



International Journal of Marketing Management

ISSN 2454 - 5007



www.ijmm.net

Email ID: editor@ijmm.net , ijmm.editor9@gmail.com

Machine Learning Algorithms for Edge-Assisted Crime Prediction and Evaluation

¹ Shereen Uzma, ² Mohammed Sannanuddin, ³ Syed Zabiulla Hussaini, ⁴ Nouman Iftaqar Hussain

¹ Assistant Professor, Department of CSE-AIML, Lords Institute of Engineering & Technology.

^{2,3,4} Student Department of CSE-AIML, Lords Institute of Engineering & Technology.

Abstract—

The optimization and management of public safety has become more complex as a result of the new challenges posed by the world's expanding urban populations. Consequently, this study presents a method for forecasting a city's crime rate using demographic data and past events. Specifically, this research suggests a paradigm for evaluating and predicting criminal activity using edge machine learning techniques. Therefore, the effectiveness of the suggested framework is confirmed by a comprehensive examination of four separate types of crimes: drugs, repression of women and children, speedy trial, and murder. A graphic picture of crime in different locations of the nation has been presented by the whole research and implementation process. Choosing, evaluating, and deploying the Machine Learning (ML) model, followed by proposing the crime prediction, concludes the whole job. The use of categorization models allows for the prediction of criminal risk for a given time period and location. Decision Trees, Neural Networks, K-Nearest Neighbors, and Impact Learning are some of the ML approaches being used for event prediction, and their performance is being compared according to the data processing and modification employed. When it comes to crime prediction, the Decision Tree algorithm achieves an accuracy of up to 81%. The results show that using Machine Learning methods helps with crime event prediction, which has helped with public safety.

Index Terms—DL, Decision Tree, KNN, MLP, Crime Prediction, Edge Computing, Impact Learning

INTRODUCTION

The protection of the general public ranks high among the world's most pressing issues. The fast pace of urbanization is one of the factors that have contributed to the increase in anxiety. It is no secret that more and more people are choosing to call cities home. In fact, the United Nations projects that more

than 70% of the global population will reside in urban areas by the year 2050 [1]. As per the Global Terrorism Database, which uses the following definition of terrorist attacks: "acts of violence by non-state actors perpetrated against civilian populations, intended to cause fear, in order to achieve a political objective," the number of terrorist attacks in the past ten years broke all previous records. When used to smart city applications, machine learning (ML) methods may help with urban development and data extraction, which in turn can lead to a decrease in crime rates [2]. The crime rate in various regions of the nation, including Bangladesh, is shown graphically in this study. We presented the seniors with data that spans from 2012 to 2019. We found that area 2 had lower levels of corruption than other cities, whereas region 1 had higher levels, according to the crime prediction model. Based on this data, we developed a model to foretell criminal activity in the year 2021 (check this). Here we show you a comparison of crime rates from 2019 and 2021. As a community service and public safety measure, we have occurrences to alert locals and government authorities to the riskiest areas. More effective and efficient patrol route planning is only one example of how this kind of forecast may be useful; it can also help visitors avoid dangerous areas of the city. Predictions are made using machine learning methods as Impact Learning [3], Decision Tree [4], K-Nearest Neighbors [5], and MLP Classifier. The data processing and transformation methods applied determine the grade of these algorithms. Here is a brief overview of the main contribution: 1) To begin, a framework for the assessment and prediction of criminal activity using ML algorithms at the network's periphery has been suggested. This system is capable of learning from past data in order to identify crimes, as well as evaluating present data in order to prevent crimes. 2) Second, we have built the suggested framework using several ML techniques, including Decision Trees, Neural Networks, K-Nearest Neighbors, and Impact Learning. In which a thorough investigation is

conducted using actual national crime statistics to confirm the efficacy of the suggested structure. 3). In order to safeguard society, our studies demonstrate the effectiveness of predicting crimes like murder, rape, repression of women and children, and narcotics. Here is the remainder of the document. Section II details the relevant work, while Section III presents the Proposed Framework. Section IV then presents the findings, discusses them, and describes the machine learning techniques. Lastly, in section V, we draw a conclusion.

