TCP/IP, UDP, ICMP (IP Fragmenting) Cyber-Attacks Detection Using Machine Learning Algorithms with Jupiter Anaconda Navigator Simulation Tool.

Ch. Kodanda Ramu[1] Dr. T. Srinivasa Rao[2]

[1]Research Scholar, Department of CSE, GITAM (Deemed to be University),,izag, Andhra Pradesh, India.

[2]Professor, Department of CSE, GITAM (Deemed to be University),Vizag, Andhra Pradesh, India.

Abstract— In present scenario various types of cyber-attacks create lot of problems in between data sending from one place to another place. Due to various kind of cyber-attacks will generate traffic congestion problem in server as well as network.[1] The use of internet communication is rapidly increase at now for more web application sending everywhere.

The web application will execute only with the help of internet communication, so that lot of categories of cyber-attacks generate in recent time. The present paper presents various kind of cyber-attacks detection like TCP/IP attacks, UDP attacks, CMP attacks, using machine learning algorithms like CNN algorithms with the help of Jupiter anaconda navigator simulation tool and generate result of algorithms with accuracy graphs and tables. [4]

Index Terms— Machine learning, Cyber-attacks, Jupiter, Data set, Deep learning, Fault, Network, Server node, etc.

I INTRODUCTION

AI is a part of man-made reasoning (AI) [1] and software engineering which centers around the utilization of information and calculations to mimic the way that people learn, continuously working on its precision. IBM has a rich history with AI. There are four kinds of AI calculations: directed, semi-managed, solo and support. [2]

AI is a field of request gave to understanding and building
techniques that 'realize', that is, strategies that influence information to further develop execution on some arrangement of errands. It is viewed as a piece of man-made brainpower [3].

Machine learning is the methods of data analysis that automates analytical data building this technique also perform analytical study on crops identification using input parameters.

It is the Brach of artificial intelligence (AI) that system can learn data interpretation with less human intervention. ML is a compilation of algorithms that are further sub classified as supervised, unsupervised, and reinforcement learning. [3]

Deep learning (DL):

DL is a sub branch of ML which is based on a Neural network (NN). DL also has a subsection as supervised, unsupervised, and reinforcement learning. Convolutional NN (CNN), deep belief network (DBN), recurrent NN (RNN) are some popular supervised DL based structures. Stacked auto encoder (SAE) belongs to the unsupervised DL technique which is used for dimensionality reduction. [6]

RNN (Recurrent Neural Network):

Is the part of artificial neural network commonly used in speech recognition and NLP. This algorithm also performs data analysis on agricultural data for crops identification, crops growth development as better compare to ML techniques. [4]

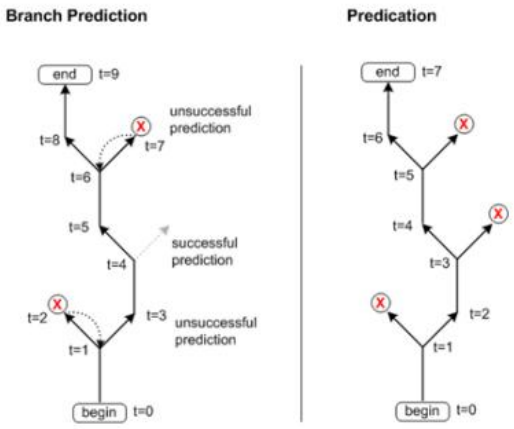

FIG 1.0 prediction task representation:

PREDICTON TECHNQUES: [5][6]

There are two types of predictive models. They are Classification models, that predict class membership, and Regression models that predict a number. These models are then made up of algorithms. The algorithms perform the data mining and statistical analysis, determining trends and patterns in data. Predictive modeling is a mathematical process used to predict future events or outcomes by analyzing patterns in a given set of input data. It is a crucial component of predictive

analytics, a type of data analytics which uses current and historical data to forecast activity, behavior and trends. [6]

MACHNE LEARNING TECHNQUES:

A regular significance of machine information is: "A PC program is said to acquire in actuality E in regards to a couple of class of endeavors T and execution measure P in case its display at tasks in T, as assessed by P, improves with experience E." Basically, machine keenness is the limit of a PC to acquire actually. Experience is ordinarily given as data. Looking at this data, the PC can find conditions in the data that are unreasonably convoluted for a human to shape. Machine information can be used to reveal a mystery class structure in an unstructured data, or it will in general be used to find conditions in a coordinated data to make figures. Last choice is the crucial point of convergence of the execution [6].

