[1]Sagil Tausif,
[2*]Mahfooz Alam,
[1]Suhel Mustajab,
[3]Bashar Alam

Dynamic Phishing Detection in Cybersecurity Method Using Machine and Deep Learning Approaches

Journal of Electrical Systems

Abstract: - Cyber phishing is one of the major threats that individual, organization, and government encounter. Intruders incessantly devise new tactics to penetrate traditional detection methods. In response, researchers and practitioners have adopted machine learning techniques to enhance phishing detection capabilities. As per many reports and our survey email phishing is the most accepted form of phishing. This paper highlights email phishing detection techniques. Fraudsters use a technique called phishing, in which they pose as reputable sources in order to trick victims into divulging important information. An email that has been phished may attempt to trick you into sending personal information. The use of machine learning algorithms is covered in this study, and the experiment looks at how to classify phishing detection problems using methods like logistic regression, SVM, decision trees, random forests, locally deep SVM, and neural networks to determine whether an email is spam or not, phishing or legitimate, and spam ham or spam. Using test data to validate the result, the detection model is trained using a split dataset, utilizing machine learning techniques. This detection model targets intrinsic characteristics of email as phishing or non-phishing. A maximum accuracy of 98.83% is achieved in classifying emails using locally deep SVM.

Keywords: Phishing, Machine Learning, Spam, Legitimate, Deep Learning.

I. INTRODUCTION

The global pandemic of 2020 led to the life of people completely dependent on technology since digitalization resulted in the increased number of cybercrimes latest reports and research point to many vulnerabilities in security that cost large sum of money or provide private information to the victims. Internet crime, especially in case of phishing that is based on technical deceit as well as social engineering, is an effort to steal financial accounts or personal identifying information credentials among victims. When it comes to phishing intruders target dependable websites as well as misdirect individuals to these websites, by which they are tricking them into distributing username credentials for credit cards, bank details as well as another delicate credential. These URLs for phishing might reach their target via text message, instant messaging, or email [1][2][37].

FBI Crime Report 2020 brought to light that phishing was the most frequent kind of cyberattack in 2020, rising from 114,702 in 2019 to 241,342, almost twice as many. According to the Verizon 2020 data breach investigation report, phishing was implicated in 22% of data breaches in 2020. Another anti-phishing work group stated phishing attacks doubled in 2020 [1][3].

In 2020's fourth quarter. Financial institutions were discovered to be the most targeted victim of the phishing attack. Phishing attacks in opposition to seas and Websites like webmail were inaccessible and attacks in opposition to e-commerce intensified. The percentage of phishing attacks against media companies dropped from 12.6% to 11.8%. The global pandemic COVID-19 was a period of boon for cyber attackers. It opened the door to deceitful people, depending on whom bogus emails are being sent by hackers and cyber scammers email even messages to those who take advantage of COVID-19. These attacks involved false employment offers, fake communications from health organizations, phishing with a COVID-19 theme, and brand impersonation [2][3].

In the next part, numerous methods for detecting phishing attacks are analyzed. The most prevalent ML algorithms applied in the instance of ML-based methodology are explored.

[1]Department of Computer Science, Aligarh Muslim University, Aligarh, U.P-202002, India. sagiltausif17@gmail.com; suhelmustajab@gmail.com

[2]Department of MCA, G.L. Bajaj Institute of Technology and Management, Greater Noida, U.P-201306, India

mahfoozalam.amu@gmail.com

[3]Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, U.P-202002, India.

basaralam925@gmail.com

Correspondence Author: mahfoozalam.amu@gmail.com; malam.cs19@gmail.com

A. Phishing Detected

Phishing attack using URL involves delivering harmful links, which appears to be acceptable to the user and deceiving them by clicking on it. Classification of URL-based phishing is done by analyzing different features of the URL. Different ML algorithms use different URL dataset feature as to mark a specific URL as phishing or legitimate [1][3].

B. Approaches for Detecting Phishing

Inside the list-based methodology, two list exits whitelists and blacklists to separate phishing and legitimate URL according to. The ability to obtain a URL is possible if the URL is present in the Whitelist alone. The phishing URL's structure in the blacklist is analyzed. A sequence of URL's they were originally categorized as phishing. URLs are identified in accordance with how well they follow this pattern. The features of the URL have played an important part in categorizing websites precisely [3].

An approach based on visual similarity focuses on considering the visual connection of web pages. A server-side view is employed to classify a website as phishing or not. These two instances of data sets are subsequently contrasted using the image analysis technique. phoney websites have a similar design which is close one to the original also, it is easy to find tiny differences using image processing techniques while the user is unable to identify things without difficulty [4][37].

Analyse using a content-based approach page content this approach draws pages' contents and outside services' functionality; such as search engine as well as DNS server. The authors suggested a detection technique in which weights are specified to words drawn between the contents of HTML and URLs. The phrase may involve the logo name used by assailants in URLweights assigned as per position in the URL. The most likely terms are selected, and Yahoo search is then asked to provide results to the domain name that appears most frequently among the top 30 results. The owner of the URL is used to determine whether it is phishing legitimate or not [1][3].

The fuzzy-based approach deals with the processing of ambiguous variables, and then it is integrated and classified by human experts, it uses a certain set of metrics and pre-established standards to categorize websites according to the degree of phishing that is present on the pages. This method is impacted by unimportant characteristics and the efficiency of this method is based on its relevant features.

In the ML-based approach, this method employs supervised learning algorithms. Trained and tested algorithms are used to measure the execution of each prototype. This method has effective methods that deliver great results for phishing detection [5][6][7]. The model's performance offers high-performance, effective solutions for phishing detection. This is an important area of study, and machine learning-based phishing detection is covered in several publications [8][9].

