American Journal of Artificial Intelligence
2021; 5(2): 66-75
http://www.sciencepublishinggroup.com/j/ajai
doi: 10.11648/j.ajai.20210502.13
ISSN: 2639-9717 (Print); ISSN: 2639-9733 (Online)

An Efficient Phishing Website Detection Plugin Service for Existing Web Browsers Using Random Forest Classifier

Adetokunbo MacGregor John-Otumu[1, *], Md Mahmudur Rahman[1], Christiana Ugochinyere Oko[2]

[1]Department of Computer Science, Morgan State University, Baltimore, USA

[2]Department of Information Technology, Federal University of Technology, Owerri, Nigeria

Email address:

adetokunbo.otumu@morgan.edu (A. M. John-Otumu), md.rahman@morgan.edu (M. M. Rahman), okokristyy@gmail.com (C. U. Oko)

[*]Corresponding author

To cite this article:

Adetokunbo MacGregor John-Otumu, Md Mahmudur Rahman, Christiana Ugochinyere Oko. An Efficient Phishing Website Detection Plugin Service for Existing Web Browsers Using Random Forest Classifier. *American Journal of Artificial Intelligence*. Vol. 5, No. 2, 2021, pp. 66-75. doi: 10.11648/j.ajai.20210502.13

Received: October 12, 2021; **Accepted**: November 1, 2021; **Published**: November 5, 2021

Abstract: An efficient phishing website detection plugin service was developed using machine learning technique based on the prevalent phishing threat while using existing web browsers in critical online transactions. The study gathered useful information from 27 published articles and dataset consisting of 11,000 data points with 30 features downloaded from phishtank. A unique architectural framework for detecting phishing websites was designed using random forest machine learning classifier based the aim and objectives of the study. The model was trained with 90% (9,900) of the dataset and tested with 10% (1,100) using Python programming language for better efficiency. Microsoft Visual Studio Code, Jupiter Notebook, Anaconda Integrated Development Environment, HTML/CSS and JavaScript was used in developing the frontend of the model for easy integration into existing web browsers. The proposed model was also modeled using use-case and sequence diagrams to test its internal functionalities. The result revealed that the proposed model had an accuracy of 0.96, error rate of 0.04, precision of 0.97, recall value of 0.99 and f1-score of 0.98 which far outperform other models developed based on literatures. Future recommendations should focus on improved security features, more phishing adaptive learning properties, and so on, so that it can be reasonably applied to other web browsers in accurately detecting real-world phishing situations using advanced algorithms such as hybridized machine learning and deep learning techniques.

Keywords: Phishing, Machine Learning, Random Forest, Web Browsers, Web Sites

1. Introduction

In general, the Internet is a hostile environment in which attacks can be easily launched and difficult to prevent, detect, or trace. However, the consequences in terms of time and money are severe. In general, ensuring the main security goals of confidentiality, integrity, and availability is difficult. There are numerous reasons for today's Internet security risks. The Internet was designed to be an open and distributed environment with no central instance controlling user communication, and mutual mistrust was not a primary concern. The existence and development of Internet security is critical to these advances in ecommerce. Consider the following scenario: if there was little or no security on the Internet, with the risk of falling victim to fraud or information theft, users would have little incentive to use it. [1]. Internet users, online banking and e-commerce applications have good protection against attacks directed towards their computer systems. Thus, the attacker has considered and instead uses “social engineering” attacks, such as phishing to gain access to the information and defraud victims [2]. Phishing is a technique used to steal private information from individuals or organizations by impersonating a reliable source (e.g., a website), usually for financial gain [3].

Unlike other deceptive information-gathering methods (such as following someone into a secure location or conversing with someone with the intent of extracting classified information), phishing is only carried out online. Phishing, which is commonly orchestrated via email, relies on exploiting human trust while avoiding email software

detection systems. It makes use of a technique known as 'Social Engineering,' in which individuals are duped into assisting the deceivers, either through actions beneficial to the deceiver or by providing confidential information [4].

The term "phishing" was derived from the analogy that early Internet criminals used email lures to "fish" for passwords and financial information from a sea of Internet users. The use of "ph" in the terminology is partly lost in the annals of time, but it is most likely related to popular hacker naming conventions such as "phreaks," which can be traced back to early hackers who were involved in "phreaking"—the hacking of telephone systems. “Phishing” first surfaced around 1996, when criminals stole American Online (AOL) accounts by “phishing” the passwords from AOL users [5]. Phishing attacks evolved from stealing AOL accounts to targeting more profitable targets, such as online banking and e-commerce services, over time. According to a global phishing survey, Apple was the most targeted brand by phishers in 2014. Attackers sent fake emails to Apple users, asking them to update their account details and providing a link to a website where they could do so; several users ended up providing their credentials on those fake sites [6]. The phishing process consists of five steps: planning, setup, attack, collection, and identity theft and fraud [7]. This is illustrated in Figure 1.

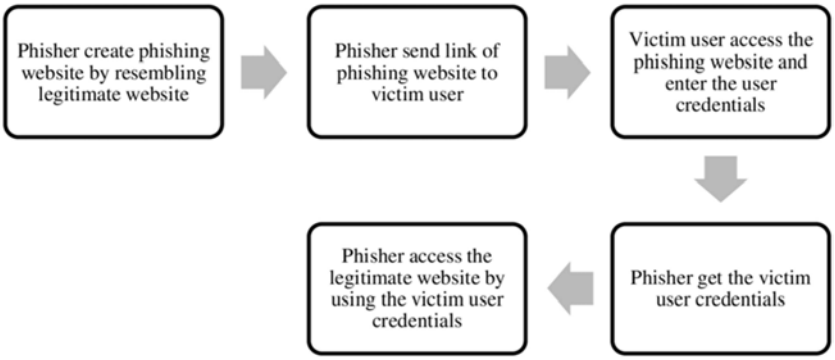

Figure 1. *Process of a phishing attack [7].*

Currently, phishing attacks target not only system end-users, but also technical employees at service providers, and may employ sophisticated techniques such as Man in the Browser (MITB) attacks. A MITB attack is one in which socially engineered messages are used to persuade victims to install MITB malware (for example, in the form of web browser ActiveX components, plugins, or email attachments), which then transfers money to the attacker's bank account whenever the victim logs in to perform his/her banking tasks, without the need to steal the victim's personal information [8]. Sometimes these messages instruct the system administrator to change their administrative passwords, and the hacker then uses key loggers to capture the password, which he will use to gain access to confidential information about the organization and its employees. Mobile applications are another new method of phishing attacks. Mobile phishing kits imitate the login screens of legitimate mobile apps.

