Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

Crime Classification Using GRU and 1CNN Techniques

[1]S. Jeya Selvakumari, [2]V. Joseph Peter

[1] Research Scholar, Research Department of Computer Science, Kamaraj College, Thoothukudi, Affiliated to Manonmanian Sundaranar University, Tirunelveli, India.
[2] Associate Professor, Research Department of Computer Science, Kamaraj College, Thoothukudi,
Affiliated to Manonmanian Sundaranar University, Tirunelveli, India.

Email: [1]jeyrobin@gmail.com, [2]peterchristy_in@yahoo.com

Abstract: Criminal activity is a widespread social problem that affects a nation's standard of living, rate of economic development, and international standing. Numerous crimes have proliferated worldwide in recent years. It has a variety of effects on people of all types. Classification of crime is necessary to create a safe society. Many researchers classify the crimes using machine learning techniques with a common dataset, which presents numerous computational opportunities and challenges. Using the learning capabilities of several deep learning architectures, we are trying to classify the criminal activity from the real-time dataset of particular city. In the proposed work, Gated Recurrent Unit (GRU) and Convolutional Neural Network (1CNN) deep learning architecture are used for crime classification. We have conducted performance comparison for classifying the crime with real-time dataset. The effectiveness of the aforementioned deep learning models was assessed by Training and Validation Loss curve. It was discovered that 1CNN models achieve better result compared to GRU. 1CNN produced 97% accuracy.

Keywords Crime, Deep Learning, GRU, 1CNN, Classification

1. Introduction

Crime is defined by the U.S. Department of Justice as all actions and behaviours for which society has established legal sanctions. Which actions are unlawful and which are not are outlined in written federal and state laws[26]. Murder, robbery, and burglary are examples of actions that have traditionally been regarded as crimes. Other behaviours, such domestic abuse or driving while intoxicated or high, have just recently been added to the list of crimes. Crime has also been impacted by societal developments in other ways. For instance, the widespread use of computers opens us new possibilities for white-collar crime and gives us a new term—"cybercrime"—to add to our lexicon. It has become more challenging to police and monitor neighbourhoods with a high likelihood of crime as a result of the recent increase in urban population. There is a rise in crime and insecurity in these regions as a result of this lack of control[27]. The rise of smart city infrastructure presents a chance to develop original solutions to these issues. It is necessary to develop the system to classify the crime for the particular city with real time data of corresponding city. Researchers now have an exceptional opportunity to investigate and conduct research on crime detection using machine learning and deep learning approaches with growing availability of crime data and the progress of existing technologies.

In order to classify the crime and forecast future crime patterns, machine learning and deep learning algorithms have been used in the field of crime classification. Neural Network[2][3][9][10][14] ,Naïve Bayes ,K-Nearest Neighbor[21],Decision tree[19][20], and Long Short Term Memory[24] for instance, have been trained on crime data from certain cities to accurately anticipate crime patterns. Many of the above methods produced good results till need the improvement in performance and technology.

The objective of the proposed system is to classify the crime based on the incident happened in the Thoothukudi city. Many crime datasets are available for predicting crime of various country. They have not provided the details for our city culture. To find the various crime have occurred depends on our culture, real time dataset is collected. The data is pre-processed and converted into vector. Deep Learning techniques GRU and

1CNN are used to classify the crime. And results are measured by Accuracy, Precision, Sensitivity, Specificity, F1-Score and Matthew's Correlation Coefficient (MCC). Analysed the results by Training and Validation Loss and bar graph is plotted. 1CNN is produced better results for real time dataset compared to GRU. Finally our 1CNN result is compared with other machine learning Techniques. 1CNN yields better accuracy.

This paper is organized as follows. In Section 2, the related works of the proposed model are presented. In Section 3, the system model has been explained in detailed. Experiments on GRU and 1CNN techniques are performed and discussion of the results are given in Section 4. Finally, a general summary and future research plan are introduced in Section 5.

2. Related Works

Modern applications for surveillance, identification, and criminal prevention have all benefited from improved machine learning capabilities. The majority of crime-based analyses [1] are currently carried out utilising one or more machine learning techniques. Authors in [2] have suggested real-time crime detection using machine learning and deep learning for crime prevention. It showed that VGG CNN models perform the best, especially in a scenario with minimal data, delivering a classification accuracy of 99.5% when detecting criminal faces[3]. According to the rational choice theory [4], a potential offender weighs the benefits of executing the crime successfully against the risk of being caught before deciding whether to actually carry it out or not. The author[5] looked into the potential of DL techniques to predict hotspots in urban environments, which are places where specific sorts of crimes are more likely to happen within a given time range. The virtual leave-one-out method's application [6] to choosing the best neural network architecture for time series prediction. The difficult work of recidivism prediction was handled by a three-layered artificial neural network. In this study[7], illustrated the improvement in Singular Race Models' predictive accuracy across a range of crime categories and examine the primary and secondary reasons of bias. We do this with the aid of several useful measures. An innovative new approach for identifying blood, knives, and firearms at a crime scene was presented[8].

