WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Classification and investigation of cyber-crime Based on Deep Learning technique

***Suvarna Unde*[1]*, Prof. Avhad G.T.*[2]**

[1]Student, Vishwabharati college of Engineering, Ahmednagar Maharashtra,India
[2]Student, Vishwabharati college of Engineering, Ahmednagar Maharashtra,India

ABSTRACT:-

In real world activities in the field of cybercrime are one of the significant problems of both private and public organizations. Building real-time control systems to prevent digital criminal activities is challenging but very useful in early and effective detection of cyber-attacks. Machine learning (ML) is a part of data science where these problems could be solved by automated processes. In cyber security, various algorithms in ML can be used to prevent attacks and minimize security risks. In the paper, how ML algorithms can be applied in detection of cybercrime activities, the mathematical background. Rapid improvements in computer systems and networks have provided a new avenue for unethical activities such as cybercrime, public safety, security and the global economy, all of which are under the threat of cybercrime. The aim of the proposed work is to classify cybercrimes into cyberbullying and IP fraud analysis. The main objectives of this thesis are through cybercrime actions to provide a brief overview of cybercrime activities, their relative foundations and offer a matching scheme.

Keywords – Machine Learning,Deep Learning,Cybersecurity,Cyber-crime,Decision Trees Prediction, Data Mining.

1. INTRODUCTION

Cybercrime involves the attempt to commit a crime through a computer and network where the computer can act as an instrument, a target, or both. Much unauthorized computer activity takes place through global electronic networks.Deep learning has emerged as the best way to derive knowledge from data with greater meaning and accuracy. Applications of Deep Neural Networks in various fields have made them an important area of research. Crime analysis is the study of the characteristics of crime and their relationships. The huge volume of data sets related to crime and various different types of crime and their different characteristics and complex relationship between them make Deep Neural Networks an ideal choice for this domain. The knowledge gained from this analysis will allow police officers to process information quickly and accurately**.**

2. LITERATURE SURVEY

Achini Adikari, Daswin De Silva, Damminda Alahakoon, and Xinghuo Yu's paper Suspicious Human Activity Recognition: a Review explored all the areas where a visual-based detection system can be used[1].

Betim Cico and Eralda Nishani's paper "Computer Vision Approaches based on Deep Learning and Neural Networks" Deep Natural Networks for Video Analysis of Human Pose Estimation explores the implementation of neural networks, specifically CNN, for HAR from the analysis of videos. Neural networks are a part of deep learning, adapted from the concept of the human nervous system in the way that they send signals in the same way as human neurons do.[2].

Khraisat et al. [3] proposed SIDS or knowledge-based detection or misuse detection that relies upon pattern matching approaches for determining a known attack.

Baole Ai, Yu Zhou, Yao Yu, and Sidan Du's paper Human Pose Estimation using Deep Structure Guided Learning shows more about the advantage of using CNN for human pose estimation. Human activity recognition is the process of classification of sequences of accelerometer data recorded by devices into well-defined movements. Convolutional Neural Networks, or CNNs, were initially developed for problems involving image classifications[4].

The study Deylami et al. proposes a cybercrime detection model that uses support vector machines to classify social network (Facebook) datasets (SVMs). The three types of classification algorithms SVMs, AdaBoostM1, and NaiveBayes were utilized to find a high percentage of classification accuracy [5].

Data mining resembled a basic area of computer science significant for identifying patterns from voluminous data. The data mining approach essentially means extracting knowledge data from databases and identifying it important relationships between voluminous data and thus enable detection of anomalous behavior. Data mining techniques such as clustering, classification, association rule, and mining are of utmost importance in IDS and are used to evaluate and observe network data and provide intrusion-related information. Natural Language Processing is defined as the processing of language by computer. NLP is mostly associated with speech recognition, natural language generation and translation. Though the interest in NLP began as early as 1950 with Turing publishing his article on Intelligence [6], the field witnessed a huge growth once machine learning algorithms, both supervised and unsupervised were introduced for language processing.

The authors identify and fix data gaps, such as the necessity to gather benign hands, which may impair the accuracy of the deployed knife threat detector. This study offers a thorough review of image-based warnings that may be used to prioritize and educate crime prevention strategies before any catastrophic results occur. Additional relevant study topics in this subject include, among others, Automated Handgun Detection Alarms in Videos Using Deep Learning, Automatic Visual Recognition of Armed Robbery, and Robust Item Detector Application for Visual Knife Detection[7].

