OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 06 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

MODEL FOR THE DETECTION OF WEB PHISHING ATTACKS BY USING MACHINE LEARNING ALGORITHMS

Husna Saleem
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

Naeem Aslam
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

Muhammad Fuzail*
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

Ayesha Faiz
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

Ahmad Naeem
Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

****Corresponding author: Muhammad Fuzail (*** mfuzail@nfciet.edu.pk***)***

Article Info

Abstract
This scholarly investigation examines the identification of web phishing attacks through the utilization of a rule-based system augmented by machine learning algorithms. Phishing represents a significant cybersecurity challenge, frequently employing misleading websites or URLs to deceive users. A dataset comprising 14 critical features characteristic of phishing behavior was employed to both train and evaluate two machine learning models: Support Vector Machine (SVM) and Random Forest. Among the assessed models, Random Forest achieved the highest level of accuracy and was subsequently chosen for rule extraction. These rules were subsequently integrated into a browser-based detection tool to facilitate real-time identification of phishing attempts. To implement this system, a Google Chrome extension called PhishNet was created using HTML, CSS, and JavaScript. PhishNet employs the derived rules to scrutinize web pages during user navigation, thereby providing immediate notifications regarding potentially suspicious sites. By directly incorporating machine learning rules into the browser environment, PhishNet significantly enhances the phishing detection mechanism within the web attack lifecycle. This solution represents a practical use of intelligent algorithms in cybersecurity, offering effective and accessible protection against phishing threats.

Keywords:
phishing attacks, web pages, machine learning, random forest, SVM.

1. Introduction

Automated Computers can acquire information without explicit programming is a great thanks to a source of Artificial Intelligence that is known as “Machine Learning”. It involves feeding data into algorithms that can recognize the pattern and make predictions on fresh data. The goal of Machine Learning is to recognize data and then fit it into models that can be understood by people and utilized by them easily. Despite being an important component of computer science, it is very different from traditional computational methods. Algorithms are the set of instructions that are used to solve a problem. Instead, Machine Learning algorithms allow computers to train from data inputs and return values that fall within a given range by using statistical analysis. Several overview articles have reviewed aspects of the learning model and its learning algorithms over the past few years, due to the hard work of many enthusiastic researchers. Huge volumes of data are first picked up by the preeminent machine learning algorithms in terms of processing and handling. This ability is valuable since the generation of data is expanding at a geometric progression [1]. For instance, machine learning in the context of security can process big logs of a network and identify suspicious signs of phishing attacks. They improve prediction and decision-making to a considerable extent for everyday commercial practices in different fields. They assess the risk of credit and predict the price of the stocks related to the financial market. In the field of healthcare, the execution of ML algorithms allows one to foresee the incidence of diseases with a precise subsequent therapeutic plan. Among the main benefits of applying machine learning in cybersecurity, it is possible to identify the possibility of calculating potential phishing attacks and avoiding them. Machine Learning is automated and does not require human interference[2-5]. ML makes it easy by giving the ability to computers learn and enables them to make predictions and refine algorithms. Previous approaches to detecting phishing are helpful but lacking in adapting to new methods employed by the attackers. This means that rule-based systems can easily be bypassed by making slight changes to the existing tactics made by the phishers, while blacklists, on average, are usually not as efficient as one would prefer, as they mostly take time to detect new sites that are involved in phishing. With phishing techniques getting more and more advanced, the call for advanced, smarter, and more preventive solutions cannot be gained. The best portion of the phishing attacks is that it spread across different computers in a single network and harms the whole system and highly secured data. One must not use unauthorized sites on the internet because they are very dangerous[6]. Thus, to protect our computer system, phishing detection is needed for all of these aspects. The security of computers and the networks that they are connected to is a significant topic in the contemporary world. In the past ten years, several 36806 approaches have been put forward to counter anti-phishing detection. These studies have mainly focused on the components of a uniform resource locator URL based on feature-selection methods for machine learning. Berners-Lee (1994) developed the URL. Thus, the format of the URL is determined by pre-existing resources and protocols. Old Systems like the domain name system of which the syntax resembles file path systems were conceptualized and proposed in 1985 [7-10]. Back slashes were employed in the path name to segregate filenames, and directories from the path of a file. The double slash was considered to separate the server names and file paths. Berners-Lee later added dots to the domain names to be separated.

Traditional methods for detection of web phishing attacks are based on rule-based systems and blacklists. Basically., the rule-based system works on predefined patterns to detect phishing attacks. But somehow, the cybercriminals introduce new methods to bypass these rules and they rapidly modify their strategies to target the users on the web. Unfortunately, many web users are targeted by phishing attacks and lose their sensitive data. On the other hand, Blacklists are another method to detect web phishing attacks that contain the databases of phishing websites, they are very active in nature, but sometimes it becomes slow to emerge new phishing domains which makes the user more vulnerable to phishing attacks [11]. Phishing is an actual threat that allows the collection of personal and sensitive data through sending emails, calls, and text messages, installing malware, and introducing people to cybercrime. Machine Learning can be considered

as a viable method for anti-phishing since it focuses on identifying patterns and applying complex data analysis. Actually, phishing scams are one of the examples of social engineering. Phishing is among the most popular cybercriminal activities. There is an example of phishing attacks [12].