Online transactions are extremely convenient and quick. Payment for goods and services can be made from any location. The existence and development of internet frauds, one of which is the phishing attack, is critical to these advances in e-commerce. According to an FBI report, phishing scams caused a minimum of $2.3 billion in damage between October 2013 and February 2016 [9]. Despite the large number of phishing attempts and the extensive attention paid to phishing, users of online applications such as e-mail and instant messaging continue to fall victim to these fraudulent efforts. Researchers and scholars have proposed various methodologies and designed models to detect and prevent phishing attacks, but people continue to fall victim to phishing attacks because the increasing sophistication of tools and techniques for protecting people from phishing attacks forces attackers to adapt and evolve their methods.

Phishing scams conducted through phishing websites can sometimes be easily deterred by observing whether a URL belongs to a phishing or legitimate website, but users do not pay attention to the URL of a website.

Nevertheless, current technologies, such as browser security indicators, are not fully capable of detecting phishing websites. According to a survey on "Why Phishing Works," 23% of respondents relied solely on the content of the webpage to determine its legitimacy. Furthermore, many users are unable to distinguish between a padlock icon in the browser and a padlock icon as a favicon or in page contents [9]. Because existing technologies have limitations in detecting a phishing website, expecting users to observe and determine whether a URL is phishing or legitimate is unrealistic, inefficient, and inaccurate.

Therefore, in addressing these challenges, this study will solve the following specific problems:

1) Protect users against phishing mistakes by notifying them if they are going to an unsafe website.
2) Protect individuals and institutions from financial losses.
3) Reduce online fraud and mistrust between organizations and their customers.

The primary objective of this study is to develop an enhanced machine learning based model for detecting phishing websites that can be integrated into existing web browsers as a plugin service. The specific objectives are to:

1) Develop a model that will extract website

characteristics that will be used in classifying websites as either phishing or legitimate.

2) Develop a machine learning based model using Random forest classifier that will classify websites based on their characteristics as either a 'phishing website' or a 'legitimate website'.
3) Further classify the level of severity of the phishing website detected as either 'high', 'moderate' or 'low'.

2. Related Works

In the existing literature, a lot of research attempts have been made towards incorporating machine learning techniques of different sorts in monitoring and detection issues phishing websites. This section systematically summarizes in a tabular form some key related works done on website phishing detection in order to establish a valid research gap.

Table 1. *Summary of the related works.*

S/N	Author/year	Techniques Used	Model Accuracy	Findings/Comments
1	Zhang et al. [10]	A Bayesian Approach to Textual and Visual Content-Based Anti-Phishing	N/A	Very few features were used to train the model. The model lacks new website URL detection
2	Martin *et al* [7]	A Neural Network-Based Framework for Predicting Phishing Websites	88.2%	The model used just 6 phishing characteristics for classifying websites as either "phishy" or "legitimate"
3	Chandan *et al.* [11]	A Neural Network-Based Machine Learning Approach for Detecting Phished Websites	86%	The model also used just 6 phishing characteristics for classifying websites as either "phishy" or "legitimate"
4	Mohammad, Mccluskey & Thabtah [12]	Using a self-structuring neural network to predict phishing websites	91%	The model used 17 phishing characteristics for classifying websites as either "phishy" or "legitimate".
5	Nanaware et al [13]	Malicious Website Detection using Visual Cryptography and one time password (OTP)	N/A	Phishing characteristics not discussed Delay in OTP and encryption/decryption
6	Jain et al [14]	Visual Cryptography and One-Time Passwords for Advanced Phishing Detection	N/A	Phishing characteristics not discussed Delay in OTP and encryption/decryption
7	Reshma [15]	Detecting Phishing Websites Based On Improved Visual Cryptography	N/A	Phishing characteristics not discussed Delay in OTP and encryption/decryption
8	Nguyen, To, & Nguyen [16]	Using a Neuro-Fuzzy Model, an Efficient Approach for Phishing Detection	94.18%	The model was trained using 4 phishing characteristics for classifying websites as either "phishy" or "legitimate
9	Swetha and Damodaram [17]	Detecting Phishing on Websites Using Neural Networks and Firefly	93%	The model was trained using 6 phishing characteristics for classifying websites as either "phishy" or "legitimate
10	Sahingoz, Saide and Bulut [18]	Artificial Neural Networks and Deep Neural Networks for detecting phishing from URLs.	92% and 94%	The model was trained using 1 phishing characteristic (URL) for classifying websites as either "phishy" or "legitimate The execution time is slow
11	Mahalakshmi, Goud, & Murthy [19]	A Survey of Phishing and Detection Techniques Using the Support Vector Method (SVM) and Software Defined Networking (SDN)	N/A	This technique is depended on 6 heuristics to detect phishing websites
12	Jain & Richarya [20]	Implementing a Web Browser with Phishing Detection Techniques	N/A	Phishing characteristics not discussed
13	Okunoye, Azeez, & Ilurimi [21]	A Web-Enabled Anti-Phishing Solution Based on Enhanced Heuristics	93.73%	Phishing characteristics not discussed Low adaptability to new phishing links
14	Yasin & Abuhasan [22]	An Intelligent Classification Model for Detecting Phishing Emails	93.1% and 92.4%	Phishing characteristics not discussed Restricted to email phishing attacks only
15	Gowda, Adithya, Prasad, and Vinay [23]	The creation of an anti-phishing browser based on a random forest and a rule-of-extraction framework.	N/A	Phishing characteristics not discussed Prone to overfitting
16	Ratnaparkhi and Jambhulkar [24]	Machine Learning Approach for Detection and Prevention of Phishing Websites	94%	Phishing characteristics not discussed This application is limited to just desktop devices Utilizes a lot of memory
17	Sonowal and Kuppusamy [25]	PhiDMA – A phishing detection model with multi-filter approach	92.72%	Phishing characteristics not discussed Lacks in new website URL detection
18	Ali [26]	Phishing Website Detection based on Supervised Machine Learning with Wrapper Features Selection	N/A	Phishing characteristics not discussed Consume more time and require extra computational overhead with some the classifiers used
19	Alyssa et al [27]	Phishing Website Detection: An Improved Accuracy through Feature Selection and Ensemble Learning	94%	Phishing characteristics not discussed The model can result to overfitting.
20	Muppavarapu et al. [28]	Phishing detection using RDF and random forests	93.4%	Phishing characteristics not discussed The technique used was able to reduce the