C. Algorithms for Machine Learning

Many machine learning algorithms exist, including Logistic regression, SVM, Random Forest, Decision tree, locally deep SVM, and Neural Network, to identify websites that are phishing. When compared to alternative approaches, this widely used strategy has been shown to be incredibly precise and efficient [5].

This paper focuses on advancing phishing detection techniques to strengthen cybersecurity. Our approach integrates diverse methodologies to effectively address phishing threats. The first method emphasizes the extraction of features from URLs to develop machine learning models. By analysing factors such as URL length, domain reputation, subdomains, and lexical patterns, we train models capable of distinguishing between legitimate and malicious URLs. This proactive approach facilitates real-time identification and mitigation of phishing attempts, enhancing overall cybersecurity measures. In addition, we introduce a second approach that examines textual features extracted from phishing emails and messages. To simulate real-world scenarios, we curate datasets with a balanced mix of spam and non-spam entities. By utilizing advanced text analysis methods, including natural language processing and sentiment analysis, we improve the detection and classification of phishing content. This method uncovers subtle indicators of phishing attempts, helping to prevent cyber threats before they cause harm and contributes to enhancing phishing detection capabilities by combining URL-based analysis with sophisticated text analysis techniques. This multi-layered defense approach offers a robust mechanism to counter evolving phishing tactics, ultimately strengthening cybersecurity for individuals and organizations.

The rest of the paper is organized as follows: Section II provides definitions relevant to phishing detection. Section III presents the background and a literature review, including phishing statistics and targeted victim groups. Section IV describes the proposed methodology in detail. Section V discusses the experimental setup and presents the results. Finally, Section VI concludes the paper and outlines directions for future work.

II. DEFINITION

Attacks using phishing techniques do not have an accurate definition in the literature, which is because phishing is a widespread issue and incorporates different situations. As per Phish Tank: "Phishing is a fraudulent attempt, usually made through email to steal your personal information" Phish Tank's meaning seems logical in many situations. which approximately covers most phishing attempts (while no reliable research has been done to consistently qualify this) [11]. This definition has its own limitations to stealing personal information only a statement that isn't always accurate. In order to transfer money to the attacker's bank account whenever the victim logs in to perform banking tasks, for instance, a socially engineered message may entice the victim to install Man in the Browser (MITB) malware. This malware can take the form of web browser ActiveX components, plugins, or email attachments. The attacker would not need to steal the victim's personal information in order to carry out this attack. Therefore, we believe that Phish Tank's meaning does not elaborate on the whole phishing problem as illustrate in figure 1 [10].

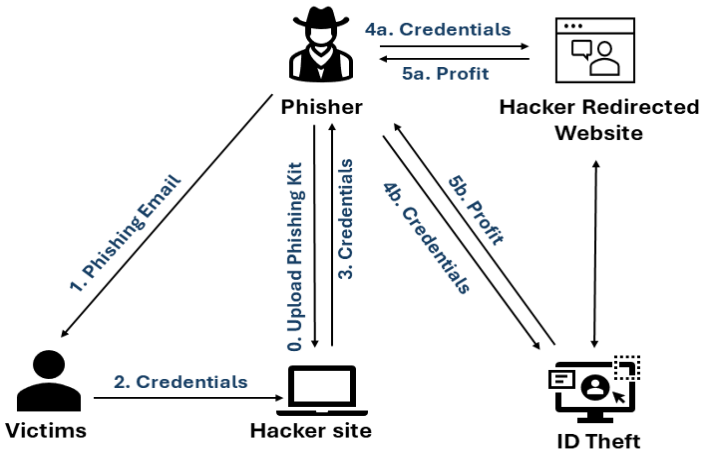


Figure 1. Phishing trap

A. Types of Phishing Attacks

There are several phishing attacks are as follows:

- Emails Phishing: Nowadays, Phishing emails are among the most common phishing attempts that are prevalent in the world. Intruders send malicious emails to targeted victims in order to snatch personal data like bank details, credit card number, password, secret pin, etc. Further the intruder utilizes stolen information to accomplish their ill will like stealing money from the victim's account [11].

- Spear Phishing: Spear Phishing is one of the phishing methods which is a specific phishing. It targets a specified or particular person. The traders gather full information about the target and they use malicious emails to trap victims [11].

- Vishing: vishing is a type of voice phishing which intruders call the victim by using a modern call ID and try to convince from different sources. It becomes very difficult for the legal organizations to trace the attackers. The main aim of the attackers is to steal confidential data of the victims [11][12].

- Clone Phishing: The intruders use original text to deceive innocent people. The intruder replaces the original email with a malicious email. When the victims click on an attachment the fake mail redirects to malicious where they are deceived. This is sent to the large no. of users and the attackers keep an eye on who clicks on the attachment that was sent as a mail. This Spread through the contact of the victim who has clicked the fake link provided in the mail [12].

III. BACKGROUND AND LITERATURE REVIEW

In this section, we have discussed the background of the phishing attacks and phishing intentions, and we have also described the literature review for the same domain.

A. History

As stated by APWG [11] [12], Phishing was the phrase used in 1996 for the first time due to an attack against American online accounts stolen from internet con artists. The origin of the word "phishing" phishing in the sense that attackers, or fishers, employ socially engineered communications as bait in order to catch fish and steal victims' personal information. The letter "ph" used in phishing is the replacement of 'f' in fishing [10]. It is because it was among the first forms of hacking that was detrimental to phone networks and it was called phone phreaking. According to APWG, data obtained through hacking by 1997, hackers were exchanging hacking software for compromised accounts, using it as a form of currency. Historically, the origins of phishing attempts were AOL and shifted to target more lucrative objectives, such as online banking as well as e-commerce offerings [13]. Right now, phishing attacks are moreover, limited to targeting end users but moreover, technical employees during services providing could be trapped in advanced techniques like the MITB attacks [13] [14].