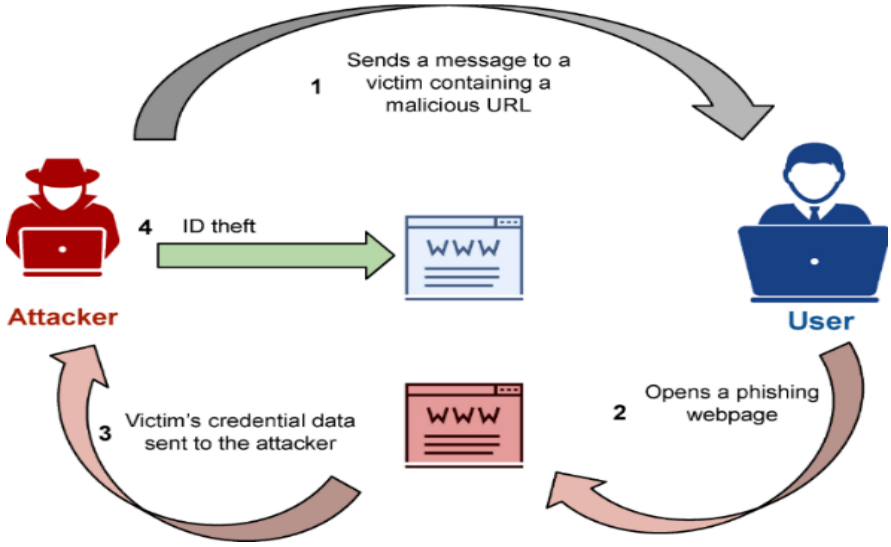

Figure 1: Attacker and User

An ordinary hacker assumes the trustworthy character of a co-worker, supervisor, or any other authoritative figure, or even a member of a reputable organization in the common and popular deception known as phishing. The target learns a code, message, or text that has been sent to them by the hacker and persuades them to pay a certain bill, click a certain link, open a certain attachment, or perform some other activity. Phishing is an attack that is carried out by technical and societal technology and hacking the user's personal information and the account's banking information. Phishing attacks occur due to target vulnerabilities and evolve as per the practices of the phishers. Phishing attacks are increasing day by day and are a growing threat to cybersecurity, which affecting millions of people and organizations worldwide[13-15]. This type of attacks involves cybercriminal fraudulent activities that mislead the targets by stealing their sensitive information such as login information, emails, passwords, bank details, and personal identification. Phishing is fundamental to breaching data and financial losses in the digital world.

Background The present world is faced with numerous and severe threats in cyberspace and among them, phishing attacks have become common. These attacks involve the use of social engineering means where the attackers lure the users into disclosing their secrets such as passwords, user names, and banking details. Several preventions have been put in place but they are still experiencing cases of phishing. They act towards individuals, corporations, and governmental bodies and result in Sizeable monetary and information losses. Description Currently the methods that aim at detecting the phishing attacks' ability to adapt to the new and increasingly complex forms of phishing is rather poor, partly because of the fact that to achieve this they often employ rule-based systems or blacklists. These traditional methods struggle to retain higher levels of accuracy as the nature of phishing techniques evolve over time and the common tools often fail to detect new instances of phishing[16]. Therefore, the requirement for advanced, adaptive, responsive, and efficient detection systems is growing. It is with this background that the purpose of this research study is to investigate and characterize the web phishing attacks using the specified machine learning-based model. By the use of Machine learning algorithms, we enhance the ability to effectively and efficiently detect the web- phishing attacks in addition to enhancing the ability to adapt the phishing detection systems.

2. LITERATURE REVIEW

In the field of cybersecurity, machine learning (ML) has become a game-changer. It provides sophisticated capabilities for threat detection and mitigation through data-driven learning and the identification of patterns that conventional methods might overlook. By evaluating enormous volumes of data, spotting subtle patterns, and adjusting to cutting-edge phishing tactics, machine learning (ML) algorithms have demonstrated considerable promise in the context of phishing detection. As a result, more resilient and dynamic detection models have been created.

The concepts, guiding principles, and frameworks that support the study should be laid out in the theoretical background section of this research. In order to lay the groundwork for research methodology and analysis, this section will explain the theories and models from the fields of machine learning and cybersecurity that are pertinent to phishing detection[17].

Phishing exploits social engineering principles, where attackers manipulate human psychology to achieve their malicious objectives. The success of phishing attacks often relies on the victim's trust, fear, or urgency, which are manipulated through carefully crafted messages or websites. There are various forms of phishing, such as email phishing, spear phishing, whaling, and pharming. Each type targets victims differently but shares the common goal of extracting sensitive information. These attacks can be understood through the lens of information security theories, which emphasize the triad of confidentiality, integrity, and availability (CIA). Phishing primarily threatens the confidentiality of information. Phishing attacks have evolved significantly over the past few decades, from rudimentary scams targeting a handful of users to sophisticated schemes that affect millions worldwide. The term "phishing" is derived from the word "fishing," where attackers metaphorically "fish" for sensitive information by luring victims with fake bait. Here's a chronological overview of how phishing has developed over time, based on recent literature[18-25]. One supervised learning technique used for classification, regression, and exterior detection is support vector machine, or SVM. But it's mostly applied to machine learning classification problems. In order to make it simple to classify new data points in the future, the SVM algorithm seeks to identify the optimal line or decision boundary that can divide n-dimensional space into classes. We refer to this optimal decision boundary as a hyperplane. SVM chooses the most extreme vectors and points to help in construction of the hyperplane. Such extreme situations are called support vectors and this is why the algorithm is called a Support Vector Machine.

2.1 Historical evolution and development of phishing attacks

Phishing schemes have also taken a rather solid turn in the course of the last few decades, starting with poor ploys that usually victimize a few users but have transformed into elaborate ones that cut across millions of people globally. The name phishing is based upon the word fishing where the attackers are said to be using a metaphor of fishing by baiting the victims with false prey. The following is a timeline following the evolution of phishing in history as told by the current literature.

