Educational Administration: Theory and Practice
2024, 30(5), 2950-2964
ISSN: 2148-2403
https://kuey.net/

Research Article

Crime Classification Using GRU, CNN And Autoencoder Techniques

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

[1*]Research Scholar[19124012282037], Research Department of Computer Science,Kamaraj College,Affiliated to Manonmanian Sundaranar University, jeyrobin@gmail.com,
[2]Associate Professor, Research Department of Computer Science,Kamaraj, College,Affiliated to Manonmanian Sundaranar University peterchristy_in@yahoo.com

Citation: S. Jeya Selvakumari, Dr. V. Joseph Peter, (2024), Crime Classification Using GRU, CNN and Autoencoder Techniques, *Educational Administration: Theory and Practice, 30(5),* 2950-2964
Doi: 10.53555/kuey.v30i5.3372

ARTICLE INFO **ABSTRACT**

An activity that violates the law is considered a crime. It is harmful to society to understand crime in order to avoid criminal action. The various research works are performed regarding crime prediction. Many researchers' work based on San Francisco Dataset and analyzed the common crime pattern. They have failed to collect real time dataset for the particular region and classify the crime against women. In this work we have collected real time dataset and classify the crime based on the activity of offender that happened in Thoothukudi district. This helps to find specific crime like crime against women and child and helps to take necessary action to reduce the crime. Using this we can analyse the reason and history of offender. For that we have used One Dimensional Convolutional Neural Network (1DCNN), Gated Recurrent Unit(GRU) and Autoencoder deep learning techniques to improve the performance of crime classification. For our dataset Autoencoder yields high accuracy, sensitivity, F1-score and MCC. GRU provides better precision over the 1DCNN and Autoencoder.

Keywords Crime, Deep Learning, 1D CNN, GRU, Autoencoder, Classification.

1 INTRODUCTION

The amount of crime in the world is rising daily. Since offence is either proficient or unplanned, it cannot be anticipated. A crime or offence is an action that causes harm to a group of people, a society, or the state as a whole ("a public wrong"). It is an illegal conduct that is subject to governmental or other punishment. According to the study, crime exhibits some patterns throughout a given time period and across various geographic locations. Crime is an economical problem that lowers life value and increases costs.
Recently, crime modeling and predictions have incorporated deep learning. We looked at real-time crime forecasting at a small geographical scale and investigated the problem by turning it into a multi classification problem. The basic concept is to use historical data as inputs to a convolutional neural network (CNN) so that it can learn the characteristics for crime pattern. A system for AI-based crime prediction is suggested to increase public safety and assist law enforcement by lowering the effort needed for human procedures when identifying crimes and assigning resources to quicken emergency response. With the help of this technology, police personnel may keep informed about both current occurrences and crimes that are anticipated to happen in the near future. This helps police personnel prepare for upcoming situations and take precautions to avoid repeat occurrences.
Crime is related to public safety and protection, and having a better understanding of crime is advantageous in a number of ways, including by encouraging law enforcement to use sensitive and targeted tactics to reduce crime and by encouraging citizens and the government to work together more closely to foster safe neighborhoods. Understanding patterns in crime from data is an active and expanding subject of research because to the Big Data age and the accessibility of quick, effective algorithms for data analysis.

The objective of the proposed system is to classify the crime based on activity performed by offender using 1D CNN, GRU and Autoencoder deep learning techniques. Gist of the case is the source for finding action. Some preprocessing techniques are used to extract the feature. After that GLoVe embedding is used to find the similar words. For example, crimes against child can be found out with the key words like "minor girl", "child pornography", "sexually assaulted", and "kidnapping". The system should be trained to identify similar words and classify the crime accurately.
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. Experimental results of 1D CNN, GRU and Autoencoder are discussed and various chart are shown in Section 4. Finally, a general summary and future research plan are introduced in Section 5.

2 RELATED WORKS

Recurrent Neural Network (RNN) with Gated Recurrent Unit (GRU) architecture was suggested as a model to predict the crime rate in Banjarmasin, taking inflation rate and discretionary income into account. The R-Squared value of 0.84 and the RMSE value of 2.21 of the GRU-RNN model demonstrated good outcomes, according to the findings[1]. We employed the Prisma flow method, a method for summarising systematic reviews and meta-analyses[2]. By utilising LSTM RNN, a new architecture was created to precisely anticipate Emergency Eents. Spatial clustering supplied the fundamental building blocks for EV prediction in the architecture[3]. To extract crime scores from street view photos, we suggest using the convolutional neural network (CNN) StreetNet. Preference learning and label ranking settings are the cornerstones around which the learning process is built. To enhance the representation of urban perception, developed a retrieval technique for street view photos. Finding impartial label mappings between the street view photos and the crime ranking records is presented using a data-driven, spatiotemporal method[4][25][30]. Prophet model, a neural network model, and the deep learning method LSTM revealed that they all outperform traditional neural network models. In order to achieve the best trend prediction in terms of RMSE and spearman correlation, we also discovered that the ideal time span for the training sample was 3 years. Additionally, the Prophet and LSTM models' ideal parameters are established5][31]. In order to provide insight into what predictive policing is, how it may be utilised and applied to forecast crime, and what is known about its success, the author offers the findings of a literature review augmented with key informant interviews. Additionally, it provides a summary of the key features and recognised uses of predictive policing[6].
The hotspots (also known as binary classifcation) approach is the most common kind of predicting inference. The majority of techniques utilised in this study were traditional machine learning techniques, along with kernel density estimation-based approaches and, less frequently, point process and deep learning techniques[7]. provided a prediction learning model using data modelling using Dynamic Windows to determine whether a certain person will commit a new crime within n window years in the future[8]. When classified into seven categories, comparable crimes are predicted by neural network models in 27 different types of crimes 16.4% of the time and 27.1% of the time, respectively. In 31.2% of cases, location prediction neural networks can accurately guess a zip code's location or a nearby location[9]. A number of indicators are presented to assess the effectiveness of the suggested deep neural network model. The best model was found by comparing the outcomes of several scenarios (tasks), and it uses the occurrences of the previous two days as its inputs to forecast the third day. The validation set's coefficient of determination in this instance was 97%. Despite using temporal data, the deep convolutional model has demonstrated excellent prediction ability, in this example using six convolutions[10]. Using the CNN-LSTM model, they tried to predict the crime rate. They used the three-year crime dataset obtained from the NCRB for this study. We settled on the following four criteria: homicide, rape, theft, and property crimes. In the beginning, we extract the attributes from the dataset using CNN, and we forecast the crime rate using LSTM. 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[11][26]. CNN's accuracy rate of 99 percent is 100% acceptable would be extremely biassed and overly pessimistic[12][27].
Crime Intention Detection System is an automated system that prevents crimes from happening by identifying a person's possession of a gun or knife using a pre-trained VGGNET19 model in less time than it would take to use a GoogleNet Inception V3 model[13]. To automatically categorise the criminal court cases, a machine learning method called naive bayes is applied. The model presented in this research may therefore identify ICT involvement in criminal court cases, as well as the features upon which it is based and the precision of this categorization[14]. When analysing historical data to discover trends in prior crime occurrences, gradient boosting algorithms demonstrated great accuracy[15]. Based on the information from the previous 24 hours, the suggested architecture forecasts the upcoming hour. The output of our proposed model can be the predicted number of crimes in any category for a specific region. Additionally, it gives law enforcement information on potential criminal activity based on category, location, and time[16]. The author made the conclusion that identifying and forecasting crime hotspots is an important procedure that requires more research. Throughout this methodical procedure, several crucial study areas are discovered, which will aid the researchers in developing a better and more reliable crime prediction system[17. Parallel crime scene analysis system based on artificial societies, computational experiments, and parallel execution (ACP), which leverages