S/N	Author/year	Techniques Used	Model Accuracy	Findings/Comments
21	Maurya, Saini, & Jain [29]	A hybrid anti-phishing framework based on browser extensions that uses feature selection	93.4%	false positive rate of the system to 1.5% Phishing characteristics not discussed
22	Shah et al [30]	Chrome Extension for Detecting Phishing Websites	89.60%.	Phishing characteristics not discussed This algorithm has a runtime of 0.59.
23	Alswailem et al [31]	Detecting Phishing Websites using different Machine Learning techniques	87.34% 89.63% 89.84%	Phishing characteristics not discussed SVM showed better classification for the phishing URL from the legitimate URLs
24	Kiruthiga & Akila [32]	Phishing Websites Detection Using Machine Learning	94%	Phishing characteristics not discussed The random forest technique obtained the highest accuracy of 94%. The paper checked only the validity of URLs.
25	Kulkarni, Leonard, and Brown [33]	Phishing Website Detection using Machine Learning Techniques.	91.5%	The model was trained using 10 phishing characteristics for classifying websites as either "phishy" or "legitimate
26	Sudha et al [34]	A Review on Phishing Website Detection using Machine Learning	92.6%	Phishing characteristics not discussed The Random forest classifier approach showed better prediction accuracy.
27	Natadimadja, Abdurohman, & Nuha [35]	A Survey on Phishing Website Detection using Hadoop	N/A	Phishing characteristics not discussed Cannot detect zero day phishing attacks

The findings from Table 1 reveals that most of the phishing website detective mechanism reviewed has fewer functional features for detecting real phishing attacks. The usage of some powerful supervised machine learning techniques such as SVM, Bayesian Classifier, SDN, and ANN were excellent idea, but the Ransom forest classifier used in a very simplified manner, coupled with the extensive features needed for the classification yielded a better result.

3. Methodology

3.1. Method of Data Collection

(a) Online Repository

The dataset consist of 11,000 data points with 30 features downloaded from phishtank was used. The features are: Having IP Address, URL Length, Shortening Service, Having @ Symbol, Double Slash Redirecting, Prefix Suffix, Having Sub Domain, Secure Socket Layer (SSL) State, Domain Registration Length, Favicon, Using Non-Standard Port, HTTPS Token, Request URL, URL of Anchor, Links in Tags, Server Form Handler (SFH), Submitting Information to Email, Abnormal URL, Website Redirect Count, Status Bar Customization, Disabling Right Click, Using Pop Up Window, Iframe, Age of Domain, DNS Record, Web Traffic, Page Rank, Google Index, Links Pointing to Page, and Statistical Report.

(b) Published Articles

Under this section, the researcher downloaded and reviewed 27 articles based on phishing detection techniques, development and deployment in order to critically assess the work done on the study area and to justify the gap in knowledge established (See Table 1).