The experimental findings showed[9] the suggested method's efficiency in classifying crime scene investigation (CSI) images as well as its adaptability to other image data with various content types. According to studies, CNNs trained with transfer learning can significantly enhance accuracy across a range of applications [10]. The initial layers of any CNN that contain low-level edge and texture characteristics and are appropriate for the majority of common machine vision tasks may typically be directly transferred between various machine vision tasks. The recommended method predicts various types of crimes by looking at text-based crime summaries. Classification techniques gave more than 90% accuracy using Bayes theorem [11]. The term frequency-inverse document frequency vectors and fuzzy c-means method are used to pre-process the data and execute the initial labelling step[12]. GloVe word embeddings were used to extract features. The system presented Real-Time Crime Detection Technique Using a Deep Learning Algorithm[13], which analyses real-time videos and alerts the nearest Cybercrime admin about the incidence of crime with current position. In the studies, we discovered that combining the CNN with the LSTM model may deliver a reliable crime prediction approach with a high forecast accuracy[14].

The system presented a graph-based crime research framework that focuses on establishing relationships between the entities found in a sizable corpus of crime data from Indian states and union territories[15]. To deal with criminal situations, big data-based data analytical approaches can be applied[16]. Additionally, many data collection techniques are used for better performance, including Web 2.0, Geographic Information System (GIS), and Volunteered Geographic Information (VGI). Evaluated the efficiency of several deep learning architectures' parameters and provided guidance on how to configure them for better performance in crime classification and, ultimately, crime prediction[17]. Illustrated a random sample of ten different crime types as well as the statistical anomalies the algorithm produced for a particular geo-spatial cell[18]. For all geospatial cells in the prediction zone, this will be calculated once more. This might lead to the type of crime and its time and space boundaries. With an accuracy of 55.03%, the Random Forest Classifier exceeded Naive Bayes and Decision Tree in the analysis of the algorithms' ability in predicting the kind of crime[19]. By examining and analysing the data on repeated events already in existence, it was possible to predict the probable crime category in a specific geographic location. On a dataset that was taken from the CLEAR database maintained by the Chicago Police Department, Decision Tree and Naive Bayes are implemented[20]. To anticipate the distribution of crime in Los Angeles at an hourly scale in parcels the size of neighbourhoods[21] adapted a spatial temporal residual network to well-represented data.

3. System Model

The crime classification based on deep learning is discussed in the Fig. 1. This work uses real dataset collected from DCRB of Thoothukudi. The dataset includes SDO(division),total crime occurred in each division, Head(crime name),determined/not, cr.no &sec. of law, complainant name, gist of the case and reason of UI. The dataset contains seven categories of crime namely murder, HBN(House Burglar by Night), rape, robbery, dowry, POSCO and Suspicious death. These six types of crime are broadly classified into 2 major categories namely property crime (HBN,robbery) and violent crimes against women and child(Murder, dowry, rape, POSCO, Suspicious death). This dataset is used for input of the system model. Figure 1 represents the architecture of the proposed method.