artificial (A) crime scenes. computational (C) experiments to compute and predict the various types of crime scene, and parallel (P) execution to direct or control the evolution of the actual criminal activity in accordance with the findings from the artificial crime scene. scene to describe the fundamental components, functions, and states of the criminals[18]. Two separate data-processing techniques are used to analyse crime statistics for Vancouver during the last 15 years. When predicting crime in Vancouver, machine-learning predictive models K-nearest-neighbour and boosted decision tree are used, and an accuracy range of 39% to 44% is obtained[19]. As a result, we were able to anticipate whether a crime will happen, how many would happen within an hour, and what kind of crime would happen most frequently[20]. With murder being the predominant criminal attribute in relation to the other categories, extensive study is conducted. On train and test data, we used the Bayesian, Levenberg, and Scaled algorithms to discover the iteration at which the best valid performance was achieved. It was found that, for the data under consideration, the Scaled algorithm produced the best result when compared to Bayesian and Levenberg . The results of the statistical analysis, which was based on correlation, ANOVA, and graphs.

The investigation revealed that using deep learning algorithms, genetic algorithms, and other algorithms, the crime rate may be lowered to 78 percent[21]. Using KNN and decision trees, Naive Bayes, Linear Regression CART (Classification and Regression Tree), and SVM algorithms utilising crime datasets from Chicago and Los Angeles, we enhanced the forecast accuracy for crimes[29]. On Chicago datasets, XGBoost obtains the highest level of accuracy among the algorithms, whereas KNN does so on Los Angeles datasets[22]. In order to assist both law enforcement officials and scientists in reducing and preventing future crime occurrences, it is our goal to gather and synthesise the necessary knowledge regarding machine learning-based crime prediction[23]. Deep convolutional autoencoder neural network for numerous temporary inputs comprising a day's worth of crime predictions[24]. The Faster R-CNN based model produced the most promising results[28]

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 system model.

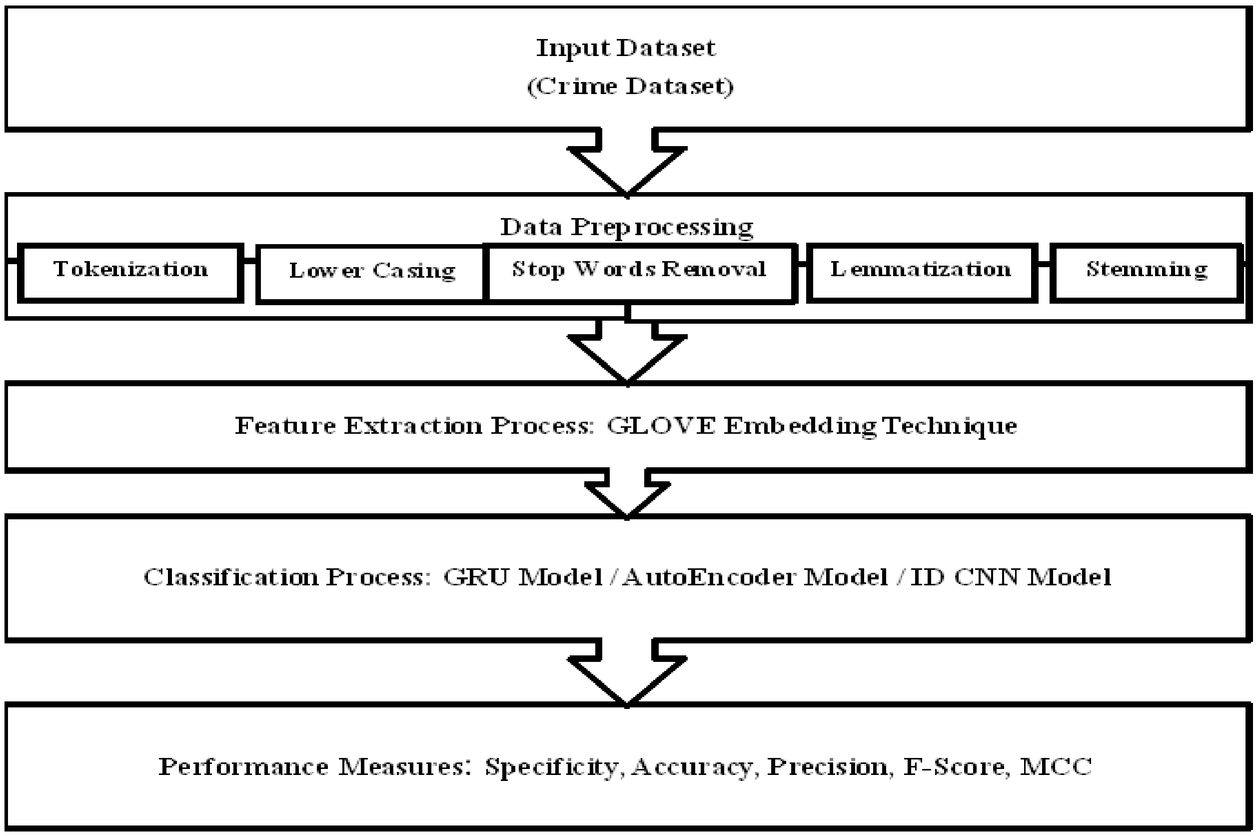


Figure 1 Architecture of the proposed method

3.1 Preprocessing

This proposed method has used real time data which is collected from DCRB of Thoothukudi. This data set contains many fields like Crime Number, section of law, date of occurance, location, sub division, reason and gist of the case. Gist of the case contains summary of the crime case. To precede the proposed work efficiently, preprocessing the raw data is mandatory. The following steps are followed to preprocess the gist of the crime. They are tokenization, lowercase conversion, removing stop words, numbers, punctuations and extra spaces. Then applied lemmatization and stemming for better performance. Lemmatization often refers to carrying out tasks correctly using a vocabulary and linguistic analysis of words to remove only inflectional endings and return the lemma, or dictionary form, of a word. Stemming is a term that often describes a fundamental computational method that removes derivational affixes from words in the hopes of attaining this goal most of the time. Both stemming and lemmatization aim to reduce a word's inflectional forms and occasionally related derivational forms to a basic form. The technique of lemmatization involves reducing a word to its most basic form. Lemmatization, as opposed to stemming, evaluates the context and changes the word to its relevant base form. Stemming just eliminates the final few characters, frequently resulting to inaccurate spellings and meanings.