The study of the composing uncovered the LA challenges about data following, data variety, and data Analysis, a relationship with acumen sciences, knowledge environment improvement, emerging development, and moral concerns concerning authentic and security issues. [7]

Particular challenges in like manner exist from the assimilation of the data Analysis because of the show association of the data. Wrong data can incline the revelations causing a mistake of the overall people. Such circumstances are common in the electronic astuteness environment. For example, an educator could make a student profile to keep an undertaking that requires checking on, test the effortlessness of convenience process, or to conclude whether there are any openings in the presentation of the instructive arrangement as it appears for students. Development of a non-existent understudy presents dull information that appears in the course without recognizing evidence. This data doesn't address student information yet rather trickery made by the teacher that streams into the huge data pool of information. While truly coordinating data Analysis, this information can be actually recognized from the general population. In any case, working with data grouping from the keenness Analysis vantage direct adds an enormous space for error toward the consequence of in everyday results. Insight sciences affiliation. [6] According to Pea, tweaked knowledge and smarts significant entryways show a feebleness to utilize mind analy-fits preferably; subsequently, "the last stage is modified Cyber information at scale for everyone on the planet for any partner region". Expressed that to progress and totally grasp acumen requires understanding how partner makes and how to assist with partner new development. Further, focus on ers ought to get a handle on the pieces of character, reputation, and impact. Focus on ers ought to find approaches to

communicating "awareness, metacognition, and informative technique" to help with additional creating information processes. With a more grounded relationship with information sciences, sharpness analy-fits can progress strong insight plan. Insight environment progression. Seen that as understudies develop the restrictions of the insight mgmt. structure into open or blended astuteness settings, focus on ers ought to find the issues looked by students and how to conclude achievement as indicated by the understudies' perspectives. This collaboration will obstruct a shift toward more testing datasets that could integrate flexible, biometric, and personality data. Other than the solitary insight part of information analy-fits, focus on ers are hoping to address another part known as amicable keenness analy-fits. In this particular circumstance, social cleverness analy-fits revolves around the joint exertion and association of understudies in a blended training environment, not just on individual information results. [8]

Emerging advancement [9]

The greatest limit of information analy-fits interfacing with astuteness requires continued and emerging advancement that at this point remains in the more energetic stages. This revelation presents a test as the advancement continues to make to stay steady with the improvement of schooling analy-fits. Further, to totally fathom the system and practice of teaching, more survey is required. Focus on focusing in on keenness analy-fits and showing strategy is still in the beginning stages. [10] [16]

Problem specification

The cyber using customers are spread although out the length and breadth of us of an even though particularly concentrated inside the city areas. Looking at the rampant use of credit score cards among the educated magnificence the researcher believes that the common use of cards wishes to be educated about the different sorts of credit cards available. This study proposes to look at the studies of the credit card users, the benefits that accrue to them as well as the troubles they face. The problems that the credit cards users stumble upon in its each day usage. The precautions to be taken via the card customers while the usage of the credit cards. The remedies and the redressed this are to be had to them in case of loss or theft of their cards. A cognizance is to be created among the credit card customers about the frauds that take region and the precautions and safety measures to be adopted.[11]

Research methodology

Inside Deep Learning, a Convolutional Neural Network or CNN is a kind of counterfeit brain organization, which is broadly utilized for picture/object acknowledgment and grouping. Profound Learning in this manner perceives objects in a picture by utilizing a CNN. fundamental distinction between convolutional brain organization (CNN) and that's what ordinary AI is, as opposed to utilizing hand-created highlights, for example, SIFT [17] and HOG, CNN can naturally gain highlights from information (pictures) and procure scores from the result of its CNN is a regulated kind of Deep learning, most ideal utilized in picture acknowledgment and PC vision.[14][16] All of the generated, consolidated and split dataset is stored in dataset folder and is open to use by any. The cleaned.csv file is KD99 generated dataset.