2.1.2 Early Beginnings (1990s)

The development of phishing began to rise popular in the mid-1990s, especially among the hackers. phishing is based on the concept of fishing with information, though with a slight modification of the spelling that makes it part of the hacker culture (such as phreaking, a term used to describe telephone hacking). Phishing is one of the oldest phishing activities observed on America Online (AOL) in the year 1996. The hackers would pose as AOL workers and write to the users in the guise of them and request their AOL passwords. Although they were primitive in the modern standards, those tactics were actually effective due to the decreased awareness of online scams and security by the users [26] .

2.1.3 Rise of Email Phishing (Early 2000s)

As use of emails became widespread in early 2000s, phishing attacks started to use it more actively. Hackers used spam messages that were disguised as issues established bodies of businesses like banks, online payment portals, or online retailers. Such mails would usually include the linkage to some fake websites whose objective was to steal personal data like usernames, passwords and credit card details. An early phishing attack was on PayPal and eBay users in the years 2003-2004 and it shows how phishing was getting increasingly articulate to online services [27].

2.1.4 Phishing and Social Media (2010s)

Social Media Exploitation: Phishing attacks began to utilize the rising popularity of social media, such as and including Facebook, Twitter and LinkedIn. Attackers would open bogus accounts or hijack genuine accounts with the aim of uploading malicious links, or sending phishing messages. Smishing and Vishing: The other forms of communication to receive phishing was through voice and SMS where terms such as voice phishing and Smishing (voice phishing through SMS) were coined. These methods exploit the trust of the people in these more un-conventional oriented channels of communication [28].

2.2. Key Machine Learning Algorithms for Phishing Detection

2.2.1 Decision Trees

Regarding regression and classification, a decision tree (DT), a supervised learning method that is non-parametric, is utilized. These are to develop a model that predicts the value of a target variable with the help of simple decision rules extrapolated based on the features of data considered. One can think of a piecewise constant approximation as being a tree. Decision trees can be given theoretical justification using information theory and especially concepts such as entropy and information gain. Decision Trees are easy to comprehend since it is similar to human decision making. It can solve either of the cases when there is discrete data or continuous data as an input[29].

2.2.2 Random Forest

The random forest algorithm is an extension of bagging because it uses feature randomness alongside bagging to create a forest of decision trees that are uncorrelated. The feature randomness which can be aptly called as the random subspace method or feature bagging (link is external to IBM.com) will produce a random selection of features that ensures that the decision trees are not strongly correlated. Random forest theory also has great influence in artificial intelligence, machine learning, and development of computers [16]. This is one of the key differences in random forests and decision trees. Random forests only select a fraction of available feature which can split, but decision trees consider all possible splits [30].

2.2.3 Support Vector Machines (SVMs)

Support vector machine or, per short SVM is one of the supervised learning methods utilized in carrying out classification, regression and exterior detection [16]. It is largely used in classification of machine learning problems[31]. To ensure that it would be easy to classify new data points in future, SVM algorithm attempts to determine the best line or decision boundary, which can be used between n-dimensional space. This optimal decision boundary is what we call as a hyperplane. SVM chooses the extreme vectors as well as points to be used in the development of the hyperplane. Such extreme cases are called support vectors and it is on these that the algorithm is called Support Vector Machine.

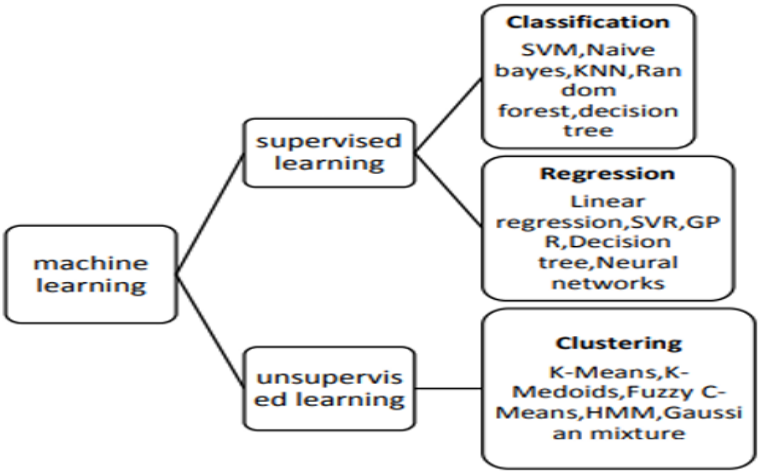

Figure 2: Types of machine learning

Machine learning (ML) techniques have become important in how malicious websites and emails are identified to offer an automated and scalable solution to detecting web phishing attacks. The methods that are applied to the latest base of phishing detection research pose a series of challenges despite their benefits. These issues must be addressed to improve phishing detection systems with regard to their accuracy, scalability, and resilience. Phishing techniques are in a never-ending evolution and machine learning models have a difficult job trying to keep up. The hackers typically manipulate the contents of the emails, the design of phish links, and layout of fake internet pages so as not to be caught. Thus, in case of emergent kinds of phishing methods, machine learning models trained on historical data can fail to work any longer[32-34]. The effectiveness of a machine learning model is highly dependent on the quality of data that it is trained using and its applicability. The majority of phishing detection models currently in use were trained using dated datasets, which do not include the most recent phishing tactics. This is a significant drawback since phishing attacks evolve over time, and data from earlier periods does not account for these changes. Furthermore, extensive feature sets and metadata are frequently absent from publicly available datasets like PhishTank, which makes it challenging for models to extract valuable insights. While web phishing attack detection using machine learning (ML) models has shown great promise, there are a number of ethical considerations that must be made to ensure that these systems are developed and implemented responsibly. These moral issues center on data misuse potential, privacy, justice, accountability, and transparency [35]. I've listed the main moral questions raised in the literature below, along with current, reliable sources. Machine learning models that contain bias may produce unfair or discriminatory results. In the context of phishing detection, the model might unjustly mark trustworthy websites from specific regions or industries as suspicious if the training data predominantly reflects phishing tactics used in those areas. Fairness concerns arise because this may result in higher rates of false positives for specific demographics or geographic areas. To prevent these biases, it is essential to make sure that the training data is diverse and balanced [36].

