



A Deep Learning Approach Against Botnet Attacks to **Reduce the Interference Problem of IoT**

www.ikas@gmail.com

💌 arushiagwl@gmail.com

(D) https://orcid.org/0000-0001-5603-5665 - Arushi Agarwal Engineer- (B. Tech., M. Tech.) (D) https://orcid.org/0000-0002-5024-1334

Scientist - (MCA, M. Tech., M. Phil., MBA)

– Vikas Dixit*



ARTICLE HISTORY

ABSTRACT Purpose: Today we are witnessing a world where hacking into a user's computer using

Paper Nomenclature: View Point (VP) Paper Code: GJEISV13I4OD2021VP1 Submission at Portal (www.gjeis.com): 01-Oct-2021 Manuscript Acknowledged: 03-Oct-2021 Originality Check: 06-Oct-2021 Originality Test (Plag) Ratio (Ouriginal): 34% Author Revert with Rectified Copy: 10-Oct-2021 Peer Reviewers Comment (Open): 28-Oct-2021 Single Blind Reviewers Explanation: 01-Nov- 2021 Double Blind Reviewers Interpretation: 13-Nov-2021 Triple Blind Reviewers Annotations: 30- Nov-2021 Authour Update (w.r.t. correction, suggestion & observation): 04-Dec-2021 Camera-Ready-Copy: 18-Dec-2021 Editorial Board Excerpt & Citation: 30-Dec-2021

Published Online First: 31-Dec-2021

botnet attacks on IoT devices, thus reducing the interference problem in the network. Methodology: In this paper, we worked on a deep learning-based botnet detection algorithm, which trains the IoT devices few ways for future research have been distinguished. To show the capacity of our created model to identify new varieties of botnets, a changed adaptation of the Torii sample will be utilized in the next phase of the work, to produce a second combined dataset and will be looked at against existing mark and stream-based oddity location strategies.

tiny bots or intercepting a group of interconnected devices is no more impossible. These tiny bots are called botnets which are a group of malicious codes that can

hamper the whole security system without the knowledge of the user. As Internet of

Things (IoT) is emerging rapidly, the interconnected devices are susceptible to breach, as one affected device can hamper the whole network. The security threat increases as

botnet attacks increase their presence to the interconnected devices. In this paper, we

are implementing Restricted Boltzmann Machine (RBM) algorithm of deep learning approach on the CTU-13 dataset (The CTU-13 is a dataset of botnet traffic that was

captured in the CTU University, Czech Republic, in 2011. The goal of the dataset was

to have a large capture of real botnet traffic mixed with normal traffic and background

traffic.) to train the algorithm about the botnet attack patterns in IoT and to prevent the

Conclusion: This paper proposes a solution to the detection of botnet activity within consumer IoT devices and networks. A novel detection algorithm was developed based on Deep learning mechanism. Earlier detection was performed at the packet level with Wireshark by creating a fake network using ApateDNS.

Paper Type: View Point

KEYWORDS Deep Learning | IOT | Interface | Attack | CTU

*Corresponding Author (Vikas Et. Al)

- Present Volume & Issue (Cvcle): Volume 13 | Issue-4 | Oct-Dec 2021 International Standard Serial Number:
- Online ISSN: 0975-1432 | Print ISSN: 0975-153X
- DOI (Crossref, USA) https://doi.org/10.18311/gjeis/2021 Bibliographic database: OCLC Number (WorldCat): 988732114
- Impact Factor: 3.57 (2019-2020) & 1.0 (2020-2021) [CiteFactor]
- Editor-in-Chief: Dr. Subodh Kesharwani
- Frequency: Quarterly

- Published Since: 2009
- Research database: EBSCO https://www.ebsco.com
- · Review Pedagogy: Single Blind Review/ Double Blind Review/ Triple Blind Review/ Open Review
- Copyright: ©2021 GJEIS and it's heirs
- Publishers: Scholastic Seed Inc. and KARAM Society
- · Place: New Delhi, India.
 - Repository (figshare): 704442/13

GJEIS is an Open access journal which access article under the Creative Commons. This CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0)promotes access and re-use of scientific and scholarly research and publishing.