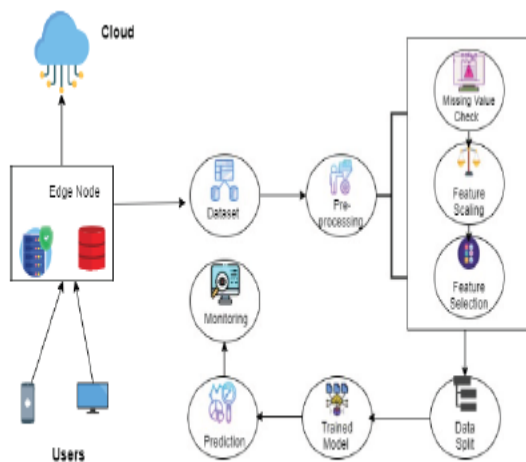


Fig. 1. A Crime Prediction and Evaluation Framework for Machine Learning Algorithms of Network Edge.

RELATED WORK

Using the open-source data-mining program WEKA, a team of researchers compared violent crime patterns in two datasets: one from the Communities and Crime Unnormalized Dataset at the University of California–Irvine, and the other from neighborhoodscout.com, which contains real crime statistics for the state of Mississippi [6]. Luis et al. [7] presented a model for crime prediction in cities using demographic and historical event data. The development of a decision tree (J48) algorithm-based crime prediction prototype model was also investigated by another research group. The testing results showed that the J48 algorithm identified the unknown category of crime data with an accuracy of 94.25287 percent [8]. Another research used a data science and machine learning approach to predict future crime rates in a dataset including information on crimes in Chicago [9]. The data for the crime statistics came from the official website of the

Chicago Police Department. Two other data-processing methods were used to analyze crime statistics from Vancouver during the last fifteen years in an additional research. There is an accuracy range of 39% to 44% for crime predicting in Vancouver [10]. An additional team of researchers in Buenos Aires put out a crime prediction model that used the Python computer language and was based on communes, the city's regions and districts. In addition to the frameworks provided in [6–10], this study also presents a full model that, given a year, can predict the crime rate for that year.

PROPOSED FRAMEWORK

We used a machine learning algorithm to do this task. We used five procedures [11] to put the collected data to use, as seen in Figure 1. Maintaining high-quality data is made easier by the large sample size and sufficient variation. There are a total of six stages to the suggested structure. Data collecting is the first step, and monitoring is the second. The data has been pre-processed after its collection. Before proceeding to feature scaling, we first ensure that no values are missing in this phase. We conclude by selecting the feature for the feature and label data. Following the completion of the

TABLE I
DATASET 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

information divided into training and testing sets. This ML method generated a forecast using the training data. At the very periphery of the network is where the suggested structure is physically set up. Section A. Network at the Network's Edge In order to decrease latency and save bandwidth, edge networking is a distributed computing architecture that brings computation and data storage as near to the point of request as possible [12], [13]. By bringing processing and data collection closer to the point of origin or intended occurrence, edge computing maximizes efficiency. It uses sensors,

computers, and other machinery to gather data, which is then sent to edge servers or the cloud [14]. Depending on the activity and intended outcome, this data may be used to feed analytic and machine learning systems, automate processes, or reveal the present state of a product, system, or gadget [15].

Source of Data (B) We obtained all of the data from the official website of the country's police force [16]. Dates from 2012 all the way up to 2019 have been used. Many crime-related data were identified; however, we selected a subset of the most relevant ones, which are both nationally widespread and showing signs of rapid increase. Here is a description of the data that was gathered for our study in Table I.

Section C: Pre-processing Data Preprocessing data on crime prediction is done once different records are acquired [13]. This dataset includes four observation values for crime prediction: narcotic, woman child repression, speedy trial, and murder. Although not all of the material was walldecorated, we did get a lot of it from this page. In order to prepare the data for this, we used three separate procedures.