B. Phishing Intentions

From the standpoint of the attacker, Weider D. et al. [11] state that the following are the main reasons for phishing attacks:

- Gaining Money: Phishing is done for financial gain by stealing credit card information, login credentials, or other financial data.
- Identity theft: Sometimes phishing is done to impersonate individuals by obtaining their social security numbers, birth dates, or other personal details and creating fake identities for fraudulent activities.
- Reputation damage: It can damage the reputation by spreading false information or leaking sensitive data to manipulate public perception.
- Cyber warfare: Phishing can be used by nation-states or hostilities actors to disrupt critical infrastructure, comprise government systems, or undermine national security.
- Psychological manipulation: Phishing is used in social engineering techniques to exploit human vulnerabilities and their exploitation.
- Ransomware: Attackers encrypt victim's data and ask for money to provide decryption keys.
- Data breaches: Phishing is also used frequently in breaching organizations' networks and accessing their valuable data, intellectual property, and other confidential information which can be exploited in many ways.
- Notoriety and fame: Phishers may target victims in an attempt to gain attention from their peers. [13] [14].

C. Literature Review

Several research has advanced in the field of email phishing detection through creative machine-learning approaches. Li, et al. 2019; study offered a feature-based approach that uses domain age and certain keywords to distinguish between phishing and authentic emails; the system demonstrated promising results in correctly detecting phishing attempts. Using these features, machine learning classifiers are trained as part of the methodology [15].

Garcia, et al., 2020; proposed the use of natural language processing (NLP) in a content-based phishing detection system. It employs linguistic and contextual signals to discern authentic emails from phishing efforts by analysing the correspondence between email content and recognized phishing patterns [16]. Zhang, et al., 2021; study uses feature engineering from URLs, such as domain reputation and URL length, to analyse URLs for phishing detection. The study found that phishing URLs could be identified with greater accuracy when machine learning classifiers were trained using these features [17]. Wang, et al., 2020; study suggested a behavioural analysis technique to enhance overall phishing detection capabilities by tracking user click behaviour within email communications to identify suspicious activity suggestive of phishing attempts [18]. Abu Nimesh, et al., 2018 studied the results of a deep learning strategy by analysing email headers and content using convolutional neural networks (CNNs). According to the study, using deep neural networks increased the precision of identifying

authentic emails from fraudulent ones [19]. Further, an intelligent cybersecurity phishing detection system using deep learning techniques has been proposed. In this study, they use algorithms for machine learning phishing system of detection and applied three datasets where both datasets are dependent upon multifeatured and the third is text-dependent features they are concerned about imbalanced datasets and give more importance to numeric features dataset [27].

D. Phishing's perilous Grip or Statics and Targeted Victim Groups

In the year 2023, each month, the cyber attackers fabricate 1.4 million phishing websites roughly. The most common cybercrime is phishing with an estimated 3.4 billion spam emails sent daily. In 2022, over 48% of emails sent were spam. As per the survey, 92% of the Atlas VPN team organizations were the victim of phishing attempts throughout the previous 12 months and 86% faced dreadful outcomes as a result [21].

More specifically, 72% of cybersecurity executives are concerned about AI being used in email phishing attacks is a matter of concern for cybersecurity leaders are conscious of the use of AI in phishing scams and they are the most alert about phishing campaigns 80% of them showed concern [22]. Approximately 323,972 individuals globally were trapped in phishing in 2021. This accounts for half of all victims of cybercrimes even though practically all phishing efforts are blocked by Google. The act of phishing was the most typical form of attack against Asian organizations in 2021. The typical monetary loss of an astounding $44.2 million was pilfered by crooks, with the victim costing about $136 in 2021. An organization's average loss from a data breach is greater than $4 million. A whaling attack costs a company $47 million.

The most favourable method to trap victims is done via email. In 2021, 16.5 out of every 100 internet users had their emails compromised. These hacked emails act as remuneration for the attackers and they sell them on the black market, where thieves purchase them to employ in their schemes. Phishing is a serious menace even after 1 billion emails were compromised. Businesses must safeguard themselves, especially those in delicate fields like law and banking. According to a 2019 survey, 65% of criminal groups' primary attack tactic was spear phishing, a targeted kind of phishing that is primarily employed for information collection [21][22].

The year 2022, '.com' was the most prevalent email scam that targeted people constituting 54 percent of the whole, as shown in figure 2. The second most typical was '.net', only 8.9%. Initial '.com' adobe, my portfolio, Backblazeb2, and Weebly Google were the involved domain names. Globally, it has been found that phishing fraud has augmented in 2022, by 48.63% in the first quarter and 46.16% in the fourth. A report from 2022 stated that email-based phishing assaults were the primary threat for about 80% of the organizations polled which shows a detrimental effect on organizations, they have noted a 96% rise in email-based phishing attacks in 2022 compared to 2021. December 2022 saw a 45.2% decrease in the email tariff phishing scam, which peaked in February 2022 at 52.78%. The year 2022 witnessed a huge increase in phishing attempts which is around 220%. In this section, we have discussed how phishing scammers target different industries to differing degrees. This is listed below [23].

- Financial organizations: Of all the targets of phishing attempts, 34% are banks, lenders, and other financial organizations. This is probably a result of these organizations' easy access to capital and insightful financial data.
- Healthcare Services: 14% of phishing attempts target the healthcare industry, which includes clinics, hospitals, and manufacturers of medical devices. Cybercriminals find the medical industry to be a lucrative target due to the potential value of confidential medical data on the illicit market.
- Food Retail Market: Eleven percent of phishing attempts target this industry, which includes food manufacturers, pharmacies, and supermarkets. Fake order confirmations or delivery notifications may be part of these schemes.
- Other: 20.5% of phishing attempts target companies in the software services sector, which is in charge of the platforms and apps that many of us use regularly. This might be because these businesses possess significant user data. The opportunity to take advantage of users' confidence in these platforms and the abundance of personal data available make social media sites the target of 12.5% of phishing assaults. 5% of phishing attempts target payment service providers, which manage transactions and retain financial data. Via these schemes,

cybercriminals frequently hope to obtain consumers' payment information. 3.8% of phishing attempts target the logistics industry, which includes delivery and transportation businesses. Scammers frequently.