Hamid Zolfi, Hamidreza Ghorbani, M. Hossein Ahmadzadegan* In this paper, how to implement a method for classification of cybercrimes in cyberattacks was discussed. The dataset containing information about cyberattacks in the petrochemical company was explained briefly. [8]

3. CRIME PREDICTION USING DEEP LEARNING TECHNIQUES

3.1.1 Deep Learning based Regression Methods for Crime Prediction

Deep learning has become a popular method for crime prediction in recent years. deep learning algorithms, such as convolution neural networks (CNN), deep neural networks, and sentiment analysis, to analyze various types of data, including text, images, audio, and social media. These algorithms are capable of detecting patterns and anomalies in the data that can indicate criminal activity. One of the key strengths of deep learning is its ability to handle large and complex datasets, making it well-suited to the task of crime prediction. Deep learning algorithms in regression analysis are used as a tool for crime prediction to identify the factors most strongly associated with crime and use these relationships to make predictions about future crime patterns. The research articles in this area highlight the strengths of regression in modeling the relationship between multiple variables, including crime data, weather data, demographic data, social media data, and location data The regression model is part of an unsupervised domain adaptation technique designed to predict the likelihood of crimes in new cities. The authors claim that using regression in combination with the unsupervised domain adaptation technique results in improved accuracy in crime prediction. The regression model adopts two deep learning models, a Long Short-Term Memory (LSTM) network and a Spatio-Temporal Graph Convolutional Network (ST-GCN), to predict the likelihood of theft crimes in urban communities. It is not just the attacker who can produce delay and chaff perturbations rather such kind of perturbations can be introduced by the network too. There are chances that the packets face propagation delay when they traverse through the network. In an unknown network, it is quite possible that the attack connection is caught at some place along with several other connections, unable to distinguish it from the attack flow.

Methodology	**DL Algorithm**
A Graph Convolution network in combination with ST-ResNet is used to perform spatiotemporal analysis and then LSTM is used to detect crimes in each community. GBDT is used to combine outputs of GCN and LSTM.	**Long Short Term Memory (LSTM) in combination with Spatio Temporal Graph Convolution Network (ST-GCN) and Gradient.**
Developed Attention-LSTM to process categorical temporal data andStacked Bidirectional LSTM model to process spatial information. The two were fused using feature anddecision-level fusion.	**ATTN-LSTM, St-Bi-LSTM, Fusion Models**
LA Crime data is separated into block-wise information based on the hour of the day, area, and city. These blocks are used to train CNN model.	**Mixed Spatio Temporal Neural Network based on CNN**
A model framework based on three phases i.e., intercity similar-grid matching, auxiliary features construction, and crime risk prediction using a dense CNN-based unsupervised domain adaptation	**CNN based on unsupervised domain adaptation model (UDAC)**

Table 3.1: Crime Detection Using Deep Learning Regression Techniques.

3.1.2 Deep Learning based classification methods for crime prediction

Deep learning models can accurately organize criminal activity and detect criminal intent by analysing vast amounts of data, including images, audio, text, and social media data. For example, image-based data can provide detailed information about crime scenes, including the presence of weapons and other objects that may indicate criminal intent. Similarly, audio-based data can provide valuable insights into the tone and context of a conversation, helping to identify potential illegal activities. Another advantage of deep learning for classification problems in crime detection is the ability to identify hidden patterns in the data that traditional methods may miss. For example, deep neural networks can be trained to analyze crime related tweets, uncovering patterns that indicate a potentially criminal act. The results of deep learning Models In Crime Detection Have Been Awe-Inspiring.

Methodology	**DL Algorithm**
Classify object or person in video feed and track abnormal activity using Deep CNN (DCNN) and Recurrent Neural Network (RNN).	**DCNN AND RNN**
Extract tweet data and process text to identify text keywords that are relevant to weapons used for crime or criminal activity.	**CNN, KNN, NB, Support Vector Machine (SVM), DT AND RF**
Audio data is used to extract Mel-Frequency Cepstrum Coeficients from sound waves and a CNN-RNN classier is deployed and text data, features are extracted for BERT. Both models are combined by fusion model.	**CNN AND BERT**
Video data is used to extract spatio temporal features and gestures to train DRNN for classification (Hostility and Violence).	**DEEP REINFORCEMENT NEURAL NETWORK**
Deep Learning and Machine learning models are applied to predict dierent types of crime and their relation to weather in Newyork city	**SVM, LR, DT, RF, CNN, RNN, LSTM and GRU**
Twitter data is processed by keyword ltering and labeled. The features vectors are generated to be inputted into models for training.	**SVM and ANN**
Live CCTV data is used to detect faces and weapons using CNN- GRU model.	**RNN, Gradient Recurrent Unit (GRU) and LSTM.**
Preprocess crime data to extract features and train both traditional and Deep Learning models to predict crime region and type.	**Traditional ML classifiers and Artificial Neural Network (ANN)**