Introduction

Today we are living in an era of IoT, robotics and artificial intelligence where sneaking into a remote system is not impossible anymore. Malwares has become so advanced today that they even do not need to write anything into the system files. The concept of dropping a malicious file into a target system is long gone now. Malwares are now able to use legitimate inbuilt applications to access the mainframe of the system. These advanced classes of malwares are called fileless malwares. The concept of botnet is somewhat like fileless malwares. Botnets do not need to drop a malicious file into the target system. Their main goal is to hamper as much devices in the network as possible without leaving any trace of the source. In IoT, where every device is connected, botnet attacks have proven themselves to be an alarming challenge. Until the IoT concerns improve their security mechanisms, botnet attacks have the potential to become the new weapon of future cyber-attacks. Since last decade, the statistics of the botnet attacks have increased rapidly. As much as technology is advancing the botnets are becoming strong and more intelligent.

Present technology of botnet detection mechanisms like signature or flow-based error detection are providing sound results when there is a small-scale botnet attack. In larger scale, they are not meeting the desired output. Now-a-days cyber world is witnessing the rise of artificial intelligencebased botnets which can learn the pattern of the users' usage mechanism and deploy themselves accordingly. Therefore, there is an unavoidable need of learning-based mechanisms to be implemented in the field of botnet attack prevention.

Here in this paper we are implementing a deep learning approach (Restricted Boltzmann Machine) to train the algorithm more accurately with a higher accuracy level about the attack patterns of the botnets. The dataset used in this work is the CTU-13 dataset of botnet attacks which shows several botnet attacks patterns and their feature-set and parameters.

This paper proposes a solution to the detection of botnet activity within consumer IoT devices and networks. A novel detection algorithm was developed based on Deep learning mechanism. Earlier detection was performed at the packet level with Wireshark by creating a fake network using ApateDNS.

Literature Review

Many researchers have been done work on providing an efficient solution against botnet attacks. Torres et. el. Proposed a methodology to compare several Recurrent Neural Network (RNN) models and the efficiency are analyzed through behavior of the traffic by designing it as a sequence of states that changes [1]. The focus of this work is to analyze the behavioral characteristics of the botnets. This work analyses the Recurrent

Neural Network (RNN) methodologies because of the two main reasons – unbalanced network traffic and length of the data sequence. To perform the proposed concept, a K-fold cross validation (a resampling procedure used to evaluate machine learning models on a limited data sample. The procedure has a single parameter called \mathbf{k} that refers to the number of groups that a given data sample is to be split into.) and test was conducted on unseen data extracted from different botnet sample. From the results, it was found that the Recurrent Neural Network (RNN) model is capable of efficiently differentiating the botnet attacks with a high detection rate but it is not as much efficient when using imbalanced network traffic.

The ongoing development of the Internet of Things (IoT) has brought about an ascent in IoT based DDoS assaults. McDermott et. el. exhibits an answer for the identification of botnet action inside purchaser IoT gadgets and systems [2]. A tale use of Deep Learning is utilized to build up an identification model dependent on a Bidirectional Long Short-Term Memory based Recurrent Neural Network (BLSTM-RNN). This paper shows that although the bidirectional methodology adds overhead to every age and expands preparing time, it demonstrates to be a superior dynamic model after some time. A named dataset was created as a component of this examination and is accessible upon solicitation.

Botnets comprise an essential risk to Internet security. The capacity to precisely recognize botnet traffic from nonbotnet traffic can help essentially in moderating pernicious botnets. Roosmalen et. el. presented a novel way to deal with botnet discovery that applies profound learning on streams of TCP/UDP/IP-parcels [4]. The test results with a huge dataset, they acquired 99.7% exactness for arranging P2Pbotnet traffic. This is practically identical to or superior to customary botnet recognition approaches, while lessening endeavors for highlight designing and highlight choice to a base.

As of late, botnets have turned out to be one of the significant dangers to data security since they have been continually developing in both size and advancement. A few botnet identifications measures, for example, honeynet-based and Intrusion Detection System (IDS) - based, have been proposed. Nonetheless, IDS-based arrangements that utilization marks appear to be insufficient on the grounds that ongoing botnets are outfitted with advanced code update and avoidance strategies. A few investigations have appeared unusual botnet recognition strategies are more successful than mark-based techniques since oddity-based botnet discovery strategies don't require pre-constructed botnet marks and henceforth they have the capacity to recognize new or obscure botnets. Toward this path, Hoang X. & Nguyen Q. proposed a botnet discovery model dependent on AI utilizing Domain Name Service question information and assesses its viability utilizing famous AI strategies [5]. Exploratory outcomes demonstrate that AI

54



calculations can be utilized successfully in botnet recognition and the irregular timberland calculation creates the best by and large identification precision of over 90%.

