techniques are taken for selecting the key features from the dataset. The experimental results confessed that the SGD model more accurately predicts the crime than other models as well as Linear Regression technique selects more salient features than other feature selection techniques. Consequently, the experimental results showed the Linear Regression with SGD model outperforms without Linear Regression for crime prediction based on time and place of crime committed.

References

[1] Kiani, Rasoul, Siamak Mahdavi, and Amin Keshavarzi. "Analysis and prediction of crimes by clustering and classification." International Journal of Advanced Research in Artificial Intelligence 4, no. 8 (2015): 11-17.

[2] Akansha A Chikhale, Ankita K Dhavale, Aparna P Thakre, Diksha B Herode, Nikita D Nasre, Pracheta D Patrikar, Prof. Milind Tote. “A Review on Crime Rate Analysis Using Data Mining” International Journal of Scientific Research in Science, Engineering and Technolgoy 5, no. 5 (2019): 119 – 125.

[3] J. Zhou, Z. Li, J. J. Ma and F. Jiang, "Exploration of the Hidden Influential Factors on Crime Activities: A Big Data Approach," in IEEE Access, vol. 8, pp. 141033-141045, 2020, doi: 10.1109/ACCESS.2020.3009969.

[4] Das, Priyanka, Asit Kumar Das, Janmenjoy Nayak, Danilo Pelusi, and Weiping Ding. "A graph based clustering approach for relation extraction from crime data." IEEE Access 7 (2019): 101269-101282.

[5] Li Ding, Dana Steil, Matthew Hudnall, Brandon Dixon, Randy Smith, David Brown, Allen Parrish, "PerpSearch: An integrated crime detection system," 2009 IEEE International Conference on Intelligence and Security Informatics, 2009, pp. 161-163, doi: 10.1109/ISI.2009.5137289.

[6] Tushar Sonawanev, Shirin Shaikh, Shaista Shaikh, Rahul Shinde, Asif Sayyad “Crime Pattern Analysis, Visualization And Prediction Using Data Mining”, International Journal of Advance Research and Innovative Ideas in Education, Vol 1, no.5 (2015): 681 – 686

[7] Birks, Daniel, Alex Coleman, David Jackson. "Unsupervised identification of crime problems from police free-text data." Crime Science 9, no. 1 (2020): 1-19.

[8] Felson, Marcus, Shanhe Jiang, and Yanqing Xu. "Routine activity effects of the Covid-19 pandemic on burglary in Detroit, March, 2020." Crime Science 9, no. 1 (2020): 1-7.

[9] Wang, Zengli, and Hong Zhang. 2020. "Construction, Detection, and Interpretation of Crime Patterns over Space and Time" ISPRS International Journal of Geo-Information 9, no. 6: 339. https://doi.org/10.3390/ijgi9060339

[10] Aarthi, S., M. Samyuktha, and M. Sahana. "Crime hotspot detection with clustering algorithm using data mining." In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 401-405. IEEE, 2019.

[11] Chen, Peng, and Justin Kurland. "Time, place, and modus operandi: a simple apriori algorithm experiment for crime pattern detection." In 2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1-3. IEEE, 2018.

[12] M. Sharma, "Z - CRIME: A data mining tool for the detection of suspicious criminal activities based on decision tree," 2014 International Conference on Data Mining and Intelligent Computing (ICDMIC), 2014, pp. 1-6, doi: 10.1109/ICDMIC.2014.6954268.

[13] Taha, Kamal, and Paul D. Yoo. "SIIMCO: A forensic investigation tool for identifying the influential members of a criminal organization." IEEE Transactions on Information Forensics and Security 11, no. 4 (2015): 811-822.

[14] David, H., and A. Suruliandi. "Survey on Crime Analysis and Prediction using Data Mining Techniques." ICTACT journal on soft computing 7, no. 3 (2017).

