DOI: https://dx.doi.org/10.4314/gjpas.v30i2.10

GLOBAL JOURNAL OF PURE AND APPLIED SCIENCES VOL. 30, 2024: 203-219
COPYRIGHT© BACHUDO SCIENCE CO. LTD PRINTED IN NIGERIA ISSN 1118 – 0579, e-ISSN: 2992 - 4464
www.globaljournalseries.com.ng, Email: globaljournalseries@gmail.com

PREDICTIVE CRIME ANALYSIS USING MULTI-LAYER PERCEPTRON ARCHITECTURE

O. B. IKWEN, I. E. ETENG, AND F. U. OGBAN

(Received 15 April 2024; Revision Accepted 3 June 2024)

ABSTRACT

In an extended period, crime and statistical professionals' analyses have channeled their skills, knowledge, and expertise to anticipate the timing and locations of future criminal incidents, although with varying degrees of success. The surge in criminal activities and the evolving strategies adopted by modern offenders have strained the efficacy of existing predictive methods. This study introduces a novel approach by leveraging the Multi-Layer Perceptron (MLP) architecture, a cutting-edge technology that uses the back-propagation algorithm to develop a predictive model for analyzing crime data. A total of 4,748 records were collected from the Cross River State Police Command. Data training was conducted using MLP, and the dataset was divided into 70% for training and 30% for testing. The outcomes of the MLP model, characterized by a precision of 0.84, an accuracy of 74%, a recall rate of 0.73, and an F1-score of 0.79, underline the suitability and effectiveness of employing MLP as an invaluable tool in crime prediction.

KEYWORDS: Multi-Layer Perceptron, Crime Prediction, Machine Learning, Neural Network

INTRODUCTION

Predictive analytics is the scientific equivalent of a crystal ball by allowing analysts to glean insights from vast datasets using statistical methods; such as modeling, and machine learning algorithms. The efficacy of these techniques hinges on several factors, including data quality, analyst expertise, and algorithm suitability. For law enforcement analysts, deciphering the complexities within ever-growing volumes of criminal data poses a formidable challenge, mainly when conducted by a team. This research, conducted in close collaboration with law enforcement agencies, is a testament to the collective effort and dedication to advancing the field of crime analytics and the practical application of our findings. Cesario et al. [1].

Global law enforcement teams, including those in Nigeria, face challenges from criminals who are better equipped and more sophisticated than government crime fighters.

As criminal activities increase, there is a growing need for advanced geographical information systems and innovative data mining techniques to improve crime analytics and protect communities. This research illustrates the practical application of machine learning, a process where computer systems learn from examples, in addressing these challenges. It offers a promising approach to enhancing crime analytics and safeguarding communities.

Supervised and unsupervised learning algorithms are two primary categories of machine learning algorithms widely employed in crime prediction and analysis. Supervised learning algorithms utilize labeled training data to infer the correct answers and make predictions based on specific attributes or qualities. In crime prediction, sophisticated techniques such as multilayer perceptron (MLP) machine learning models are essential.

O. B. Ikwen, The Nigeria Police Force, Cross River State Police Command, Diamond Hill Calabar, Nigeria
I. E. Eteng, Department of Computer Science, Faculty of Science, University of Calabar, Calabar, Nigeria
F. U. Ogban, Department of Computer Science, Faculty of Science, University of Calabar, Calabar, Nigeria

Cross River State, situated in southern Nigeria, comprises 18 Local Government Areas (LGAs), each with unique socio-economic dynamics and crime challenges. With its vast boundaries and dense population, the state experiences frequent crime occurrences, necessitating the deployment of advanced predictive models like MLP. MLP, a feedforward neural network comprising multiple layers of interconnected artificial neurons, holds significant potential for crime prediction and proactive crime prevention measures.

(NO 5 Correction is captured in this paragraph) MLP's key advantages over other algorithms are its ability to capture complex relationships between various crime-related features and pattern recognition. By training the network on historical crime data and relevant contextual factors, MLP can effectively predict crime rates and identify potential crime hotspots.it is important to note that MLP Furthermore, MLP's flexibility and scalability make it suitable for handling large-scale datasets and adapting to evolving crime patterns.

Some of the outstanding features of MLP include:

- **Non-Linear Pattern Detection:** MLP identifies intricate, non-linear patterns in historical crime data. This is essential for recognizing crime incidents' complex relationships and dependencies, enabling law enforcement to anticipate and address evolving criminal behaviors effectively.
- **Scalability:** MLP can easily accommodate the extensive historical crime datasets typically encountered by law enforcement agencies. Its scalability allows for including a broad spectrum of crime-related variables and parameters, which is crucial for comprehensive crime analysis. This research collected a wide array of datasets resulting in four thousand seven hundred and forty-eight (4748) records.
- **Generalisation and Adaptation: MLP's ability to generalise** from historical data to predict future scenarios aligns with crime detection's dynamic nature. The model can adapt to shifting crime patterns and emerging trends by continuously learning from updated historical data.
- **Interpretability:** MLP's hidden layers can be analysed to reveal which features and feature combinations most influence crime predictions. This interpretability aids law enforcement in understanding the drivers of crime and making data-informed decisions.
- **Feature Extraction and Selection:** MLP can automatically learn and extract relevant features from historical data, reducing the reliance on manual feature engineering. In the context of the Nigerian Police Force, this capability aids in uncovering latent indicators of criminal activity.
- **Parallel Processing:** Parallel processing can further enhance MLP's efficiency, making it capable of swiftly handling vast datasets often present in historical crime records.
- **Multi-Class Classification:** MLP supports multi-class classification, which is invaluable for categorising crimes into various types and identifying trends within the Nigerian context, such as property crimes, violent crimes, and white-collar crimes.
- **Regularisation Techniques:** MLP can be reinforced with regularisation techniques to counteract overfitting and enhance model robustness, ensuring the reliability of predictions in real-world crime detection scenarios.
- **Ensemble Learning:** Ensembling multiple MLP networks can amplify the accuracy and resilience of crime rate predictions by combining the insights of diverse models and mitigating potential biases.
- **Continuous Learning and Updating:** MLP's adaptability allows the incorporation of new historical data, permitting the model to remain current with evolving criminal dynamics and emerging threats within the Nigerian police force's jurisdiction.

Researchers have applied MLP to Predict in various fields. However, very little research has been carried out with Multi-Layer Perceptron for crime Prediction. A review of related works revealed that several were more concerned with combining MLP with other algorithms and applying MLP in different Fields. A number of the reviewed papers are presented below. Section. First and foremost, we review papers related to crime prediction employing MLP and various other algorithms. Subsequently, we examine documents that have employed MLP in diverse domains beyond crime prediction.

In their study, Kang HW and Kang HB [4] propose a feature-level data fusion method incorporating environmental context, leveraging a deep neural network (DNN). Their dataset, drawn from various online crime statistics, demographics, meteorological sources, and images of Chicago, Illinois, is augmented with critical data from multiple domains. Experimental findings demonstrate the superior predictive accuracy of the DNN model compared to alternative prediction models.

Mandalapu, Varun, et al. [5] identify potential gaps and suggest future avenues for refining crime prediction accuracy. Their study provides an extensive overview of machine learning and deep learning applications in crime prediction, aiming to deepen the understanding of predictive techniques in this domain.

On the other hand, Lin Y-L et al. [6] integrate the concept of the criminal environment into grid-based crime prediction modelling. They establish a diverse array of spatial-temporal features by leveraging 84 types of geographic information obtained through the Google Places API, applied to theft data from Taoyuan City, Taiwan.

Employing Deep Neural Networks, their approach outperforms traditional models such as Random Decision Forests, Support Vector Machines, and K-nearest neighbour algorithms. Experimental results underscore the importance of geographic feature design in enhancing model performance and interpretability. Furthermore, testing for crime displacement reveals superior performance of the proposed model compared to the baseline.