- Check for missing values.
- 2. Feature Sizing.
- 3. Feature Picking.

Check for Missing Value: Ethics in Leadership For the most part, people think of a missing value as a value that was not preserved in the example, but that's not always the case. Having no value at all is a common occurrence in data. In addition, the majority of systems used for presenting forecasts cannot handle data that is missing. Accordingly, this issue needs fixing before modeling may start. To fill up a blank, we used the mean. In order to determine the average, or mean, of a set of data,

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

Scaling Features: Normalizing or scaling features is a crucial step in machine learning methods; without it, objective systems won't work. Feature scaling techniques include Standardization, Mean Normalization, Unit vectors, Min-Max Scaling, Variance Scaling, and Standardization. Here, min-max normalization was used. Both the basic formula of min-max in [0,1] and the range of normalization in min-max in [0,1] or [-1,1] are provided by equation:2.

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

Another task that needs doing before a model can be deployed is feature selection. The main goal of this method is to establish a relationship between the feature and the target variables. While retaining the

essential highlights, we have eliminated a few that do not contribute significantly to the aim variable. Reducing the amount of highlights lowers the computational cost. Our dataset has thirteen feature columns.

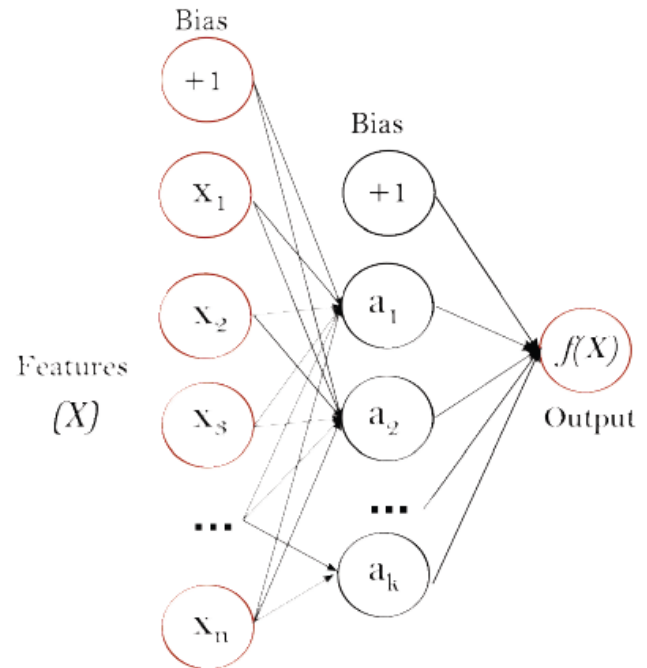


Fig. 2. Diagram of MLP Classifier [17].

Division of Data

All algorithms used in machine learning have training data as their basis. The data that was collected has been divided into two parts. A training set is one, while a data set is the other.

TABLE II
DETAILED TRAINING PARAMETERS OF IMPLEMENTED ALGORITHMS.

Name of Algorithms	Description	Initial Parameters
K-Nearest Neighbors	The majority of votes from its fixed neighbors classify the new data point.	Nearest Neighbours: 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 dataset, a huge number of Epoch are run. Iteration is another name for Epoch.	Epoch: 2000

constitutes a set for testing purposes. We trained using 70% of the data and tested with 30%. This is the last stage before training a machine learning model on this dataset. Section E on the Trained Model In order to discover the most astounding accuracy, we used four machine learning techniques. Every one of these methods works well with this dataset. To find the top-performing model, we look at how many algorithms have the lowest error rate. All of the algorithms that were used and their parameters are summarized in Table II. Educational Influence: Using supervised classification and linear or polynomial regression, impact learning is a knowledge learn technique. It is also useful for comparing different data systems that are available. Using this method to infer the effect of separate contest characteristics is unusual. Put another way, students learn how the RNI [intrinsic rate of natural growth] comes into play. Here, equation:3 represents the RNI.