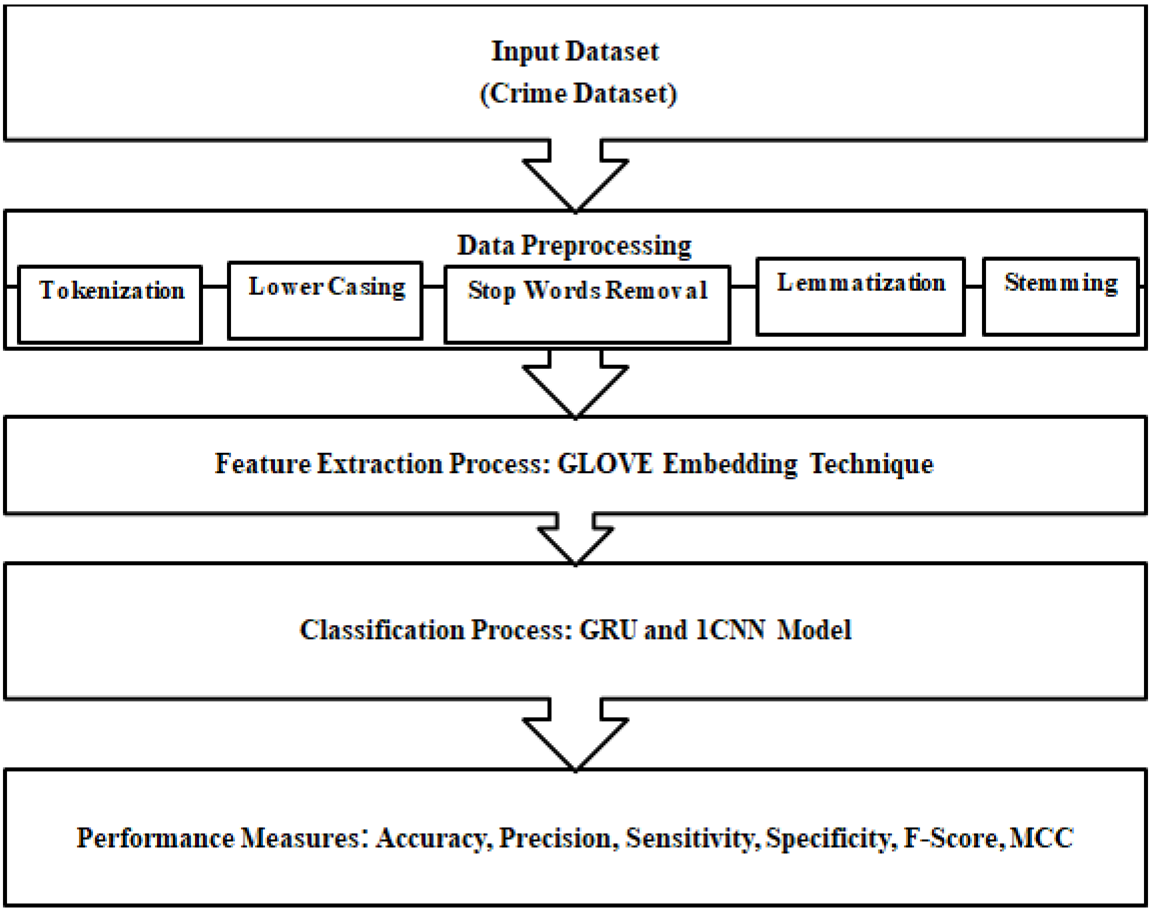


Fig 1: Architecture of the proposed method

3.1 Preprocessing

This suggested methodology has made use of real-time data gathered from DCRB in Thoothukudi. The Crime Number, Section of Law, Date of Occurrence, Location, Sub Division, Reason, and Gist of the Case are just a few of the many fields in this data collection. The crime case is summarised in the case's gist. Preprocessing the raw data is necessary to efficiently prepare for the proposed work. The gist of the crime is preprocessed using the techniques below. Tokenization, lowercase conversion, eliminating stop words, digits, punctuation, and superfluous spaces are some of their techniques. Lemmatization and stemming were then used for improved performance.

When a task is correctly completed utilising a vocabulary and linguistic analysis of words with the aim of deleting only inflectional endings and returning the lemma, or dictionary form, of a word, the term "lemmatization" is frequently used. Stemming is a phrase that frequently refers to a fundamental computational technique that aims to achieve this goal most of the time by removing derivational affixes from words. A word's inflectional forms and occasionally associated derivational forms are intended to be reduced to a fundamental form through stemming and lemmatization. Lemmatization entails stripping a word down to its simplest form. Unlike stemming, lemmatization analyses the context and converts the word to its appropriate base form. Stemming only gets rid of the last few characters, which commonly leads to incorrect spellings and interpretations.

3.2 Feature Extraction

There are various procedures to follow for feature extraction after preprocessing in order to get better outcomes. Feature extraction is a process that converts raw data into manageable numerical features while preserving the original data set's information. In this work, a counter vectorizer for numerical conversion is developed, data is fitted, text is converted to features, the features are shaped (128 by 1572), a counter vectorizer is selected, and the features are saved for the model's training. GloVe Embeddings are also utilised for improved performance.

3.2.1 GloVe Embeddings

It is employed to locate the unusual words. It is a publicly accessible pre-trained word embeddings approach [23]. The "glove.6B.300d.txt" package has an 822 MB file size. The sizes of the embedding vectors range from 50, 100, 200, and 300 dimensions. A co-occurrence matrix for the supplied term is generated using Glove [24]. Eq. (1) depicts the GloVe model for the co-occurrence matrix.

$$\boldsymbol{w_i^T \tilde{w}_j + b_i + b_j} \approx \log(1 + \boldsymbol{X_{ij}}) \tag{1}$$

This method generates two sets of word vectors w_i and $\tilde{w}_i$. Typically, the left and right contexts are distinguished, so Xij is asymmetric. There are two distinct word vectors here. A single word vector is obtained as $w_i' = w_i + \tilde{w}_i$. The co-occurrence matrix has few values with high probabilities and most of its entries are very close to or equal to zero, making it sparse.

3.3 Classification Process

Data can be categorized using the process of classification into a specified number of classes. Classifying the crime can able to predict and reduce the crime in a particular area. It's crucial to categorize the various types of assaults. In this work, two deep learning methods are introduced for better classification. They are GRU and CNN. In the real dataset the following crime are available. They are, Murder, HBN(House Burglar by Night), rape, robbery, dowry, POSCO and Suspicious death. These are considered to be crimes and each one has been labeled 1 through 5 based on its nature.

3.3.1 GRU

The calculations carried out within the GRU cell to create the hidden state are what allow the GRU to retain long-term dependencies or memory[24]. The cell state and hidden state, which contain the long and short-term memories, respectively, are passed between the cells in LSTMs, whereas just one hidden state is passed between time steps in GRUs as shown in Fig 2. Due to the computations and gating procedures that the hidden state and input data undergo, this hidden state is able to hold both the long-term and short-term dependencies at the same time.