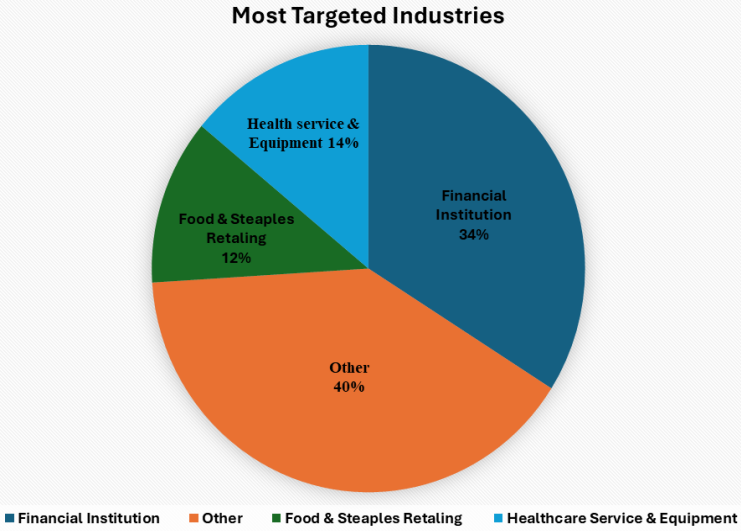


Figure 2. Most targeted industries

E. Targeted Victim's Groups

In this section, we have discussed phishing attacks, which are dishonest attempts to obtain private information by impersonating reliable sources. Phishing attacks target victims according to a number of criteria, such as:

- Employees: Phishing often attempts to impersonate employees because employees are the source to penetrate an organization. The intruders can easily get huge amounts of data which will be further used to deceive employee chief executive officer. According to recent surveys, approximately 72% of phishing attempts specifically target employees [24]

- Senior Citizens: Senior citizens lack cognizance of technological literacy. Which provides opportunities for phishers to impersonate old age people of society. Approx. 15% of phishing attempts are directed towards old age of people because there is more possibility of impersonality senior citizens because they are unaware of digital knowledge [20][24][25].

- Technology Users: Online users, particularly those active in technology such as shoppers, travel agents, and social media users, are common targets for phishing attacks. These attacks are often aimed at collecting personal information, financial details, or login credentials. Survey data indicates that around 60% of phishing attempts target technology users, who are most vulnerable due to their frequent online engagement [25].

- Women: Women are known for the management of personal finance and online shopping. These interests of women are exploited by phishing attacks. Approximately 45% of women are trapped in phishing attacks due to their blind trust in online interactions. They are impersonated on the ground of fashion, parenting, and health [26][27][28].

- Teenagers: Teenagers are immature and inexperienced in the field of technology. They are feeble and easily targeted by phishers. Teenagers use social media and online platforms for social validation and entertainment. The survey findings suggest that around 30% of phishing attempts are aimed at teenagers. Teenagers are trapped by fake social media messages or offers pertaining to gaming or entertainment [20][29][30].

IV. PROPOSED METHODLOGY

Our proposed phishing detection approach follows a structured methodology with five key stages: dataset selection, data preprocessing, dataset splitting, model selection, and model evaluation, as shown in figure 3. We used three publicly available phishing datasets, chosen for their diversity and relevance to phishing detection. The datasets include web page phishing, email spam detection, and spam-ham email classification [38][39][40]. The Web Page Phishing Detection dataset from Kaggle provides information on phishing and legitimate websites,

including features like redirects, HTTPS usage, and suspicious keywords. The Email Spam Detection dataset focuses on classifying spam emails, with features related to suspicious links and attachments, while the Spam-Ham Emails dataset helps detect phishing based on email metadata, such as header anomalies and sender authentication. These datasets were stored in a structured format compatible with our machine learning tools.