Walczak S. [7] presents a comprehensive analysis of current research on applying neural networks for crime forecasting and other decision-making processes in law enforcement. These predictive models utilising neural networks for crime detection leverage geographical data to offer timely insights into criminal activities, enhancing law enforcement decision-making. Notably, these models demonstrated the capability to forecast the category of criminal activity with an accuracy of 16.4% for 27 different crime types or 27.1% when categorised into seven crime categories. Furthermore, location prediction neural networks exhibited a 31.2% accuracy in predicting zip codes or adjacent locations. Bappee et al. [8] introduce a study proposing identifying points derived from crime-prone areas using Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN). Subsequently, a spatial distance characteristic is calculated according to the distance of these points for various crime types, serving as a feature for classifiers. The crafted attributes were tested on crime data obtained from the Royal Canadian Mounted Police of Halifax, NS, resulting in a notable enhancement in the accuracy of crime forecasting.

Shah et al. [9] present the outcomes of their research investigating the integration of machine learning (ML) and computer vision techniques for crime detection, prevention, and resolution by law enforcement agencies. Their study aims to explore how this combination can significantly improve the accuracy and speed of crime-related tasks. The findings underscore the potential of ML and computer vision techniques to revolutionise law enforcement practices.

Ramirez-Alcocer et al. [10] introduce an innovative approach based on deep learning to classify incidents related to community security crimes through predictive analysis. Their proposed predictive algorithm employs a Long Short-Term Memory (LSTM) neural network, trained using a concise set of features, facilitating accurate class label prediction during validation. The efficiency of this method is assessed using an extensive collection of open data containing historical information about crimes in an intelligent city. In their study, Cortes et al. [11] analyse and validate the utility of Machine Learning applications within the Mexican and Latin American contexts, focusing on algorithmic fairness and ethical concerns surrounding employing Artificial Intelligence in predicting criminal activities. Their findings from neural network tests reveal a high accuracy rate in predicting urban crimes, predominantly leveraging historical crime data from the cities under study. Chatzimichail et al. [12] developed a Multi-Layer Perceptron model for predicting persistent respiratory disorders in youngsters using advanced machine intelligence methodology. By discrimination partial least square regression, 9 out of 48 extrapolative aspects related to the tenacious respiratory disorder are classified. The multilayer perceptron algorithm was used. The researchers specify that the system effectively forecasts asthma outcomes with high prediction accuracy. Karthik Kalyan et al. [13]. A Multilayer Perceptron model was developed to assess the presence of medical conditions in various texture patterns extracted from ultrasonic liver images. By analysing these images, an echo texture pattern was identified. Liver infection conditions such as fatty liver, cirrhosis, and abnormalities are known to produce distinctive echo patterns in abdominal imaging. Zavvar et al. [14] explored spam mail detection by employing a combination of machine-learning algorithms. They collected sample emails from the UCI Database, pre-processing them to remove redundant and root words. Utilising particle swarm optimisation (PSO) and the Multi-Layer Perceptron (MLP), they identified relevant features from the initial set and constructed a spam detection model using support vector machines (SVM). Similarly, Pilita et al. [15] conducted a detailed technical analysis of Twitter data to prepare for prediction using Multilayer Perceptron (MLP) in Weka algorithms. By leveraging the Twitter developer’s account, they legally extracted 1000 tweets, validating 931 tweets essential for exploration. The study optimised parameters, tuning the epoch to 200 and evaluating eight features with a 70:30 split for testing, providing insights on parameter exploration to enhance reliability.

Rozos et al. [16] employed a multilayer perceptron network to generate synthetic daily rainfall data in their study. This method was implemented in two areas with distinct weather conditions and extensive daily measurement records spanning over a century and more than four decades. The findings suggest that their proposed methodology effectively preserves crucial statistical characteristics. A notable advantage of this model is its ease of application post-training and adaptability for analysing various hydrological time series. Meanwhile, Ke KC and Huang [17] introduced a Multilayer Perceptron (MLP) neural network model integrated with quality indices to swiftly and automatically predict finished product geometry. Their study incorporated pressure curves detected by an in-mould pressure sensor, which reflect changes in a melt flow state, various indicators, and moulding quality. The results demonstrate the accuracy of training and testing concerning geometric widths, affirming the feasibility and efficacy of the proposed approach.

Tamim and Elshrkawey [18] introduced a supervised learning approach based on a multi-layer perceptron neural network and a carefully curated feature vector. Specifically, they formulated a 24-dimensional feature vector for each pixel of a retinal fundus image, incorporating information on local intensity, morphology transformation, and principal moments of phase congruency, Hessian, and difference of Gaussian values. The proposed method underwent rigorous testing on three well-established datasets: Digital Retinal Image for Extraction (DRIVE), Structure Analysis of the Retina (STARE), and CHASED_DB1 datasets. The experimental results, encompassing both visual and quantitative evaluations, validate the robustness of the proposed methodology. Furthermore, they demonstrate its superiority over seven comparable state-of-the-art methods.

Bikku [19] emphasises the deep learning mechanism for disease prediction based on historical medical data. The study utilises Multilayer Perceptron (MLP) for disease prediction, feature extraction, and medical data classification. The outcomes of the MLP Neural Network classifier reveal a sparse distribution of the selected feature subset number, leading to lower overall performance.

Huang et al. [20] propose a semi-supervised multiple-layer perceptron (SSMLP) approach for landslide susceptibility mapping (LSM). The method involves generating an initial LSM using MLP based on recorded landslide samples and environmental factors, then classification into five susceptibility levels. New landslide samples are selected from highly susceptible areas, while non-landslide samples are chosen from low susceptibility regions. The expanded dataset is then used in MLP to predict the final LSM. Comparative analysis with conventional supervised (MLP) and unsupervised (K-means clustering) machine learning models shows the superior performance of the SSMLP model in LSM for Xunwu County, Jiangxi Province, China.

Savalia and Emamian [21] carried out a research endeavour to differentiate between various arrhythmias utilising deep neural network algorithms, including the multi-layer perceptron (MLP) and convolutional neural network (CNN). The investigation employed the TensorFlow library, developed by Google for deep learning and machine learning, implemented in Python. Training, testing, and validation of the MLP and CNN algorithms were executed using ECG databases accessible at PhysioBank.com and kaggle.com. The proposed algorithm encompasses four hidden layers with weights and biases in MLP and four-layer convolutional neural networks that map ECG samples to distinct arrhythmia classes. Results demonstrate that the algorithm's accuracy surpasses current algorithms devised by other cardiologists, particularly in sensitivity and precision metrics.

Mogan and Jashila [22] integrated a pre-trained VGG-16 model with a multilayer perceptron in this study to optimise performance across different covariates. Initially, relevant images were obtained by some transformation on silhouettes over a gait cycle. Finally, a classification layer was utilised to identify the subject. Experiments were conducted to evaluate the proposed method's performance on the CASIA-B dataset, the OU-ISIR dataset D, and the OU-ISIR large population dataset. Analysis against well-established methods showed superior performance across all datasets.

In their research, Baressi [23] leveraged a publicly accessible dataset spanning 51 days (from January 22, 2020, to March 12, 2020), comprising data on ailing, recovered, and dead patients across 406 locations. Designed initially as a time-series dataset, it was restructured into a regression dataset and employed to train a multilayer perceptron (MLP) artificial neural network (ANN). During cross-validation, the achieved scores revealed robust performance metrics, with scores of 0.94 for confirmed cases, 0.781 for recovered cases, and 0.986 for deceased cases. This indicates a high level of robustness in predicting deceased cases, moderate robustness in predicting confirmed cases, and relatively lower robustness in predicting recovered cases.

In their study, Li et al. [24] developed a particle-swarm-optimized multilayer perceptron (PSO-MLP) model to overcome the constraints of the traditional gradient descent algorithm and identify the most effective structural parameters for implementing MLP in LSP. Their research revealed that the PSO-MLP model achieved superior prediction accuracy, surpassing MLP-only models, as evidenced by the area under the receiver operating characteristic curve. Furthermore, their findings indicate that the PSO-MLP model holds promise for enhanced prediction and classification performance in diverse domains compared to conventional ANNs and statistical models.