The updated hidden state serves as the basis for calculating the GRU's output. The following formulas are used to determine a GRU's reset gate(rt), update gate(u_t), and hidden state(h_t)

$$rt = sigmoid(wr \times [ht - 1, xt]) \tag{2}$$

$$ut = sigmoid(wu \times [ht - 1, xt]) \tag{3}$$

$$\mathrm{hht} = \tanh(wh \times [rt \times ht - 1, xt]) \tag{4}$$

$$ht = (1 - ut) \times ht - 1 + ut \times ht' \tag{5}$$

Here Wr,Wu and Wh are learnable weight matrices. Gradient descent is utilized to train the GRU network, which additionally produces an output at each time step.

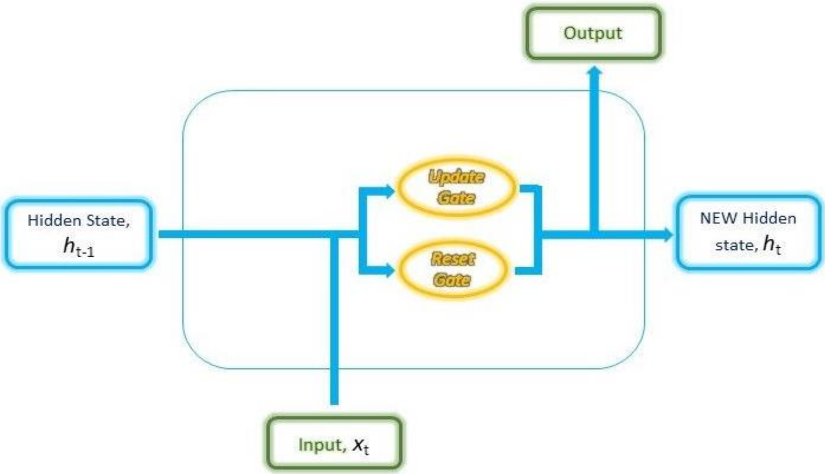


Fig 2: GRU Architecture

3.3.2 1CNN

1CNN is used to classify the crime using gist of the case. It takes lot of data, processing it in a grid structure, and then extracting crucial granular features for detection and classification. Convolutional, pooling, and fully connected layers are the three types of layers that make up a CNN as shown in Fig 3. A network that makes use of the convolution mathematical technique is referred to as a "convolutional neural network". Convolution is a specific form of linear processing[22]. Simple neural networks called convolutional networks use convolution rather than traditional matrix multiplication in at least one of their layers.

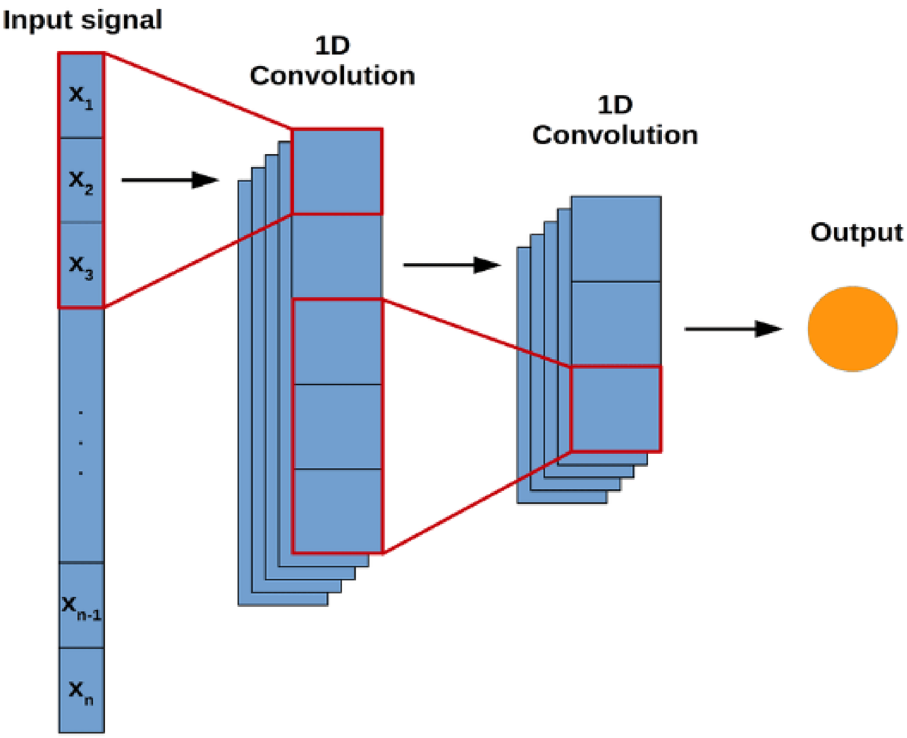


Fig 3: CNN Architecture

In order to process an entire string of words, convolutional kernels will iteratively cycle through a set of word embeddings [23]. This is referred to as a 1D convolution since the kernel is only moving in one dimension—time. One kernel will run through each word embedding on a list of input embeddings individually, starting with the first one (and a brief window of next-word embeddings), then moving on to the next, and the next, and so on. A feature vector with nearly the same number of values as the input embeddings will be the output. A variable dimension Q×N kernel is used in each 1D convolutional layer, where Q represents the temporal window that the filter covers and N = 2 because the kernel does not slide over the channels. The notation used in mathematics for a 1D convolutional is shown in Eq.6