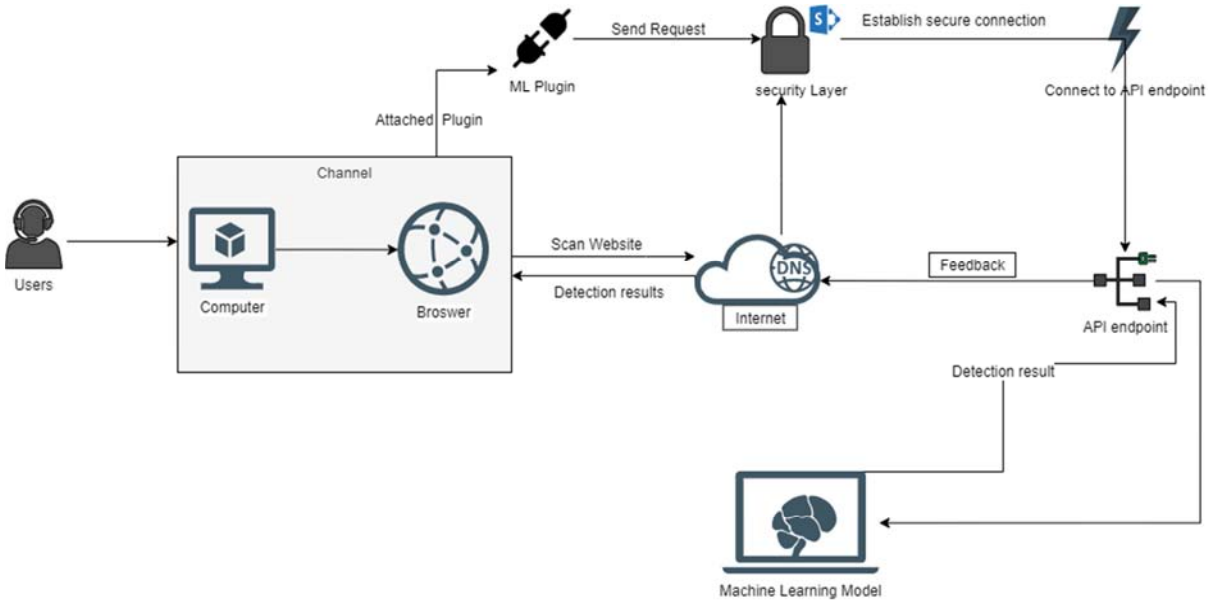

Figure 2. *Proposed System Architectural Framework.*

3.2. System Architecture

Figure 2 describes the overview of proposed system framework architecture. The framework reveals the interaction between the user's and the components of the system. The user opens a browser in order to connect to the Internet. A user loads a URL which immediately activates the machine learning plugin. The user then sends a web request which passes through the webserver to the authentication module. The authentication modules checks if the request is valid or not. If the request is valid the request connects to the API endpoint of the machine model. The model analyzes the website based on extracted features and then classifies it as a phishing or not a phishing website using random forest algorithm.