THE CYBER ATTACKS DETECTION PERFORM FOLLWING POINTS MENTION BELOW:

The ids system performs the following attacks like 1. DDOS, 2. R2L.3. L2R. 4. PROBE. 5. DDOS. 6.TCP/P. 7.UDP. 8.ICMP. 9.IP.[27]

The CNN algorithm performs the supervised clustering algorithm the detection of cyber-attacks system using JUPYETR tool. [19][20] The decision tree algorithm also supervised classification algorithm perform the following points. WEKA simulation tool detail of representation. Graph and plotted valued design for performing various operation.[32]

MACHINELEARNING SIMULATION TOOL- PYCHARM SIMULATION TOOL AND PYTHON LANGUAGE:

The simulation performs the prediction of cross scripting attacks which is also kind of cybercrime attacks accrued on web application domain. So, we are describing in this research work proposed the python language and PYCHARM simulation tool for count and prediction of cross scripting attacks [26][29][32]. This kind of both implementations of used in this research work. We are used to detection of two types of attacks.[25]

Cyber-attacks on network system.

Cross scripting attacks on web application.

These are performing by WEKA and PYCHARM simulation tool work done. For using implementation, the decision tree and J 48 algorithms are implemented on WEKA tool. The cross-scripting attacks detection and prevention done by PYCHARM simulation tool.[26]

All these types of attacks which is found of cyber law categories under act IT 2000.

SIMULATION AND RESULT ANALYSIS:

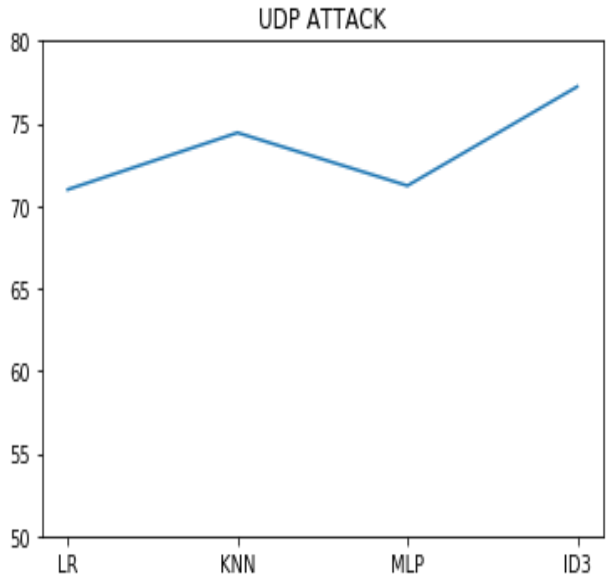

Fig2. UDP attacks detection by CNN algorithms using Jupiter simulation tool.

In this figure we have found one UDP attacks with the help of CNN algorithm and compare with another technique like LR and MLP. The UDP attacks determine on JUPYETR anaconda navigator simulator tool using CNN algorithm and determine attacks with the level of attacks ratio is also given in this presentation. [29]