Problem Statement

Many technologies have been developed over the years to cope with the increasing threat of the cyber-security, but the ration of cyber-crime is increasing instead of decreasing. Heuristic-based tools use rules to examine suspicious codes and classify them as malware. This approach is limited, however, due to the fact that it relies on the sequence of repeated code that is indicative of malicious intent. Hence, *in this work, we are presenting a view on the combined approach of static and dynamic analyses with tools based on real-time extraction*. collapse as much IoT devices as possible and to gain access to the privileged information of a target system. The effects of botnets are increasing in a rapid manner which was anticipated in the early ages. Statistics shows that most of cyber-attacks somehow involve in botnet implementations.

Deep Learning Against Botnets

Now we are living in the world of IoT where almost every device is interconnected to ease our lives but now this IoT network is sometimes used against the security of the cyber-space. As we know, botnets target the most vulnerable devices and connections in an IoT network to spread into other devices, the question remains that how we are going

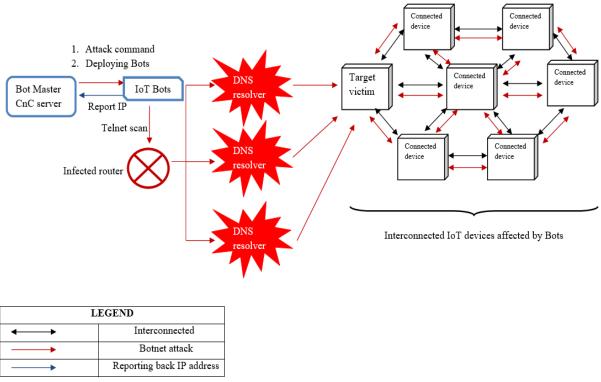


Fig. 1. Existing problem of botnet attack on IoT network

Botnets in Internet of Things

A botnet is a network of inter-connected small group of computers which runs on one or more bots. Their main target is to deploy botnet framework into the target system's mainframe to gain access to the system's administrative privileges. Once this step is done, they use their framework to manipulate the user's behavior and intervention to the system processes. For example, the Torii botnet attack (September 20, 2016) was one of the most severe botnet attacks which took over the IoT network. The "Botmaster" uses the C&C server gain access to the victim system and transfer the information via IP trafficking through the botnet's framework. Botnet is an abbreviation which was taken from "Robot" and "Network" (Bot + Net). The main motive behind a botnet attack is to to stop this attempt of cyber-security invasion. Malwares, ransomwares, botnets are becoming advanced and growing much faster than we can improve our defense against them. AI enabled malwares, ransomwares, botnets are not new today. This threatens cyber-security and raises the question that are we secure enough and ready to defend our-selves in case in future any more devastating security threats raises concern.

Deep learning mechanism (DL) is an advanced extension of the Machine learning mechanism (ML) which includes following main types – Boltzmann Machine (BM), Neural Network (NN), and Stacked Auto Encoders. In this work, we are comparing machine learning approaches with the



deep learning Neural Network Sequential algorithm to find out whether machine learning or deep learning does more accurate job in learning the attack patterns accurately.

The following figure shows the implementation of Restricted Boltzmann Machine (RBM) on the CTU-13 dataset in our secure lab environment by creating an imitated network using ApateDNS tool (*ApateDNS*TM *is a tool for controlling DNS responses though an easy-to-use GUI*) to restrict the botnet sample from gathering information about the virtual environment. ApateDNS starts a VPN (Virtual Private Network) inside the Virtual Machine to limit the botnet's access to the public DNS.

As we can see from the result, the accuracy of the algorithm is 64.88% when we are using epoch = 5. Epoch is the count of iterations that are implied on the dataset to train the algorithm in a recursive manner. The more the count of epoch, the better the accuracy level is.