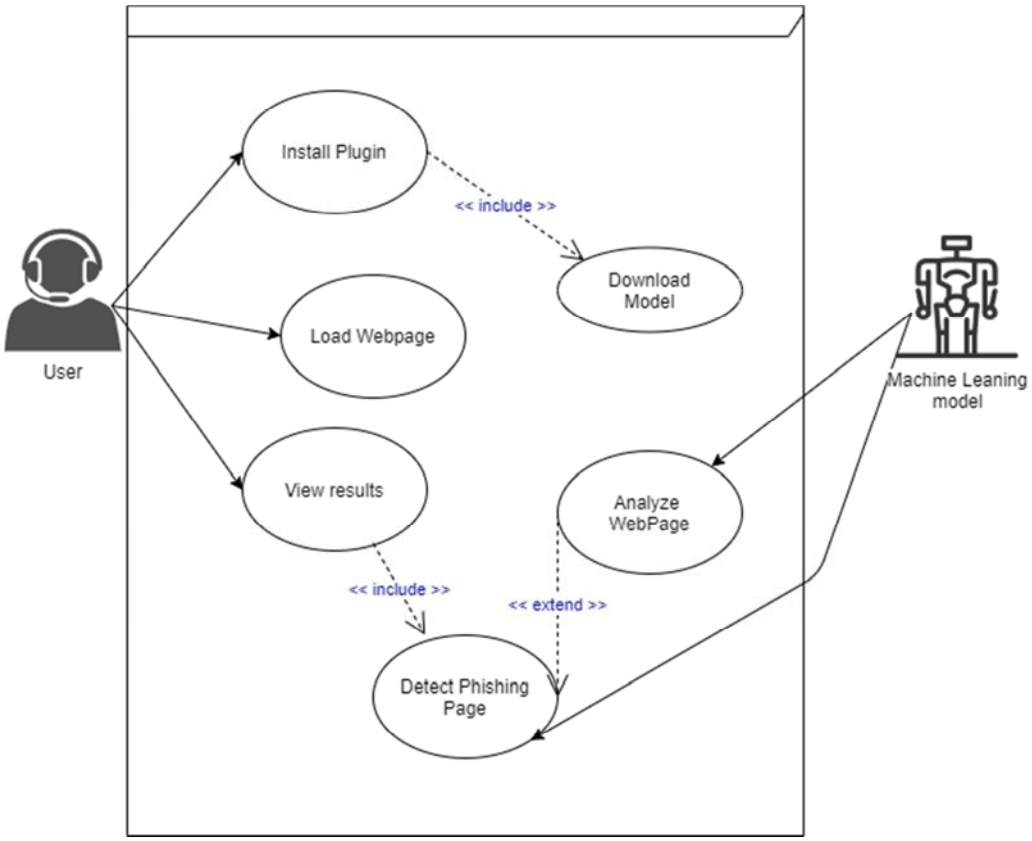

Figure 3. *Use case diagram of the Proposed System.*

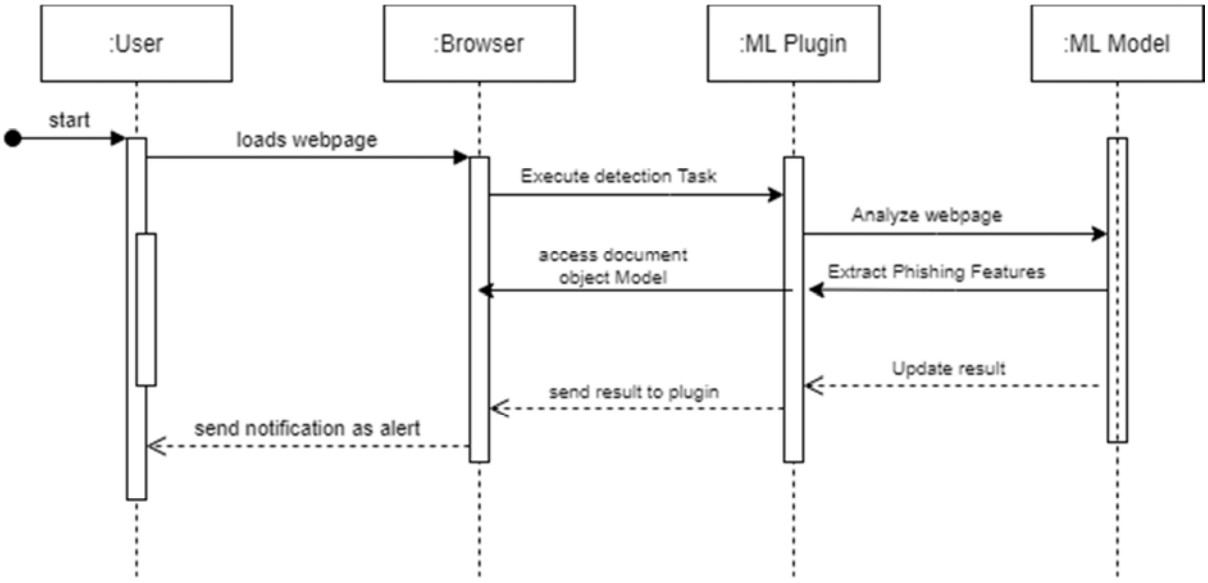

Figure 4. *Sequence Diagram of the Proposed System.*

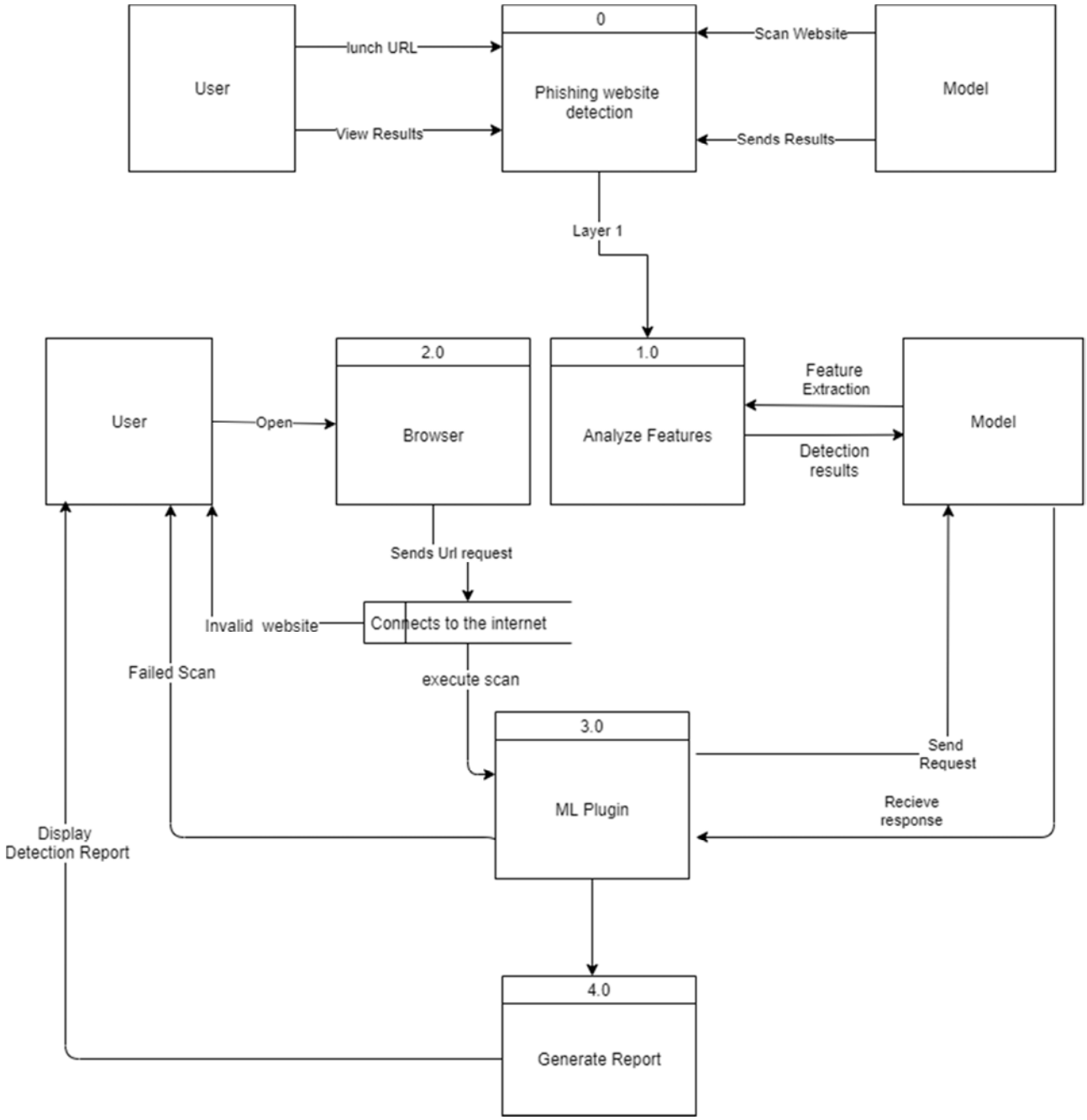

Figure 5. *Data Flow Diagram of the Proposed System.*

3.3. The Proposed Algorithm

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees.

3.4. Experimental Setup

This section provides a summary of the performed experiment in this research. A set of 30 features which are content-based, URL-based and domain information. Content-based features are mostly derived from the technical (HTML) contents of webpages e.g., counting external and internal links, counting IFRAME tags, and checking whether IFRAME tags source URLs are present in blacklists and search engines, checking for password field and testing how the form data is transmitted to the servers (whether Transport Layer Security is used and whether the "GET" or "POST" method is used to send form data with a password field, for example. URL-based features include checking for lexical properties of URLs such as the // and @ symbols in various parts of URLs, determining whether an IP address is used and what type of notation is used to represent the IP address in place of a domain name. It also looks at the domain's age, website traffic, page rank, and other factors. The proposed methodology was written in the

Python programming language.

```
Step 1: function Random Forest(S, F)
Step 2: H ← Ø
Step 3: for i ∈ 1, . . . , B do
Step 4: S (i) ← A bootstrap sample from S
Step 5: hi ← RandomizedTreeLearn(S (i), F)
Step 6: H ← H ∪ {hi}
Step 7: end for
Step 8: return H
Step 9: end function
Step 10: function RandomizedTreeLearn(S, F)
Step 11: At each node:
Step 12: f ← very small subset of F
Step 13: Split on best feature in f
Step 14: return the learned tree
Step 15: end function
```

Figure 6. *Random Forest Classifier.*

3.5. Dataset Preparation and Preprocessing

Data preparation refers to a set of procedures that make a dataset more suitable for machine learning. In a broader sense, data preparation entails establishing the appropriate data collection mechanism. The majority of the time spent on machine learning is spent on these procedures. After preprocessing of dataset, it was divided into two for training and testing purposes. At first it was divided into 70% training, 30% testing dataset. To better improve the model, the test data was reduced to 10% while the training data was increased to 90%.