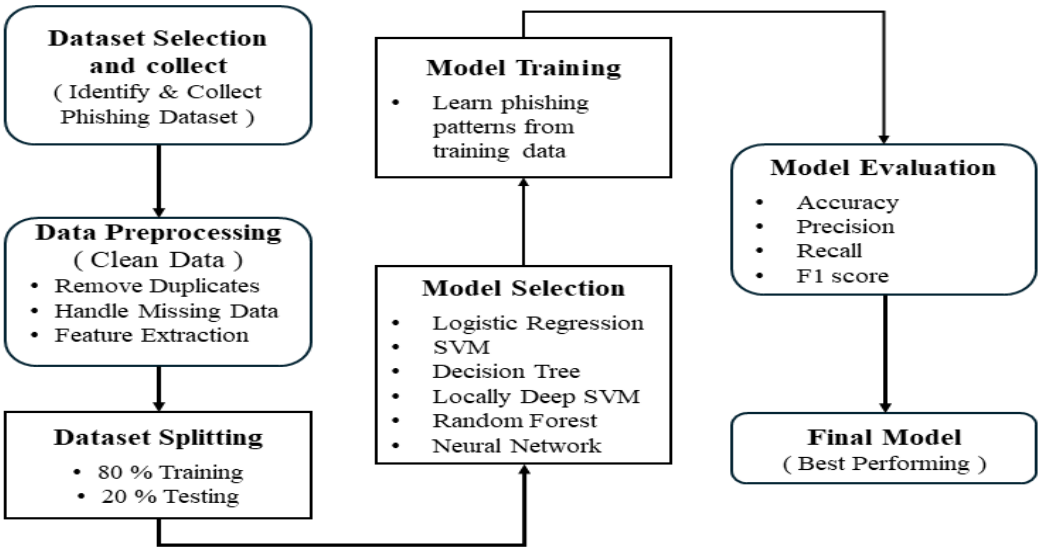


Figure 3. Proposed detection model

The preprocessing phase involved cleaning the data by removing duplicates and handling missing values. We imputed non-critical missing data and eliminated critical data when necessary. Key features were extracted from the datasets, including URL characteristics, email header analysis, content analysis, and metadata features such as WHOIS data and IP tracking. After preprocessing, the data was split into training and testing sets in an 80:20 ratio, using random shuffling and stratified sampling to ensure balanced representation of phishing and legitimate samples.

We evaluated several machine learning and deep learning models, including traditional models like Random Forest, Decision Tree, SVM, and Logistic Regression, as well as advanced models such as Locally Deep SVM. The models were trained using TensorFlow and Scikit-learn, with hyperparameter tuning performed through Grid Search and Random Search to optimize performance. Model evaluation was based on key metrics: accuracy, precision, recall, and F1-score, which allowed us to assess how well the models distinguished between phishing and legitimate emails.

A. Techniques Used

We chose six different algorithms for training and testing how accurately we can detect phishing emails using grouped features. We picked these methods because they use various training strategies to figure out the rules and how learning and testing work. The algorithms listed below are widely recognized.

- **Logistic regression:** It predicts the probability of an observation belonging to a class, such as spam or not spam, using a logistic function. It optimizes parameters to minimize classification error and assigns labels based on a threshold, making it a simple, efficient, and interpretable classification method [31].

- **Support Vector Machines (SVM):** It classifies data by finding the optimal hyperplane that maximizes the margin between classes, as shown in figure 4. Using the kernel trick, SVM handles both linear and non-linear separation. It formulates classification as a convex optimization problem, ensuring robust and efficient performance [32].

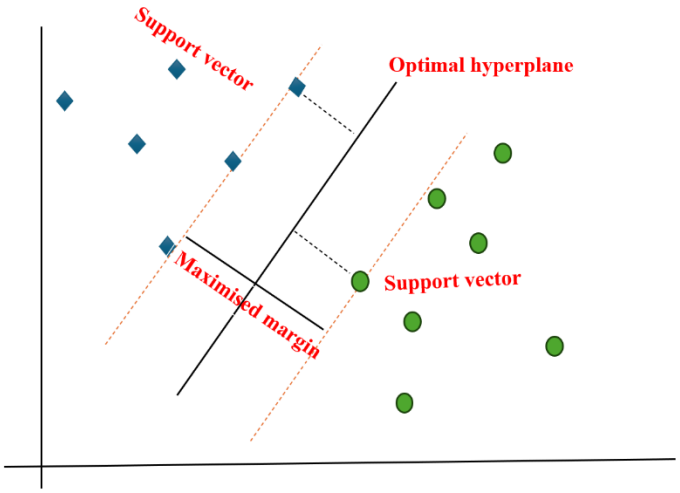


Figure 4. Support Vector Machine (SVM)

- **Decision Tree:** Decision trees are versatile machine learning models for classification and regression. They split data recursively based on features that maximize information gain or minimize Gini impurity, as shown in figure 5. The process continues until a stopping criterion is met. Decision trees are interpretable, making them useful for understanding decision paths [33].

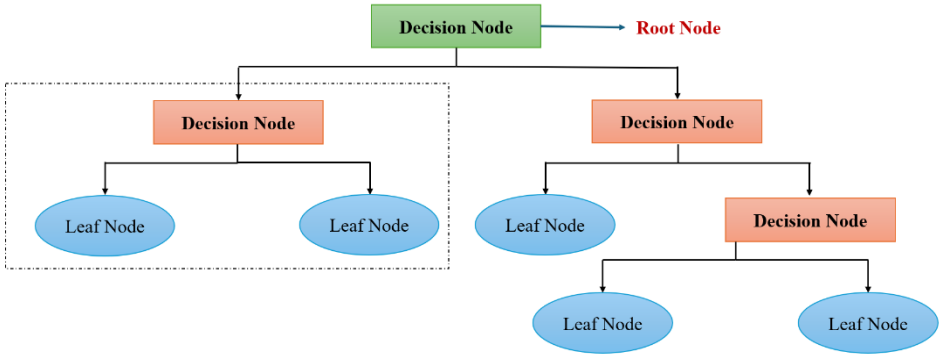


Figure 5. Decision Tree

- **Random forest:** It is an ensemble learning algorithm for classification and regression that constructs multiple decision trees using random subsets of data and features, as shown in Figure 6. It reduces overfitting by averaging predictions, enhancing accuracy and robustness. Known for scalability and resilience, it effectively handles high-dimensional data [34].