3.2 Feature Extraction

After preprocessing , some steps to be followed for feature extraction in order to produce better results. The technique of turning raw data into numerical features that can be handled while keeping the information in the original data set is known as feature extraction. In this work, countervectorizer is built for numerical conversion, data is fitted, transforming text to features, shaping the features(128 by 1572), picking countervectorizer and saved the featurs for training the model. Further, GloVe Embeddings is also used.

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.

$$\boldsymbol{w}_{\boldsymbol{i}}^{\boldsymbol{T}}\widetilde{\boldsymbol{w}}_{\boldsymbol{j}}+\boldsymbol{b}_{\boldsymbol{i}}+\boldsymbol{b}_{\boldsymbol{j}} \approx \log(1+\boldsymbol{X}_{\boldsymbol{ij}}) \qquad (1)$$

This method generates two sets of word vectors w_i and $\widetilde{w}_i$. 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 $w_i'=w_i + \widetilde{w}_i$. 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. Finding the class is the main goal of a grouping problem. It's crucial to categorize the various types of assaults. There are three classification methods: ICNN, GRU3, and AE. Murder, HBN(House Burglar by Night), rape, robbery, dowry, POSCO and Suspicious death are all considered to be crimes and each one has been labeled 1 through 5 based on its nature. The collection is arranged similarly in Table 1 according to the nature of the offence.

3.3.1 1D CNN

The phrase "convolutional neural network" refers to a network that uses the convolution mathematical technique. A specialised kind of linear processing is convolution. Convolutional networks are simple neural networks that, in at least one of their layers, employ convolution rather than standard matrix multiplication[33].

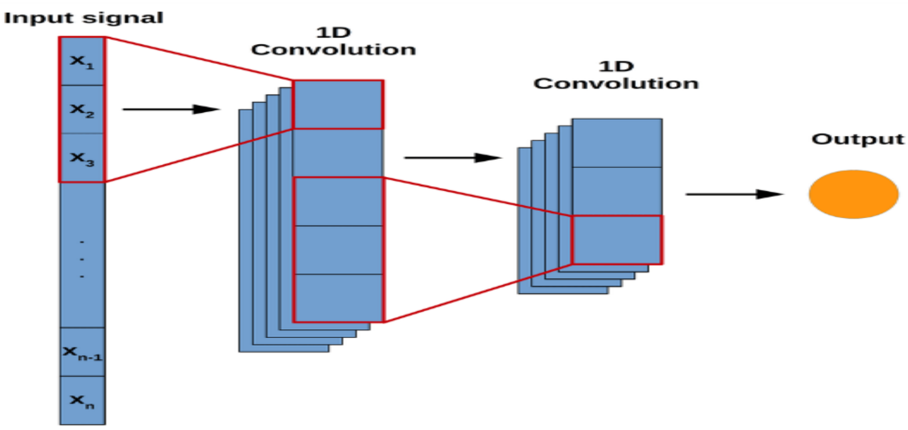


Fig 2. Structure of ID CNN

Convolutional kernels will slide through a list of word embedding in order to process a complete string of words as shown in figure 2. Due to the fact that the kernel is only moving in one dimension time, this is known as a 1D convolution. A single kernel will go through each word embedding on a list of input embeddings one at a time, starting with the first one (and a short window of next-word embeddings), then moving on to the next, and the next, and so forth. The output will be a feature vector with roughly the same number of values as the number of input embeddings.

3.3.2 GRU

The ability of the GRU to hold on to long-term dependencies or memory stems from the computations within the GRU cell to produce the hidden state[33]. While LSTMs have two different states passed between the cells — the cell state and hidden state, which carry the long and short-term memory, respectively — GRUs only have one hidden state transferred between time steps as shown in figure 3. This hidden state is able to hold both the long-term and short-term dependencies at the same time due to the gating mechanisms and computations that the hidden state and input data go through[32].

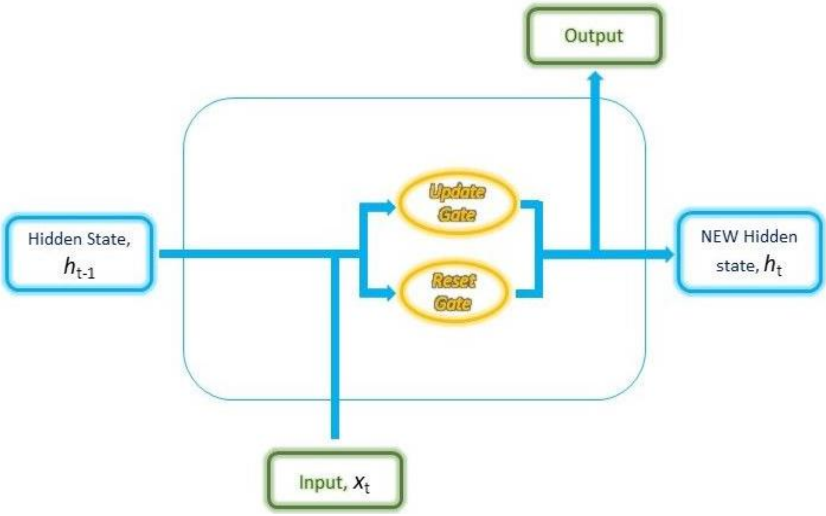


Figure 3 Structure of GRU

3.3.3 AUTOENCODER

An autoencoder is a kind of neural network that may be trained to learn how to rebuild images, text, and other types of data from compressed copies of itself. Typically, an autoencoder has three layers: Encoder, Code and Decoder[33]. To decrease dimensionality, we can use an autoencoder to compress input data, encode it, and then reconstruct the data in a different format. Autoencoders let you to concentrate solely on the most important aspects of your data. The autocoder for one-class categorization in our work has a straightforward architecture with just one layer.

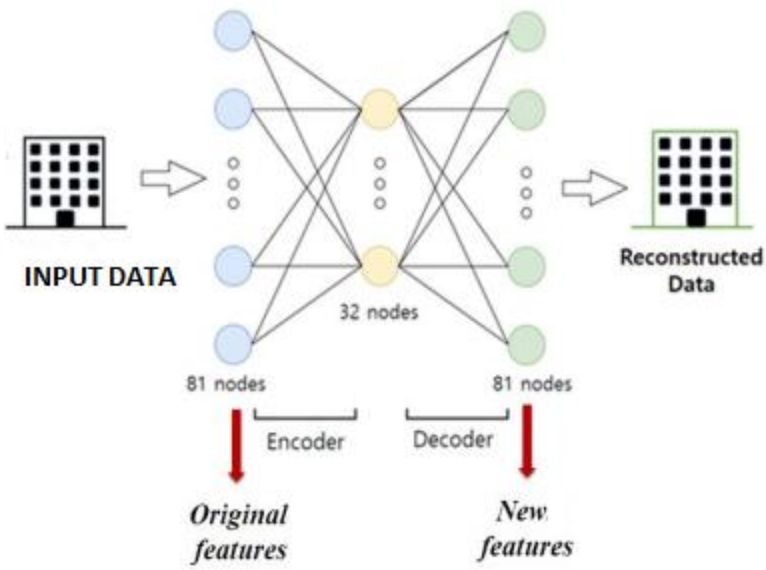


Figure 4 Autoencoder Structure

According to Figure 4, the autoencoder takes 81-dimensional input data and decreases it to 32 dimensions in the encoder stage before increasing it once again to 81 dimensions in the decoder stage. 2,500 data from the training set are used to teach the autoencoder[35].