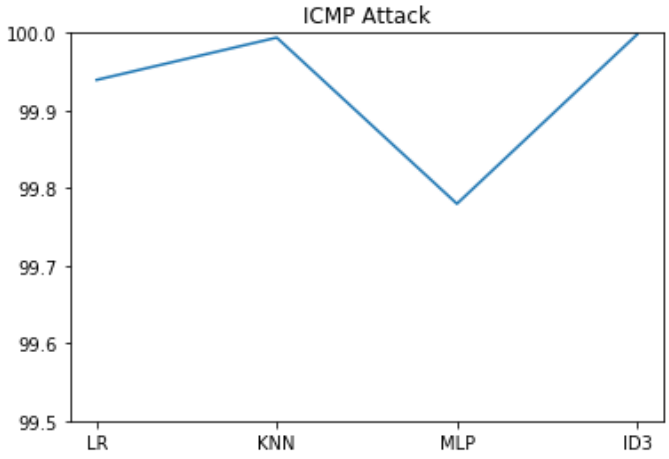

Fig3. ICMP IP attacks detection by CNN algorithms using Jupiter simulation tool.

In this graph representation of ICMP and IP fragmenting attacks detection using CNN algorithm on JUPYETR anaconda navigator simulation tool with compare to another algorithm like LR and

MLP along with intrusion detection ID3. The ratio of ICMP attacks determined is very high in network congestion system file. [30]

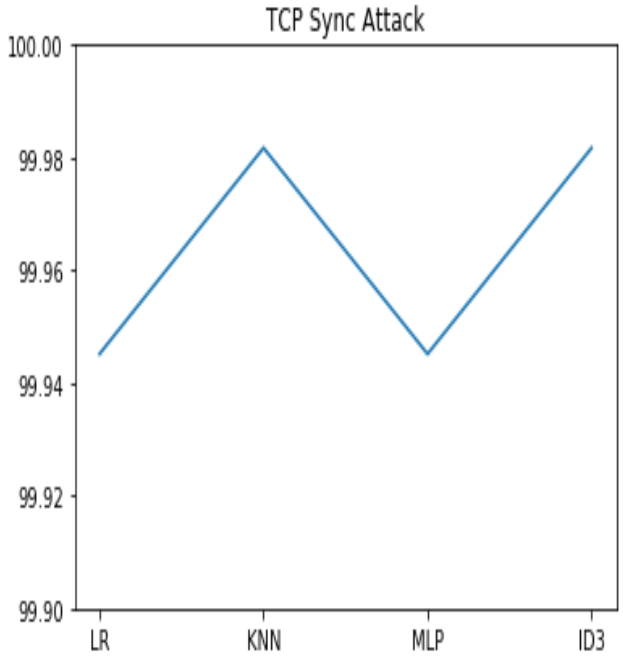

Fig4.TCP/IP attacks detection by CNN algorithms using Jupiter simulation tool.

In this graph representation of TCP/IP attacks detection using JUPYETR anaconda navigator simulation tool using

CNN algorithm with compare to another technique like LR, MLP, KNN and IID3. The ratio of TCP attacks determined is very high value shown.[32]

CONCLUSION

The result shows the plotted graph analysis on the basis of CNN algorithm implementation on KDD99 cup data set and get result in form of F1 score, ROC curve and confusion matrix. The accuracy of CNN algorithm is better than data mining algorithms like decision tree, J48, H tree, M-tree and random forest. [36] The result graph shows detection of various TCP/IP attacks, UPD attacks, CMP attacks detection using machine learning algorithms. [38]

FUTURE SCOPES

In the space of CYBER ATTACKS location, there are number of ways of distinguishing misrepresentation exchanges. Here we proposed one of them. In our proposed calculation we use

for oversampling in ongoing we can change testing calculation for over examining or under inspecting. In ongoing we can further develop AUC and Recall boundary of brain organization. Here we utilized ReLU actuation work, we can work on the consequence of it by utilizing other enactment work. And furthermore, we can further develop exactness by changing Hybrid 4 classifier.

REFERENCES

[1]. Ahmad Z, Khan AS, Shiang CW, Abdullah J, Ahmad F (2021) Network intrusion detection system: a systematic study of machine learning and deep learning approaches. Trans Emerg Telecommun Technol ,156:456.78.87.