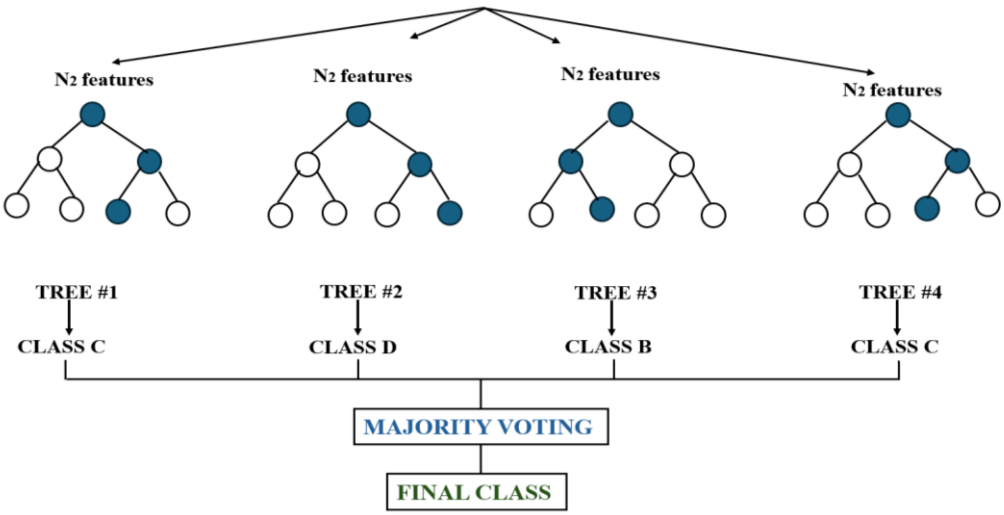


Figure 6. Random Forest

- **Locally deep SVM:** Locally Deep SVM integrates deep neural networks with traditional SVM, enhancing its ability to capture complex nonlinear patterns while maintaining interpretability, as shown in figure 7. By embedding deep learning locally around support vectors, it improves classification performance, particularly for high-dimensional and nonlinear data [35].

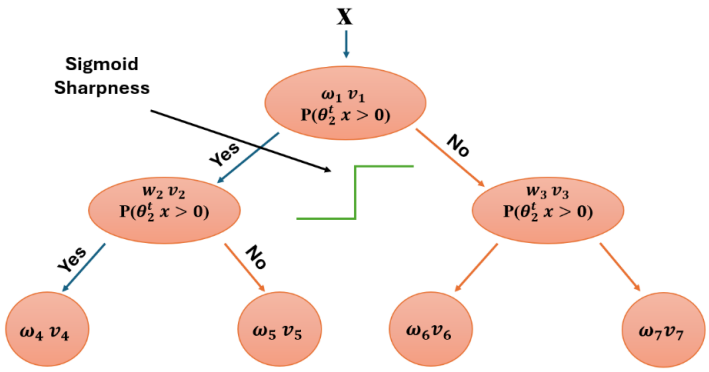


Figure 7. Locally deep SVM

- **Neural Network:** Neural networks mimic the human brain using interconnected layers of artificial neurons to process data and make predictions, as shown in figure 8. They learn by adjusting weights through backpropagation, enabling them to capture complex patterns. Their versatility makes them effective for tasks like classification, regression, image recognition, and natural language processing [36].

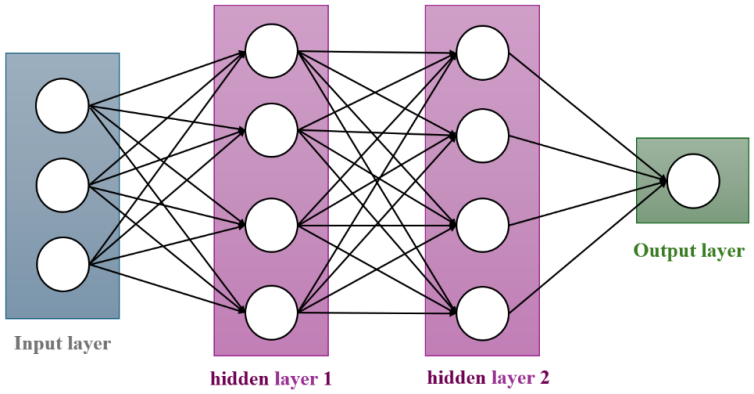


Figure 8. Neural Network

B. Problem Statement

As phishing assaults become more frequent and sophisticated, there is a pressing need for efficient detection techniques to safeguard individuals and organizations from cyber threats. Phishing attacks, which typically involve deceptive emails, websites or messages created to deceive people into disclosing private information or downloading harmful software pose a significant risk to cybersecurity. Existing detection methods, while somewhat successful frequently find it difficult to stay up with changing phishing tactics and techniques.

The challenge lies in developing robust and adaptive phishing detection techniques that can accurately identify fraudulent content and activities in real-time. This involves addressing several key issues including the detection of previously unseen phishing attempts the classification of legitimate versus malicious URLs, the identification of social engineering tactics, and the integration of contextual information to enhance detection accuracy. Additionally, the detection techniques must be scalable, efficient, and compatible with diverse platforms and environments.

Furthermore, as phishing attacks continue to evolve in complexity and sophistication there is a need for continuous research and innovation to stay ahead of cyber threats. This includes exploring new technologies including natural language processing, deep learning and machine learning, and behavioral analysis to enhance the efficacy of phishing detection systems.

V. EXPERIMENT AND RESULTS

Environment and Tools: This study utilized Python and Jupyter Notebook for data analysis, model development, and experimentation, chosen for their flexibility, extensive machine learning support, and widespread adoption in research.

- **Python:** Python is a versatile, high-level programming language widely used in scientific computing and machine learning. Its simple syntax enhances readability, while extensive libraries like NumPy, Pandas, Scikit-learn, TensorFlow, and PyCaret support data analysis and model development. Python's cross-platform compatibility ensures smooth execution across operating systems, and its popularity in academia provides strong community support, facilitating research and collaboration.

- **Jupyter Notebook:** Jupyter Notebook is an interactive computing environment that supports iterative experimentation and model development. It enables step-by-step execution, integrates with visualization tools like Matplotlib and Plotly, and facilitates data exploration. Additionally, it supports reproducibility and collaboration, allowing easy sharing via platforms like GitHub for transparency and peer review.