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%).

Data Source and Selection

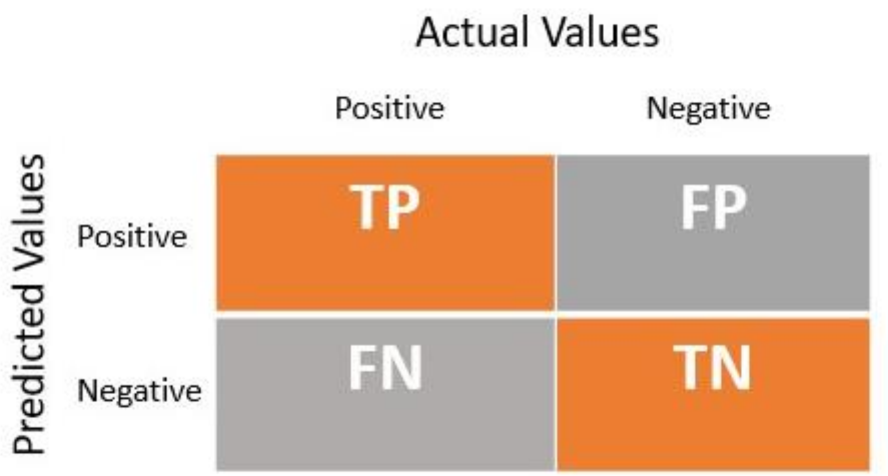
The DCRB (District Crime Record Bureau) in Thoothukudi is the source of the actual dataset. Factual reports on decoity, daytime and nighttime home invasions, murder, robberies, dowries, POSCO, rape, and theft are among the data that have been gathered and are thought to be the major components. The DCRB 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.

4.1 Data set used

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.

4.2 Evaluation Metrics used

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. 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 (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 as shown in Eq. (1).

$$FPR=\frac{FP}{FP+TN} \quad (1)$$

False Negative Rate (FNR) also called as miss rate. It shows the false negative value divided by sum of false negative and true positive as shown in Eq.(2).

$$FNR=\frac{FN}{FN+TP} \quad (2)$$

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. (3).

Accuracy=$\frac{TP+TN}{TP+TN+FP+FN}$ (3)

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.(4).

Precision = $\frac{TP}{TP+FP}$ (4)

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.(5).

Sensitivity(Recall) = $\frac{TP}{TP+FN}$ (5)

The number of accurate negative predictions divided by the total number of negatives is used to calculate specificity as shown in Eq.(6). Another name for it is true negative rate (TNR).

Specificity = $\frac{TN}{TN+FP}$ (6)

F1 score is the harmonic mean of the precision and recall as shown in Eq. (7).

F1 = 2× $\frac{Precision \times Recall}{Precision+Recall}$ (7)

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.(8)

MCC=$\frac{TN\times TP-FN\times FP}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ (8)

4.3 Results and Discussions

The illustration below provides a good explanation of the data's highest crime rate from 2018 to 2022. These reucords of crimes collected from DCRB. This work implemented by using ID CNN, GRU and Autoencoder techniques for crime classification. The confusion matrix of 1D CNN is obtained for our dataset is shown in Fig.5.

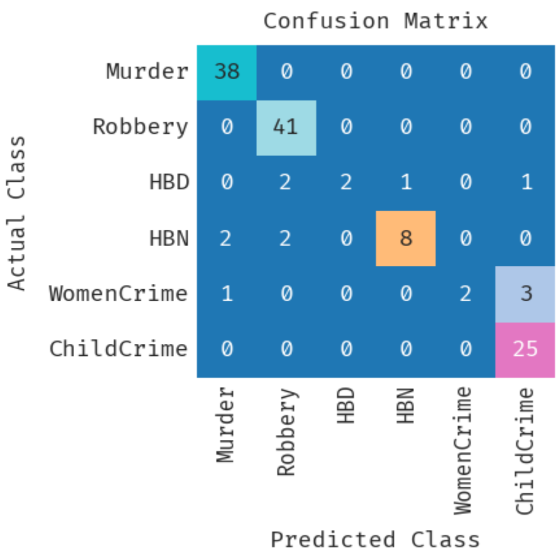


Fig 5. The confusion matrix of 1D CNN

The appropriate threshold for a given problem can be chosen using a precision-recall curve, which also aids in visualising how the threshold selection influences IDCNN classifier performance. It is often used in situations where classes are heavily imbalanced. Generally, the higher the AUC-PR score, the better a classifier performs for the given task. AUC-PR stands **for** area under the (precision-recall) curve as shown in Fig 6**.** Different crime categories are compared by using Precision-Recall curve. Average Precisiopn (AP) is high in childcrime.

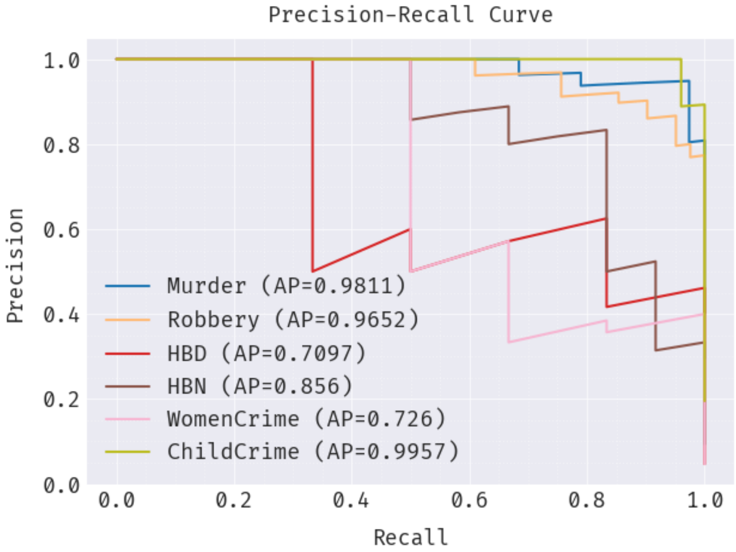


Fig 6. AUC – PR graph

A Receiver Operating Characteristic (ROC) curve is a plot of the true positive rate (Sensitivity) is function of the false positive rate (100-Specificity) for different cut-off points of a parameter. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. ROC curve shown in Fig.7.