2.3 Research Gap:

The literature shows that there is no much conversation in the literature regarding the identification of web phishing attacks through the use of machine learning algorithms. Phishing attacks have become more advanced over the past years due to the improved technology. This has increased the demand to find proper anti-phishing solutions since users should also be safeguarded against such smart threats. References to machine learning (ML) methods In this regard, several works have used the ML methods, with features based on hyperlinks and URLs to detect the phishing websites. Raw data contained special characters and strings that were to be preprocessed and converted to a form that can be used by ML models. The

comparison and testing of the effectiveness of two feature extraction (FE) strategies, i.e. URL-based FE and SVM, in a number of experiments were done. The findings were that application of the URL-based FE together with SVM was a good approach in phishing site detection. In a like manner, APuML (Anti-Phishing using Machine Learning) system was suggested to lure out features according to the criterion of static features and also site popularity compose features through URLs to form the feature vectors. These sets of features were then fed into a suitable ML classification algorithm after which the model updating the database was done to enhance better accuracy. The study emphasized that compared to multiple classifiers, the Random Forest algorithm exhibited the utmost level of accuracy in the detection of 93.85 percent. Also, the study of developing a machine learning approach in detection of a phishing URLs showed the application of decision tree, the random forest, and supporting vector machine algorithms. This paper proposed to assess phishing URLs through accuracy, false positive and false negative examination of every algorithm. Another notion presented in the study is PhishTransformer a type of deep learning model, that by analyzing both the URL property and subsequently the contents of the page locates phishing attacks. This susceptibility uses these properties to teach a provider with the ability to make judgments about phishing and non-phishing websites. The machine learning based solution in this paper is in line with the general area of phishing detection with an increased accuracy of reporting phishing threat and a better general security of online spaces.

Table 1: Research gap for different Research

Sr	Title	Year	Methodology	Dataset	Performance Measures
1	Machine learning Model for Identifying Phishing Websites	2023	Used machine learning methods like K Nearest Neighbor (KNN), Support Vector Machine (SVM), and Naive Bayes (NB) to identify phishing websites on their own.	Developed a phish crawler to collect phishing URLs from the PhishTank website. Crawled 10,000 phishing URLs and 10,000 non-phishing URLs from the dataset.	Accuracy, Specificity, Precision, Recall, F1-Score.
2	A Deep Learning Based Phishing detection System using CNN, LSTM, and LSTM-CNN	2023	Deep Learning LSTM, CNN, and LSTM-CNN Approach	URL dataset (ISCX-URL2016)	Precision, Recall, Accuracy, F1-Score.
3	PhishCatcher: Client-Side Defense Against Web Spoofing Attacks Using Machine Learning	2023	Machine Learning Algorithm: Random Forest	Dataset available on at the UCI Machine Learning Repository, The set of 310 blacklisted URLs	Precision, Recall, Latency, Accuracy.

				from the Phish Tank. The set of 310 genuine URLs from moz.com/top500.	
4	Development of a Novel approach to Phishing detection Using Machine Learning	2024	Decision Tree, Multilayer Perceptron, Random Forest, Autoencoder Neural Network, XGBoost, Support Vector Machines:	Phish Tank.	Accuracy, Precision, recall (R) and f1 value (F1).
5	Towards a Lightweigh URL-Based Phishing Detection	2021	Supervised Machine Learning Algorithms: Naïve Bayes, Decision tree, Random Forest, Support Vector Machine.	PhishTank.	Precision, Sensitivity, F-Measure, Accuracy, Receiver Operating Characteristic (ROC) curve, and Confusion matrix.
6	PhishSKaPe: A Content Based Approach to Escape Phishing Attacks	2020	TF-IDF Term Frequency–Inverse Document Frequency Algorithm	Alexa dataset, OpenPhish, Phish Tank.	Sensitivity, Specificity, Accuracy.
7	Performance evaluation of machine learning tools for detection of phishing attacks on web pages	2022	Supervised machine learning algorithms to train models: Support Vector Machine (SVM), Random Forest, and k-Nearest Neighbor (k-NN)	Phish Tank.	Accuracy, Confusion Matrix, Precision, Recall, F-score
8	PhishTransformer: A Novel Approach To Detect Phishing Attacks	2023	PhishTransformer combines convolutional neural networks and transformer encoders to extract	Phisharmy, PhishTank, Alexa.	F1-score, precision, and recall.

			features from website URLs and page content.		