[2] Assis MV, Carvalho LF, Lloret J, Proença ML (2021) A GRU deep learning system against attacks in software defined networks. J Netw Comput Appl 177:102942

[3] Manso, P.; Moura, J.; Serrao, C. SDN-Based Intrusion Detection System for Early Detection and Mitigation of DDoS Attacks. Information 2019, 10, 106.

[4] Aldweesh A, Derhab A, Emam AZ (2020) Deep learning approaches for anomaly-based intrusion detection systems: a survey, taxonomy, and open issues. Knowl Based Syst 189:2020

[5] Xu, R.; Cheng, J.; Wang, F.; Tang, X.; Xu, J. A DRDoS Detection and Defense Method Based on Deep Forest in the Big Data Environment. Symmetry 2019

[6] Anwar, S.; Mohamad Zain, J.; Zolkipli, M.; Inayat, Z.; Khan, S.; Anthony, B., Jr.; Chang, V. From intrusion detection to an intrusion response system: Fundamentals, requirements, and future directions. Algorithms 2017,

[7] A. R. Wani, Q. Rana, U. Saxena, and N. Pandey, "Analysis and Detection of DDoS Attacks on Cloud Computing Environment using Machine Learning Techniques," in 2019 Amity International Conference on Artificial Intelligence (AICAI), pp. 870-875, 2019.

[8] D. Peraković, M. Periša, I. Cvitić, and S. J. T. J. Husnjak, "Model for detection and classification of DDoS traffic based on artificial neural network," vol. 9, no. 1, p. 26, 2017.

[9] J. Bakker, B. Ng, W. K. Seah, and A. Pekar, "Traffic Classification with Machine Learning in a Live Network," in 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), pp. 488-493, 2019.

[10] Ye Zheng-Wang, “The Research of Intrusion Detection Algorithms Based on the Clustering of Information Entropy”, Elsevier, Procedia Environmental Sciences, Vol. 12, pp: 1329-1334, 2012.

[11] Z. Muda, W. Yassin, M. N. Sulaiman and N. I. Udzir, “A K-Means and Naïve Bayes Learning Approach for Better Intrusion Detection”, Information Technology Journal, Vol. 10 No. 3, pp: 648-655, 2011.

[12] Zachary Miller, William Deitrick, Wei Hu, “Anomalous Network Packet Detection Using Data Stream Mining”, Journal of Information Security, Vol. 2, pp: 158-168, 2011.

[13] Zainal, A. Maarof, M.A. ; Shamsuddin, S.M., “Feature Selection Using Rough Set in Intrusion Detection”, In the Proceedings of the IEEE Region 10 Conference TENCON, pp. 1-4, 2006.

[14] Li, Z.; Qin, Z.; Huang, K.; Yang, X.; Ye, S. Intrusion detection using convolutional neural networks for representation learning. In Proceedings of the International Conference on Neural Information Processing, Guangzhou, China, 14–18 November 2017; pp. 858–866.

[15] Chris Clifton and Gary Genko” Vol. XLVIII, No. 83 ROAD TOWN, TORTOLA TUESDAY 16 DECEMBER 2014.

[16] Halder, D., &Jaishankar, K. (2011) Cybercrime and the Victimization of Women: Laws, Rights, and Regulations. Hershey, PA, USA: IGI Global. ISBN 978-1-60960-830-9.

[17] Lambert, Glenn M. II, "Security Analytics: Using Deep Learning to Detect Cyber Attacks" (2017). UNF Graduate Theses and Dissertations. 728, https://digitalcommons.unf.edu/etd/728.

[18] ManjeetRege& Raymond Blanch K. Mbah, Machine Learning for Cyber Defense and Attack , DATA ANALYTICS 2018 : The Seventh International Conference on Data Analytics, Copyright (c) IARIA, 2018. ISBN: 978-1-61208-681-1 , pp.73–78.

[19] Dmitri Koteshov, How Can Ai Change The State Of Cybersecurity, March 7, 2018, https://www.elinext.com/industries/financial/trends/aiand-security/.

