iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VII **Month of publication:** July 2020

DOI: http://doi.org/10.22214/ijraset.2020.7011

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

A Review on Application of Deep Learning in Cyber Forensics

Sri Vishva. E[1]
Student, School of Computer Science and Engineering (SCOPE), VIT UNIVERSITY, Vellore, India[1]

Abstract: ***In recent times, most of the data such as books, personal materials and genetic information digitally. This transformation gives rise to a field of Cyberspace. This newly created space, gave rise to a new set of crime Cybercrime, this lead to the development of securing the cyberspace and protecting against cybercrime. As more people started using cyberspace, more number of cybercrime were registered. As the number of crime increases, we are in need of help from Machine Learning. Machine Learning in the field of Cyber forensics, is a boon. In this paper we have an overview of Machine Learning in the field of Cyber Forensics and various method of it implementation.***

I. INTRODUCTION

1) *Deep Learning and Machine Learning:* The field of Artificial Intelligence and Machine Learning has been around since a long time but it is now that we have enough computational power to effectively develop strong artificial neural networks (ANN) in a reasonable time frame with the help of strong hardware and software support. The most important aspect of Cyber security involves protecting key data and devices from cyber threats. It's an important part of corporations that collect and maintain large databases of client data, social platforms wherever personal information were submitted and also the government organizations wherever secret, political and defence information comes into measure. Unlike the traditional machine learning algorithm that uses feature engineering and illustration ways. They will chose the best options by themselves.
2) *Cyber Security:* Cyber security is that the set of framework and processes designed to protect computers, networks, programs, and data from attack, unauthorized access, change, or destruction. These frameworks are consist of network security and host security systems, every of those has a minimum firewall, antivirus computer code, associated an intrusion detection system.
3) *Deep Learning and Cyber Security:* This survey summarizes the association of cyber security and Deep learning techniques (DL). Deep learning technique are being used by researchers in recent days. Deep learning can be used alongside the prevailing automation ways like rule and heuristics based and machine learning techniques. This survey helps is understand the benefits of deep learning algorithms to classify and tackle malicious activities that perceived from the varied sources like DNS, email, URLs etc. In recent days, non-public firms and public establishments are dealing with constant and complicated cyber threats and cyberattacks. As a precaution, organizations should build and develop a cybersecurity culture and awareness so as to defend against cyber criminals.
4) *Shared Task:* In this shared task conference, the train data set will be distributed among the participants and the train model will be evaluated based on the test data set. This is most common in NLP area recently shared task on identifying phishing email has been organized by. The details of the submitted runs are available throughout. Followed by shared task on detecting malicious domain organized intruders. These two shared tasks enables the participants to share their approach through working notes or system description paper. Each year there is one more shared task conducted by CDMC. But they don't provide us an option to submit system description papers. But recently they are giving an option to submit system description papers (CDMC 2018). One significant issue was that the available data sets are very old and each data set has their own limitations. The main issue we face now is due to the non-maintenance of Cybercrime data. To overcome such issues a brief investigative issue made to understand the need of Security domain, datasets and key feature of data sciences is discussed in for problems employing the data science towards cyber security. The need for such dataset in to be promoted.

II. LITERATURE SURVEY

1) *Intrusion Detection:* The traditional machine learning approach (shallow models) for Intrusion Detection System primarily include the artificial neural network (ANN), support vector machine (SVM), K-nearest neighbour (KNN), naïve Bayes, logistic regression , decision tree, clustering, combined and hybrid methods. Some of these methods have been studied for several years, and their methodology is mature. Their focus not only limited to the detection effect but also on practical problems, e.g., detection efficiency and data management. Deep learning models contains of diverse deep neural networks. Among them, deep

©IJRASET: All Rights are Reserved

brief networks , deep neural networks , convolutional neural networks , and recurrent neural networks are supervised learning models, while autoencoders, restricted Boltzmann machines , and generative adversarial networks are unsupervised models. The various number of studies of deep learning-based IDSs has increased rapidly from 2015 to the present. Deep learning models directly interact with feature representations from the original data, such as images and texts, without requiring manual feature engineering. Hence, deep learning methods can be execute in an end-to-end manner. For huge datasets, deep learning methods have a huge advantage over shallow models. In the field of deep learning, the main emphases are network architecture, hyperparameter selection, and optimization strategy.