3. RESEARCH METHODOLOGY

The main goal of our research A detection model development required the implementation of this procedure: A supervised learning machine system applied trained models by leveraging the extracted features to achieve precise phishing site recognition. The rule-based approach implements these methods to enhance both phishing detection precision along with security protocols. For creating an efficient phishing detection system three separate models received training from Support Vector Machine (SVM) and Random Forest classification algorithms. The Scikit-learn library of Python was used for implementing these algorithms. To construct and enhance the detection system the following activities were implemented: Tests were conducted to evaluate the classifications of the two trained models. The evaluation of multiple classifiers led to the identification of an optimal model by analyzing their precision, recall and accuracy values combined with F1-score [38]. The detection system selected its most effective model following model evaluation. This synthetic dataset incorporates predictions which the selected model executed correctly. The Decision Tree model received training data from this dataset to generate if-else decision rules. The established rules create a system for interpreting how phishing attempts can be detected. The research team created PhishNet as a browser extension which implements the extracted rules for live phishing web page detection. PhishNet received development treatment as a Chrome extension through HTML and CSS and JavaScript programming. The script runs inside web pages through its content status to inspect features and applies rules which detect potential website threats.

3.1 Data Collection

The initial part of this project required obtaining both phishing and legitimate webpage URLs to support feature extraction activities. The model received effective training by utilizing data consisting of both legitimate and phishing websites. One thousand phishing webpages came from PhishTank (https://www.phishtank.com) whose purpose is to assist in identifying and verifying phishing sites. A total of 400 legitimate webpages originating from internet banking and financial sectors were acquired from directories such as Jasmine Directory along with The Financial Brand, Intechnic, Business, and SimilarWeb. Web scraping tools enabled the automation of the data collection procedure. The tools accessed webpage information through DOM to extract needed data for further analysis and feature extraction. An organized data acquisition process was established to collect information that would later be processed in the phishing detection model [39-40].

3.2 Feature Extraction

The obtained data underwent a transformation process which produced a set of identification features to represent important webpage characteristics. The set of extracted features includes a total of 14 elements which include: The analysis includes six features which derive from webpage URLs. A total of 8 features obtained through examination of the webpage Document Object Model (DOM). The extracted features from web content provide the base which enables the detection between phishing websites and genuine ones. The model obtains a stronger understanding of phishing characteristics through its evaluation of both URL composition and webpage content analysis.

3.3 IP Address Usage

Phishing websites prefer to present their domain names through IP addresses rather than traditional URLs. The URL "http://125.94.3.135/site.html" would appear instead of displaying familiar domain names. Two basic reasons exist which motivate attackers to use this method:

1. Phishing sites evade detection techniques through IP addressing rather than using domain names because IP addresses help avoid detection systems that identify suspicious domains.

2. The absence of domain name purchase enables phishers to lower their expenditure costs through cost avoidance methods. The detection of phishing sites relies on this important feature which is marked as Feature 1 in the model. The system detects phishing threats more easily through the identification of websites that employ IP addresses instead of domain names [41].

$$f1 = \begin{cases} 1, & IP\ address \\ 0, & not\ IP\ address \end{cases}$$

3.4 SSL security

SSL security is the main way to detect phishing websites since hackers are inclined to forego SSL certificates to make a saving. The websites under security exist in the form of SSL encryption that is revealed by including the identification mark of secure connection to the site under a prefix in the URL of the site in the form of the letter's https. Phishing sites are also very sensitive to the SSL security checks since they tend to omit them to reduce operating cost. A system can use string manipulation to determine whether websites communicate on HTTP instead of HTTPS to establish secure connections. It is depicted as Feature 2 because users need to associate websites that offer credible security with secure websites and dangerous oriented sites [42].

$$f2 = \begin{cases} 1, & SSL\ security \\ 0, & no\ SSL\ security \end{cases}$$

3.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a classification machine learning algorithm. It is a linear model, that is, it attempts to classify data points in various classes by incorporating decision boundaries given by straight lines, planes or even higher dimensional hyperplanes. These borders can be used to categorize new data points into the two categories using which side of the boundary they are on. In binary classification task SVM finds the optimal separating hyperplane that best separates the data points per classification. The equation which forms the decision boundary can be given as something like this:

$$\mathbf{y = w[0] * x[0] + w[1] * x[1] + ... + w[p] * x[p] + b > z}$$

So, what are the constituents of such an equation:

- w = weights of each feature.

- x [i] denotes vectorized input features of the classification.
- b is the intercept, a constant added to the equation.
- y is the prediction made by the model.
- z is a threshold value that helps to determine which class the observation belongs to.

The greater-than sign (> z) at the end of the equation signifies that the model predicts the class for a data point depending on whether its value is greater than z or not [7]. If the prediction is greater than z, the data point is classified into one class. If the value is less than or equal to z, it is classified into the other class. In simpler terms, SVM tries to find the most optimal line (or hyperplane) that divides the data into two distinct classes, making it a very powerful tool for classification tasks.

3.6 Random Forest

A Random Forest is an ensemble method that combines multiple Decision Trees to make more accurate predictions. Each tree in the forest is slightly different from the others, and the final prediction is made by averaging the results from all individual trees. This approach helps to mitigate overfitting, a common issue with Decision Trees, where the model becomes too complex and specific to the training data[25]. Why it is called Random Forest has a reason in the method introduced by randomness at the tree building process. At each of those steps, the algorithm introduces random variations to the given tree, thus guaranteeing its uniqueness and lowering the rate of model being too much pattern-specific. The randomness assists in enhancing generalization of the model and its overall performance.