Yun et al. [25] devised a robust classification approach harnessing magnetic resonance imaging (MRI) alongside feature extraction and selection combinations facilitated by human and machine learning methodologies. They explored the integration of radionics or deep features within a limited dataset. Their findings revealed comparable high performance between the generalised linear model (GLM) classifier and Multi-Layer Perceptron (MLP) employing radiomics features during internal validation. Notably, the MLP network exhibited resilience during external validation conducted using diverse MRI protocols, whereas the Convolutional Neural Network (CNN) showed minor efficacy in both validation sets.

Moreover, their study underscores the effectiveness of combining radiomic features with an MLP network classifier, presenting a high-performing and adaptable model for classification tasks, particularly with datasets characterised by heterogeneous MRI protocols.

Mokbal et al. [26] introduce a robust artificial neural network-based multilayer perceptron (MLP) scheme enhanced with a dynamic feature extractor to detect XSS attacks. Their detection methodology harnesses a significant real-world dataset, incorporates a dynamic feature extraction mechanism, and utilises the MLP model. The scheme presents promising and cutting-edge outcomes through rigorous experimentation across multiple metrics, including accuracy, detection probabilities, false positive rate, and AUC-ROC scores. Consequently, it exhibits potential for deployment in identifying XSS-based attacks, whether on the client or server sides.

Gaikwad et al. [27] present a specialised hardware-based Human Activity Recognition (HAR) system tailored for smart military wearables, employing a multilayer perceptron (MLP) algorithm for activity classification. Their approach achieves a versatile and practical hardware design by integrating the inherent MLP architecture with parallel computation on Field-Programmable Gate Array (FPGA) platforms. Evaluation conducted using the UCI human activity dataset, comprising 7767 feature samples from 20 subjects, underscores the system's performance. Through training, validation, and testing across ten distinct MLP models with varied topologies, the study reveals that the MLP designed with 16-bit fixed-point data precision yields the most efficient implementation regarding classification accuracy, resource utilisation, and power consumption.

Lorencin et al. [28] present a novel approach employing genetic algorithms (GA) to optimise the design of multi-layer perceptron (MLP) for estimating power output in combined cycle power plants. Utilising a UCI Machine Learning Repository dataset comprising 9568 data points divided into training (7500) and testing (2068) sets, they validate MLP configurations generated by GA using Bland-Altman analysis. The optimised MLP, with five hidden layers of 80, 25, 65, 75, and 80 nodes, outperforms complex algorithms like KStar and tree-based approaches.

In a study by Pham et al. [29], the authors compare the predictive performance of Functional Trees (FT), Multilayer Perceptron Neural Networks (MLP Neural Nets), and Naïve Bayes (NB) for assessing landslide susceptibility in the Uttarakhand Area of India. They employ feature selection with the Linear Support Vector Machine (LSVM) algorithm to evaluate the influence of conditioning factors on landslide models. Using a training dataset, they construct NB, MLP Neural Nets, and FT models and validate and compare their predictive capabilities using success rate and predictive rate curves. Overall, all three models demonstrate high performance in landslide susceptibility assessment.

Wang et al. [30] introduce a novel Alzheimer's disease (AD) detection system demonstrating superior performance compared to existing methods. Utilising data from the OASIS dataset, comprising 28 AD patients and 98 healthy controls (HCs), the researchers employed the inter-class variance criterion to select a single slice from 3D volumetric data. The classification system integrates three key components: wavelet entropy, multilayer perceptron (MLP), and biogeography-based optimization. Comparative analysis revealed that the proposed pathological brain detection system outperforms the six most recent approaches.

Meha and Manan [31] aim to review Artificial Neural Networks (ANNs), specifically Multi-Layer Perceptron Neural Networks (MLP) and Convolutional Neural Networks (CNNs) for the early diagnosis of breast cancer by detecting breast malignancies. Their study compares the accuracy of MLP and CNN in identifying breast cell malignancies. CNN exhibits slightly higher accuracy than MLP for diagnosing and detecting breast cancer. Other research works that have used police-related datasets and analytics methods to solve problems relevant to developing economies include [32], [33], and [34].

This study introduces a Multilayer Perceptron Neural Network (MLPNN) algorithm tailored for forecasting future crime occurrences and trends within Cross River State. Utilizing comprehensive crime datasets sourced from the Nigerian Police across all eighteen (18) local Government areas of the State, the algorithm is expertly developed and calibrated to provide insightful predictions.

1. MATERIAL AND METHODS

The methodology adopted in developing this multilayer perceptron model for crime prediction includes dataset collection and variable identification, dataset preprocessing, hyperparameter turning, dataset splitting, the multilayer perceptron model development for crime prediction, and model evaluation.

a) Dataset collection and variable identification: The dataset used in this study was all original data obtained from criminal records of the Nigeria Police Force (NPF) in Cross River State. It was collated in Excel format. The data source covers 12 years from 2010 to 2022 and contains 4748 records captured under 11 attributes and 1 class variable, "crime type." The dataset attributes and their description are shown in Table 1.

Table 1: police crime dataset description

S/No.	Attribute	Description	Data Type
1	Year	Year the crime was committed	Numeric
2	District Code	The district code of the police division that recorded the crime	Numeric
3	CaseID	The case identification number of the recorded crime	String
4	Gender	The gender of the offender	Char
5	Age	The age of the offender	Numeric
6	Date	The day the offences occurred.	DateTime
7	Time	The time offences occurred in hours, minutes and seconds	DateTime
8	Crime_City	The city where the crime was committed	String
9	Crime_LGA	The local government area where the crime was committed	String
10	Suspects_address	The address of the crime suspect	String
11	Suspects_LGA	The local government area of the suspect	String
12	Crime_Type	The type of crime committed	String

	year	District code	CaseID	Gender	Age	Date	Time	Crime_City	Crime_LGA	Suspect_Address	Suspect_LGA	Crime
0	2016	20	M/20	M	44	30/04/2016	600	Bitiah Bush	Boki	Bitiah Village, Irruam	Boki	Murder
1	2016	11	A/11	M	22	24/04/2016	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	Armed Robbery
2	2016	11	A/11	M	22	24/04/2016	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	Armed Robbery
3	2016	11	A/11	M	18	24/04/2016	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	Armed Robbery
4	2016	11	A/11	M	19	24/04/2016	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	Armed Robbery

Table 2 Dataset of Criminal Records extracted from the Nigeria Police Force

A depiction of the criminal records extract shows that the Crime class variable is categorical. There are nine crimes categorised under this variable, which include armed robbery, armed robbery and murder, armed robbery and rape, defilement, malicious damage, murder, rape, stealing, and stealing and malicious damage. Table 3 shows a pandas data frame showing these 11 attributes and class variables.