$$yr = f(\sum_{q=1}^{Q} \sum_{n=1}^{N} w_{qn} x_{r+q,r+n} + b) \tag{6}$$

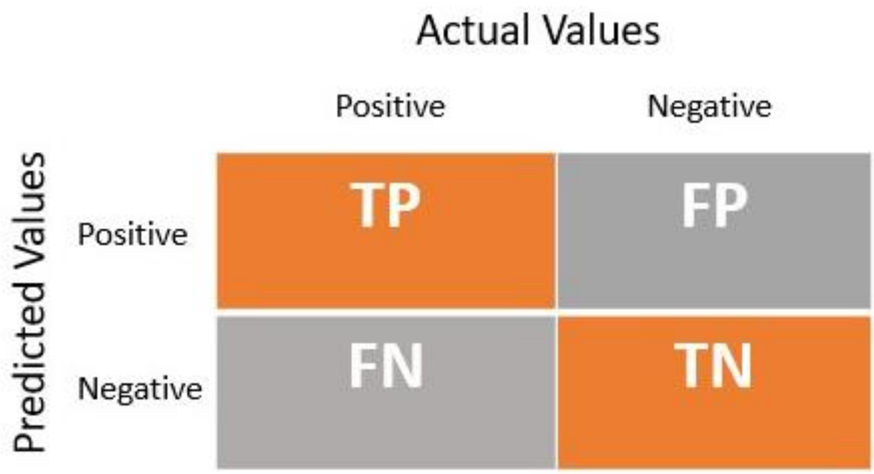
where x is the two-dimensional input portion that overlaps the filter; w is the convolutional filter's connection weight; b is the bias term; f is the filter's activation function; and yr is the output of the unit r of the filter feature map of size R. After the convolution operation (R), we can use the following formula to get the filter feature map's dimension:

$$R = \frac{M-(K-1)+2\times P}{S} \tag{7}$$

where *P* is the padding and *S* is the stride (number of positions skipped by each shift of the filter during convolution).

3.4 Performance Measures

A confusion matrix is used to evaluate the model's performance in a classification problem. Finding three crucial factors, namely accuracy, sensitivity, and specificity, is done by using the confusion matrix's elements[25]. Four results are produced by this categorization (or prediction): True Positive(TP), True Negative(TN), False Positive(FP), and False Negative(FN).



True positive (TP): correct positive prediction
False positive (FP): incorrect positive prediction
True negative (TN): correct negative prediction
False negative (FN): incorrect negative prediction

In this model, totally seven evaluation metrics are used, namely accuracy, precision, sensitivity(recall), specificity, F1 score and Matthew's Correlation Coefficient (MCC) to measure the performance of the models.

A false positive shows an error in binary classification. Test result incorrectly indicates the presence of a condition such as a crime type when the crime name is not present. A false negative is the opposite error where the test result incorrectly fails to indicate the presence of a condition when it is present. False Positive Rate (FPR) also called as fall-out is the probability of false alarm. It is the false positive value divided by sum of false positive and true negative is shown in Eq.(8).

$$FPR = \frac{FP}{(FP+TN)} \tag{8}$$

False Negative Rate (FNR) also called as miss rate. It shows the false negative value divided by sum of false negative and true positive is shown in Eq.(9).

$$FNR = \frac{FN}{(FN+TP)} \tag{9}$$

4. Experimental Results

This section reports the findings of the data analysis. Data from crimes like murder, robbery, kidnapping, and dower are included in the dataset. The initial data set was separated into two parts: training data (80%) and testing data (20%).

4.1 Data Source and Selection

The DCRB (District Crime Record Bureau) in Thoothukudi is the source of the actual dataset. Factual reports on decoity, daytime and nighttime home invasions, murder, robberies, dowries, POSCO, rape, and theft are among the data that have been gathered and are thought to be the major components. The DCRB has generated crime statistics that are ready for further processing. The Thoothukudi District has kept track of all violent crimes for the past five years, from 2018 to 2022, according to crime data. There have been roughly 2556 criminal cases overall, broken down into Thoothukudi District's eight divisions and 56 police stations.

The crime data were collected from the DCRB 'Crimes - 2018 to May 2022 present' dataset, a real-data which contains criminal cases happening in Thoothukudi District from 2018 to May 2022. We collected all criminal cases 8 divisions over six years from January 1, 2018, to May, 2022. Based on text analysis crime prediction is found. Some crime type and sample keywords used for crime prediction is given in Table 1.