3.7 Model Assessment

After training the models, it is then time to test them; that is model assessment, where we test the models and decide their effectiveness in terms of numerous statistics. Such metrics as True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) are estimated per model [13]. Together with them, other measures are also calculated which include Precision, F-score, Accuracy and Recall to give a clear picture on the extent the models are performing. The formulae of two most important values which shall be employed in our analysis are:

- **Precision** $= \frac{TP}{TP+FP}$
- **Sensitivity (Recall)** $= \frac{TP}{TP+FN}$
- **F-Score=** $2 \times \frac{Precision \times Recall}{Precision+Recall}$
- **Accuracy** $= \frac{TP+TN}{TP+TN+FN+FP}$

4. RESULT AND DISCUSSION

4.1 Experimental Results

After the URLs have been gathered, the second step was extraction of features. A similar one was also made with Python script and Beautiful Soup library that is intended specifically to navigate through the web pages and extract the information. Every feature that was going to be extracted had its own method in the script, which made it easy to modify the script and add new ones whenever necessary. It was going through list of URLs and fetching each of them one by one. After visiting a site through a URL, the script

then used the web page Document Object Model (DOM), also known as the representation of a site that reflects the content of the page. Based on this we were able to derive the desired characteristics. As an example, the page-level attributes, including the presence of the SSL-security, length of the URL, any mentions of certain keywords, etc. were retrieved out of the DOM of the page. When the corresponding features have been obtained, the data were stored in a form of a structured CSV file to be further analyzed [41]. This file had all the information required to train and test the machine learning models at a later stage. Fig. 3 contains a sample of the output of the feature extraction and gives a feeling of how those features were stored and arranged in order to use it later on. Automated data collection and feature extraction process allowed us to construct a small and well-organized dataset effectively, which became the foundation of the whole phishing detection system. The True Positive and True Negative rates of the model are also presented as performance metrics in Figs. 3-4 and Table 2 and Table 3 as well. These numbers allow getting a more visual and complete picture of how the model will behave, including its strengths and weaknesses Although its True Positive output was excellent, the fact that its True Negative was not so high makes it reasonable to assume that there is an area to focus on improving the model and making it code more efficiently distinguishing between phishing and genuine sites.

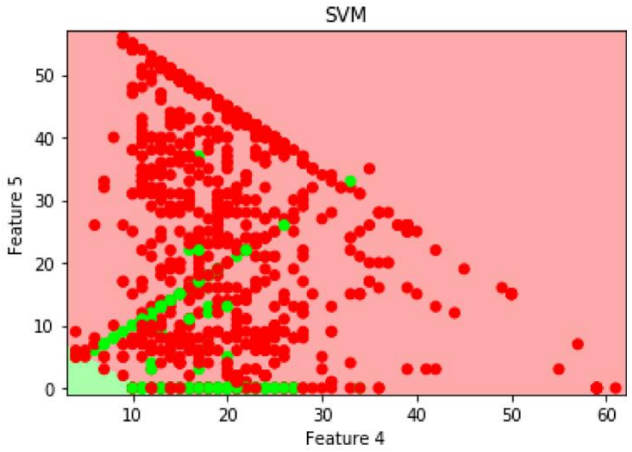

Figure 3: Decision boundary of SVN model of two features.

Predicted Classes

	Phishing	**Legitimate**
Phishing	97.0%	3.0%
Legitimate	33.33%	66.66 %

Actual Classes

Figure 4: Confusion matrix Random Forest.

Random Forest was trained with five decision trees. To remedy the situation of the lack of feature values, a dummy value of -1 was put when there was a missing feature. This was to make sure that the model would have itself fully informed. This model was observed to be the finest model among the various algorithms which were tested. It also recorded a perfect 100 percent True Positive, which translates to the fact that it did not fail to identify any phishing site. Also, it behaved reasonably in the detection of genuine websites, its True Negative value was 90.48, and it mainly labelled the non-phishing web sites. Consequently, the overall percentage of the correctness of the model turned out to be 98.35% on the testing dataset, which proves its high effectiveness. These excellent values are also reflected in Figs. 2-4 and

Tables 1 and 3 that allow one to better understand the functioning of the model. The great accuracy of the Random Forest, as well as sufficiently high True Positive and True Negative rates, make that an optimal choice in detecting phishing websites in the specified context. The effectiveness of this model to differentiate between phishing and legitimate sites in this model underlines its reliability and strength in the task.

Table 1: Random Forest classifications report 1.

	precision	**recall**	**F1- Score**	**Support**
Legitimate	1.00	0.90	0.95	21
Phishing	0.98	1.00	0.99	100
Micro Avg	0.98	0.98	0.98	121
Macro Avg	0.99	0.95	0.97	121
Weighted Avg	0.98	0.98	0.98	121

Table 2: Random forest summary table 2.

	precision	**recall**	**F1- Score**	**Support**
Legitimate	0.82	0.67	0.74	21
Phishing	0.93	0.97	0.95	100
Micro Avg	0.92	0.92	0.92	121
Macro Avg	0.88	0.82	0.84	121
Weighted Avg	0.91	0.92	0.91	121

Table 3: Random Forest outline of results.

Statistics	**Value**
Accuracy	98.35%
Error	1.65%
True Positive	100%
True Negative	90.48%
False Positive	9.52%

False Negative	0%
Precision	0.98
Recall	0.98
F-Score	0.98

5. CONCLUSION

In this thesis, to assess the effectiveness of PhishNet in this thesis we tested and trained two machine learning constructs, Random Forest, and Support Vector Machine (SVM). The role of each model was evaluated in terms of identifying phishing and genuine sites properly. Random Forest hence proved the most effective and accurate among the two models with a huge margin over SVM in phishing detection. Random Forest had the highest True Positive rate of 100% and overall accuracy of 98.35% which makes it the best option to carry out phishing detection. The Support Vector Machine (SVM) model performed best under certain set of conditions but its performance overall was the lowest at only 66.67 percent accuracy in detection of phishing sites. The findings support the main outcome that Random Forest is the best selection in phishing detection because besides its high level of accuracy, it has low false positives and negatives. The capacity of the model to decide using various decision trees has been chosen so that a solid and steady classifying structure is provided, thus the model is considered a perfect choice in the phishing prevention framework in real life. PhishNet was written as a web browser plug-in to offer instantaneous phishing recognition through study of substance and design of web sites. The extension also has an automatic scan feature that retrieves the feature of webpages visited by the user and classify them as a legitimate or a phishing site using machine learning models. PhishNet gets to the user via an alert message that comes up in the event that there is a phishing site identified and this may help to avert a security attack. Accommodating the model of the best-performing Random Forest model, the PhishNet is an efficiency tool that guarantees high accuracy and fast detection speed which will prove useful for daily web navigation. Instead, it provides a convenient user experience, which does not require manual handling, as the users could be safeguarded without knowing much about cybersecurity.