Table 3: Pandas data frame showing the attributes and class variables of the crime dataset.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4748 entries, 0 to 4747
Data columns (total 12 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   year             4748 non-null   int64
 1   District code    4748 non-null   object
 2   CaseID           4745 non-null   object
 3   Gender           4748 non-null   object
 4   Age              4748 non-null   object
 5   Date             4748 non-null   object
 6   Time             4748 non-null   object
 7   Crime_City       4748 non-null   object
 8   Crime_LGA        4747 non-null   object
 9   Suspect_Address  3950 non-null   object
 10  Suspect_LGA      4748 non-null   object
 11  Crime            4748 non-null   object
```

b. **Dataset preprocessing:** This cleans the original police crime dataset obtained from police crime records to ensure it is suitable for analysis. The collected data was cleaned and preprocessed to ensure it was suitable for analysis. Missing values were handled, features were normalised and standardised, categorical variables were encoded, and outliers were removed. To preprocess the data, the categorical class variable is label encoded to convert the categorical values to discrete numeric labels. These numeric labels are used to identify each crime type and are helpful for the classification of the crime records in the dataset.

In Table 4, the encoded categorical class variable is

	year	District code	CaseID	Gender	Age	Time	Crime_City	Crime_LGA	Suspect_Address	Suspect_LGA	Crime
0	2016	20	M/20	M	44	600	Bitiah Bush	Boki	Bitiah Village, Irruam	Boki	5
1	2016	11	A/11	M	22	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	0
2	2016	11	A/11	M	22	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	0
3	2016	11	A/11	M	18	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	0
4	2016	11	A/11	M	19	1430	No. 18 Orok Effiom Street	Calabar South	14 Mount Zion Lane	Calabar South	0

Table 4: The crime dataset showing the encoded crime class variable Similarly, the standard scaling method is applied to the crime records in the dataset to normalise the data and make it suitable for the model to understand the relationship in the data. The normalised data is shown in Table 5

	year	District code	CaseID	Gender	Age	Time	Crime_City	Crime_LGA	Suspect_Address	Suspect_LGA	Crime
0	2016	20.00	73	2	44.00	600.00	655	18	694	382	5
1	2016	11.00	21	2	22.00	1,430.00	1630	26	110	402	0
2	2016	11.00	21	2	22.00	1,430.00	1630	26	110	402	0
3	2016	11.00	21	2	18.00	1,430.00	1630	26	110	402	0
4	2016	11.00	21	2	19.00	1,430.00	1630	26	110	402	0
...	...	...	...	...	...	...	...	...	...	...	...
4743	2013	14.67	15	2	33.00	1,230.00	427	8	2311	323	7
4744	2013	14.67	3	2	22.00	2,230.00	783	23	2311	347	0
4745	2013	14.67	2	2	24.00	1,000.00	1762	25	2311	438	0
4746	2013	14.67	2	2	24.00	1,000.00	1762	25	2311	438	0
4747	2013	14.67	2	2	36.00	1,000.00	1762	25	2311	438	0

[4748 rows x 11 columns]

Table 5 The normalised crime data for the Prediction Model.

The variations in the ranges of variables are evident, underscoring the importance of data normalisation. Normalization is achieved following the equation as $\text{Xnormalized} = \frac{x-\min(x)}{\max(x)-\min(x)}$ Follows:

(1)

X_{max} is the maximum value, X_{min} is the minimum value, and X_{mon} is the normalised value.

The pre-processed class variable is encoded as follows, as shown in Table 6

TABLE 6: Crime Types and Encoded Values

Crime Type	Encoded Value
Armed Robbery	0
Armed Robbery and Murder	1
Armed Robbery and Rape	2
Defilement	3
Malicious Damage	4
Murder	5
Rape	6
Stealing	7
Stealing and Malicious Damage	8

Establishing the groundwork for data preprocessing in machine learning in Python involves importing essential libraries tailored for specific tasks. As outlined by Cruz et al. [35], The essential Python libraries foundational for this project include NumPy, Pandas, Matplotlib, Plotly, Folium, Seaborn, Scikit-learn (SKlearn), Statsmodels, and DESlib. The datasets, conveniently stored in .csv file format, are seamlessly imported into the Python environment utilising the read_csv() function within the code.

C. **Data Split:** The provided data undergoes a split into training and testing data using the train-test split method, which evaluates the performance of machine learning algorithms. This process divides the entire dataset into two parts. The first subset, the training dataset, fits the model. Meanwhile, the second subset, the test dataset, is not involved in model training. Instead, it serves as input to the model, enabling predictions to be made and compared against the expected values (Rácz et al. [36]). The preprocessed datasets were divided into 70% for training and 30% for testing. The training set was utilised to train the ensemble models, the validation set assisted in tuning hyperparameters, and the testing set was employed to evaluate the final model's performance.

Table 7: shows the data split data;

Data Set	Proportion (%)
Training Set	70
Test Set	30

The designed system ensures the accurate prediction of crime types based on the data recorded at the crime scene. The data is visualised to analyse the crime that may have occurred in a particular location of Cross River State across the 18 local government areas. Consequently, this research is targeted at assisting law enforcement agencies in predicting crimes to reduce the rate of crimes in the state.

Multi-Layer Prediction Model

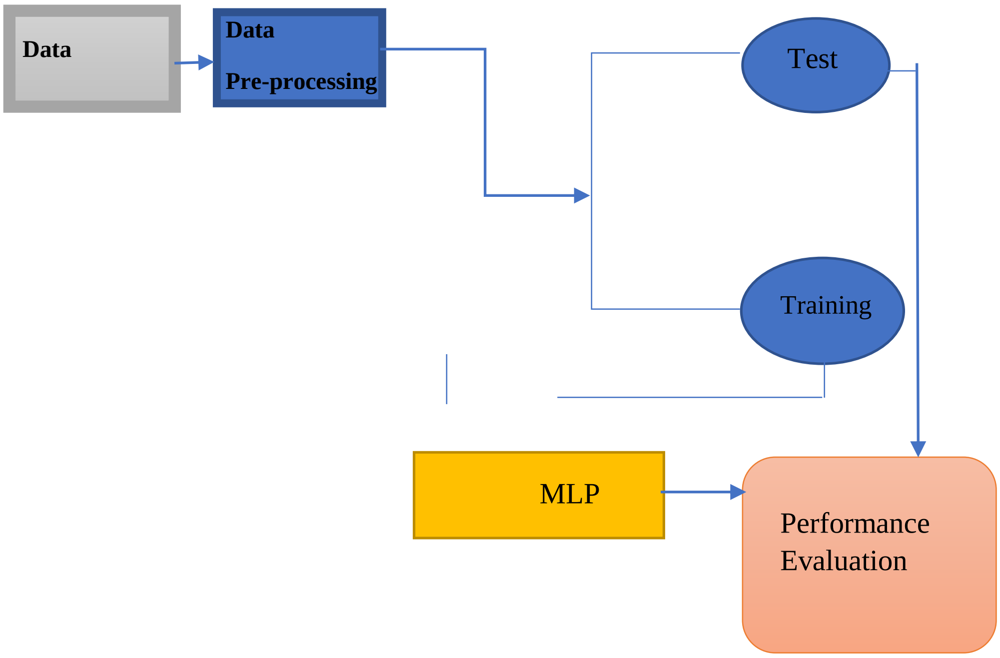

Figure 1 Architecture of the Prediction Model using Multi-Layer Perceptron

A Multi-Layer Perceptron (MLP) is a type of Artificial Neural Network (ANN) frequently employed in machine learning. It comprises multiple layers of interconnected nodes, each conducting a weighted sum of its inputs, followed by an activation function. The equations for the MLP algorithm can be described as follows;
Consider a standard MLP architecture with L layers. The mathematical equations for a forward pass in an MLP can be represented as: Algorithm 1: Multi-Layer Perceptron (MLP) Forward Propagation

Input:
- x: Input data
- L: Number of layers
- W[l]: Weight matrix at layer l
- b[l]: Bias vector at layer l
- activation_function (): Activation function

Output:
- a[L]: Output of the MLP

Procedure:
1. Initialize a[0] = x as the input data.
2. For l = 1 to L do:
   a. Compute the weighted sum of inputs at layer l:
      z[l] = W[l]a[l-1] + b[l]
   b. Apply the activation function to obtain the output of layer l:
      a[l] = activation_function(z[l])
3. Output the final activation output a[L].

Pseudocode:
Input: x, L, W, b, activation_function()

1. Set a[0] = x

2. For l = 1 to L do:
   a. Compute the weighted sum of inputs at layer l:
      z[l] = W[l] * a[l-1] + b[l]
   b. Apply the activation function to obtain the output of layer l:
      a[l] = activation_function(z[l])

3. Output a[L]

The activation function is non-linear and applied element-wise, such as the sigmoid function, ReLU, or tanh. The input from the datasets (crime record) represents vector x of size n, where
$x = [x_1, x_2, ..., xn]$. The input layer passes the input values to the next layer without computation. A single hidden layer with m nodes. For the jth neuron in the hidden layer, denoted as h_j, the calculation is described as follows:
$a_j = \Sigma_i(w_{ji} * x_i) + b_j$
Where:
w_{ji} represents the weight connecting the ith node to the jth hidden neuron.
x_i represents the input value from the previous layer.
b_j is the bias term associated with the jth hidden neuron.
a_j is the weighted sum of inputs and biases.
After computing a_j, an activation function f() sigmoid, we applied to introduce non-linearity, resulting in the output of the hidden neuron:
$h_j = f(a_j)$
This process is repeated for each neuron in the hidden layer. The output layer consists of a single node or multiple nodes. For the output neuron, denoted as y, the computation is similar to the hidden layer:
$z = \Sigma_i (w_{ji} * h_j) + b$
Where:
w_{ji} represents the weight connecting the ith node to the jth output neuron.
h_j represents the output value from the jth previous layer.
b_j is the bias term associated with the jth output neuron.
a_j is the weighted sum of inputs and biases.
Finally, an activation function g () was what we applied to obtain the output of the MLP:
y=g(z)
These equations describe the basic mathematical operations in an MLP. The choice of activation functions (f) and g ()) is sigmoid. The weights (w_{ji}) and biases (b) are typically learned through backpropagation, which adjusts them to minimise the error between the predicted and desired outputs during the model's training.
- The algorithm propagates forward in a multi-layer perceptron (MLP) neural network.
- It iterates Through each layer, computing the weighted sum of inputs (z[l]) and applying the activation function (activation_function()) to produce the output of each layer (a[l]).

The algorithm's final output is the activation output of the last layer (a[L]), representing the output of the MLP.
To efficiently predict crime using the Nigeria police crime record dataset, about the functions above, the following algorithmic steps were taken:

Algorithm: Multi-Layer Perceptron (MLP) for Crime Problem Solving

Input:
- Dataset: Raw data containing relevant information about the crime
- Training dataset: Subset of the dataset used for training the MLP model
- Test dataset: Subset of the dataset used for evaluating the performance of the trained model

Output:
- Predicted result: Predictions made by the MLP model based on the input data
- Evaluation metrics: Performance metrics used to assess the accuracy of the model's predictions

Procedure:
1. Import the Relevant Libraries:
   - Import necessary Python libraries such as NumPy, Pandas, Scikit-learn, etc.

2. Import Dataset:
   - Load the dataset containing information about the crime.

3. Read the Dataset:
   - Read the dataset into a Pandas data frame for further processing.

4. Pre-process the Dataset:
   - Clean the dataset by handling missing values, encoding categorical variables, and performing feature scaling or normalisation if necessary.

5. Split the Dataset:
- Split the pre-processed dataset into training and test sets to train and evaluate the MLP model.

6. Create the MLP Model Object:
- Define the architecture of the MLP model, including the number of layers, number of neurons in each layer, activation functions, and optimisation algorithms.

7. Fit the Training Dataset into the MLP Model:
- Train the MLP model using the training dataset by fitting the data to the model.

8. Predict the Result:
- Use the trained MLP model to make predictions on the test dataset.

9. Evaluate the MLP Model Result:
- Assess the performance of the trained MLP model using evaluation metrics such as accuracy, precision, recall, and F1-score.

10. Reporting:
- Generate a report summarising the results of the MLP model and any insights gained from the analysis

(i) Accuracy, a pivotal metric in assessing predictive models, quantifies the ratio of correct predictions to the total number of predictions made. Mathematically, it is computed by dividing the number of correct predictions of crime types by the total dataset size. Comparative analysis of accuracy entails evaluating the performance of four classification algorithms.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \qquad (2)$$

(ii) Precision measures the proportion of correctly predicted instances of a positive class out of all the cases predicted as positive. A high precision value indicates consistent or repeatable measurement results where obtained readings align closely. Conversely, low precision suggests variability in measurement values. Mathematically, precision is calculated as the ratio of true positives to the number of optimistic predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (3)$$

(iii) Recall, also known as the True Positive Rate (TPR), represents the proportion of true positives correctly identified from all actual positive instances in the ground truth. In detecting heart disease, recall signifies the test's effectiveness in correctly labelling individuals with the condition as positive. A high recall value indicates a test with minimal false negative results, thereby reducing the likelihood of missing cases of heart disease.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (3)$$

(iv) F1-score: The F1-score metric combines precision and recall. The F1 score is the harmonic mean of the two. A high F1 Score indicates perfect precision and recall of the proposed model.

$$\text{F1} - \text{score} = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall+Precision}} \qquad (4)$$

Where TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative