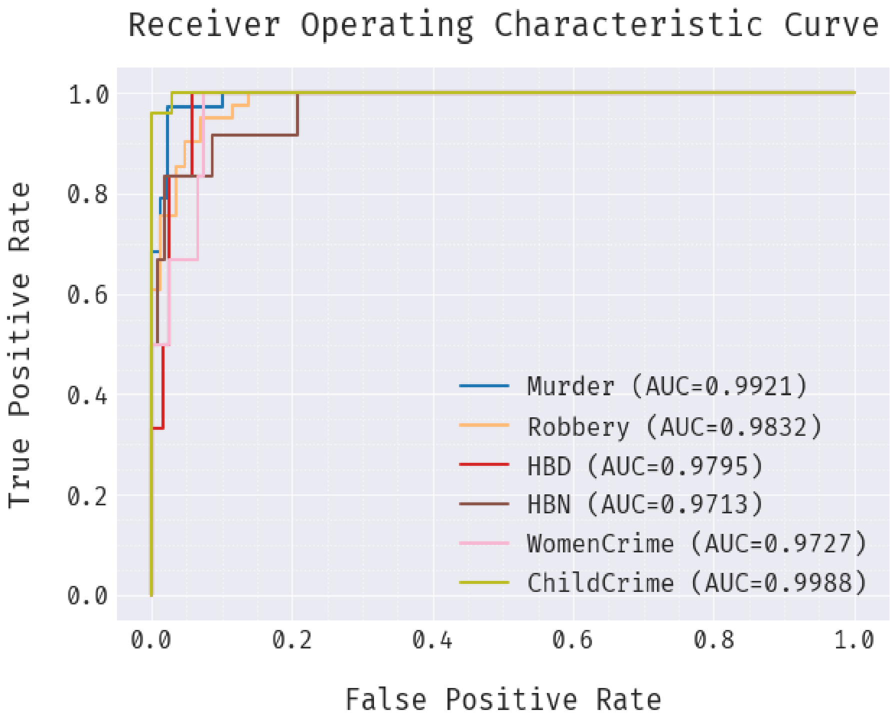


Fig 7. ROC Curve

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 1D CNN is shown in Fig.8.

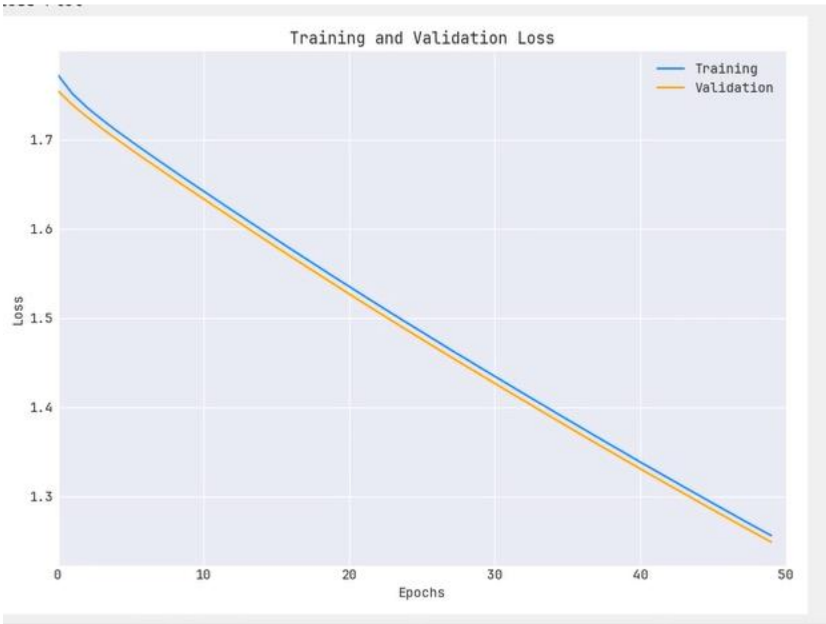


Fig 8. Training and Validation Loss Function of 1D CNN

The confusion matrix of GRU is obtained for our dataset is shown in Fig. 9

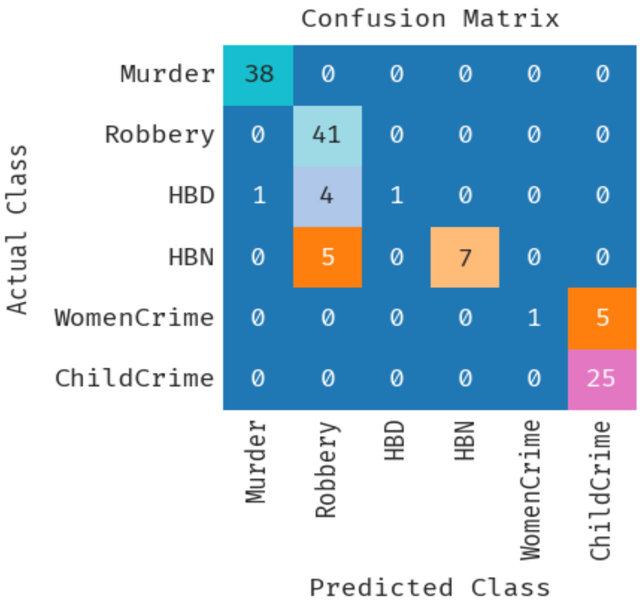


Fig 9. The Confusion Matrix of GRU

The appropriate threshold for a given problem can be chosen using a precision-recall curve, which also aids in visualising how the threshold selection influences GRU classifier performance. It is often used in situations where classes are heavily imbalanced. Generally, the higher the AUC-PR score, the better a classifier performs for the given task. AUC-PR stands for **area under the (precision-recall) curve.** Average Precision (AP) is high in murder and child crime.