$$\frac{wP}{dt} \approx rP \quad (3)$$

The impact learning equation is shown by equation:4:

$$Imp = (y' - (\frac{k \sum_{i=1}^n w_i x_i}{r - w_y k} + b))^{2/N} \quad (4)$$

K- Nearest Neighbor Classifier: The K-Nearest Neighbors

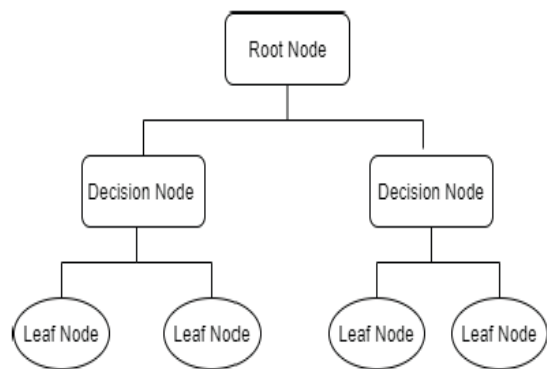


Fig. 3. Diagram of Decision Tree Classifier.

One of the most basic Machine Learning algorithms, K-NN is based on the Supervised Learning methodology. It stores all the data that is already available and uses it to classify new data points according to how similar they are to previous data. This allows new data to be quickly classified into a well-defined category using the K-NN method [5]. The KNN calculates the shortest distance between the selected neighbors. It uses the Euclidean distance

function to calculate the distances between each new data point and the existing data points. Equation:5 can be used to determine the value of the distance.

$$EuclideanDistance = \sqrt{\sum_{i=0}^k (x_i - y_i)^2} \quad (5)$$

A Neural Network that employs Multi-layer Perceptron Classifiers is known as an MLP Classifier. When it comes to classification, MLP Classifier is different from the rest of the pack since it relies on an underlying Neural Network. A completely linked input layer and an output layer comprise the perceptron. The input layer in Figure 2 is X1, X2... Xn, while the output layer is f(x). The input and output layers of an MLP are identical, but as shown in Figure 2, there is room for more than one hidden layer in between them. The Decision Tree: This forecasting approach is very adaptable and may be used for a variety of circumstances. In a nutshell, decision trees are an algorithmic tool for finding different ways to divide a dataset according to certain criteria. It is among the supervised learning strategies that are most often used [4]. The objective is to build a model that, given some basic decision-tree instructions, can learn to predict the value of a target variable. Since it doesn't need changing any parameters, it's perfect for discovering new information. Two nodes make up a decision tree: the decision node and the leaf node. In a decision-making process, a choice node may have several branches, but a leaf node is the end result of that decision and can only have one. The Decision Tree technique is shown in Figure 3. It's broken down into three parts. Every other node is a sibling of the root node, which is the first step. Section F. Supervision The Infraction The responsible agency can keep tabs on or assess the outcomes based on the forecasts produced by the prediction framework using network edge ML algorithms [16], [17]. After keeping an eye on the forecast, the government, police, or security agency can take the necessary steps. Another feature is the ability to set the monitoring time and send an emergency message to the appropriate agency, be it the government, the police, or the security agent, through the network. Thus, this strategy can help bring down a country's crime rate.

RESULT AND DISCUSSION

In order to provide the most accurate results for crime prediction, the crime dataset was tested extensively. The first step is to use Google Colab to preprocess

the crime dataset. Then, split 30% of the dataset into separate training and test sets.

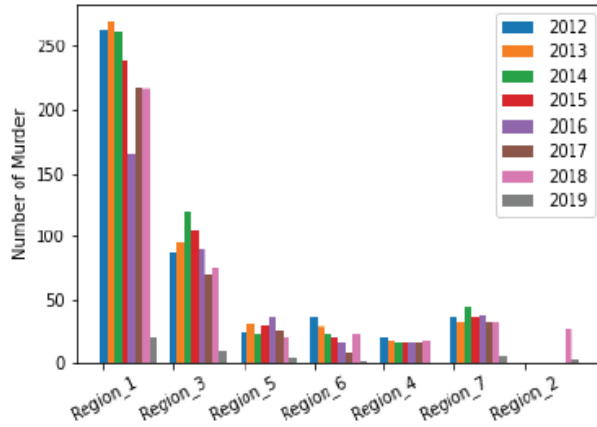


Fig. 4. The rate of Murder in various region.