In the crime case and pattern analysis function, data is captured from a panda frame and entered into the system for analysis. Crime details captured include date, division, police station, crime category, status, etc. When these data are fed into the system, all of these data will be analysed using several functions that model the relationships between the attributes in the data. The pattern of crime and distribution is then graphically shown, including the location, to allow the law enforcement agents to understand the crime distribution based on location information.

Furthermore, the model classifies the crimes recorded at the time of data capture based on the number of occurrences, the date and time of the crime, and the location information. Further analysis is done on the raw inputs to remove noise from the data, such as missing and inconsistent values, to enhance prediction accuracy on the crime data. At the same time, the data can be incremented based on recurring crime instances with similar results when fed into the model. This means that the model can generate enough knowledge using historical data so that new cases of crime data are correspondingly learned and classified into different crime categories using available knowledge generated by the model during the training phase. In this manner, any crime type can be predicted based on the data recorded to represent the crime committed.

Results and Discussion

1) Exploratory data analysis

Implementing the crime prediction model produced outstanding results based on the dataset used. The distribution of the crime patterns per year from 2010 to 2022 is given in Figure 2.

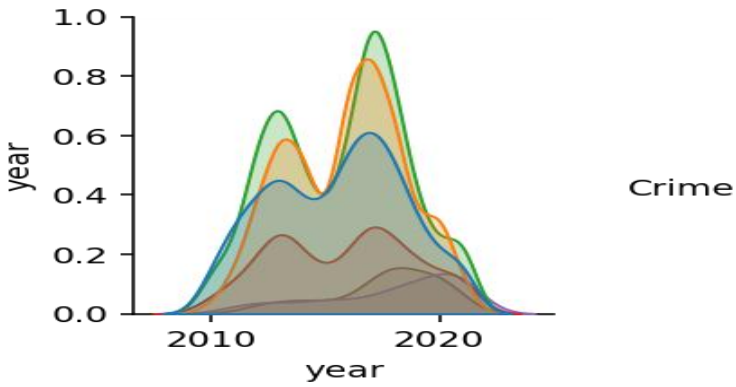

Figure 2 Pattern of crime rate from 2010 to 2022

Figure 2 shows a significant increase in the crime rate from 2010 to 2022. The crime patterns depicted show that the lack of crime prediction systems may have impacted the number of criminal cases recorded over the last decade. This trend in rising crime rate is also illustrated in Figure 3.

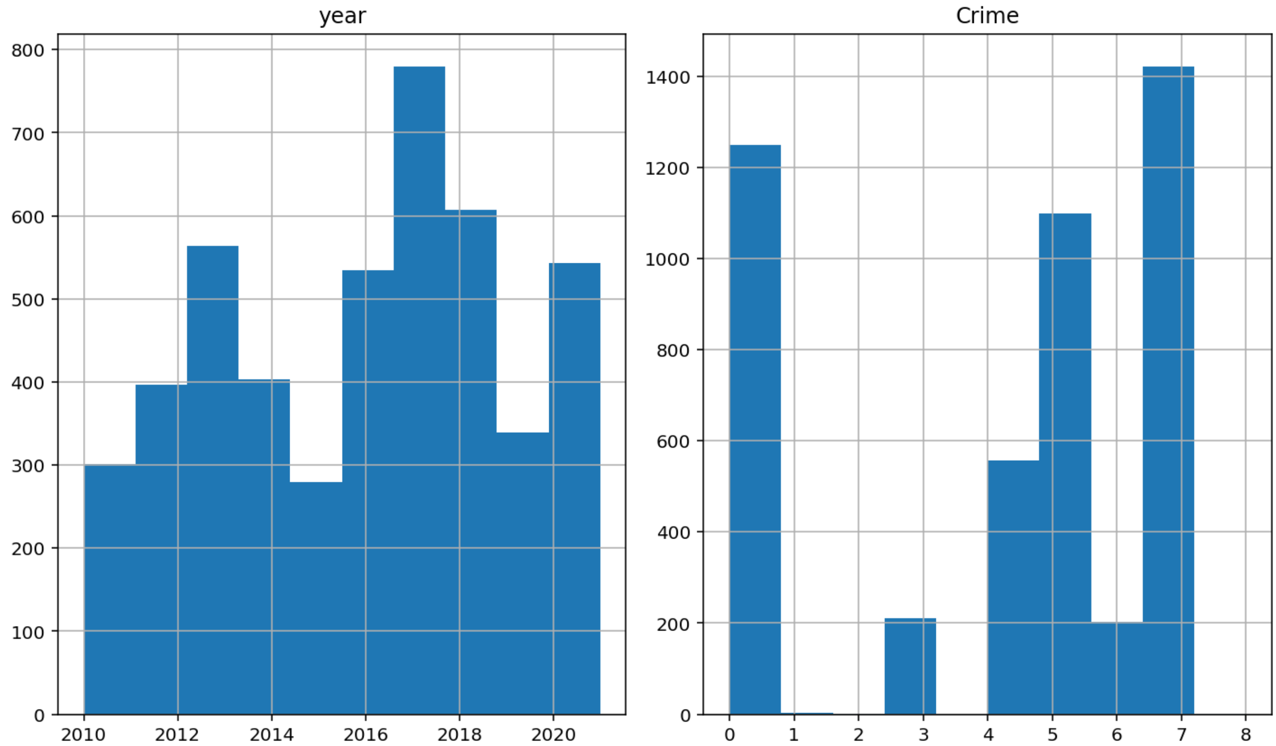

Figure 3 (a and b): Distribution of crime rate per year and per crime type

In Figure 3a, the bar plot shows that since 2010, there has not been a significant decrease in the crime rate, with the lowest recorded crime rate in 2015. 2017 recorded the highest crime rate, and succeeding years, 2018 to 2022, recorded lower crime rates, with 2019 recording the weakest over these four years. Similarly, using Figure 3b, it is observed that stealing was the highest crime during the last decade, with about 1422 reported cases. This is followed by 1249 reported cases of armed robbery and murder, with about 1099 reported cases. There were about 210 reported cases of heresy and 202 reported rape cases. In contrast, crimes such as armed robbery and rape, malicious damage, as well as stealing and malicious damage were the lowest reported crimes over the last decade. Furthermore, there are rising cases of defilement and murder reported in recent times, which strengthens the relevance of a predictive system to allow for quick interventions by the Nigeria Police Force.

Analysing crimes by location, Table 8 shows that armed robbery was prevalent from 2017 to 2018 in Abi and Akamkpa Local Government Areas, respectively. Stealing was primarily reported in Yala, Obudu, and Odukpani local government areas between 2015 and 2017 while stealing and malicious damage was expected in Calabar Municipality and Calabar South local government areas in 2017.

Table 8. Crime distribution by location

Crime	Crime_LGA	year	District code	CaseID	Gender	Age	Date	Time	Crime_City	Suspect_Address	Suspect_LGA
Armed Robbery	Abi	2018	19`	A/19`	M	21	14/01/2018.	2129	Itigidi/Ugep Highway	Itigidi/Ugep High Way	Abi
	Abi	2018	19	A/19	M	25	8/10/18 0:00	2040	Ugep/Ediba Highway	Anong	Abi
	Akamkpa	2017	13	A/13	M	25	11/10/17 0:00	100	Awi	Awi	Akamkpa
	Akamkpa	2017	13	A/13	M	27	25/03/2017	545	Ayakbam Village	Oborito Village	Akamkpa
	Akpabuyo	2017	24	A/24	M	20	16/11/2016	200	idebe Road	Ikot Nakanda	Akpabuyo
...	...	...	...	...	...	...	...	...	...	...	...
Stealing	Yala	2015	S/25	25	M	68	2/8/15 0:00	1430	Ugaga Village	Ugaga Village	Yala
	obudu	2017	26	S/26	M	22	31/05/2017	1200	Utukwane	Utukwane	obudu
	odukpani	2016	22	S/22	M	25	10/12/16 0:00	930	Itu Bridge	Pamol 22/20 Camp	Calabar Municipality
Stealing and Malicious Damage	Calabar Municipality	2017	12	S/12	M	55	19/05/2017	1600	No. 101 MCC Road	3 Johnson Ishie Street	Calabar Municipality
	Calabar South	2017	11	S/11	M	21	19/10/2017	230	Ekpenyong Ekpe Street	No. 15 Ekpenyong Ekpe Street	Calabar South

From Table 8, the age range of suspects is 21 to 68 years, all males. The distribution of crime types is given in Figure 5

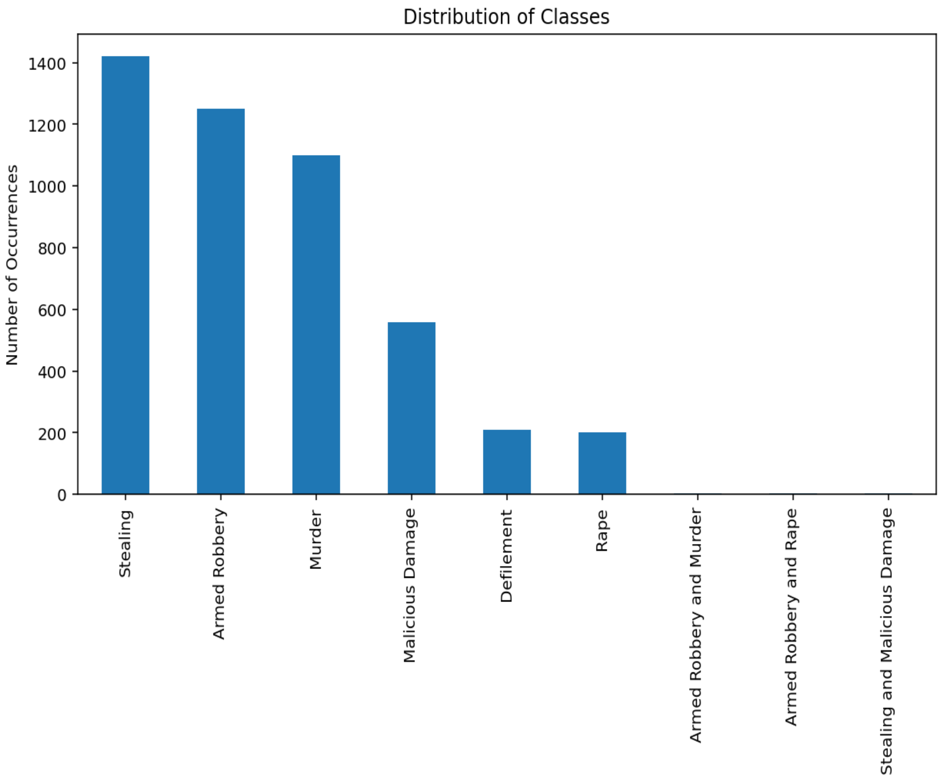

Figure 5 Distribution of classes of crime types committed between 2010 and 2022.

The class distribution of crime types shows that stealing and armed robbery. Murder was the most prevalent crime committed over the last decade, while there were very few incidents of armed robbery and murder, armed robbery and rape, and stealing and malicious damage.

Performance Evaluation Metrics

Table 9 provides the performance metrics used to evaluate the model (Rácz *et al*., 2019).

Table 9: Evaluated result of MLP

Matrix	**Precision**	**Accuracy**	**Recall**	**f1 score**
Results	**0.86**	**0.75**	**0.73**	**0.79**

Discussion: In conclusion, the MLP model showcased commendable performance metrics, including an accuracy of approximately 75%, a precision of 86%, a recall of 73%, and an F1 score of 79%. Despite a moderate loss value of 0.7193, indicating a degree of prediction error, the model's detailed evaluation through the confusion matrix and classification report offered valuable insights into its performance across different classes. Notably, specific courses, such as class 1 and class 9, exhibited challenges with lower precision and recall.

The model architecture, comprising four layers with varying neuron dimensions and activation functions, demonstrates a thoughtful design strategy tailored to the dataset's categorical output. Leveraging adaptive moment estimation (Adam) as the optimiser and categorical cross-entropy as the loss function, the model underwent rigorous evaluation against the crime dataset. Choosing a training batch size of 64 samples and 100 epochs ensured robust training and convergence. Additionally, a prudent train-test split of 70:30, with 15% reserved for testing, further validated the model's generalisation capabilities.

Overall, the MLP model's comprehensive architecture and meticulous evaluation underscore its efficacy in predicting crime classes within the dataset, making it a promising tool for crime prediction and prevention.