[15] Alkesh Bharati, Dr Sarvanaguru RA.K, "Crime Prediction and Analysis Using Machine Learning". International Research Journal of Engineering and Technology, 5, no. 9: 1037 – 1042 (2018).

[16] Emmanuel Ahishakiye, Elisha Opiyo Omulo, Danison Taremwa, and Ivan Niyonzima. "Crime Prediction Using Decision Tree (J48) Classification Algorithm." International Journal of Computer and Information Technology, 6 no. 3: 188 - 195 (2017).

[17] G. Anderson (2008) "Random relational rules", PhD thesis (The University of Waikato).

[18] Sri, Linga Akhila, Kalluri Manvitha, Gorantla Amulya, Ikkurthi Sai Sanjuna, and V. Pavani. "FBI Crime Analysis and Prediction using Machine Learning." Journal of Engineering Sciences, 11, no. 4: 441 – 448, (2020).

[19] Ivan Kholod, Andrey Shorov, and Sergei Gorlatch, "Improving Parallel Data Mining for Different Data Distributions in IoT Systems", In International Symposium on Intelligent and Distributed Computing, Springer, Cham 2019, pp. 75-85, 2019, doi: https://doi.org/10.1007/978-3-030-32258-8_9

[20] Kianmehr, Keivan, and Reda Alhajj. "Effectiveness of support vector machine for crime hot-spots prediction." Applied Artificial Intelligence 22, no. 5 (2008): 433-458.Angelina Tzacheva, Jaishree Ranganathan, and Sai Yesawy Mylavarapu, "Actionable Pattern Discovery for Tweet Emotions", In International Conference on Applied Human Factors and Ergonomics, Springer, Cham, 2019, pp. 46-57, 2019, doi: 10.1007/978-3-030-20454-9_5.

[21] Antolos, Daniel, Dahai Liu, Andrei Ludu, and Dennis Vincenzi. "Burglary crime analysis using logistic regression." In International Conference on Human Interface and the Management of Information, pp. 549-558. Springer, Berlin, Heidelberg, 2013.

[22] Kumar, Vinod. "Evaluation of computationally intelligent techniques for breast cancer diagnosis." Neural Computing and Applications 33, no. 8 (2021): 3195-3208.

[23] Reier Forradellas, Ricardo Francisco, Sergio Luis Náñez Alonso, Javier Jorge-Vazquez, and Marcela Laura Rodriguez. "Applied Machine Learning in Social Sciences: Neural Networks and Crime Prediction." Social Sciences 10, no. 1 (2020): 4.

[24] Galuzzi, Bruno G., Ilaria Giordani, Antonio Candelieri, Riccardo Perego, and Francesco Archetti. "Hyperparameter optimization for recommender systems through Bayesian optimization." Computational Management Science 17, no. 4 (2020): 495-515.

[25] Awal, Md Abdul, Jakaria Rabbi, Sk Imran Hossain, and M. M. A. Hashem. "Using linear regression to forecast future trends in crime of Bangladesh." In 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), pp. 333-338. IEEE, 2016.

[26] Hasan, Md Abid, Md Kamrul Hasan, and M. Abdul Mottalib. "Linear regression–based feature selection for microarray data classification." International journal of data mining and bioinformatics 11, no. 2 (2015): 167-179.

[27] Shariff, Nurul S. Mohamad, and H. M. B. Duzan. "An application of proposed Ridge Regression Methods to real data problem." International Journal of engineering and technology 7 (2018): 106-108.

[28] Liu, Yih-Wu, and Richard H. Bee. "Ridge regression: A multivariate analysis of criminal activity." Sociological Spectrum 3, no. 2 (1983): 143-157.

[29] Biswas, Al Amin, and Sarnali Basak. "Forecasting the trends and patterns of crime in Bangladesh using machine learning model." In 2019 2nd international conference on intelligent communication and computational techniques (ICCT), pp. 114-118. IEEE, 2019.