4.2 Evaluation Metrics used

The number of all valid predictions divided by the total number of the dataset is used to calculate accuracy. Accuracy is the number of correctly predicted data points out of all the data points as shown in Eq. (10).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

Precision is also called as Positive Predictive Value (PPV). It is the number of correct positive results divided by the number of positive results predicted by the classifier as shown in Eq. (11).

$$\text{Precision} = \frac{TP}{TP+FP} \tag{11}$$

The number of accurate positive predictions divided by the total number of positives is used to calculate sensitivity. It is also known as recall or true positive rate (TPR). TPR is the true positive values divided by total number of actual yes. It is the number of correct positive results divided by the number of all relevant samples taken from data science as shown in Eq. (12).

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

The number of accurate negative predictions divided by the total number of negatives is used to calculate specificity as shown in Eq. (13). Another name for it is true negative rate (TNR).

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{13}$$

F1 score is the harmonic mean of the precision and recall as shown in Eq. (14).

$$\text{F1} = 2\times \frac{Precision \times Recall}{Precision+Recall} \tag{14}$$

The statistical measure known as Matthew's Correlation Coefficient (MCC) is employed to assess models. It performs the same function as chi-square statistics for a 2 x 2 contingency table, which is to assess the difference between the expected values and actual values as shown in Eq.15.

$$\text{MCC}=\frac{TN \times TP - FN \times FP}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (15)$$

4.3 Results and Discussions

The illustration below provides a good explanation of the data's highest crime rate from 2018 to 2022. These records of crimes collected from DCRB. This work implemented by using GRU and 1CNN techniques for crime classification.

The performance of model is measured by accuracy, precision, sensitivity(recall), specificity, F- score and MCC. 1CNN yields better performance compared to GRU. The comparison is performed with 50 epoches. Table 1 and Fig. 4 represent the performance comparison of GRU and 1CNN.

Table 1: Comparison of Performance between GRU and 1D CNN

	GRU	**CNN**
Accuracy	0.9609	0.9788
Precision	0.9379	0.9385
Sensitivity	0.6528	0.7222
Specificity	0.9728	0.9789
F-Score	0.6842	0.7672
MCC	0.7129	0.7793

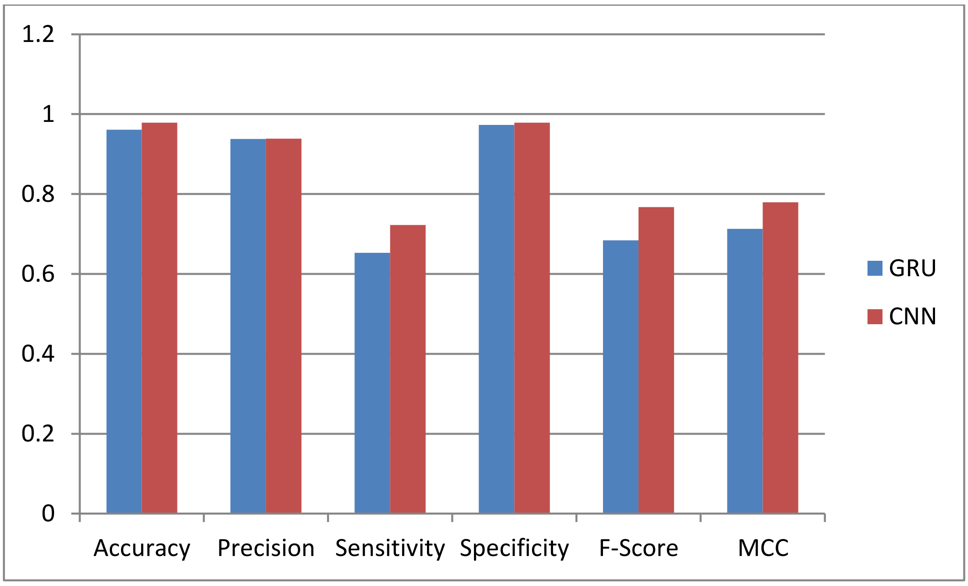


Fig 4: Comparison of Performance between GRU and 1CNN

With the help of metrics, 1CNN is a best classifier for our real dataset. Specificity gives better result. Crime against women and Child are high according to ROC curve. Next level crime is Murder. Mostly womens are murdered in Thhothukudi District. Specificity provides better result. Due to lack of information in our real time dataset we couldn't find family background of accused and reason for committing offense.

The consequence of a poor prediction is loss. In other words, loss is a measure of how poorly the model predicted a single case. The loss is zero if the model's forecast is accurate; otherwise, the loss is higher. Training and Validation Loss Function of GRU and 1CNN is shown in Fig.5 and Fig.6.