```
45/45 [==============================] - 0s 1ms/step - loss: 0.7701 - acc: 0.7439
Loss of the model is -  0.7701186537742615
45/45 [==============================] - 0s 1ms/step - loss: 0.7701 - acc: 0.7439
Accuracy of the model is -  74.38596487045288 %
```

Table 10 MLP crime prediction model performance in terms of accuracy and loss.

With an accuracy of 74.4% on the validation set, the model's performance is significantly high, given the novelty of the dataset used in the study. Also, a loss of 0.77 shows that the model can understand the relationships in the dataset attributes to make predictions on the validation set. The predictions made on each class in terms of the crime types are given in the confusion matrix of Table 11

Table 11: The classification report of the crime prediction model for MLP

In the classification report of Table 11, the prediction of each crime type demonstrated exemplary performance. This performance is summarised using Table 12

	precision	recall	f1-score	support
0	0.86	0.73	0.79	397
1	1.00	1.00	1.00	1
2	0.00	0.00	0.00	1
3	0.97	0.71	0.82	51
4	0.93	0.55	0.70	177
5	0.72	0.73	0.73	305
6	0.81	0.61	0.70	64
7	0.78	0.75	0.76	429
8	0.00	0.00	0.00	0
9	0.00	0.00	0.00	0
micro avg	0.81	0.71	0.75	1425
macro avg	0.61	0.51	0.55	1425
weighted avg	0.82	0.71	0.75	1425
samples avg	0.71	0.71	0.71	1425

TABLE 12: Crime Model's Predictions for MLP

Crime Label	Crime Type	Precision (%)	Recall (%)	F1-score (%)
0	Armed Robbery	83	73	79
1	Armed Robbery and Murder	100	100	100
2	Armed Robbery and Rape	0	0	0
3	Defilement	97	71	82
4	Malicious Damage	93	55	70
5	Murder	72	73	3
6	Rape	81	61	70
7	Stealing	78	75	76
8	Stealing and Malicious Damage	0	0	0

Using Table 12, the performance of the model based on precision was 100% for armed robbery and murder, 97% for defilement, 93% for malicious damage, 83% for armed robbery, 81% for rape, 78% for stealing, and 72% for murder. Furthermore, the model's performance in terms of recall shows 100% for armed robbery and murder, 75% for stealing, 73% for armed robbery, 73% for murder, 71% for defilement, 61% for rape, and 55% for malicious damage. For f1-score, the model produced 100% for armed robbery and murder, 82% for defilement, 79% for armed robbery, 76% for stealing, 73% for murder, 70% for malicious damage, and 70% for rape. However, armed robbery and rape, as well as stealing and malicious damage, performed poorly in the model's predictions due to the few samples available in the dataset, which were used in the training and evaluation of the model.

CONCLUSION

This research aimed to forecast nine distinct crime types occurring across various regions within Cross River State. A comprehensive crime dataset was curated from authentic criminal records from diverse districts and divisions of the Nigeria Police Force within the state, ensuring the utmost accuracy and reliability.

To accurately predict crimes, the availability of historical data about existing crime records is indispensable, serving as a foundational basis for anticipating future occurrences over a projected timeframe. A dataset comprising four thousand seven hundred and forty-eight (4748) entries was meticulously compiled from the Cross River State Police Command. This dataset was meticulously partitioned, allocating 70% for the training phase and reserving 30% for rigorous testing.

Numerous evaluation metrics, including specificity, sensitivity, precision, false positive rate (FPR), false negative rate (FNR), F1 score, and accuracy, were rigorously applied throughout this study. Following exhaustive experimentation, the Multi-Layer Perceptron (MLP) algorithm exhibited commendable performance, achieving a notable accuracy rate of 74%. This achievement marks a significant milestone in crime prediction within Cross River State and Nigeria.

Our predictive system can identify regions with a heightened likelihood of crime, facilitating the visualisation of crime-prone areas. Analysis of crime patterns has unveiled the prevalence of offences such as theft, armed robbery, homicide, and malicious damage. This invaluable insight equips the Nigeria Police Force (NPF) with actionable intelligence to implement proactive measures aimed at mitigating the pervasive crime menace, particularly within Local Governments, including Abi, Akamkpa, Yala, Odukpani, Obudu, Calabar South, and Calabar Municipality.

Looking ahead, our research aims to further augment crime prediction accuracy and overall performance by exploring additional classification models. Future extensions will encompass the incorporation of income data for neighbourhoods to discern potential correlations between income levels and crime rates and ascertain underlying causes of crime occurrences within urban settings. Additionally, including longitude and latitude coordinates as features will enhance spatial analysis capabilities. Moreover, the study intends to explore crime datasets from emerging towns alongside their corresponding demographic profiles, fostering a comprehensive understanding of crime dynamics across diverse local Government Areas.(NO 4 correction) We recommend the creation and implementation of two well-equipped, well-trained crime units across States and Local Government Areas in the nation. Selected members of these units should undergo training on a quarterly and yearly basis to help mitigate and decrease the alarming crime rate.

REFERENCES

Cesario, E., Catlett, C. and Talia, D., 2016. "Forecasting crimes using autoregressive models." In: Proceedings of the IEEE 14th International Conference on Dependable, Autonomic and Secure Computing, pp. 795–802, IEEE, Sydney, Australia.

Tastle, W. J., 2013. "Introduction to artificial networks and law enforcement analytics." In: Intelligent data mining in law enforcement analytics

Borrion, H., Ekblom, P. and Alrajeh, D., 2020. “The problem with crime problem-solving: Towards a second generation pop?” The British Journal of Criminology, 60(1), 219–240.

Kang HW, Kang HB, 2017. Prediction of crime occurrence from multi-modal data using deep learning. PLOS ONE 12(4): e0176244. https://doi.org/10.1371/journal.pone.0176244

Varun Mandalapu, Lavanya Elluri , Piyush Vyas and Nirmalya Ro, 2023. “Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions.” IEEE Access, vol. 11, 2023, pp. 60153–70. Crossref, https://doi.org/10.1109/access.2023.3286344

Lin Y-L, Yen M-F, Yu L-C., 2018. Grid-Based Crime Prediction Using Geographical Features. ISPRS International Journal of Geo-Information.; 7(8):298. https://doi.org/10.3390/ijgi7080298

Walczak S., 2021. Predicting Crime and Other Uses of Neural Networks in Police Decision Making. Frontiers in Psychology, 12, 587943. https://doi.org/10.3389/fpsyg.2021.587943

Bappee, F. K., Soares Júnior, A., Matwin, S., 2018. Predicting Crime Using Spatial Features. Lecture Notes in Computer Science, 367–373. doi:10.1007/978-3-319-89656-4_42