6. FUTURE WORK

This paper presents an assessment of the utility of a machine learning-based classification model to the detection of threats of phishing using lightweight Google chrome extension named PhishNet. PhishNet is a security tool, designed to work in real-time, that allows its users to get information about fraud sites and avoid their use in order to make a lot of money in phishing. Phishing is still a significant cybersecurity threat and culprits are developing fake websites that target valuable information like passwords, credit cards, and personal information. PhishNet focuses on this problem with the help of machine learning algorithms that can analyze the contents of the websites and their URLs, detecting the possible threat identified in them and preventing user access to it. PhishNet allows online users to be confident in navigating the web by offering them an automated and easy-to-use phishing detecting system that provides sufficient protection to users. It can identify and issue a warning on the phishing attempt thus limiting the chances of cyber fraud through an additional protection to both individuals and organizations. The causes of the phishing scams and the users can be prevented by creating an efficient system to fight phishing by using the machine learning models which are discussed in this paper. This implies that the model presented today is not perfect, but there is need to improve it. In the future, the work is intended to reveal how one can improve detection options by combining other methods, including behavioral analysis, real-time tracking, or using deep learning principles. The additional important suggestion to future work associated with the project will be to increase the size of the accumulated dataset. The bigger and richer dataset will

enable one to check the precision of the model when dealing with various forms of phishing attacks on a larger scale. The model can be trained to understand the behavior of new attacks by gathering data in an increased number of phishing and genuine websites. Also, two models of machine learning were studied in this paper Random Forest and Support Vector Machine (SVM). Although Random Forest was most effective, it is possible that other machine learning algorithms could be considered in their future research including deep learning models, neural network, or ensemble learning methods to find out whether they are more effective. Lastly, bridging the gap between real-life performance and PhishNet browser extension is also another positive move. Future work can also involve the creation of cross-browser compatibility thus PhishNet can also be accessed across other browsers such as the Firefox, Edge and Safari on top of Chrome. Improving the user experience by making better use of more intuitive ways of alerting, security control customization and better integration of the PhishNet with existing cybersecurity systems would also enable PhishNet to be more user-friendly and useful. As these aspects will be taken care of, future studies will assist to develop an even stronger, adaptive and accurate phishing detection system, and finally will make internet a safe place to be used by the whole world.

References

[1] B. Mahesh, "Machine Learning Algorithms - A Review," International Journal of Science and Research (IJSR), vol. 9, no. 1, 2020.

[2] M. . U. Elamathi and M. . A. V. M. B. Aruna, "An Effective Secure Mechanism For Phishing, " Journal of Pharmaceutical Negative Results, pp. 1-2, 2023.

[3] S. Asiri, Y. Xiao and T. Li, "PhishTransformer: A Novel Approach to Detect Phishing Attacks, " Electronics, 2023.

[4] "IBM," 17 May 2024. [Online]. Available: https://www.ibm.com/topics/phishing#:~:text=Phishing%20is%20a%20type%20of,a%20form%20of%20social%20engineering..

[5] A. Mughaid, . S. AlZu'bi, A. Hnaif, S. Taamneh, A. Alnajjar and E. A. Elsoud, "An intelligent cyber security phishing detection system using deep learning techniques," Cluster Computing, vol. 25, p. 3819–3828, 2022.

[6] S. Alnemari and M. Alshammari, "Detecting Phishing Domains Using Machine Learning," applied sciences, vol. 13, 2023.

[7] B. Naqvi, K. Perova, A. Farooq, I. Makhdoom , S. Oyedeji and J. Porras, "Mitigation strategies against the phishing attacks: A systematic literature review.," Computers and Security, vol. 132, 2023.

[8] A. Petrosyan, "Number of global phishing sites Q3 2013- Q1 2024," 2023-2024.

[9] A. KARIM, M. SHAHROZ, K. MUSTOFA, . S. B. BELHAOUARI and S. K. JOGA, "Phishing Detection System Through Hybrid Machine Learning Based on URL," IEEE Access, vol. 11, 2023.

[10] R. . S. Rao and A. R. Pais , "Jail-Phish: An improved search engine based phishing detection system," Computers & Security, vol. 83, pp. 246-267, 2019.

[11] M. Jakobsson and S. Myers, Phishing and Counter-Measures: Understanding the Increasing Problem of Electronic Identity Theft, 2018.

[12] A. W. G. (APWG), "Phishing Activity Trends Report," 2004.

[13] Z. Alkhalil, C. Hewage, L. Nawaf and I. Khan, "Phishing Attacks: A Recent Comprehensive Study and a New Anatomy," Frontiers in Computer Science, vol. 3, 2021.

[14] B. B. Gupta, N. A. G. Arachchilage and . K. E. Psannis , "Defending against phishing attacks: taxonomy of methods, current issues and future directions," Telecommunication Systems, vol. 67, p. 247–267, 2018.