3.6. System Performance Evaluation

This section discusses the proposed random forest classifier performance evaluation in terms of its ability to detect and properly classify phishing websites and the level of severity of phishing attacks on websites using confusion matrix. It also highlights performance evaluation of the predictive model in terms of accuracy, precision, recall and f1-score.

Note: 11,000 dataset; Training 9,900 (90%), while testing 1,100 (10%).

Table 2. *Confusion Matrix Evaluation.*

TP	TN	FP	FN
1050	10	35	05

(a) Computation of Accuracy

The number of correct classifications made out of all instances in the test data is simply referred to as accuracy. Eq. 1 gives the formula for calculating accuracy.

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Accuracy}=\frac{1050+10}{1020+10+35+05}$$

$$\text{Accuracy}=\frac{1060}{1100}$$

$$\text{Accuracy}=0.96=96\%$$

(b) Computation of Error Rate

$$\text{Error Rate}=1-\frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Error Rate}=1-\frac{1050+10}{1050+10+35+05}$$

$$\text{Error Rate}=1-\frac{1060}{1100}=1-0.96$$

$$\text{Error Rate}=0.04=4\%$$

(c) Computation of Precision

Precision is a measure of the classifier's exactness and measures the number of instances that have been correctly classified. It is calculated by dividing the number of positive predictions by the total number of positive instances predicted. Eq. 3 gives the precision calculation formula.

$$\text{Precision}=\frac{TP}{TP+FP} \quad (3)$$

$$\text{Precision}=\frac{1050}{1050+35}$$

$$\text{Precision}=\frac{1050}{1085}$$

$$\text{Precision}=0.967=97\%$$

(d) Computation of Recall

Recall is the number of positive instances correctly identified by the classifier from a set of all positive instances. In other words, recall counts the number of missed opportunities. Recall is a measure of how complete the classifier is. Eq. 4 gives the formula for calculating recall.

$$\text{Recall}=\frac{TP}{TP+FN} \quad (4)$$

$$\text{Recall}=\frac{1050}{1050+05}$$

$$\text{Recall}=\frac{1050}{1055}$$

$$\text{Recall}=0.99=99\%$$

(e) Computation of F1-Score

The weighted average of precision and recall scores is the F1-score. Eq. 5 gives the formula for calculating the F1-score.

$$\text{F1-scores}=\frac{2\ x\ Precision\ x\ Recall}{Recall+Precision} \quad (5)$$

$$\text{F1-scores}=\frac{2\ x\ 97\ x\ 99}{99+97}$$

$$\text{F1-scores}=\frac{19{,}206}{196}$$

$$\text{F1-scores}=97.98\%$$

4. Results and Discussion

This section displays the various outcomes of this research

project in order to achieve its goals and objectives.

4.1. Developed Interface

Figure 7. *Developed Plugin Interface for showing results.*

Figure 7 depicts the service interface for the developed phishing website detector plugin. It employs approximately 30 features to perform binary and multi-class classifications from live Internet traffic, and it also performs live analysis for each site.

Table 3. *Phishy websites.*

SN	URL
1	https://survivalfund.online/CBN/
2	http://www.civil-service.xyz/
3	https://eifi.com/register/F0Fbj0xHd7
4	https://www.npower-gov.com.ng/login/
5	https://npower-fmhds-gov-ng.web.app/
6	http://full-scholarships.xyz/dubai/
7	https://n-power-list.bid/
8	https://gramfree.world/?r=10656123
9	http://nigeria.anonymoushack.xyz/ or https://bit.ly/get-free-2million-endsars
10	https://h5.ng.o-kash.com/invite/html/invte.html?adChannel=H5invite&inviteCode=sz4j6q
11	https://freeinternetoffer.xyz/5GB
12	http://bit.ly/MTN_GIFT
13	https://www.full-scholarship.online/
14	https://kvoes.cn/v2/free.html?1619030
15	https://sci-hub.se/
16	https://cac-registration.get-noww.xyz/
17	https://bit.ly/ATIKU-YOUTH-EMPOWERMENT
18	https://a.aliexpress.com/_msKukJZ
19	https://racksterli.com/post/SPONSORED-POST-FOR-23RD-OF-FEBRUARY-2021
20	https://luckthebag.com/#1627046207674

Table 4 shows the four major web browsers used in testing the 20 known phishing web sites as represented in Table 3. The level of accuracy of the different modern browsers alongside the proposed plugin for any existing browser in detecting phishing website were also revealed. The proposed plugin service yielded a higher accuracy of 99.99% as compared to other existing web browsers.

Table 4. *Phishing website detection rate.*

S/N	Browser	Detection Rate (%)
1	Google Chrome	20%
2	Mozilla Firefox	20%
3	Microsoft Edge	10%
4	Flock Browser	45%
5	Proposed Plugin	99.99%

Table 5. *Summary of the comparative analysis.*

Researchers	Accuracy (%)
Zhang et al. (2011)	94
Kulkarni et al. (2019)	90.39
Sahingoz et al. (2018)	94
Marjan et al. (2016)	91
Nguyen et al. (2015)	94
Ratnaparkhi et al. (2020)	94
Sonowal et al. (2020)	92.72
Proposed Phishing Detection Model	96

4.2. System Comparison Evaluation

This section compares the proposed phishing website plugin service accuracy to other existing research-based models. Table 5 displays the accuracy of the various models.