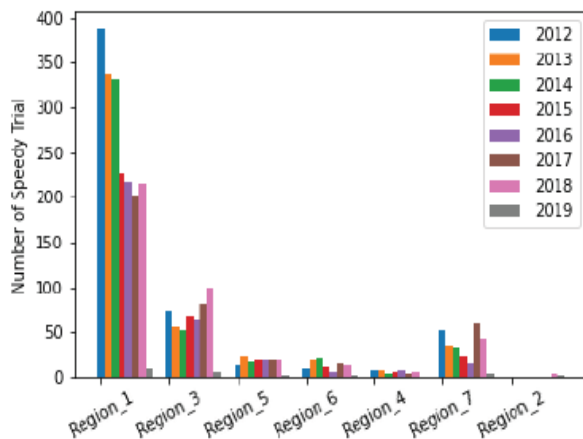


Fig. 5. The rate of Speedy Trial in various region.

We choose ML methods [18] and trained a classifier model for each algorithm using the training data. Then we tested the models. The results show how well each classifier performed and which one was the best for the provided dataset in terms of several metrics including recall, accuracy, precision, and F-measurement. Accuracy is determined using Equation 6.

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

The following equation represents the algorithm's sensitivity or recall as a proportion of all relevant results accurately classified by the algorithm: 7.

$$Recall = \frac{TP}{TP + TN} \quad (7)$$

The proportion of correctly identified occurrences is known as precision or samples from the group that was determined to be positive. You can see this in Equation: 8.

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

In order to get the F1 score, the harmonic mean of precision and memorization. The formula for the F1 score is 9.

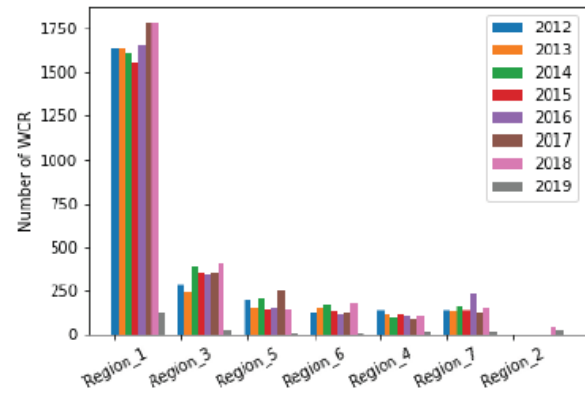


Fig. 6. The rate of WCR in various region.

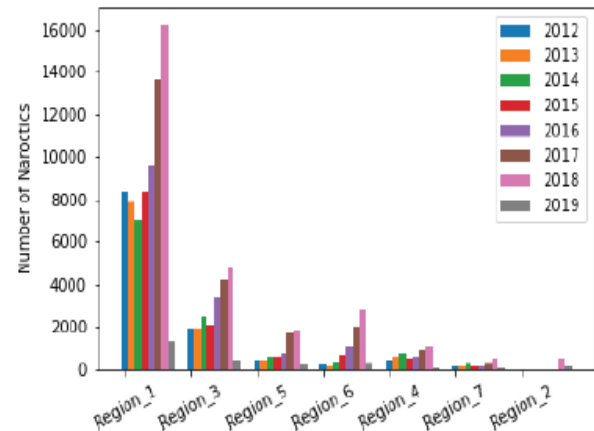


Fig. 7. The rate of Narcotics in various region.

$$F1Score = \frac{2TP + TN}{2TP + FP + FN} \quad (9)$$

When evaluating the performance measures, we took into account True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives

(FN). The categorization of crime prediction severity was used to assess the prediction abilities of four artificial intelligence approaches. The pre-classifier is shown in Table III.