Table.3.2 Crime detection using deep learning classification techniques

4. SYSTEM ARCHITECTURE

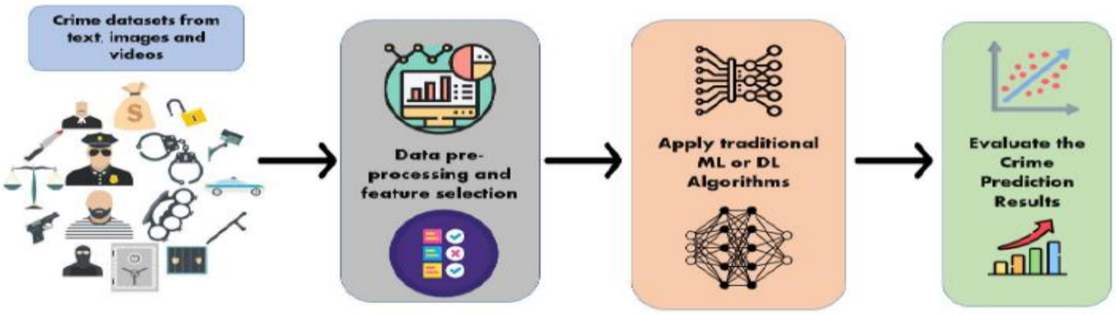
Crime datasets from text, images and videos
Data pre-processing and feature selection
Apply traditional ML or DL Algorithms
Evaluate the Crime Prediction Results

Fig. 4.1 Architecture flow of crime prediction.

Random Forest (RF) models can be used to analyze a wide range of features and make predictions about crime patterns. In addition to these techniques, traditional machine learning models can also be used for anomaly detection and outlier analysis in crime data. By identifying unusual patterns or outliers in the data, law enforcement agencies can detect potential criminal activity and take action to prevent it. Random Forest (RF) models can be used to analyze a wide range of features and make predictions about crime patterns. In addition to these techniques, traditional machine learning models can also be used for anomaly detection and outlier. analysis in crime data. By identifying unusual patterns or outliers in the data, law enforcement agencies can detect potential criminal activity and take action to prevent it.The latest research on using machine learning model-based regression and classification for crime prediction.

5. Conclusion

In this paper explain how to implement a method for classification of cybercrimes in cyber-attacks was discussed. The dataset containing information about cyber-attacks in the petrochemical company was explained briefly. Then, preprocessing and normalization were discussed and implemented. For implementation, Rapid Miner which was described in the previous section s was used. For the implementation of the algorithms, Support Vector Machine, Na¨ve Bayes, Decision Tree, and logistic regression were used, respectively. All of these algorithms were implemented and the obtained results were given in different tables. The best classification model was Support Vector Machine with a precision of 99precision for classification of cybercrimes in cyber-attacks.

REFERENCES

[1] Hamid Zolfi, Hamidreza Ghorbani, M. Hossein Ahmadzadegan* , "Investigation and classification of cyber-crimes through IDS and SVM algorithm IEEE Xplore Part Number:CFP19OSVART; ISBN:978-1-7281-4365-1

[2] Shah, N., Bhagat, N. Shah, M. Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention. Visual Computing for Industry, Biomedicine, and Art 4, 114 (2021) .

[3] Sergio Gomez and Francesc Moll. "Lithography aware regular cell design based on a predictive technology model." J. Low Power Electronics, 6(4):1–14, 2010

[4] Chun, S. A. et al. Crime prediction model using deep neural networks, 512514 (2019).

[5] Kshatri, S. S. et al. An empirical analysis of machine learning algorithms for crime prediction using stacked generalization: An ensemble approach. IEEEAccess 9, 6748867500 (2021)

[6] Agarwal, S., Yadav, L. Thakur, M. K. Crime prediction based on statistical models, 13 (IEEE, 2018). Bandekar, S. R. and Vijayalakshmi, C. Design and analysis of machine learning algorithms for the reduction of crime rates in india. Procedia Computer Science 172, 122127 (2020) .

[7] Sivanagaleela, B. and Rajesh, S. Crime analysis and prediction using fuzzy c-means algorithm, 595599 (IEEE, 2019).

[8] Shermila, A. M., Bellarmine, A. B. and Santiago, N. Crime data analysis and prediction of perpetrator identity using machine learning approach,107114 (IEEE, 2018).

[9] Catlett, C., Cesario, E., Talia, D. and Vinci, A. Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments. Pervasive and Mobile Computing 53, 6274 (2019) .

[10] Yi, F., Yu, Z., Zhuang, F., Zhang, X. and Xiong, H. An integrated model for crime prediction using temporal and spatial factors, 13861391 (IEEE, 2018).

[11] Dash, S. K., Safro, I. and Srinivasamurthy, R. S. Spatio-temporal prediction of crimes using network analytic approach, 19121917 (IEEE, 2018).

[12] Tasnim, N., Imam, I. T. and Hashem, M. A novel multi-module approach to predict crime based on multivariate spatio-temporal data using attention and sequential fusion model. IEEE Access 10, 4800948030 (2022) .

[13] Araujo, A., Cacho, N., Bezerra, L., Vieira, C. and Borges, J. Towards a crime hotspot detection framework for patrol planning, 12561263 (IEEE, 2018).

[14] Almuhanna, A. A., Alrehili, M. M., Alsubhi, S. H.and Syed, L. Prediction of crime in neighbourhoods of new york city using spatial data analysis, 2330 (IEEE, 2021).

[15] Algefes, A., Aldossari, N., Masmoudi, F. and Kariri, E. A text-mining approach for crime tweets in saudi arabia: from analysis to prediction, 109114 (IEEE, 2022).