ZKG INTERNATIONAL

ISSN: 2366-1313

Edge Assisted Crime Prediction and Evaluation of Framework for Machine Learning Algorithms

[1] Mr.K.Vamshee Krishna, [2] K Prasad,[3] Korivipalli Sai Sharanya,[4] Johirul Islam, [5]Mudavath Sharath

[1] Assistant Professor, Dept. Of CSE, Samskruti College of Engineering & Technology, TS.

[2,3,4,5]B. Tech Student, Dept. Of CSE, Samskruti College of Engineering & Technology, TS.

Abstract: *The growth of the global population, especially in large cities, has created new problems, especially in terms of legislation and the implementation of effective civil protection. Therefore, in this paper, a method is proposed to estimate crime occurring in a city, based on all past events and public observations. In particular, this study presents the crime detection and evaluation framework for the system to gain knowledge of the community algorithms. Thus, a comprehensive analysis of four types of serious crimes, such as murder, speeding trials, abuse of girls and boys, and drugs, leads to positive results among the preparation process. All the research and its implementation showed a visible example of crime in many parts of the United States. The general work is done by selecting, analyzing and using the machine learning (ML) model and finally reporting the crime. Violation risk is estimated by using distribution patterns for time, programming language and region. To anticipate what will happen, ML methods that include Decision Trees, Neural Networks, K-Nearest Neighboring, and Functional Learning are used, and their overall performance is is compared by processing data and using modifications. The highest accuracy of eighty-one percent is achieved for the decision tree rule set throughout the crime prediction. The results showed that the use of machine learning techniques made it easier to predict crimes, which helped to improve public safety.*

Keywords—Machine Learning, Edge Computing, Crime Prediction, Impact Learning, Decision Tree, KNN, MLP

I INTRODUCTION

One of the main concerns of citizens around the world is civil protection. Many reasons, including the rapid growth of cities, have led to growing concerns. The migration of people to cities is well known in recent years and, according to UN

ISSN: 2366-1313

estimates, more than 70% of the population will live in cities by 2050 [1]. Furthermore, according to the Global Crime Database, which defines terrorist acts as "acts committed with the assistance of non-state actors against civilians, this means that objective concern, as a means to achieve the political goal", is diverse. of criminals over the past decade is the highest on record. Machine control (ML) techniques are important for smart city projects and can be used to reduce crime as they help solve smart city problems.

This image was adapted for support from the Korean government-funded Institute of Information Technology Planning Evaluation (IITP) (MSIT) (No.2019-zero-01287, Evolvable Deep Learning Model Generation Platform for Edge Computing) and MSIT (Ministry of Science). and ICT), Korea, under the Grand Information Technology Research Center Support Software (IITP-2020-2015-0-00742) supervised by IITP. *Dr. CS Hong is the founder.

Urban development and vaporization of the obtained fact [2].

This document can provide an overview of crime in your area. S. A., like Bangladesh. We use facts from 2012 to 2019 to show adults. Using the crime prediction version, we find that zone 1 is more damaged than different cities while zone 2 is not damaged. Using all these statistics, we created a 2021 crime forecast (check this). We present the political comparison between 2019 and 2021. We have a situation to inform city residents and local authorities about the most dangerous places, thus presenting value to the community and increasing public protection. In this case, this type of forecast can be useful for influencing the manner, such as controlling the effectiveness and efficiency of work, as well as for vacationers who have blind visibility to the most dangerous areas of the city .

Machine learning techniques, including expertise learning [3], decision tree [4], K-Nearest Neighbors [5], and MLP classifier, are used to make predictions. These algorithms are scored based on the number of correct actions and strategies used. We summarize the main points of the program as follows:

1) First, we proposed a fault prediction and evaluation framework for network-side ML algorithms. It is not easy to identify an old measure of crime detection, but it can also be measured as a set of modern information intended to prevent crime.

ISSN: 2366-1313

2) Second, we implemented several ML algorithms including decision trees, neural networks, K-Neighborhood, and constraint analysis at the end of the planning process. In this, really the criminal record of the u. S. It is used to confirm the effectiveness of the proposed system through comprehensive analysis.