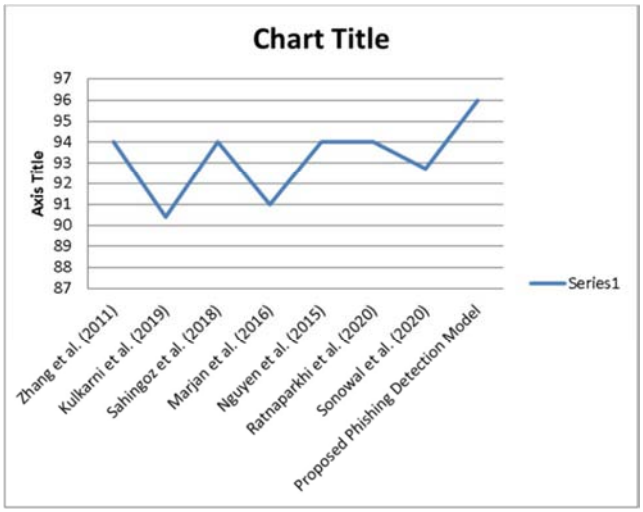

Figure 8. *Line graph showing accuracy results for detecting phishing website.*

Figure 8 depicts a line graph showing the accuracy results from the detection of phishing website from the different researchers observed. The curve shows that the proposed model outperform other recent models developed by different researcher scholars.

5. Conclusion

In this article, the enhanced phishing website plugin service developed for existing web browsers is presented, which can

efficiently monitor real live phishing website traffic based on the added features to the proposed detective mechanism using random forest machine learning classifier for both binary classification as either "Phishy" or "Legitimate" and also multi-class classification of the website as either "high", "medium", or "low". The most interesting thing about this approach is its ability to protect a user from an attacker in real-time. The model provides a fast, reliable, and secure browsing experience for the users. From the results it is seen that this proposed model achieved high results compared to existing models. Therefore, we can conclude that the research objectives were successfully achieved. Results from the system testing shows that the model developed matched the functional requirements. This model will help to reduce Internet scams. Overall, we were able to show that Random Forests is a good technique in predicting phishing websites. The proposed model is highly recommended for integration into existing web browsers so that it can notify users whenever they are on a phishing website.

In future, the proposed framework can be further enhanced for inclusion of more security features, and phishing adaptive properties and so on, so that it can be reasonably applied to other web browsers in accurately detecting real-world phishing situations using advanced algorithms such as deep learning techniques.

Acknowledgements

Special thanks to the management of the Federal University of Technology, Owerri, and the Tertiary Education Trust Fund (TETFUND) Nigeria for providing the funds to publish the results of this research work, which will be extremely useful in mitigating the phishing cybersecurity threat in Nigeria and around the world.

References

[1] Konakalla A. and Veeranki B. (2013), Evolution of Security Attacks and Security Technology, *International Journal of Computer Science and Mobile Computing*, 2 (11): 270–276.

[2] Persson A., Persson A., and Boldt M. (2007), Exploring Phishing Attacks and Countermeasures, (September).

[3] Bendovschi A. (2016), Cyber-Attacks – Trends, Patterns and Security Countermeasures Cyber-attacks – trends, patterns and security countermeasures, 5671 (January). Retrieved from https://doi.org/10.1016/S2212-5671(15)01077-1.

[4] Kleitman S., Marvin K. H. L. and Kay J. (2018), It's the deceiver and the receiver: Individual differences in phishing susceptibility and false positives with item profiling, *Individual Differences in Phishing Susceptibility and False Positives with Item Profiling*, 13 (10): 1–29.

[5] Martino, A. S., & Perramon, X. (2010). Phishing Secrets: History, Effects, and Countermeasures. *International Journal of Network Security*, 11 (3), 163–171.

[6] Gupta, B. B., Arachchilage, N. A. G., & Psannis, K. E. (2018a). Defending against Phishing Attacks: Taxonomy of Methods, Current Issues and Future Directions. *Telecommunication Systems*, 67, 1–32.

[7] Martin, A., Anutthamaa, N. B., Sathyavathy, M., Manjari, M., & Francois, S. (2015). A Framework for Predicting Phishing Websites Using Neural Networks, (September 2011).

[8] Khonji, M., Iraqi, Y., Member, S., & Jones, A. (2013). Phishing Detection: A Literature Survey, (May 2014). Retrieved from https://doi.org/10.1109/SURV.2013.032213.00009.

[9] Ubing, A. A., Kamilia, S. J. B., Abdullah, A., Jhanjhi, N. Z., & Supramaniam, M. (2019). Phishing Website Detection : An Improved Accuracy through Feature Selection and Ensemble Learning. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 10 (1), 252–257. Retrieved from https://doi.org/10.14569/IJACSA.2019.0100133.

[10] Zhang, H., Liu, G., Chow, T. W. S., Member, S., Liu, W., & Member, S. (2011). Textual and Visual Content-Based Anti-Phishing: A Bayesian Approach. *IEEE TRANSACTIONS ON NEURAL NETWORKS*, 22 (10), 1532–1546.

[11] Chandan, C. J., Chheda, H. P., Gosar, D. M., Shah, H. R., & Bhave, P. U. (2014). A Machine Learning Approach for Detection of Phished Websites Using Neural Networks. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2 (12), 4205–4209.

[12] Mohammad, R. M. A., Mccluskey, T. L., & Thabtah, F. (2013). Predicting Phishing Websites using Neural Network trained with Back - Propagation, (January).