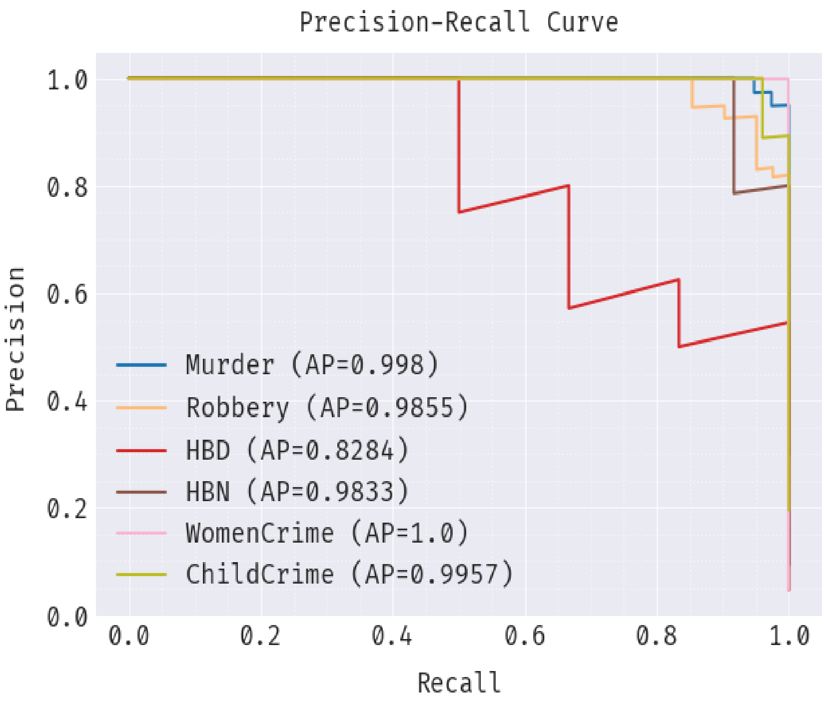


Fig 10. AUC – PR graph

A Receiver Operating Characteristic (ROC) curve is a plot of the true positive rate (Sensitivity) is function of the false positive rate (100-Specificity) for different cut-off points of a parameter. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. The classifier can successfully differentiate between all Positive and Negative class values when AUC = 1. However, if the AUC had been 0, the classifier would have predicted that all Positives were Negatives and all Negatives were Positives. ROC curve shown in Fig 11.

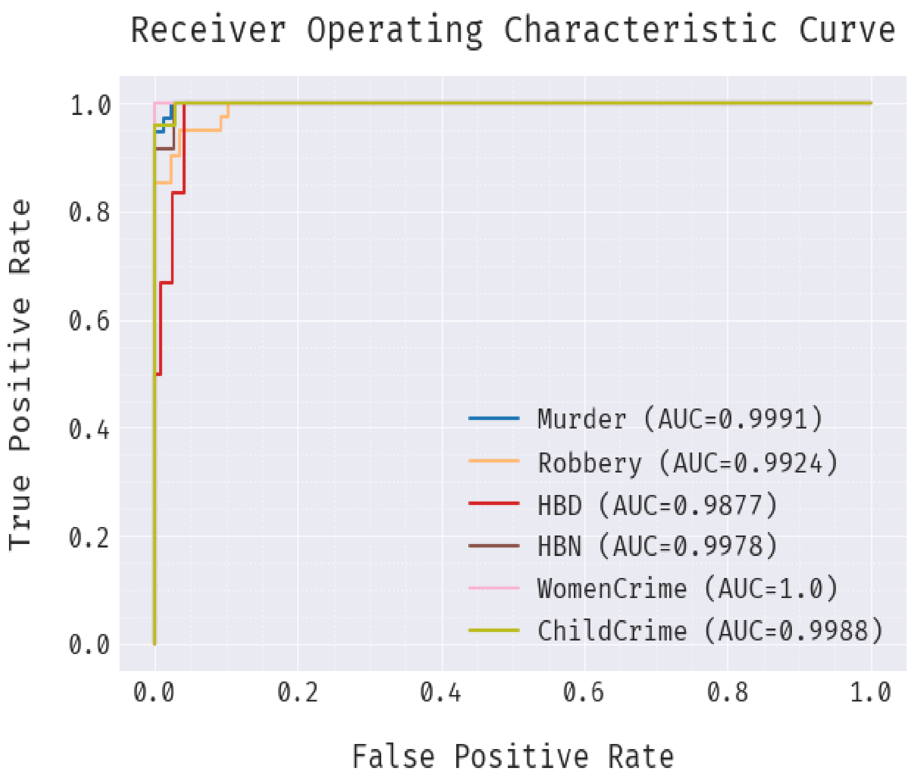


Fig. 11 ROC Curve

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 is shown in Fig.12

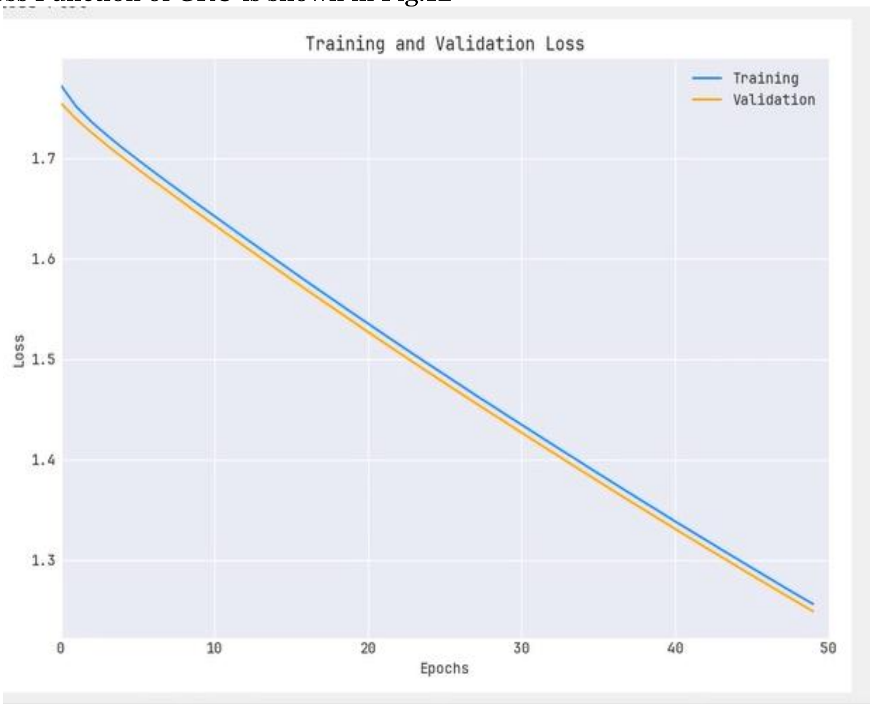


Fig 12. Training and Validation Loss Function of GRU

The confusion matrix of Autoencoder is obtained for our dataset is shown in Fig. 13.

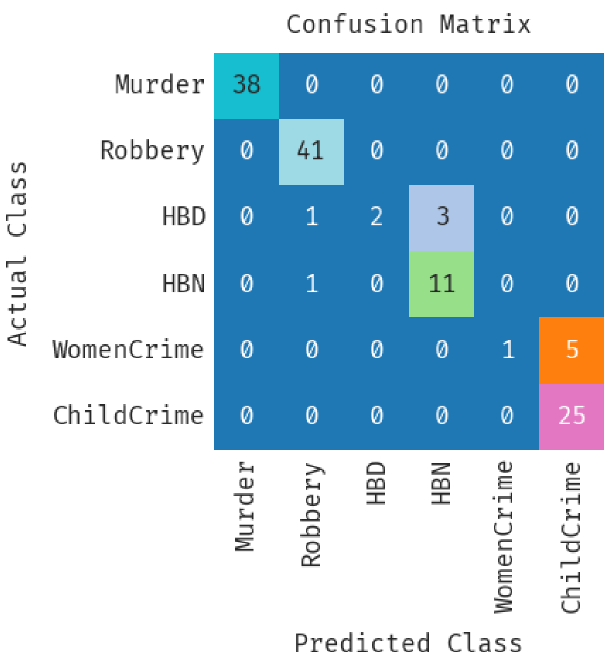


Fig. 13 Confusion Matrix of AutoEncoder

Different crime categories are compared by using Precision-Recall curve. Average Precisiopn(AP) is high in murder, robbery and child crime.