Count of epoch ∞ Accuracy level

```
Successfully downloaded train-images-idx3-ubyte.gz 9912422 bytes.
Extracting /tmp/data/train-images-idx3-ubyte.gz
Successfully downloaded train-labels-idx1-ubyte.gz 28881 bytes.
Extracting /tmp/data/train-labels-idx1-ubyte.gz 1648877 bytes.
Extracting /tmp/data/t10k-images-idx3-ubyte.gz 1648877 bytes.
Extracting /tmp/data/t10k-images-idx3-ubyte.gz
Successfully downloaded t10k-labels-idx1-ubyte.gz 4542 bytes.
Extracting /tmp/data/t10k-labels-idx1-ubyte.gz
Epoch 0 completed out of 10 loss: 1736390.3468322754
Epoch 1 completed out of 10 loss: 410037.0322418213
Epoch 2 completed out of 10 loss: 223206.3439064026
Epoch 3 completed out of 10 loss: 131157.39508372545
Epoch 4 completed out of 10 loss: 86741.48155975342
Accuracy: 0.6488
```

Fig. 2. Implementation of Restricted Boltzmann Machine (RBM) using epoch = 5

Using the CTU-13 dataset, we extracted the features from the dataset and applied Restricted Boltzmann Machine (RBM) algorithm on it.

We adjusted the number of epochs in different phases (for e.g., epoch =1, 2, 3 and 4) and found that, the model gives under fitting results till epoch = 9. In epoch = 10, the model provides optimal result with an accuracy of 95.13%. Again, in epoch = 11, the model becomes over fitted. Hence, we selected number of iterations of epoch = 10.

The following figure shows the implementation result of the RBM model on the CTU-13 dataset with epoch = 10 (0 to 9). This model provides an optimal result with a minimum processing time. Also, this model analyzes the dataset and its patterns more accurately than other phases performed.

```
Successfully downloaded train-images-idx3-ubyte.gz 9912422 bytes
Extracting /tmp/data/train-images-idx3-ubyte.gz
Successfully downloaded train-labels-idx1-ubyte.gz 28881 bytes.
Extracting /tmp/data/train-labels-idx1-ubyte.gz
Successfully downloaded t10k-images-idx3-ubyte.gz 1648877 bytes.
Extracting /tmp/data/t10k-images-idx3-ubyte.gz
Successfully downloaded t10k-labels-idx1-ubyte.gz 4542 bytes.
Extracting /tmp/data/t10k-labels-idx1-ubyte.gz
Epoch 0 completed out of 10 loss: 1736390.3468322754
Epoch 1 completed out of 10 loss: 410037.0322418213
Epoch 2 completed out of 10 loss: 223206.3439064026
Epoch 3 completed out of 10 loss: 131157.39508372545
Epoch 4 completed out of 10 loss: 86741.48155975342
Epoch 5 completed out of 10 loss: 51217.02963626385
Epoch 6 completed out of 10 loss: 38083.232104536146
Epoch 7 completed out of 10 loss: 28213.538931260107
Epoch 8 completed out of 10 loss: 22449.897190473974
Epoch 9 completed out of 10 loss: 21965.627538987996
Accuracy: 0.9513
```

Fig. 3. Implementation of Restricted Boltzmann Machine (RBM) using epoch = 10

Conclusion and Future Scope

In this paper, we worked on a deep learning-based botnet detection algorithm, which trains the IoT devices few ways for future research have been distinguished. To show the capacity of our created model to identify new varieties of botnets, a changed adaptation of the Torii sample will be utilized in the next phase of the work, to produce a second combined dataset and will be looked at against existing mark and stream-based oddity location strategies.

This work will be extended using a sample analysis of Torii botnet and combining the parameter set extracted with the CTU-13 dataset. A novel comparison will be done in the next phase on the dataset between machine learning approach and deep learning approach to find out the which algorithm provides best result and creating a model which can predict the attacks.