[15] P. P. Harsh H. Patel*, "Study and Analysis of Decision Tree Based Classification Algorithms," International Journal of Computer Sciences and Engineering , 2018.

[16] A. M. A. D. Q. Z. Dildar Masood Abdulqader, "Machine Learning Supervised Algorithms of Gene Selection: A Review," Technology Reports of Kansai University, 2020.

[17] D. P. B. S. D. S. S. G. Uday Bhaskar Penta, "MACHINE LEARNING MODEL FOR IDENTIFYING PHISHING WEBSITES," Journal of Data Acquisition and processing, 2023.

[18] R. A. 1. J. A.-M. 1. Q. E. U. H. 3. K. S. 2. a. M. Zainab Alshingiti 1, " A Deep Learning-Based Phishing Detection System Using CNN, LSTM,andLSTM-CNN," electronics, 2023.

[19] A. B. A. 2. K. 1. M. A. Z. H. K. A. A. 3. MUZAMMIL AHMED1, " PhishCatcher: Client-Side Defense Against Web Spoofing Attacks Using Machine Learning," IEEE Access, 2023.

[20] S. F. N. T. Agboola Olayinka Taofeek, "Development of a Novel Approach to Phishing Detection Using Machine," JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION , 2024.

[21] A. M. 2. a. N. P. 3. Andrei Butnaru 1, " Towards Lightweight URL-Based Phishing Detection," future internet, 2021.

[22] S. AnkitKumarJain, " PhishSKaPe: A Content based Approach to Escape Phishing Attacks," Procedia Computer Science , 2020.

[23] G. O. b. ,. ∗. B. O. a. ,. O. F. b. ,. S. F. c. ,. N. O. a. T.O. Ojewumi a, "Performance evaluation of machine learning tools for detection of phishing attacks on web pages," Scientific African, 2022.

[24] "APhishing-Attack-Detection Model Using Natural Language Processing and Deep Learning," Eduardo Benavides-Astudillo 1,2,* , Walter Fuertes 2, Sandra Sanchez-Gordon 1, Daniel Nuñez-Agurto 2, GermánRodríguez-Galán 2, 2023.

[25] D. S. F. F. L. I. S. R.-E.-. U. S. H. Mohammad Nazmul Alam, " Phishing Attacks Detection using Machine Learning Approach," in Smart Systems and Inventive Technology , 2020.

[26] 2. ·. H. A. ·. M. ·. A. ·. A. Umer Ahmed Butt1 ·Rashid Amin1, " Cloud-based email phishing attack using machine and deep learning," Complex &Intelligent Systems, 2022.

[27] L. L. B. I. Arun Kulkarni1, "Phishing Websites Detection using Machine Learning," International Journal of Advanced Computer Science and Applications,, vol. 10, 2019 .

[28] I. S. Rishikesh Mahajan, "Phishing Website Detection using Machine Learning Algorithms," International Journal of Computer Applications , 2018.

[29] Y. C. B. W. Chidimma Oparaa, " Look before you leap:Detecting phishing web pages by exploiting raw URL and HTML characteristics," ExpertSystemsWithApplications, 2024.

[30] A. K. J. ·. N. D. ·. A. K. Jain2, " APuML: An Efficient Approach to Detect Mobile Phishing Webpages using Machine Learning," Wireless Personal Communications, 2022.

[31] A. K. J. ·. B. B. Gupta1, " A machine learning based approach for phishing detection using hyperlinks information," Journal of Ambient Intelligence and Humanized Computing, 2018.

[32] S. M. I. A. M. L. F.-S. S. M. A. Nureni Ayofe Azeez, "Adopting automated whitelist approach for detecting phishing attacks," Computers & Security, 2021.

[33] A. A. Aleroud and L. Z. Zhou, "Phishing environments, techniques, and countermeasures: A survey," Computers & Security, vol. 68, pp. 160-196, 2017.

[34] R. M. Mohammad, F. Thabtah and L. McCluskey, "An assessment of features related to phishing websites using an automated technique," in 2012 international conference for internet technology and secured transactions, 2012.

[35] A. . K. Jain and B. B. Gupta, "Comparative analysis of features based machine learning approaches for phishing detection," in 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 2016.

[36] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman and A. Galstyan, "A Survey on Bias and Fairness in Machine Learning," ACM Computing Surveys, vol. 54, pp. 1 - 35, 2022.

[37] R. A. M. M. A. M. 3. ,. 4. a. J. X. 5. MuhammadWaqasShaukat1, " A Hybrid Approach for Alluring Ads Phishing Attack Detection Using Machine Learning," sensors, 2023.

[38] *. A. A.-S. M. R. A.-M. H. K. G. S. M. A. N. A.-S. a. S. A. R. Alazaidah1, "Website Phishing Detection Using Machine Learning Techniques," Journal of Statistics Applications & Probability , 2024.

[39] G. B. Gururaj Harinahalli Lokesh, " Phishing website detection based on effective machine learning approach," JOURNAL OF CYBER SECURITY TECHNOLOGY , pp. ,, 2020.

[40] ,. S. (. I. O. K. E. H.-V. (. I. H. F. (. S. M. I. NGUYET QUANG DO, " Deep Learning for Phishing Detection: Taxonomy, Current Challenges and Future Directions," IEEE Access, 2022.

[41] G. F. ,. J. W. P.A. Barraclough, "Intelligent cyber-phishing detection for online," Computers & Security, 2021.

[42] A. S. A. A. A.A. Orunsolu, " A predictive model for phishing detection," Journal of King Saud University Computer and Information Sciences, 2022.