Shah, N., Bhagat, N., and Shah, M., 2021. Crime forecasting: a machine learning and Computer vision approach to crime prediction and prevention. Visual computing for industry, biomedicine, and art, 4(1), 9. https://doi.org/10.1186/s42492-021-00075-z.

Ramirez-Alcocer, U.M., and Tello-Leal, E., 2019. Predicting Incidents of Crime through LSTM Neural Networks in Smart City Domain.

Cortes, Ana Laura Lira and Silva, Carlos Fuentes, 2021. Artificial Intelligence Models for Crime Prediction in Urban Spaces Machine Learning and Applications: An International Journal (MLAIJ) Vol.8, No.1, March 2021, Available at SSRN: https://ssrn.com/abstract=3822346

Chatzimichail, E., Paraskakis, E. and Rigas, A., 2013. “Predicting asthma outcome using partial least square regression and artificial neural networks.” Hindawi Publishing Corporation Advances in Artificial Intelligence, Article ID 435321, 7 page.

Karthik, K., BinalJakhia, R., DattatrayaLele, M. and Abhay, C., 2014. Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images”. Hindawi Publishing Corporation Advances in Bioinformatics, Article ID 708279, 14 pages

Zavvar M, Razaei M, Garavand, 2016. S. E-Mail spam detection uses a combination of particle swarm optimisation, an artificial neural network, and a support vector machine. International Journal of Modern Education and Computer Science. 7: 68-74

Pilita A. Amahan, Mia V. Villarica, and Albert A. Vinluan. 2021. Technical Analysis of Twitter Data in Preparation of Prediction using Multilayer Perceptron Algorithm. In 2021 4th International Conference on Data Science and Information Technology (DSIT 2021). Association for Computing Machinery, New York, NY, USA, 109113. https://doi.org/10.1145/3478905.3478928

Rozos E, Dimitriadis P, Mazi K, Koussis AD. A Multilayer Perceptron Model for Stochastic Synthesis. Hydrology. 2021; 8(2):67. https://doi.org/10.3390/hydrology8020067

Ke KC, Huang MS. Quality Prediction for Injection Molding by Using a Multilayer Perceptron Neural Network. Polymers (Basel). 2020 Aug 12;12(8):1812. doi: 10.3390/polym12081812. PMID: 32806786; PMCID: PMC7464357.

Tamim, N., Elshrkawey, M., Abdel Azim, G., and Nassar, H., 2020. Retinal Blood Vessel Segmentation Using Hybrid Features and Multi-Layer Perceptron Neural Networks. Symmetry, 12(6), 894. MDPI AG. Retrieved from http://dx.doi.org/10.3390/sym12060894.

Bikku, T. Multi-layered deep learning perceptron approach for health risk prediction. J Big Data **7**, 50, 2020. https://doi.org/10.1186/s40537-020-00316-7

Huang, F., Cao, Z., Jiang, SH. et al., 2020. Landslide susceptibility prediction based on a Semi-supervised multiple-layer perceptron model. Landslides **17**, 2919–2930 https://doi.org/10.1007/s10346-020-01473-9

Savalia, S., and Emamian, V., 2018. Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks. Bioengineering (Basel, Switzerland), 5(2), 35. https://doi.org/10.3390/bioengineering5020035

Mogan, Jashila Nair, Chin Poo Lee, Kian Ming Lim, and Kalaiarasi Sonai Muthu., 2022."VGG16-MLP: Gait Recognition with Fine-Tuned VGG-16 and Multilayer Perceptron" Applied Sciences 12, no. 15: 7639. https://doi.org/10.3390/app12157639

Car, Z., Baressi Šegota, S., Anđelić, N., Lorencin, I., and Mrzljak, V., 2020. Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron. Computational and mathematical methods in medicine, 2020, 5714714 https://doi.org/10.1155/2020/5714714

Li, D., Huang, F., Yan, L., Cao, Z., Chen, J., and Ye, Z., 2019. Landslide Susceptibility Prediction Using Particle-Swarm-Optimized Multilayer Perceptron: Comparisons with Multilayer-Perceptron-Only, BP Neural Network, and Information Value Models. Applied Sciences, 9(18), 3664. MDPI AG. Retrieved from http://dx.doi.org/10.3390/app9183664

Yun, J., Park, J. E., Lee, H., Ham, S., Kim, N., and Kim, H. S., 2019. Radiomic features Multilayer perceptron network classifier: a robust MRI classification strategy for distinguishing glioblastoma from primary central nervous system lymphoma. Scientific reports, 9(1), 5746. https://doi.org/10.1038/s41598-019-42276-w

F. M. M. Mokbal, W. Dan, A. Imran, L. Jiuchuan, F. Akhtar and W. Xiaoxi, 2019."MLPXSS: An Integrated XSS-Based Attack Detection Scheme in Web Applications Using Multilayer Perceptron Technique," in IEEE Access, vol. 7, pp. 100567-100580 doi: 10.1109/ACCESS.2019.2927417.

N. B. Gaikwad, V. Tiwari, A. Keskar and N. C. Shivaprakash, 2019. "Efficient FPGA Implementation of Multilayer Perceptron for Real-Time Human Activity Classification," in IEEE Access, vol. 7, pp. 26696-26706, doi: 10.1109/ACCESS.2019.2900084.

Lorencin, I., Anđelić, N., Mrzljak, V., and Car, Z., 2019. Genetic Algorithm Approach to Design of Multi-Layer Perceptron for Combined Cycle Power Plant Electrical Power Output Estimation. Energies, 12(22), 4352. MDPI AG. Retrieved from http://dx.doi.org/10.3390/en12224352

Pham, B.T., Tien Bui, D., Pourghasemi, H.R. et al., 2017. Landslide susceptibility Assessment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. Theor Appl Climatol **128**, 255–273 https://doi.org/10.1007/s00704-015-1702-9

Wang, SH., Zhang, Y., Li, YJ. et al., 2018. Single slice-based detection for Alzheimer's Disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. Multimed Tools Appl **77**, 10393–10417. https://doi.org/10.1007/s11042-016-4222-4

Meha Desai, Manan Shah, 2021. An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN), Clinical eHealth, Volume 4, Pages 1-11, ISSN 2588-9141, https://doi.org/10.1016/j.ceh.2020.11.002.

Gideon Magesa, Emmanuel E. Sinkwembe, Nyimvua Shaban, Triphonia Ngailo. 2023. Bivariate Discrete Time Series Model for Assessing the Effects of Rainfall and Temperature on Road Accidents: The case of Morogoro and Pwani Regions in Tanzania, Scientific African, Volume 19, e01522, ISSN 2468-2276, https://doi.org/10.1016/j.sciaf.2022.e01522.https://www.sciencedirect.com/science/article/pii/S2468227622004264

Nina Molovčáková, Veronika Adamová, Viktor Šoltés. 2021. Analysis of Traffic Accidents in Selected Period in the Žilina Region and Proposal of Security Measures, Transportation Research Procedia, Volume 55, 2021, Pages 1585-1592, SSN 2352-1465, https://doi.org/10.1016/j.trpro.2021.07.147. https://www.sciencedirect.com/science/article/pii/S2352146521005652

C. L. Udeze, I. E. Eteng, A. E. Ibor. 2022. Application of Machine Learning and Resampling Techniques to Credit Card Fraud Detection. 2022. Journal of the Nigerian Society of Physical Sciences, 4(3), 769. https://doi.org/10.46481/jnsps.2022.769

Cruz, R. M. O., Hafemann, L. G., Sabourin, R. and Cavalcanti, G. D. C., 2020. DESlib: A dynamic ensemble selection library in Python. Journal of Machine Learning Research, 21.

Rácz, A., Bajusz, D. and Héberger, K., 2021. Effect of dataset size and train/test split ratios in qsar/qspr multiclass classification. Molecules, 26(4), https://doi.org/10.3390/molecules26041111