3) Finally, our interventions have proven effective in anticipating crime, which includes murders, speedy trials, rape of women and children, and truth-telling to protect the community.

The remainder of this article is presented as follows. The relevant work is described in Chapter II and the reporting process is contained in Chapter III. The results and discussion are then presented in Section IV, with details on the machine learning techniques. Finally, we conclude in section V.

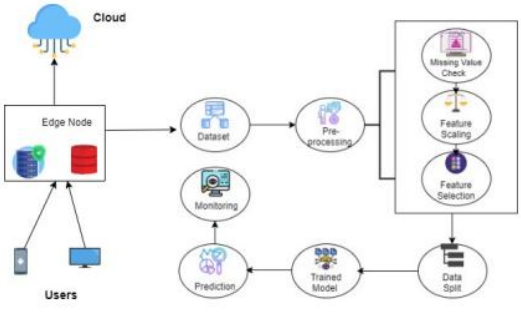

Fig. 1. A Crime Prediction and Evaluation Framework for Machine Learning Algorithms of Network Edge.

II Related Work

A group of researchers used WEKA, an open source statistical mining software, to analyze crime rates from non-standardized community and crime data. was provided by the University of California-Irvine with information on criminal activity for the Mississippi Kingdom received from the community scout.Com [6]. Based on old and public records, Luis et al. The concept of a model to predict the crime situation in the city [7]. Another group of studies has determined the direction of the standard model of violation prediction using the decision tree algorithm (J48). The rule J48 detected the unknown crime with an accuracy percentage of 94.25287 according to the test results [8]. In another study, the authors used cognitive and scientific techniques to predict crime in the Chicago crime data [9]. Details of the crime were obtained from the Chicago Police Department's website. In a unique study, crime data from Vancouver over the past 15 years was analyzed using advanced data processing techniques. When it comes to predicting crime in Vancouver, the accuracy ranges from 39% to 44% [10]. Other schools have published legal prediction models, mostly based on cities (regions or regions that make up the city of Buenos Aires) and using the Python programming language for predictions first. In addition to [6-10], we provide a general

ISSN: 2366-1313

guide to this work that enters the year, this model will provide the crime for the past year.

III Proposed Framework

We carried out this work using a working machine. To use the obtained data, we followed five steps [11], as shown in Figure 1. The number of samples and the number of variables are sufficient to maintain good accuracy. The proposed framework consists of six methods. The first is the collection of information, which is accompanied by control. We were the first to come to the truth after we gathered together. In this step, we will start by checking for missing values before moving directly to the characteristic test. Finally, we select the function for the label and the working file. We rent the ML route once completed

TABLE I
DATA SET ATTRIBUTES DETAILS INFORMATION.

Attributes	*Description*	*Type*
Murder	Number of Murder in different city of the country.	Numerical
Speedy trial	Speedy trial in five Metropolitan areas of the country.	Numerical
Woman and Child Repression	Up and down of this crime is shown in number.	Numerical
Narcotics	Number of Narcotics in different city of the country.	Numerical

This information is broken for education and exit. According to the university literature, this ML approach has led to predictions. The proposed framework is to use the body in the community.

A. Peripheral network

Edge networking is a distributed computing model that makes computing and data storage closer to the need to reduce latency and save bandwidth [12], [13]. Edge captures and processes data as close as possible to the data or events planned. It collects statistics on the use of sensors, computing gadgets, and machines before sending them to other servers or the cloud [14]. These statistics can be used for power measurement and physical understanding, enabling automation, or providing insight into the current status of a device, gadget, or device. product [15], depending on preferences and needs.

B. Data sources

We collected all the data from the United States Police website [16]. We have used the date from 2012 to 2019. We have seen a wide range of facts about crime here, but we have selected some of the most important, which are rising and very old in the United States. Table I describes the data collected for our study.