[20] Anti-Phishing Working Group, “Phishing and Fraud solutions”. [Online], [Accesses: March 18, 2019] http://www.antiphishing.org/.

[21] S. Abu-Nimeh, D. Nappa, X. Wang, and S. Nair, “A Comparison of Machine Learning Techniques for Phishing Detection”, APWG eCrime Researchers Summit, October 4-5, 2007, Pittsburg, PA.

[22] N. Lu, S. Mabu, T. Wang, and K. Hirasawa, “An Efficient Class Association Rule-Pruning Method for Unified Intrusion Detection System using Genetic Algorithm”, in IEEJ Transactions on Electrical and Electronic Engineering, Vol. 8, Issue 2, pp. 164 – 172, January 2, 2013.

[23] Knowledge Discovery and Data Mining group, “KDD cup 1999” [Online], [Accessed:March 18, 2019], http://www.kdd.org/kddcup/index.php.

[24] Denning D E, “An Intrusion-Detection Model,” In IEEE Transaction on Software Engineering, Vol. Se-13, No. 2, pp. 222-232, February 1987.

[25] Lee, W, Stolfo S and Mok K , “Adaptive Intrusion Detection: A Data Mining Approach,” In Artificial Intelligence Review, Kluwer Academic Publishers, 14(6), pp. 533 - 567, December 2000.

[26] Satinder Singh, Guljeet Kaur, “Unsupervised Anomaly Detection In Network Intrusion Detection Using Clusters,” Proceedings of National Conference on Challenges & Opportunities in Information Technology RIMT-IET, Mandi Gobindgarh. March 23, 2007.

[27] Eric Bloedorn , Alan D. Christiansen , William Hill , Clement Skorupka , Lisa M. Talbot , Jonathan Tivel, “Data Mining for Network Intrusion Detection: How to Get Started,” CiteSeer, 2001.

[28] L. Portnoy, “Intrusion Detection with Unlabeled Data Using Clustering,” Undergraduate Thesis, Columbia University, 2000.

[29] Theodoros Lappas and Konstantinos Pelechrinis, “Data Mining Techniques for (Network) Intrusion Detection Systems,” http://citeseerx.ist.psu.edu/viewdoc/ download? doi=10.1.1.120.2533&rep=rep1&type=pdf..

[30]. Dewan Md. Farid, Nouria Harbi, Suman Ahmmed, Md. Zahidur Rahman, and Chowdhury Mofizur Rahman, “Mining Network Data for Intrusion Detection through Naïve Bayesian with Clustering”, World Academy of Science, Engineering and Technology, 2010.

[31] X. Li and N. Ye., “A supervised clustering algorithm for computer intrusion detection,” Knowledge and Information Systems, 8, pp498-509, ISSN 0219-1377, 2005

[32] Kruegel C., Mutz D., Robertson W., Valeur F., “Bayesian event classification for intrusion detection,” In: Proceedings of the 19th Annual Computer Security Applications Conference; 2003.

[33] Portnoy L., Eskin E., Stolfo S.J., “Intrusion detection with unlabeled data using clustering,” In: Proceedings of The ACM Workshop on Data Mining Applied to Security; 2001

[34] Paxson V., “Bro: A System for Detecting Network Intruders in Real-Time”, Computer Networks, 31(23-24), pp. 2435-2463, 14 Dec. 1999.

[35] D.Barbara, J.Couto, S.Jajodia, and N.Wu, "ADAM: A test bed for exploring the use of data mining in intrusion detection”, SIGMOD, vol30, no.4, pp 15-24, 2001.

[36] Ahmad R, Alsmadi I (2021) Machine learning approaches to IoT security: a systematic literature review. Internet Things 14:100365.

[37] F. Provost, and T. Fawcett, “Robust classification for imprecise environment,” Machine Learning, vol. 42/3, 2001, pp. 203-231.

[38] KDD. KDD CUP. Available online: https://kdd.ics.uci.edu/databases/kddcup99/task.html (accessed on 17 March 2020).