2) *Traffic Analysis:* By training an algorithm against history of production traffic data, the mathematical model gets a boost start in creating a data classification. Subsequently data can be classified with existing class, and new clusters can be derived that are more efficiently group the data. As models for differentiating traffic emerge, they can be tested against new data and compared against existing dataset for security rulesets for accuracy. Improvements that can be derived from the current model can be integrated into the framework, and the process can be repeated continuously.

3) *Malware Analysis:* Malware refers to malicious software attackers dispatch to infect individual computers or an entire organization's network. It exploits target user system vulnerabilities, such as a bug in legitimate software (e.g., a browser or web application plugin) that can be hijacked. A classical fully connected neural network and recurrent neural network (RNN) model of deep learning was traditionally employed to detect malware with 300 bytes information from the Portable Executables header file . Subsequently, we have employed convolutional neural network (CNN) on a large number of byte long executables and obtained consistent results across 2 different tests based on a previous study. Using domain level knowledge, we have extracted several features and showed that its performance is comparable to the MalConv deep learning approach. The performance metrics of MalConv model was improved by making modification to the existing architecture. I believe that the deep learning potential have not been fully realized, and the work proposes the application of Windows-Static Brain-Droid (WSBD) model for incorporating deep learning.

III. CONCLUSION

This Survey help us to conclude the need of automated algorithm against Cyberattacks. In comparison between Machine Learning Algorithms and Deep Learning Algorithm, Deep Learning Algorithm turns out to be the optimal for Cyber Security. The main set back we encounter is the unavailability of updated dataset. This review also urges to maintain a Cyberattack Dataset, for further usage.