C. Prior information

Crime prediction data is the first priority when collecting various statistics [13]. Murder, Speedy Trials, Abuse of Women and Children, and Narcotic are the top four in this database for predicting crime. We have put many disks on this website, but not everything is put on the wall anymore.

ISSN: 2366-1313

We use 3 different steps to advance the statistics for this purpose.

"It's worthless, look."

“Measuring the scale.

“Special selection.

Pay attention to the missing value: The holding value The missing value is incorrect in the highest cases as the tax that is not stored in the example. In the information industry, lack of value is a daily occurrence. In addition, most future presentation methods cannot handle missing data. Therefore, this problem must be solved before the model begins to evolve. We use the expression to update the missing value. To find the optimal solution, this is equal to the average value of a to create reality, calculation is necessary. Subtract the total number of digits from the total range of values in the data collection [15]. Equation: 1 is used to get the implicit.

$$Mean = \frac{\sum_{i=0}^{n} x_n}{n} \qquad (1)$$

Feature Scaling: Feature scaling or normalization is one of the most important techniques in machine learning techniques; Without it, the target systems will not function properly. Min-Max scaling, variance scaling, normalization, mean normalization, and unit vectors are a number of characteristic scaling techniques available. For this task, we use min-max normalization. The normalization range in min-max in [0, 1] or [-1, 1] is given by equation: 2 and the min-max values in [0,1].

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \qquad (2)$$

Feature Selection: Another task that must be done before deploying a release is feature selection. The main purpose of this method is to find ways to influence the work for different purposes. We removed a few important points that were not important for different purposes, while keeping the important points. The calculation cost is reduced when the range of key points is reduced. There are thirteen unique systems in our database.

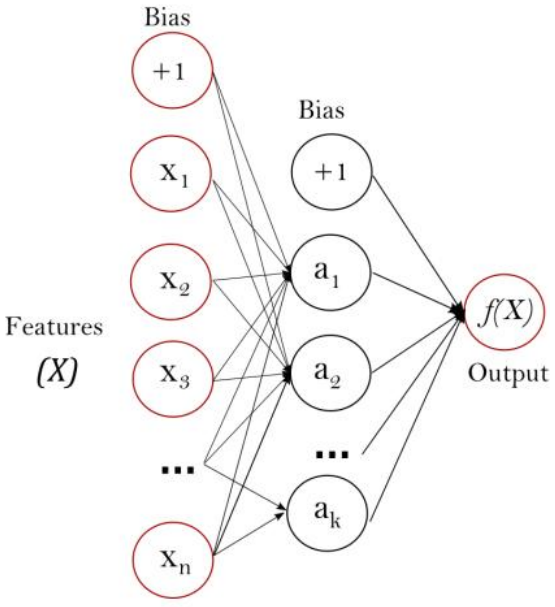

Fig. 2. Diagram of MLP Classifier [17].

D. Data Split

Training information is the muse for all system learning algorithms. All of the statistics collected has been split into

ISSN: 2366-1313

sections. The first is a education set, while the second one

TABLE II
DETAILED TRAINING PARAMETERS OF IMPLEMENTED ALGORITHMS.

Name of Al-algorithms	*Description*	*Initial Parameters*
K-Nearest Neighbors	The majority of votes from its fixed neighbors classify the new data point.	Nearest Neighbour: 5
MLP Classifier	All input is processed through a series of hidden layers, and a final output is anticipated as a result.	Random state: 1
Decision Tree	The more splits there are in a tree, the more information it captures about the data.	Class weight: None
Impact Learning	On the training data set, a huge number of Epoch are run. Iteration is another name for Epoch.	Epoch: 2000

F. Training model

We use four control algorithms to find the most accurate results. For this information, all methods are similar to praising bureaucracy. Based on the lowest error rate of the algorithms, a particularly efficient model was detected. Table II contains the results of all algorithms in addition to their limitations.