A. The chosen phishing datasets

We have used three types of datasets [38][39][40] as follows to evaluate the effectiveness of various machine learning algorithms in phishing detection, we conducted three experiments with different datasets. Each experiment aimed to assess the performance of multiple models, and the results were compared using common classification metrics: Accuracy, Precision, Recall, and F1-Score. Below, we describe the experiments and provide insights into the comparative performance of the models.

1) Web Page Phishing Detection Dataset

The first dataset [38] we utilized is the 'phishing' dataset. It consists of 11430 instances. Out of which 5715 are phishing URLs and 5715 are legitimate URLs as shown in Table 1. 87 different features are extracted from the URLs to explore and understand the relationship between the structure of the URL and its authenticity.

Table 1 Phishing Dataset

	Train	**Test**
Legitimate	4586	1157
Phishing	4558	1129
Total	9144	2286

2) Email Spam Detection Dataset

The second 'spam' dataset [39] we used to consist of 5572 emails. Out of them, 4825 are ham messages and 747 are spam messages as shown in Table 2. The ratio of ham vs spam messages is 6:1. The dataset has only 2 columns, the other column contains the binary label ham and spam.

Table 2 Spam or ham Dataset

	Train	**Test**
Spam	593	154
Ham	3864	961
Total	4457	1115

3) Spam-Ham Emails Dataset

The third 'email' dataset [40] we used to consist of 5728 emails. Out of them, 4360 are non-spam emails and 1368 are spam emails as shown in Table 3. The ratio of non-spam vs spam mail is 3:1. The dataset has only 2 columns, the other column contains the label of spam as binary 0 for non-spam and 1 for spam.

Table 3 Email Dataset

	Train	Test
Spam	1091	277
Non-spam	3491	869
Total	4582	1146

B. Results

1) Experiment 1

For the first experiment, we used the Web Page Phishing Detection Dataset from Section 5.A.1). This dataset consists of 89 variables, of which many were not highly correlated with the dependent variable, 'Status' (phishing or not). To improve the model's performance, we calculated the correlation of all variables with the target variable and discarded those with correlations below a predefined threshold. After this process, 10 variables were retained for modelling. Applied different ML algorithms to the dataset and the resultant accuracies are represented in Table 4:

Table 4 A Comparative Study Based on Experiment 1

Algorithm	Accuracy	Precision	Recall	F1 Score
Logistic Regression	92.08	92.43	91.87	92.15
SVM	70.42	76.39	60.15	67.31
Decision Tree	93.13	93.47	92.91	93.19
Random Forest	95.1	95.15	95.15	95.15
Locally Deep SVM	90.33	92.62	87.89	90.19
Neural Network	90.76	91.93	89.62	90.76

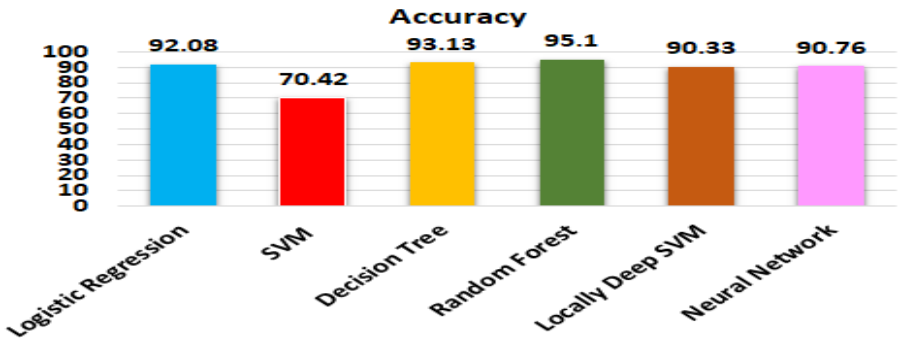


Figure 9. The first experiment outcomes

In this experiment, Random Forest outperformed all other models with the highest accuracy of 95.10%. On the other hand, SVM showed the lowest accuracy of 70.42%, highlighting its suboptimal performance for phishing detection in Figure 9. Locally Deep SVM, which combines the benefits of deep learning and SVM, performed relatively well with an accuracy of 90.33%, but it did not surpass Random Forest. The decision tree performed similarly to Random Forest but showed slightly lower precision and recall values. The relatively higher performance of Logistic Regression compared to SVM suggests that linear models might still be effective for certain phishing datasets, despite being less complex.

2) Experiment 2

For the second experiment, we used the Email Spam Detection Dataset from Section 5.A.2), which consists of text data (independent variables) and binary labels ('ham' and 'spam'). We applied natural language processing

(NLP) techniques to convert the text data into a format suitable for machine learning algorithms. After preprocessing, we applied multiple machine learning algorithms to the dataset. The resultant accuracies of all models are represented in Table 5:

Table 5 A Comparative Study Based on experiment 2

Algorithm	Accuracy	Precision	Recall	F1 Score
Logistic Regression	98.11	98.12	98.12	98.07
SVM	97.13	97.22	97.13	96.99
Decision Tree	97.48	97.47	97.49	97.42
Random Forest	97.48	97.53	97.49	97.39
Locally Deep SVM	98.83	98.85	98.83	98.81
Neural Network	98.65	98.68	98.65	98.63

In this experiment, Locally Deep SVM achieved the highest accuracy (98.83%) and demonstrated superior precision and recall values compared to all other algorithms. Neural Network also performed very well with an accuracy of 98.65%, as it is clear from Figure 10. SVM, although performing reasonably well, did not show the same level of accuracy as Locally Deep SVM or Neural Networks. The high performance of Locally Deep SVM can be attributed to its ability to combine the strengths of deep learning (for feature extraction) with the robust decision boundaries of SVM.