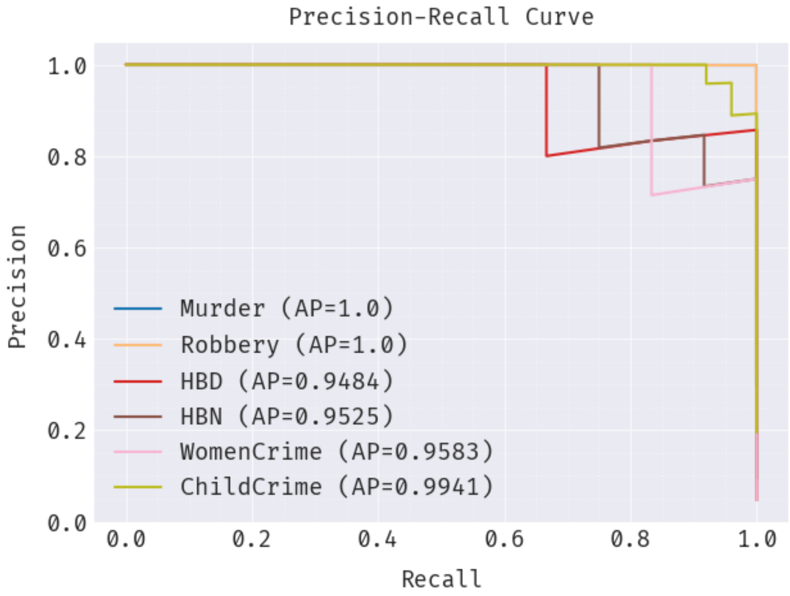


Fig 14. AUC – PR graph

Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. AUC=1 in Murder and Robbery crime.

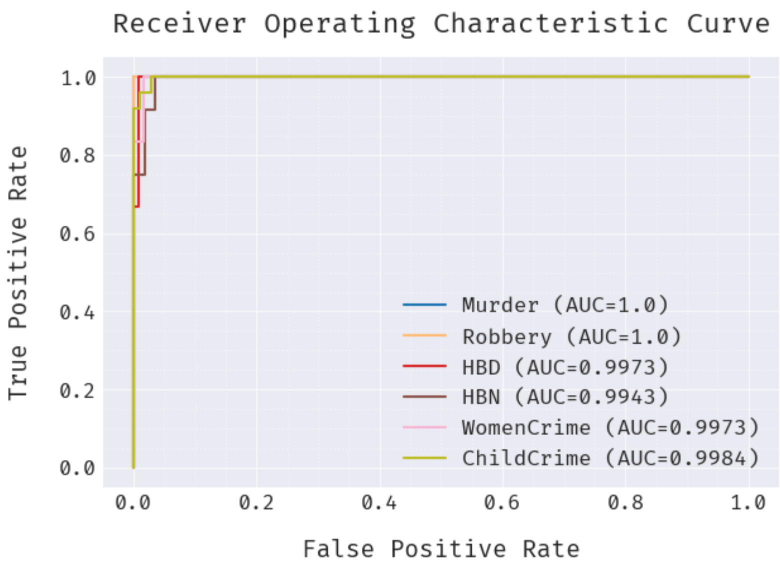


Fig.15 ROC Curve

This comparison is performed with 50 epoches. The performance of model is measured by accuracy, precision, sensitivity(recall), specificity, F1 score and MCC. Autoencoder yields better performance compared to 1D CNN and GRU.

Table 1 Performance Comparison of Different Deep Learning Techniques

	1D CNN	GRU	AutoEncoder
Accuracy	0.9609	0.9609	**0.974**
Precision	0.9323	**0.9379**	0.9288
Sensitivity	0.6528	0.6528	**0.7361**
Specificity	0.9728	0.9728	**0.9838**
F-Score	0.6842	0.6842	**0.7529**
MCC	0.7129	0.7129	**0.776**

The 1D CNN, GRU and AutoEncoder Techniques are compared and shown the performance in percentage. Except Precision, AutoEncoder produced good results as shown in Fig. 16.

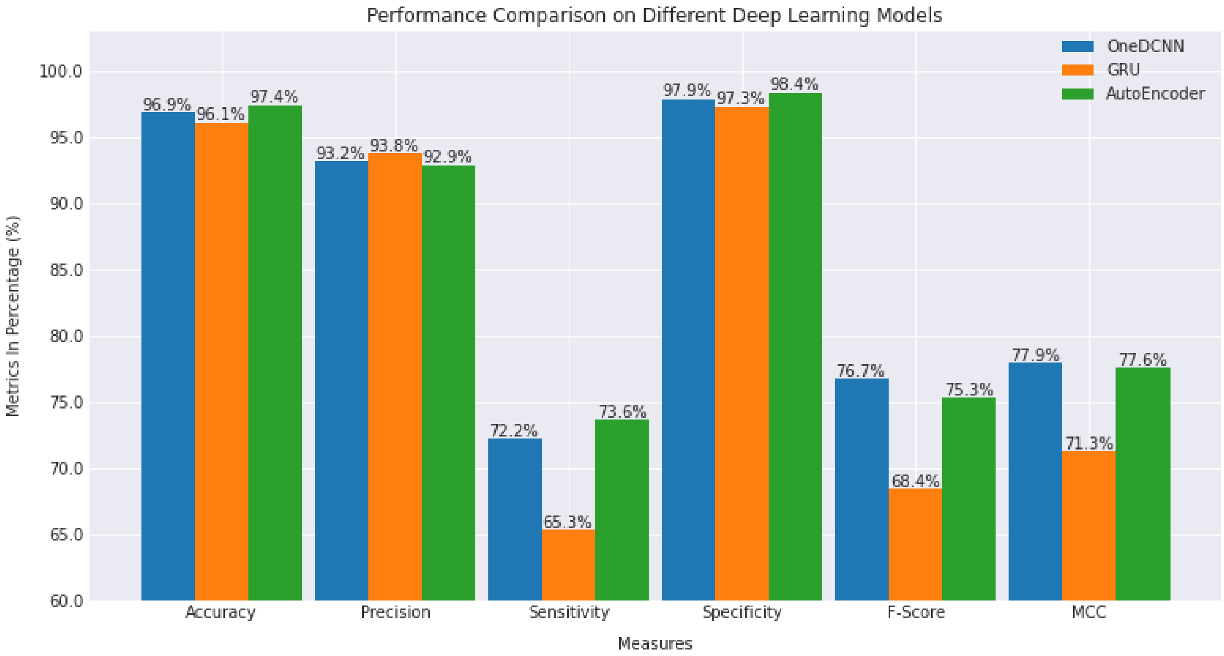


Fig 16. Performance Comparision on Different Deep Learning Methods

With the help of metrics, we consider Autoencoder is a best classifier for my 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.