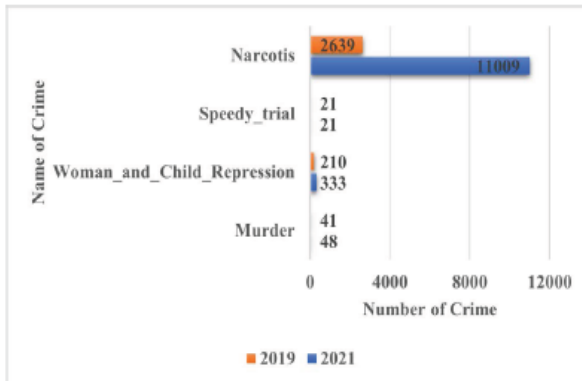


Fig. 8. Crime prediction in 2021 using Decision Tree and compression this crime what was happened 2019.

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

word lists. In order to classify crimes according to their severity, researchers examined the predictive capabilities of four AI methods. Each method's performance indicators are compared in Table III, including accuracy, precision, recall, and F1 score. Using four strong approaches, we achieved excellent accuracy with all of these algorithms. The decision tree achieved an accuracy of 81%, according to Table III, while the KNN classifier achieved an accuracy of 73%, which is the lowest. Decision tree algorithm also achieved the highest values for all other performance parameters, including 78% Precision, 73% Recall, and 71% F1 score. However, when it came to performance indicators like accuracy (73%), F1 Score (66%), recall (70%), and precision (70%), the KNN Classifier algorithm came out on top. Additionally, using the MLP classifier and impact learning algorithms, we had excellent results across all performance criteria. Comparatively, MLP classifier accuracy is 77% and impact learning accuracy is 76%. While impact learning outperformed the MLP classifier in terms of recall and F1 score, the former was more accurate.

The murder rate in area 1 is quite high, as seen in Figure 4. Region 2 has a lower murder rate. The murder rate is second highest in the nation in Region 3. The rate is also typical for four more metropolitan regions. The murder rate in the metro region is steadily declining, which is excellent news. Figure 5 shows a similar graph for a Speedy Trail. Region 1 has a much higher incidence compared to other urban regions. For Region 3, the second post is here. Both Region 4 and Region 2 have a lower rate of Speedy trials. The other three major cities are in the middle of the rate's progression. But each year, fewer and fewer quick trials are being conducted. There were 544 quick trials in 2012, but just 21 this year. According to Figure 6, the rate of WCR (Woman and Child Repression) was higher in 2018 compared to 2019. All major cities may relate to this situation. The crime rate has always been higher in region 1. The same holds true for WCR in this situation. There has never been a region with a higher crime rate than Region 3. In area 5, the rate is much higher. The rate is lower for the following four areas. The most aggravating part is that the WCR situation doesn't alter over time, although this ratio has dropped significantly this year. It is clear that there were much more drugs involved in the other three cases. Even if the rate was much higher in 2018, it remains at risk. Not only that, but unlike other forms of criminality, the amount of narcotics is actually growing with time, which is the most shocking part. The prevalence of drugs is increasing everywhere. The rate is higher than in other cities since region 1 has a higher population density. Additionally, as seen in Figure 7, opioids pose a threat in area 3. On the other hand, we found a promising situation for Region 2, where the narcotics population is much lower than in other regions. The national crime rate in 2021 is shown in Figure 8. The number of homicides is expected to rise somewhat in 2021 when compared to 2019. Our best guess puts the number at 48 by year's end, up from 41 in 2019. In both 2019 and 2021, we were able to get the same quick trial outcome, which is 21. We found a horrible outcome for Narcotic. By the end of 2021, Narcotic will have risen to 11009, a much higher level than in 2019. Many of my fellow citizens may be quite worried about this discovery. Furthermore, we disclosed that 2021 would see an increase in the oppression of women and children. According to our prediction, the overall crime rate will reach an all-time high by the year 2021. We examined four distinct types of criminal activity and found that three of them would

see an increase in 2021. Our study examines four different crimes throughout the nation, while their [19] article focuses on only one. In the end, it's possible to say that this research will help get people thinking about environmental issues.

