Kabeer_Updated_Litrature_Review.docx

by

Submission date: 03-Apr-2022 07:23PM (UTC+0500)
Submission ID: 1800184737
File name: Kabeer_Updated_Litrature_Review.docx (22.05K)
Word count: 1277
Character count: 7233

TABLE 1

Paper Title	Technique/Proposed Methodology	The Issue Highlighted/Addressed	Proposed Architecture: Positive and Negative Points	Network Security Strong/Weak Also Named Security Schemes Applied	Technology
[1]	They employ a number of methods to carefully analyze the models' performance. After changing the algorithms and using the F1 score to measure performance, the result increases considerably from 0.14 to around 0.90. Using the ultimate well-trained model, we can correctly predict the safety of any URLs on the Internet, protecting our personal information and data.	Not all URLs are completely safe. We must exercise caution while clicking on those URLs, as you may unintentionally enter some hazardous URLs. With the recent outbreak of COVID-19, an increasing number of people have been forced to stay at home. To keep up with what's going on outside, the majority of them utilize their technical devices to access the Internet.	As Internet traffic rose during the COVID-19 pandemic, hackers exploited malicious URLs to attack Internet users and steal their information. I use machine learning to develop a high-performing prediction model to discover them.	In this paper, no such network security mechanism is discussed or used.	Fine Tuned Classification Model
[2]	To detect hazardous URLs, this study recommends using a multilayer Convolutional Neural Network (CNN). In the proposed model, one layer of CNN was first investigated. Following that, a two-layer CNN will be used to boost accuracy. The accuracy of detecting bogus websites rises from 89 percent to 91 percent when two layers of CNN are added in the algorithm.	Most of the time, people trust the website and input personal information without considering if it is genuine or not. A malicious website can be used by cybercriminals to launch ransomware attacks, steal passwords, and commit financial fraud. At the Centre of these cybercriminal operations is the social engineering-based Phishing attack.	CCN architecture is proposed. The proposed detection method was initially put to the test on one CNN layer. After that, two layers of CNN are used to increase the detection accuracy. The model's accuracy as well as performance were evaluated, and positive results were revealed.	In this paper, no such network security mechanism is discussed or used.	Multilayer CNN
[3]	The proposed technique proved successful in revealing the patterns of various sorts of malicious domain names in groups. It outperformed prior techniques and gave a clearer view of the data's common patterns when used to evaluate a blacklisted collection of URLs in a genuine business network.	Monitoring the URL list in network traffic data might reveal anomalous behavior. Malware is commonly transmitted through email and rogue websites.	It dynamically visualizes the dynamism of the attack pattern. It contributes to a damaging website ecology by reducing the number of black listings. It leverages open-source threat information to validate the risk level and damaging category.	In this paper, no such network security mechanism is discussed or used.	Blacklist and malicious URL extension with FQDNs
[4]	The author of this paper has created a strong foundation for swiftly and automatically detecting phishing URLs. They tested their method on an actual dataset and were able to reach an accuracy of 87 percent in real-time.	A phishing attack impersonates a trustworthy third party in order to get sensitive information from a victim. In such an attack, users are typically routed to a fake website that appears to be legitimate. The URL of the phishing website is frequently communicated by email or instant messaging.	They've laid down a solid foundation for detecting phishing URLs automatically. To deal with the limitless increase of URL space, they employed online learning.	In this paper, no such network security mechanism is discussed or used.	Lexical, Host Based, Domain WHOIS Based, and GeoIP Based Features.
[5]	The research suggests employing a convolution neural network and a Recurrent Neural Network with Extended Present Moment Memory as models to find dangerous Uniform Resource Locators. A recurrent neural	Spam mail and phishing websites are widely used to aid certain assaults, which trick the client into revealing credentials and other sensitive information. Email is frequently thought to be the major mechanism of propagating	Deep learning methods such as RNN and RNN-LSTM are preferred to AI approaches since they may give outstanding component portrayal utilizing only raw URLs as data.	In this paper, no such network security mechanism is discussed or used.	RNN and RNN-LSTM models

	network with extended short-term memory achieved the highest accuracy of roughly 98 percent for categorizing phishing Uniform Resources.	a wide range of malicious assaults.

Table 2. Current work analysis.

Table 2

Paper Title	Strengths	Weaknesses	Future Scope for Improvement
[1]	They use a variety of characteristics to train and evaluate the model as well as optimize hyperparameters. The F1 score is close to 0.92, and the model's accuracy is 97 percent. To aid consumers in spotting bogus URLs, the concept might be implemented as software or browser extensions.	N/A	The different features that improve the efficiency and performance of system even further.
[2]	The multilayer CNN will extract relevant patterns in the data from given URLs through using several convolutions with different kernel sizes. The model is more efficient and independent of feature engineering because to the self-extraction process in the multistage trainable neural network architecture.	To achieve a better model for the job, the CNN model parameters such as number of layers, size of kernels, number of kernels, and optimizer may be modified.	It may be compared against standard machine learning-based classification models like Naive Bayes, Random Forest, Support Vector Machine, Logistic Regression, and others to evaluate if the automated feature extraction-based deep learning model is relevant.
[3]	Even if antivirus programs do not identify it, every unfamiliar website or URL that is grouped in a cluster with harmful websites should be explored further.	On the basis of the obtained groupings, no clustering or community analysis was attempted. Other security visualization techniques to properly portray distinct harmful patterns were not used.	It should design and implement a group aggregation technique that is buttoned up.
[4]	They used selective sampling and delayed feature capture to increase the system's performance. Their program can find URLs that have never been seen before with an accuracy of 87 percent.	They didn't include n-grams, DNS query results, web page network traffic, bag of words, black list presence, web page content, and other features.	It should add time-varying URL characteristics in the future.
[5]	They are able to provide assurance. Based on their findings, they think that AI and deep learning-based vindictive URL recognition can replace boycotting and traditional articulation approaches in discovery frameworks.	The extraction of web page code and content was not included.	It should turn their technique into a module for use in a web application

Reference:

[1] C. Ding, "Automatic Detection of Malicious URLs using Fine-Tuned Classification Model," 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT), 2020, pp. 302-320, doi: 10.1109/ISCTT51595.2020.00060.

[2] A. Singh and P. K. Roy, "Malicious URL Detection using Multilayer CNN," 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2021, pp. 340-345, doi: 10.1109/3ICT53449.2021.9581880.

[3] S. -Y. Huang, T. -H. Chuang, S. -M. Huang and T. Ban, "Malicious URL Linkage Analysis and Common Pattern Discovery," 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 3172-3179, doi: 10.1109/BigData47090.2019.9006145.

[4] F. Sadique, R. Kaul, S. Badsha and S. Sengupta, "An Automated Framework for Real-time Phishing URL Detection," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), 2020, pp. 0335-0341, doi: 10.1109/CCWC47524.2020.9031269.

[5] M. Arivukarasi and A. Antonidoss, "Performance Analysis of Malicious URL Detection by using RNN and LSTM," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 454-458, doi: 10.1109/ICCMC48092.2020.ICCMC-00085.

Kabeer_Updated_Litrature_Review.docx

ORIGINALITY REPORT

15%	8%	14%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	mdpi-res.com Internet Source	6%
2	Ashish Singh, Pradeep Kumar Roy. "Malicious URL Detection using Multilayer CNN", 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2021 Publication	5%
3	M. Arivukarasi, A. Antonidoss. "Performance Analysis of Malicious URL Detection by using RNN and LSTM", 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020 Publication	3%
4	cps-vo.org Internet Source	1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

Kabeer_Updated_Litrature_Review.docx

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4