Impact Learning: Impact for knowledge is a professional learning technique that uses supervised classification and linear or polynomial regression. It is also useful for evaluating the true pattern of the competition. There is rarely a good way to understand the impact of independent versus competitive capabilities using this method. In a separate article, the result of the herb cultivation value (RNI) [3] is presented. RNI is represented by the equation: 3 in this case.

$$Imp = (y' - (\frac{k \sum_{i=1}^{n} w_i x_i}{r - w_y k} + b)^{2/N} \quad (3)$$

K- Nearest Neighbor Classifier: The K-Nearest Neighbors

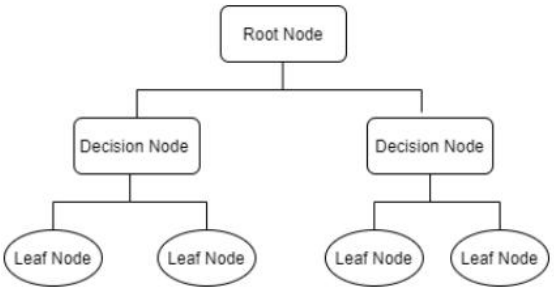

Fig. 3. Diagram of Decision Tree Classifier.

This technique is one of the most important machine learning algorithms and it is based on supervised learning. The K-NN algorithm collects all available data and classifies new data based on their similarity to previous data. This means that by using the K-NN method [5], the new data can be quickly classified into a good group. The maximum distance between the selected neighbors is calculated using the KNN method. KNN uses the Euclidean distance function to calculate the distance between the existing data and each new data. Equation 5 can be used to calculate the Euclidean distance.

$$EuclideanDistance = \sum_{i=0} (x_i - y_i)^2 \quad (4)$$

MLP Classifier: The term MLP classifier refers to a neural network that uses multi-layer Perceptron classifiers. Unlike other classifiers such as Support Vectors or Naive Bayes Classifier, MLP Classifier uses the underlying neural network to

ISSN: 2366-1313

perform classification. The perceptron is composed of two layers: a fully connected input layer and an output layer. In Figure 2, the first layer X1, X2... Xn is the input layer and f(x) is the output layer. MLPs have the same concept and publishing process, but they can have several hidden layers between them, as shown in Figure 2.

Decision Tree: This is a versatile forecasting technique that can be used in many situations. In general, decision trees are algorithmic methods for determining alternative ways to classify a data set based on certain criteria. It is one of the most used educational monitoring methods [4]. The goal is to create a model that can learn and predict the value of different objectives using instructions from a decision tree. This is good for knowledge discovery because there is no parameter change. The two nodes of a decision tree are the decision node and the leaf node. Selected nodes are used to make decisions and have many branches, while leaf nodes are the result of decisions and have no additional branches. Figure 3 describes the decision tree process. It is divided into 3 stages. The first stage is the root node, and all others are siblings of the root node.

IV Result and Discussion

The crime data has been subjected to extensive testing to achieve the best results for crime prediction. To begin with, the data breach is first in Google Co lab, and 30% of it is

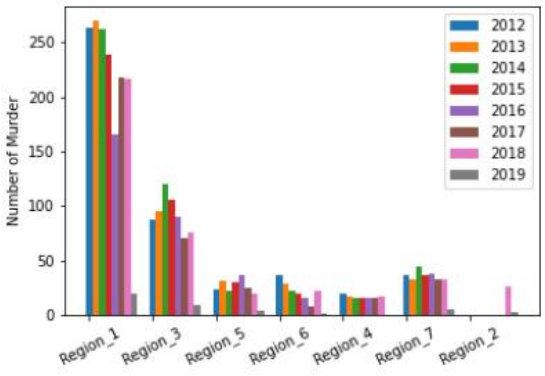

Fig. 4. The rate of Murder in various region.