CONCLUSION

Here, we provide a system for evaluating and predicting criminal activity using network edge machine learning methods. In order to assess and analyze our prediction, we gathered data from 2012 all the way up to 2019. The application of machine learning techniques to predict future criminal activity has the potential to greatly enhance city public safety, an issue that plagues several urban centers worldwide. I found it really interesting to see how the model's output may change depending on the pre-processing and transformation, especially when dividing the day into many time periods. The data's origin meant that this solution could only be tailored to one city in the nation. It is possible to apply the approach to other cities provided comparable data is made available. We discover that the decision tree technique is the most effective and precise in predicting crime statistics, using the training set input for the four algorithms. Because there is a low correlation coefficient among the four algorithms, we can infer that the Decision Stump algorithm performs poorly. On the other hand, the KNN branches are more strict and produce accurate results only if the test set matches the pattern that was modelled.

REFERENCES

- [1] "68% of the world population projected to live in urban areas by 2050, says UN — UN DESA — United Nations Department of Economic and Social Affairs." <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html> (accessed Oct. 14, 2021).
- [2] 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.
- [3] 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.
- [4] 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.
- [5] L. Jiang, Z. Cai, D. Wang, and S. Jiang, "Survey of improving Knearest- neighbor for classification," *Proceedings - Fourth International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2007*, vol. 1, pp. 679–683, 2007.
- [6] 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.
- [7] Fonseca, Luis, F. C. Pinto, and S. Sargento. "An Application for Risk of Crime Prediction Using Machine Learning." *International Journal of Computer and Systems Engineering* 15.2, pp. 166–174, 2021.
- [8] E. Ahishakiye, D. Taremwa, E. O. Omulo, and I. Niyonzima, "Crime Prediction Using Decision Tree (J48) Classification Algorithm," 2017. [Online]. Available: www.ijcit.com188
- [9] S. K. Senthil Kumar, G. Adarsh, J. Shashank, and A. Sameer, "CRIME PREDICTION AND ANALYSIS USING MACHINE LEARNING", [Online]. Available: <http://ijte.uk/>
- [10] 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.
- [11] 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.
- [12] 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.
- [13] M. S. Munir, S. F. Abedin and C. S. Hong, "Artificial Intelligencebased Service Aggregation for Mobile-Agent in Edge Computing," *2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 1-6, 2019.
- [14] M. S. Munir, S. F. Abedin, N. H. Tran and C. S. Hong, "When Edge Computing Meets Microgrid: A Deep Reinforcement Learning Approach," in *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7360-7374, Oct. 2019.
- [15] M. S. Munir, S. F. Abedin, D. H. Kim, N. H. Tran, Z. Han and C. S. Hong, "A Multi-Agent System toward the Green Edge Computing with

Microgrid,” 2019 IEEE Global Communications Conference (GLOBECOM), pp. 1-7, 2019.

[16] “Bangladesh Police”.
<https://www.police.gov.bd/en/crime-statistic/year/2019> (accessed Oct. 13, 2021).

[17] “sklearn.tree.DecisionTreeClassifier-scikitlearn 1.0 documentation.”<https://scikitlearn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html> (accessed Oct. 13, 2021).

[18] A. J. M. Muzahid, S. F. Kamarulzaman, and M. A. Rahman, “Comparison of PPO and SAC Algorithms Towards Decision Making Strategies for Collision Avoidance Among Multiple Autonomous Vehicles,” pp. 200–205, Sep. 2021.

[19] M. A. Awal, J. Rabbi, S. I. Hossain, and M. M. A. Hashem, “Using linear regression to forecast future trends in crime of Bangladesh.” 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV). IEEE, 2016.