References

- McDermott, C. D., Majdani, F., & Petrovski, A. V. (2018, July). Botnet detection in the internet of things using deep learning approaches. In 2018 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.
- Yan, Q., Yu, F. R., Gong, Q., & Li, J. (2018). Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges. *IEEE Communications Surveys & Tutorials*, 18(1), 602-622.
- Torres, P., Catania, C., Garcia, S., & Garino, C. G. (2016, June). An analysis of Recurrent Neural Networks for Botnet detection behavior. In 2016 IEEE biennial congress of Argentina (ARGENCON) (pp. 1-6). IEEE.
- Roosmalen, J. V., Vranken, H. P. E., van Eekelen, M. C. J. D., & Haddad, H. H. (2018). Applying Deep Learning on Packet Flows for Botnet Detection. In *Haddad, HH (ed.), SAC18: The 17th edition of the Computer Security track at the 33rd ACM Symposium on Applied Computing, 9-13 April 2018, Pau, France* (pp. 1629-1637). New York: ACM.
- Zin, H., Kim, C., Wu, M., & Kim, S. (2017). Avoidance of channel interference in polygonal IoT networks. *Concurrency and Computation: Practice and Experience*, *29*(11), e4060.
- Hoang, X., & Nguyen, Q. (2018). Botnet detection based on machine learning techniques using DNS query data. *Future Internet*, 10(5), 43.



- Mathur, L., Raheja, M., & Ahlawat, P. (2018). Botnet Detection via mining of network traffic flow. *Procedia computer science*, *132*, 1668-1677.
- Homayoun, S., Ahmadzadeh, M., Hashemi, S., Dehghantanha, A., & Khayami, R. (2018). BoTShark: A deep learning approach for botnet traffic detection. *Cyber Threat Intelligence*, 137-153.
- Mohammadi, M., Al-Fuqaha, A., Sorour, S., & Guizani, M. (2018). Deep learning for IoT big data and streaming analytics: A survey. *IEEE Communications Surveys & Tutorials*, 20(4), 2923-2960.
- Kudugunta, S., & Ferrara, E. (2018). Deep neural networks for bot detection. *Information Sciences*, *467*, 312-322.
- Qi, Y., Wu, J., Gong, G., Fan, J., Orlandi, A., Yu, W., ... & Drewniak, J. L. (2018). Review of the EMC Aspects of Internet of Things. *IEEE Transactions on Electromagnetic Compatibility*, 60(5), 1152-1160.
- Bakshi, A., Chen, L., Srinivasan, K., Koksal, C. E., & Eryilmaz, A. (2016, April). EMIT: An efficient MAC paradigm for the Internet of Things. In *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications* (pp. 1-9). IEEE.
- Conti, M., Dehghantanha, A., Franke, K., & Watson, S. (2018). Internet of Things security and forensics: Challenges and opportunities.
- Li, H., Ota, K., & Dong, M. (2018). Learning IoT in edge: deep learning for the internet of things with edge computing. *IEEE Network*, 32(1), 96-101.
- Kasprzyk, R., Paź, M., & Tarapata, Z. (2017). Modeling and simulation of botnet based cyber-threats. In *MATEC Web of Conferences* (Vol. 125, p. 03013). EDP Sciences.
- Meidan, Y., Bohadana, M., Mathov, Y., Mirsky, Y., Shabtai, A., Breitenbacher, D., & Elovici, Y. (2018). N-BaIoT— Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders. *IEEE Pervasive Computing*, 17(3), 12-22.

GJEIS Prevent Plagiarism in Publication

The Editorial Board had used the Ouriginal – a Swedish anti-plagiarism software tool which is a fully-automatic machine learning textrecognition system made for detecting, preventing and handling plagiarism and trusted by thousands of institutions across worldwide. Urkund is GDPR compliant with privacy by design and an uptime of 99.9% and have trust to be the partner in academic integrity. https://www.urkund.com] tool to check the originality and further affixed the similarity index which is $\{34\%\}$ in this case (See below Annexure-I). Thus, the reviewers and editors are of view to find it suitable to publish in this Volume-13, Issue-4, Oct-Dec 2021

Annexure 1

Submission Date	Submission Id	Word Count	Character Count
06-Oct-2021	D126535859 (Ouriginal)	2317	14817