[13] Nanaware, K., Kanade, K., Bhat, M., Patil, R., & Deokar, A. S. (2014). Malicious Website Detection using Visual Cryptography and OTP. *International Journal of Current Engineering and Technology*, 4 (5), 3310–3313.

[14] Jain, N. R., Ujwal, K., Apsara, S., Nikhil, P., & Tejashri, D. (2016). Advance Phishing Detection Using Visual Cryptography And One Time Password. *International Journal of Advanced Research in Science, Engineering and Technology*, 3 (4), 1808–1812.

[15] Reshma, R. T. (2015). Detecting Phishing Websites Based On Improved Visual Cryptography. *International Journal Of Engineering And Computer Science*, 4 (8), 14009–14014. Retrieved from https://doi.org/10.18535/ijecs/v4i8.67.

[16] Nguyen, A. T. L., To, B. L., & Nguyen, H. K. (2015). An Efficient Approach for Phishing Detection Using Neuro-Fuzzy Model. *Journal of Automation and Control Engineering*, 3 (6), 519–525. Retrieved from https://doi.org/10.12720/joace.3.6.519-525.

[17] Swetha, B. K. P., & Damodaram, R. (2016). Phishing Detection in Websites Using Neural Networks and Firefly. *International Journal Of Engineering And Computer Science*, 5 (9), 18197–18204. Retrieved from https://doi.org/10.18535/ijecs/v5i9.

[18] Sahingoz, O. K., Saide, I., & Bulut, D. (2018). Phishing Detection from URLs by Using Neural Networks, 41–54.

[19] Mahalakshmi, A., Goud, N. S., & Murthy, G. V. (2018). A Survey on Phishing and It's Detection Techniques Based on Support Vector Method (SVM) and Software Defined Networking (SDN). *International Journal of Engineering and Advanced Technology*, 8 (2), 498–503.

[20] Jain, A., & Richarya, V. (2011). Implementing a Web Browser with Phishing Detection Techniques. *World of Computer Science and Information Technology Journal (WCSIT)*, 1 (7), 289–291.

[21] Okunoye, O. B., Azeez, N. A., & Ilurimi, F. A. (2017). A Web Enabled Anti-Phishing Solution Using Enhanced Heuristic Based Technique. *FUTA Journal of Research in Sciences*, 13 (2), 304–321.

[22] Yasin, A., & Abuhasan, A. (2016). An Intelligent Classification Model for Phishing Email Detection. *International Journal of Network Security & Its Applications*, 8 (4), 55–72. Retrieved from https://doi.org/10.5121/ijnsa.2016.8405.

[23] Gowda, M., Adithya, Prasad, G., & Vinay. (2020). Development of anti-phishing browser based on random forest and rule of extraction framework. *Cybersecurity*, 3 (1), 1–14. Retrieved from https://doi.org/10.1186/s42400-020-00059-1.

[24] Ratnaparkhi, P. V., & Jambhulkar, S. S. (2020). Framework for Detection and Prevention of Phishing Website Using Machine Learning. *JOURNAL OF CRITICAL REVIEWS*, 7 (7), 2108–2125.

[25] Sonowal, G., & Kuppusamy, K. S. (2020). PhiDMA – A phishing detection model with multi-filter approach. *Journal of King Saud University - Computer and Information Sciences*, 32 (1), 99–112. Retrieved from https://doi.org/10.1016/j.jksuci.2017.07.005.

[26] Ali, W. (2017). Phishing Website Detection based on Supervised Machine Learning with Wrapper Features Selection. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 8 (9), 72–78. Retrieved from https://doi.org/10.14569/IJACSA.2017.080910.

[27] Alyssa A. U., Syukrina K. B. J., Azween A., Jhanjhi N. Z., and Mahadevan S. (2019), Phishing Website Detection: An Improved Accuracy through Feature Selection and Ensemble Learning, International Journal of Advanced Computer Science and Applications (IJACSA), 10 (1), 252–257.

[28] Muppavarapu, V., Rajendran, A., & Vasudevan, S. K. (2018). Phishing detection using RDF and random forests. *International Arab Journal of Information Technology*, 15 (5), 817–824.

[29] Maurya, S., Saini, H. S., & Jain, A. (2019). Browser extension based hybrid anti-phishing framework using feature selection. *International Journal of Advanced Computer Science and Applications*, 10 (11), 579–588. Retrieved from https://doi.org/10.14569/IJACSA.2019.0101178.

[30] Shah, B., Dharamshi, K., Patel, M., & Gaikwad, V. (2020). Chrome Extension for Detecting Phishing Websites. *International Research Journal of Engineering and Technology (IRJET)*, 7 (3), 2958–2962.

[31] Alswailem, A., Alabdullah, B., Alrumayh, N., & Alsedrani, A. (2019). Detecting Phishing Websites Using Machine Learning. *2nd International Conference on Computer Applications and Information Security, ICCAIS 2019*, 7 (2), 1–9. Retrieved from https://doi.org/10.1109/CAIS.2019.8769571.

[32] Kiruthiga, R., & Akila, D. (2019). Phishing Websites Detection Using Machine Learning. *International Journal of Recent Technology and Engineering*, 8 (2), 111–114. Retrieved from https://doi.org/10.35940/ijrte.B1018.0982S1119.

[33] Kulkarni, A., Leonard, L., & Brown. (2019). Phishing Websites Detection using Machine Learning. *International Journal of Advanced Computer Science and Applications*, 10 (7), 8–13.

[34] Sudha, M., Jaanavi, V. R., Blessy, I. G. S., & Priyadharshini. (2020). A Review on Phishing Website Detection using Machine Learning. *Journal Of Critical Reviews*, 7 (19), 4847–4853.

[35] Natadimadja, M. R., Abdurohman, M., & Nuha, H. H. (2020). A Survey on Phishing Website Detection Using Hadoop. *Jurnal Informatika Universitas Pamulang*, 5 (3), 237–246.