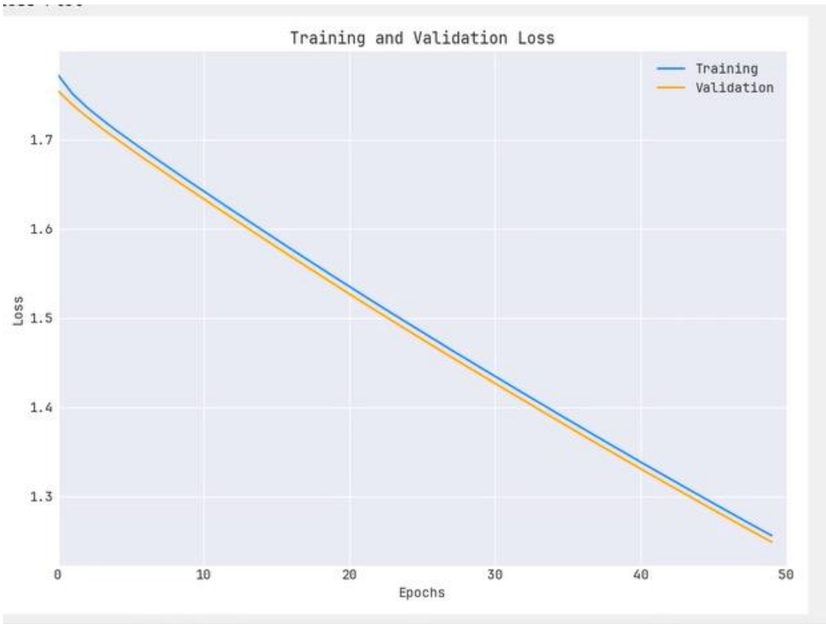


Fig. 5: Training and Validation Loss Function of 1CNN

Fig. 6: Training and Validation Loss Function

Table 2 and Fig. 7 represent the classification accuracy of the GRU and 1CNN techniques with other ML models [20]. The results highlight that the 1CNN technique gains a higher accuracy of 97.88%. At the same time, GRU, Decision Tree and Naïve Bayes models obtained decreased accuracy values of 96.09%, 91.68% and and 83.40% respectively.

Table 2: Accuracy of Various Techniques

Techniques	**Accuracy(%)**
1CNN	**97.88**
GRU	96.09
Decision Tree	91.68
Naïve Bayes	83.4

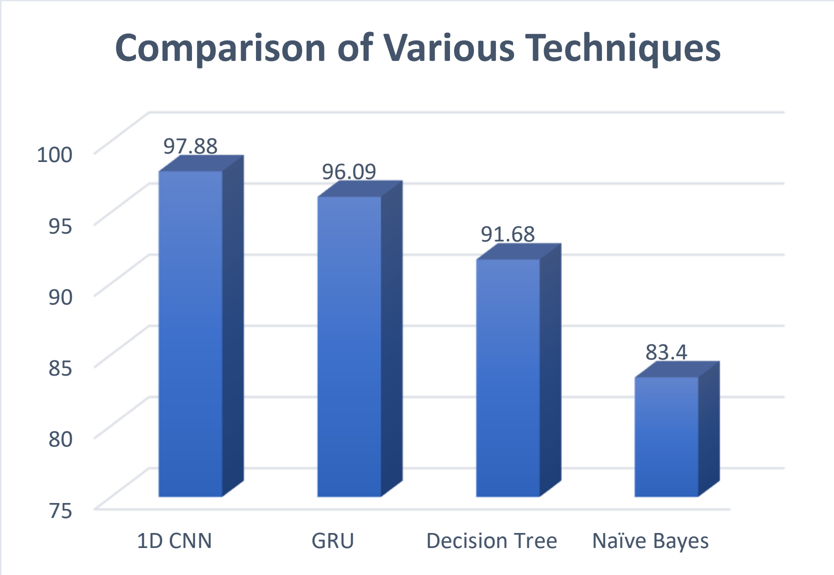


Fig 7: Comparison of Accuracy for Various Techniques

5. Conclusions And Future Work

In an effective manner, deep learning approaches are being created for the classification of crimes. On the DCRB 'Crimes - 2018 to May 2022 current' dataset, this research evaluated the effectiveness of 1CNN and GRU algorithms. 1CNN is regarded as one of the greatest strategies for classifying crimes in this work. When compared to GRU, the 1CNN exhibits great accuracy (0.9788). These two techniques are compared with Naïve Bayes and Decision tree techniques.

The proposed method will be enhanced by applying cutting-edge methodologies and a better set of crime data.

References

[1] Tamilarasi, P., and R. Uma Rani. ”Diagnosis of Crime Rate against Women using k-fold Cross Validation through Machine Learning.” 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2020.

[2] Rajapakshe, Charuni, et al. ”Using CNNs RNNs and Machine Learning Algorithms for Real-time Crime Prediction.” 2019 International Conference on Advancements in Computing (ICAC). IEEE, 2019.

[3] H. Verma, S. Lotia and A. Singh, "Convolutional Neural Network Based Criminal Detection," 2020 IEEE REGION 10 CONFERENCE (TENCON), Osaka, Japan, 2020, pp. 1124-1129, doi: 10.1109/TENCON50793.2020.9293926.

[4] Derek B Cornish and Ronald V Clarke. The reasoning criminal: Rational choice perspectives on offending. Transaction Publishers, 2014.

[5] Panagiotis Stalidis, Theodoros Semertzidis and Petros Daras,”Examining Deep Learning Architectures for Crime Classification and Prediction” ,, IEEE and Petros Daras, Senior Member, IEEE.