	Document Information	on				
	Analyzed document	VP-1 Vikas arushi-Jan 2022.docx (D126535859)				
	Submitted	2021-10-0	6T12:55:00.0	000000		
	Submitted by	Dr. Subodh	Kesharwani			
	Submitter email	skesharwani@ignou.ac.in		n		
	Similarity	34%				
	Analysis address	skesharwar	ni.ignou@ana	ilysis.urkund.com		
our	rces included in the report					
	URL: http://vikasdixit.in/about.aspx	88	1 W	URL: https://rgu-repository.worktribe.com/output/249108 Fetched: 2022-01-30T12:55:49.6470000	88	1
w	Fetched: 2021-11-12T20:13:18.7770000	00				
			1 W	URL: https://www.semanticscholar.org/paper/Applying-deep-learning-on-packet-flows-for-botnet- Roosmalen-Vranken/25e183e3493sfe5a796f318bc1ed55d2bae8a90f Fetched: 2022-01-30T12:55:47.4070000		1
w	Fetched: 2021-11-12T20:13:18:7770000 URL: https://www.stratosphereips.org/datasets-ctu13/			Roosmalen-Vranken/25e183634936563796(318bc1ed55d2bae8a90f Fetched: 2022-01-30T12:55:47.4070000		

57



Reviewers Memorandum

Reviewer's Comment 1: Botnets attacks are increasing day by day and have the potential to become the new weapon of future cyber-attacks. These malicious attacks aim to stop the cyber security invasion, which in turn can have threatening results. By keeping that in mind, the study proposes an algorithm based on Deep learning mechanism.

Reviewer's Comment 2: Paper is planned in a phased manner. Earlier detection of botnet activity was performed at the packet level with Wireshark, this paper proposes a solution to the detection of botnet activity within consumer IoT devices and networks and it provides a future scope of adapting the Torii sample in the next phase, which will also allow the comparison based on the previous work done.

Reviewer's Comment 3: The paper is quite technical in nature, yet presented very strategically, also the choice of topic is very appropriate. Yet a more strengthened review of literature could be included to further improve the quality of the work done.



Vikas Dixit and Arushi Agarwal "A Deep Learning Approach Against Botnet Attacks to Reduce the Interference Problem of IoT" Volume-13, Issue-4, Oct-Dec 2021. (www.gjeis.com)

https://doi.org/10.18311/gjeis/2021 Volume-13, Issue-4, Oct-Dec 2021 Online iSSN : 0975-1432, Print iSSN : 0975-153X Frequency : Quarterly, Published Since : 2009

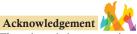
> **Google Citations:** Since 2009 **H-Index** = 96 **i10-Index:** 964

Source: https://scholar.google.co.in/citations? user=S47TtNkAAAAJ&hl=en

Conflict of Interest: Author of a Paper had no conflict neither financially nor academically.



The article has 34% of plagiarism which is the accepted percentage as per the norms and standards of the journal for the publication. As per the editorial board's observations and blind reviewers' remarks the paper had some minor revisions which were communicated on a timely basis to the authors (Vikas and Arushi) and accordingly all the corrections had been incorporated as and when directed and required to do so. The comments related to this manuscript are noticeably related to the "A Deep Learning Approach Against Botnet Attacks to Reduce the Interference Problem of IoT" both subject-wise and research-wise. This paper proposes a solution to the detection of botnet activity within consumer IoT devices and networks. A novel detection algorithm is developed based on Deep learning mechanism. Overall, the paper promises to provide a strong base for the further studies in the area. After comprehensive reviews and editorial board's remarks, the manuscript has been categorized and decided to publish under "View Point" category.



The acknowledgment section is an essential part of all academic research papers. It provides appropriate recognition to all contributors for their hard work and effort taken while writing a paper. The data presented and analyzed in this paper by (Vikas and Arushi) were collected first handily and wherever it has been taken the proper acknowledgment and endorsement depicts. The author is highly indebted to others who had facilitated in accomplishing the research. Last but not least endorse all reviewers and editors of GJEIS in publishing in a present issue.



All views expressed in this paper are my/our own. Some of the content is taken from open source websites & some are copyright free for the purpose of disseminating knowledge. Those some We/I had mentioned above in the references section and acknowledged/cited as when and where required. The author/s has cited their joint own work mostly, Tables/Data from other referenced sources in this particular paper with the narrative & endorsement has been presented within quotes and reference at the bottom of the article accordingly & appropriately. Finally, some of the contents which are taken or overlapped from open source websites for the knowledge purpose. Those some of i/we had mentioned above in the references section. On the other hand opinions expressed in this paper are those of the author and do not reflect the views of the GJEIS. The author has made every effort to ensure that the information in this paper is correct, any remaining errors and deficiencies is solely the responsibility of the author.

On the other hand author had 34% plag as the paper overlaps the research work which was actually published in his past research and played an important role.