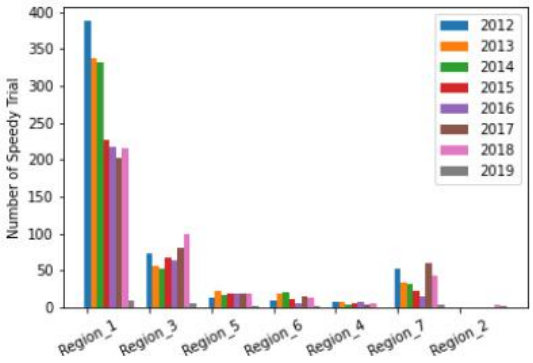

Fig. 5. The rate of Speedy Trial in various region.

of the data set is divided into training and testing. We selected machine learning algorithms [18] and used the training data to create separate models for each algorithm used for testing. The results obtained show the performance of each classifier and the best classifier based on several metrics such as precision, accuracy, recall and F-test for the given data.

ISSN: 2366-1313

Equation 5 is used to calculate the precision.

$$Accuracy = \frac{TP + TN}{TP + TN.FP + FN} \quad (5)$$

The F1 score is calculated by taking the harmonic mean of accuracy and recall. *Equation:5* represents the F1 score.

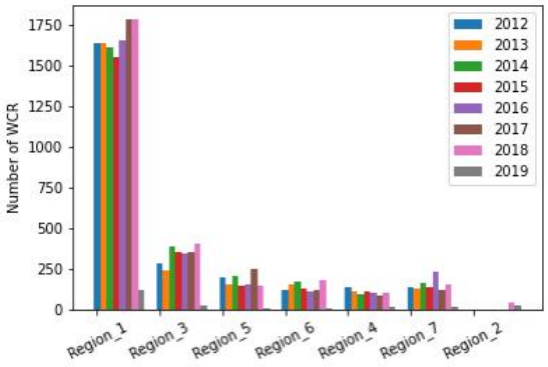

Fig. 6. The rate of WCR in various region.

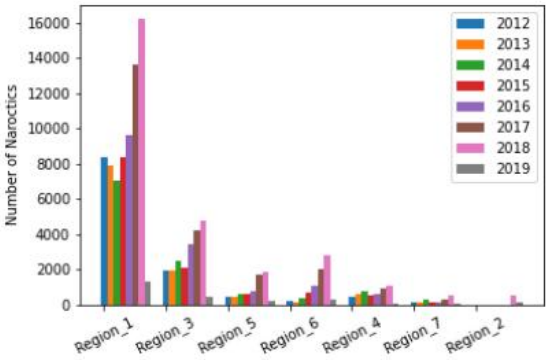

Fig. 7. The rate of Narcotics in various region.

$$F1Score = \frac{2TP + TN}{2TP + FP + FN} \quad (6)$$

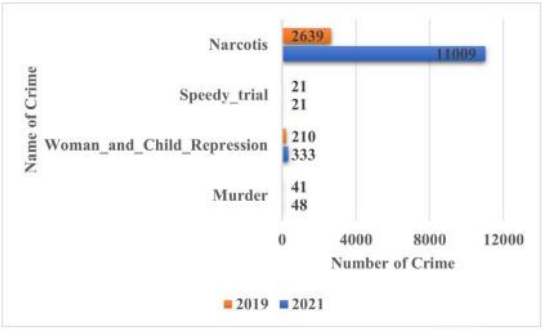

Fig. 8. Crime prediction in 2021 using Decision Tree and compression this crime what was happed 2019.

TABLE III
ALL IMPLEMENTED ALGORITHMS AND METRIC SCORE.

Algorithm	*Accuracy*	*F1 Score*	*Recall*	*Precision*
K-Nearest Neighbor	0.73	0.66	0.69	0.70
MLP Classifier	0.77	0.70	0.73	0.71
Decision Tree	0.81	0.73	0.73	0.78
Impact Learning	0.76	0.69	0.72	0.72

speech. The predictive ability of four artificial intelligence methods has been studied to identify serious crimes.