[6] Jankowski, Stanisław, et al. "Selection of neural network for crime time series prediction by virtual leave-one-out tests." Theory and Applications of Time Series Analysis: Selected Contributions from ITISE 2018 5. Springer International Publishing, 2019.

[7] Jain, Bhanu, et al. "Singular race models: addressing bias and accuracy in predicting prisoner recidivism." Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments. 2019.

[8] Jyotsna Bhargavi Karri . Classification of Crime Scene Images using the Computer Vision and Deep Learning Techniques. International Journal for Modern Trends in Science and Technology 2022, 8 pp. 01-05. https://doi.org/10.46501/IJMTST0802001.

[9] Liu Ying, Zhang Qian Nan, Wang Fu Ping, Chiew Tuan Kiang, Lim Keng Pang, Zhang Heng Chang, Chao Lu, Lu Guo Jun & Ling Nam (2021) Adaptive weights learning in CNN feature fusion for crime

scene investigation image classification, Connection Science, 33:3, 719-734, DOI: 10.1080/09540091.2021.1875987.

[10] Liu, Y., Zhang, S., Wang, F., & Ling, N. (2018). Read pattern image classification using convolutional neural network based on transfer learning. IEEE Workshop on Signal Processing Systems (SIPS, 2018), pp. 21–24.

[11] C. Chauhan and S. Sehgal, "A review: Crime analysis using data mining techniques and algorithms," 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017, pp. 21-25, doi: 10.1109/CCAA.2017.8229823.

[12] Deepak, G., Rooban, S. & Santhanavijayan, A. A knowledge centric hybridized approach for crime classification incorporating deep bi-LSTM neural network. Multimed Tools Appl 80, 28061–28085 (2021). https://doi.org/10.1007/s11042-021-11050-4.

[13] P. Sivakumar, J. V, R. R and K. S, "Real Time Crime Detection Using Deep Learning Algorithm," 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), Puducherry, India, 2021, pp. 1-5, doi: 10.1109/ICSCAN53069.2021.9526393.

[14] Stalidis P, Semertzidis T,Daras P. ,"Examining Deep Learning Architectures for Crime Classification and Prediction," Forecasting 2021, 3, 741–762. https://doi.org/10.3390/ forecast3040046.

[15] P. Das, A. K. Das, J. Nayak, D. Pelusi, and W. Ding, ''A graph based clustering approach for relation extraction from crime data,'' IEEE Access, vol. 7, pp. 101269–101282, 2019.

[16] Jha, Pranay, Raman Jha, and Ashok Sharma. "Behavior analysis and crime prediction using big data and machine learning." International Journal of Recent Technology and Engineering 8.1 (2019).

[17] Stalidis, P.; Semertzidis, T.; Daras, P. Examining Deep Learning Architectures for Crime Classification and Prediction. Forecasting 2021, 3, 741–762. https://doi.org/10.3390/ forecast3040046.

[18] J. Azeez and D. J. Aravindhar, "Hybrid approach to crime prediction using deep learning," 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015, pp. 1701-1710, doi: 10.1109/ICACCI.2015.7275858.

[19] M. N. Haider, R. Mumtaz and S. M. H. Zaidi, "Crime Classification Using Machine Learning and Data Analytic," 2022 IEEE 19th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET), Marietta, GA, USA, 2022, pp. 093-098, doi: 10.1109/HONET56683.2022.10018990.

[20] S. A. Bshayer, M. A. Futun, S. A. Noura, M. A. Manar, M. A. Razan, A. Nida, et al., "A Comparative Study of Decision Tree and Naive Bayes Machine Learning Model for Crime Category Prediction in Chicago", Proceedings of 2020 the 6th International Conference on Computing and Data Engineering (ICCDE 2020), pp. 34-38, 2020.

[21] Wang B, Yin PH, Bertozzi AL, Brantingham PJ, Osher SJ, Xin J (2019) Deep learning for real-time crime forecasting and its ternarization. Chin Ann Math Ser B 40(6):949–966.

[22] Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville published by MIT Press, 2016.

[23] https://datascience.stackexchange.com/questions/14349/difference-of-activation-functions-in-neural-networks-in-general.

[24] ArunKumar KE, Kalaga DV, Mohan Sai Kumar C, Kawaji M, Brenza TM. Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. Alexandria Engineering Journal. 2022 Oct;61(10):7585–603. doi: 10.1016/j.aej.2022.01.011. Epub 2022 Jan 6. PMCID: PMC9453185.

[25] Vanacore, A., Pellegrino, M.S. & Ciardiello, A. Fair evaluation of classifier predictive performance based on binary confusion matrix. Comput Stat (2022). https://doi.org/10.1007/s00180-022-01301-9

[26] "Crime—an Overview ." Information Plus(R) Reference Series Fall 2004. . Encyclopedia.com. 29 Jun. 2023 https://www.encyclopedia.com.

[27] S. Chackravarthy, S. Schmitt and L. Yang, "Intelligent Crime Anomaly Detection in Smart Cities Using Deep Learning," 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC), Philadelphia, PA, USA, 2018, pp. 399-404, doi: 10.1109/CIC.2018.00060.