5 CONCLUSIONS AND FUTURE WORK

Deep learning techniques are developed for crime classification in efficient manner. This paper tests the performance of 1DCNN, GRU and Autoencoder techniques on the DCRB 'Crimes - 2018 to May 2022 present' dataset. In this work, Autoencoder is considered as best techniques in crime classification. The Autoencoder shows high specificity 0.9838 when compared with 1DCNN and GRU.
The proposed method achieves better performance in crime classification. But, there is some misclassification in our techniques due to lack of information in our dataset. Hence, this proposed method will be improved by using advanced techniques and improving crime data set. Furthermore, our present crime classification method is unable to provide information regarding family background of accused and reason for committing offense.

REFERENCES

1. Tahani Almanie, Rsha Mirza, Elizabeth Lor,"Crime Prediction Based on Crime Types and Using Spatial
2. and Temporal Criminal Hotspots", International Journal of Data Mining & Knowledge Management
3. Process (IJDKP), July 2015, Boulder,USA.
4. B. Cavadas and P. Branco, "Crime Prediction Using Regression and Resources Optimization Crime Prediction Using Regression and Resources Optimization," no. September 2015, 2017.
5. Y. Lin, T. Chen and L. Yu, "Using Machine Learning to Assist Crime Prevention," 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 2017, pp. 1029-1030, doi: 10.1109/IIAI-AAI.2017.46.
6. S. Kim, P. Joshi, P. S. Kalsi and P. Taheri, "Crime Analysis Through Machine Learning," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2018, pp. 415-420, doi: 10.1109/IEMCON.2018.8614828.
7. Soon Ae Chun, Venkata Avinash Paturu, Shengcheng Yuan, Rohit Pathak, Vijay Atluri and Nabil R. Adam. 2019. Crime Prediction Model using Deep Neural Networks. In Proceedings of dg.o 2019: 20th Annual International Conference on Digital Government Research (dg.o 2019), June 18-20, 2019, Dubai, United Arab Emirates. ACM, New York, NY, USA, https://doi.org/10.1145/3325112.3328221.
8. C. Rajapakshe, S. Balasooriya, H. Dayarathna, N. Ranaweera, N. Walgampaya and N. Pemadasa, "Using CNNs RNNs and Machine Learning Algorithms for Real-time Crime Prediction," 2019 International Conference on Advancements in Computing (ICAC), 2019, pp. 310-316, doi: 10.1109/ICAC49085.2019.9103425.

9. X. Zhang, L. Liu, L. Xiao and J. Ji, "Comparison of Machine Learning Algorithms for Predicting Crime Hotspots," in IEEE Access, vol. 8, pp. 181302-181310, 2020, doi: 10.1109/ACCESS.2020.3028420.
10. X. Han, X. Hu, H. Wu, B. Shen and J. Wu, "Risk Prediction of Theft Crimes in Urban Communities: An Integrated Model of LSTM and ST-GCN," in IEEE Access, vol. 8, pp. 217222-217230, 2020, doi: 10.1109/ACCESS.2020.3041924.
11. VO. Llaha, "Crime Analysis and Prediction using Machine Learning," 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), 2020, pp. 496-501, doi: 10.23919/MIPRO48935.2020.9245120.
12. A. Shukla, A. Katal, S. Raghuvanshi and S. Sharma, "Criminal Combat: Crime Analysis and Prediction Using Machine Learning," 2021 International Conference on Intelligent Technologies (CONIT), 2021, pp. 1-5, doi: 10.1109/CONIT51480.2021.9498397.
13. 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.
14. X. Zhang, L. Liu, L. Xiao and J. Ji, "Comparison of Machine Learning Algorithms for Predicting Crime Hotspots," in IEEE Access, vol. 8, pp. 181302-181310, 2020, doi: 10.1109/ACCESS.2020.3028420.
15. W. Safat, S. Asghar and S. A. Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques," in IEEE Access, vol. 9, pp. 70080-70094, 2021, doi: 10.1109/ACCESS.2021.3078117.
16. Muthamizharasan, M., Ponnusamy, R. (2022). Forecasting Crime Event Rate with a CNN-LSTM Model. In: Raj, J.S., Kamel, K., Lafata, P. (eds) Innovative Data Communication Technologies and Application. Lecture Notes on Data Engineering and Communications Technologies, vol 96. Springer, Singapore. https://doi.org/10.1007/978-981-16-7167-8_33.
17. P. Kaur, G. Rani, T. Sharma and A. Sharma, "A Comparative Study to analyze crime threats using data mining and machine learning approach," 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), 2021, pp. 1-4, doi: 10.1109/ICSCAN53069.2021.9526489.
18. S. S. Kshatri, D. Singh, B. Narain, S. Bhatia, M. T. Quasim and G. R. Sinha, "An Empirical Analysis of Machine Learning Algorithms for Crime Prediction Using Stacked Generalization: An Ensemble Approach," in IEEE Access, vol. 9, pp. 67488-67500, 2021, doi: 10.1109/ACCESS.2021.3075140.
19. M. -S. Baek, W. Park, J. Park, K. -H. Jang and Y. -T. Lee, "Smart Policing Technique With Crime Type and Risk Score Prediction Based on Machine Learning for Early Awareness of Risk Situation," in IEEE Access, vol. 9, pp. 131906-131915, 2021, doi: 10.1109/ACCESS.2021.3112682.
20. Patel H, Prajapati D, Mahida D, Shah M (2020) Transforming petroleum downstream sector through big data: a holistic review. J Pet Explor Prod Technol 10(6):2601–2611. https://doi.org/10.1007/s13202-020-00889-2.
21. Shah, N., Bhagat, N. & Shah, M. Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention. Vis. Comput. Ind. Biomed. Art 4, 9 (2021). https://doi.org/10.1186/s42492-021-00075-z.
22. 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): 1.
23. Wang, Bao, et al. "Deep learning for real-time crime forecasting and its ternarization." Chinese Annals of Mathematics, Series B 40.6 (2019): 949-966.
24. Sangeetha and Prabha, "Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM", Journal of Ambient Intelligent Human Computing, 2020.
25. Kang H-W, Kang H-B (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.
26. Bates, Adam. "Stingray: A new frontier in police surveillance." *Cato Institute Policy Analysis* 809 (2017).
27. Chen and Jiang, "Fully-connected LSTM–CRF on medical concept extraction", International Journal of Machine Learning & Cyber, Vol. 11, pp. 1971–1979, 2020.
28. Muzammil Khan, Azmat Ali, Yasser Alharbi, and Gonzalo Farias. 2022. Predicting and Preventing Crime: A Crime Prediction Model Using San Francisco Crime Data by Classification Techniques. Complex. 2022 (2022). https://doi.org/10.1155/2022/4830411.
29. Cortes, Ana & Fuentes-Silva, Carlos. (2020). Artificial Intelligence Models for Crime Prediction in Urban Spaces. Machine Learning and Applications: An International Journal. 8. 1-13. 10.5121/mlaij.2021.8101.