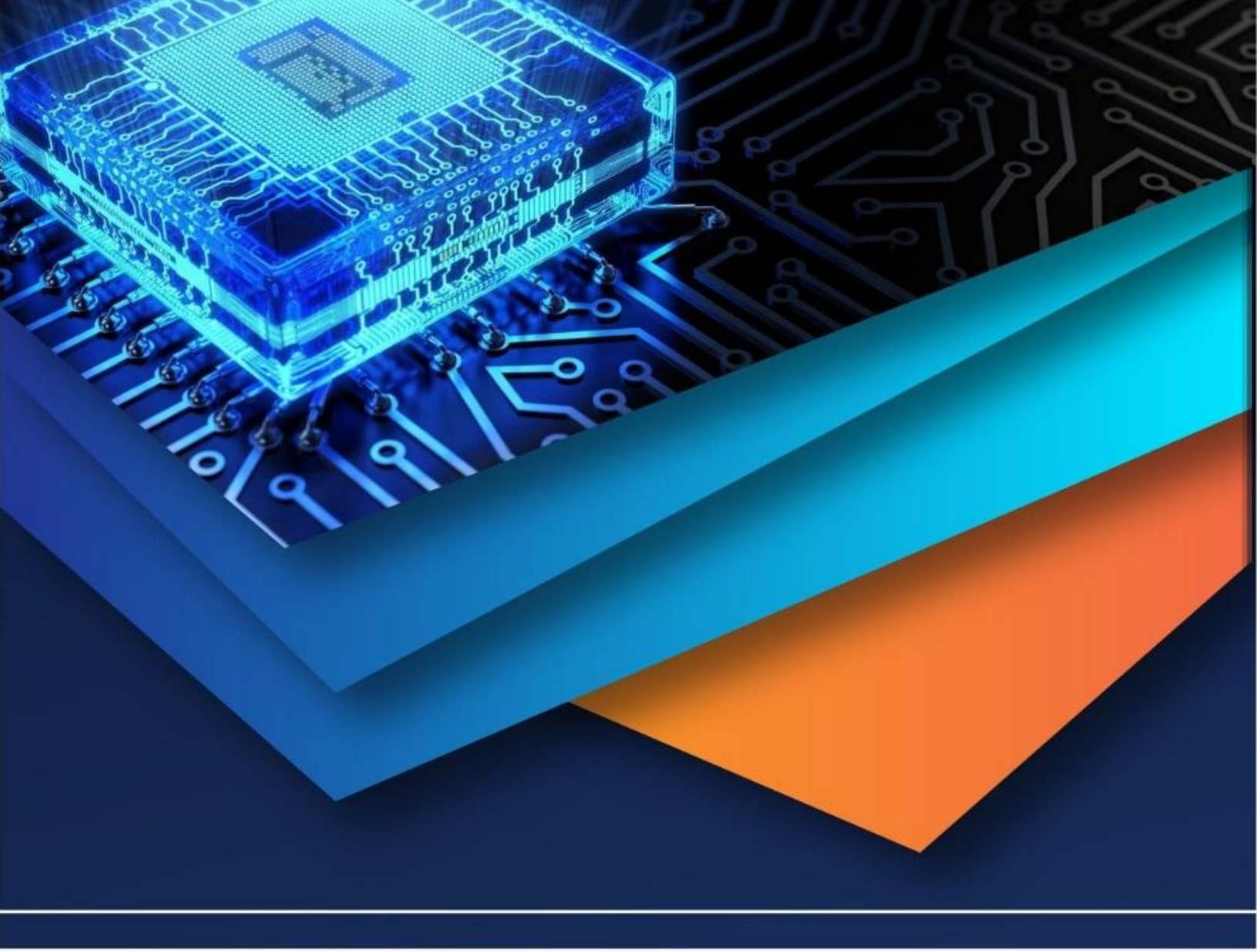
IV. ACKNOWLEDGMENT

I would like to take this opportunity for thanking my Guide Dr Aju.D, who has cultivated the knowledge on Cyberspace in me. I would also thank my Institution VIT-VELLORE , parents and friends for their continuous support.

REFERENCES

[1] Hemant Rathore, Swati Agarwal, Sanjay K. Sahay, Mohit Sewak, Malware Detection using Machine Learning and Deep Learning
[2] R. Devakunchari, Sourabh, Prakhar Malik, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075,Volume-8, Issue-7C2, May 2019, A Study of Cyber Security using Machine Learning Techniques
[3] R. Vinayakumar , Mamoun Alazab , K. P. Soman, Prabaharan Poornachandran, Sitalakshmi Venkatraman, Robust Intelligent Malware Detection Using Deep Learning
[4] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. Journal of machine learning research, 15(1), 1929-1958.
[5] Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International Conference on Machine Learning (pp. 448-456).
[6] Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2018). Detecting malicious domain names using deep learning approaches at scale. Journal of Intelligent & Fuzzy Systems, 34(3), 1355-1367.
[7] Vinayakumar, R., Poornachandran, P., & Soman, K. P. (2018). Scalable Framework for Cyber Threat Situational Awareness Based on Domain Name Systems Data Analysis. In Big Data in Engineering Applications (pp. 113-142). Springer, Singapore.
[8] Khalil, I. M., Khreishah, A., & Azeem, M. (2014). Cloud computing security: a survey. Computers, 3(1), 1-35.
[9] Karbab, E. B., Debbabi, M., Derhab, A., & Mouheb, D. (2017). Android Malware Detection using Deep Learning on API Method Sequences. arXiv preprint arXiv:1712.08996.
[10] Kapratwar, A. (2016). Static and Dynamic Analysis for Android Malware Detection.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)