Table III compares the accuracy, precision, recall, and F1 performance scores of each method. We are able to obtain high levels of accuracy for all these algorithms by using four powerful methods. Table III shows that the decision tree has the highest accuracy of 81%, while the KNN classifier has the lowest accuracy of 73%. In addition, the decision tree algorithm produces the highest value for all other performance measures, 73% of F1 score and recall, and 78% of accuracy. On the contrary, the KNN Classifier algorithm produces the lowest value for performance evaluation, 73% precision, 66% F1 gain, 69% recall, and 70% precision. We also

ISSN: 2366-1313

performed the performance of each performance measure for the other two algorithms, MLP classifier and disturbance learning. The accuracy of the MLP classifier is 77%, while the accuracy of the learning intervention is 76%, which is comparable. The MLP classifier performs well in terms of F1 score and recall compared to interference learning, but interference learning performs better in terms of accuracy.

V CONCLUSION

In this paintings, we added a criminal offense prediction and assessment framework for system gaining knowledge of algorithms at the edge of the network. We collected records from 2012 to 2019 to research and evaluate our predictions. We use system gaining knowledge of to are expecting crime events, which can be useful for enhancing city public safety, a chief trouble addressed in many towns round the world. It's interesting to see how things are finished earlier, and changes can affect the design, specifically when the date spans more than one intervals. With information proof, this solution was created for a selected town within the u . S .. However, if equivalent data may be used, the method will be applied to different cities. Based at the schooling information of the 4 algorithms, we determined that the decision tree approach was powerful and accurate in predicting crime facts. The negative overall performance of the Stump choice set of rules can be attributed to a few inconsistencies in diverse crimes and associated traits (showing the poor comparison of the four algorithms); KNN legs are tighter and only provide correct effects if the measurements comply with the standard model.

REFERENCES

1. Y. Wu, W. Zhang, J. Shen, Z. Mo, and Y. Peng, “Smart city with Chinese characteristics against the background of big data: Idea, action and risk,” Journal of Cleaner Production, vol. 173, pp. 60–66, Feb. 2018.

2. M. Kowsher, A. Tahabilder, and S. A. Murad, “Impact-learning: A robust machine learning algorithm,” in ACM International Conference Proceeding Series, pp. 9–13, Jul. 2020.

3. Priyanka and D. Kumar, “Decision tree classifier: A detailed survey,” International Journal of Information and Decision Sciences, vol. 12, no. 3, pp. 246–269, 2020.

4. L. Jiang, Z. Cai, D. Wang, and S. Jiang, “Survey of improving K- nearest-neighbor for classification,” Proceedings - Fourth

ISSN: 2366-1313

International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2007, vol. 1, pp. 679–683, 2007.

5. L. McClendon and N. Meghanathan, "Using Machine Learning Algorithms to Analyze Crime Data," Machine Learning and Applications: An International Journal, vol. 2, no. 1, pp. 1–12, Mar. 2015.

6. Fonseca, Luis, F. C. Pinto, and S. Sargent. "An Application for Risk of Crime Prediction Using Machine Learning." International Journal of Computer and Systems Engineering 15.2, pp. 166-174, 2021.

7. E. Ahishakiye, D. Taremwa, E. O. Omulo, and I. Niyonzima, "Crime Prediction Using Decision Tree (J48) Classification Algorithm," 2017. [Online]. Available: www.ijcit.com188

8. S. K. Senthil Kumar, G. Adarsh, J. Shashank, and A. Sameer, "CRIME PREDICTION AND ANALYSIS USING MACHINE LEARNING",

[Online]. Available: http://ijte.uk/

9. 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, Jan. 2019.

10. S. A. Murad, Z. R. M. Azmi, Z. H. Hakami, N. J. Prottasha, and M. Kowsher, "Computer-aided system for extending the performance of diabetes analysis and prediction," pp. 465–470, Sep. 2021.

11. L. U. Khan, I. Yaqoob, N. H. Tran, S. M. A. Kazmi, T. N. Dang, C. S. Hong, "Edge Computing Enabled Smart Cities: A Comprehensive Survey," IEEE Internet of Things Journal, Vol.7, Issue 10, pp.10200- 10232, Oct. 2020.

12. Prasadu Peddi, and Dr. Akash Saxena. "studying data mining tools and techniques for predicting student performance" International Journal Of Advance Research And Innovative Ideas In Education Volume 2 Issue 2 2016 Page 1959-1967