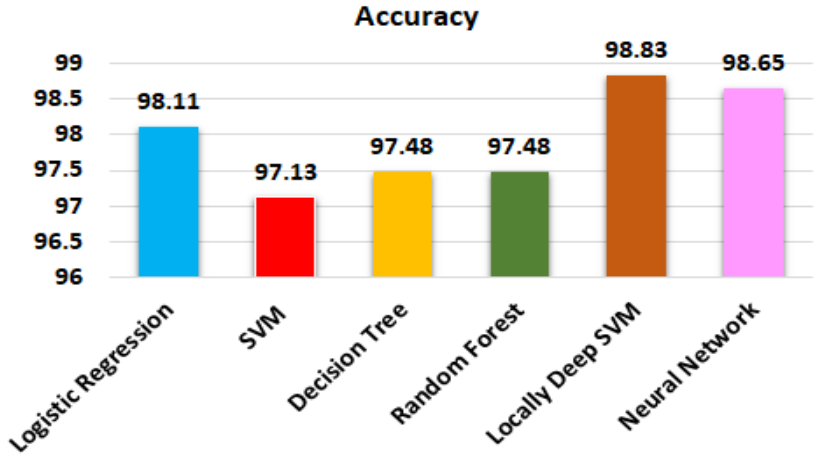


Figure 10. The second experiment outcomes

3) Experiment 3

For the third experiment, we used the Spam-Ham Emails Dataset from Section 5.A.3), which also contained text data. Similar to Result 2, we employed natural language processing techniques and pre-processed the data to apply the selected machine learning algorithms. The resultant accuracies of all models are represented in Table 6.

Table 6 A Comparative Study Based on Experiment 3

Algorithm	Accuracy	Precision	Recall	F1 Score
Logistic Regression	98.60	98.60	98.60	98.60
SVM	97.20	97.22	97.21	97.17
Decision Tree	96.68	96.76	96.68	96.71
Random Forest	97.90	97.94	97.91	97.88
Locally Deep SVM	98.08	98.07	98.08	98.08
Neural Network	98.69	98.69	98.69	98.69

In this experiment, Neural Network achieved the highest accuracy of 98.69%. Logistic Regression and Random Forest also performed well, with accuracies of 98.60% and 97.90%, respectively. Locally Deep SVM performed slightly worse than Neural Networks but still outperformed traditional models like SVM and Decision Trees and it is clear from Figure 11.

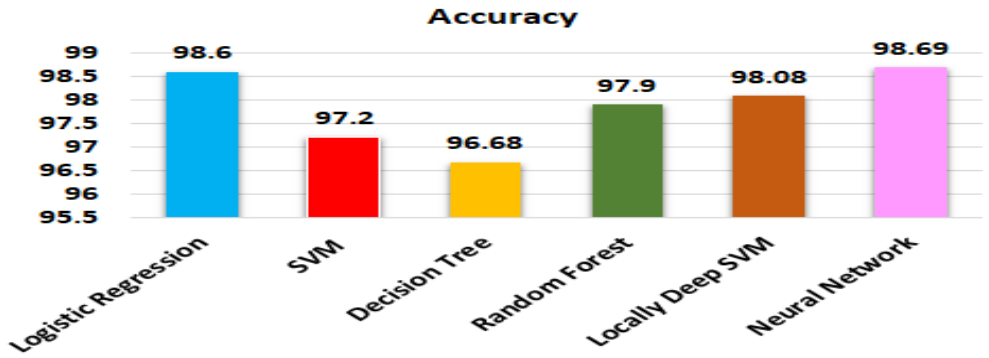
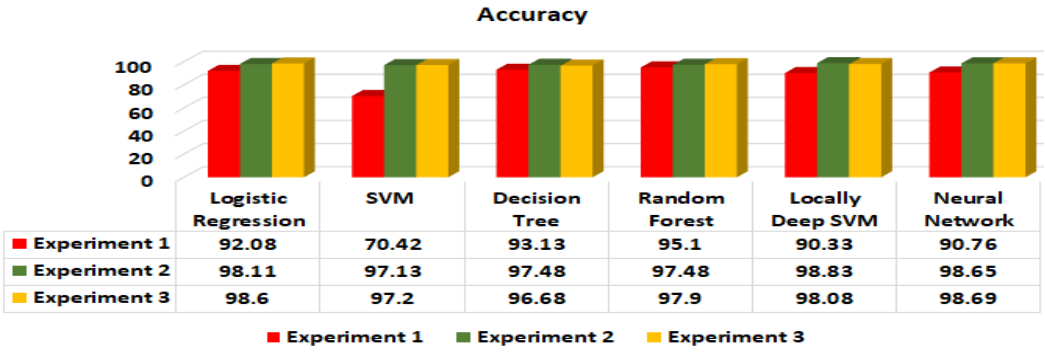


Figure 11. The third experiment's outcomes

C. Comparative Analysis of All Experiments

Model performance varied based on dataset characteristics. Random Forest achieved the highest accuracy (95.10%) for structured phishing detection, while deep learning models, particularly Locally Deep SVM (98.83%) and Neural Network (98.69%), excelled in text-based spam classification. Logistic Regression performed well in spam detection but struggled with phishing data. SVM showed the weakest performance in Experiment 1 but improved in text-based tasks. Ensemble methods, especially Random Forest, remained strong across datasets. Feature selection in Experiment 1 further enhanced model performance, the results are summarised in the following Figure 12.

Overall, deep learning models were superior for text classification, while ensemble methods were more effective for structured phishing detection. Future research should explore hybrid approaches integrating deep learning and ensemble methods for enhanced classification.



	Logistic Regression	SVM	Decision Tree	Random Forest	Locally Deep SVM	Neural Network
Experiment 1	92.08	70.42	93.13	95.1	90.33	90.76
Experiment 2	98.11	97.13	97.48	97.48	98.83	98.65
Experiment 3	98.6	97.2	96.68	97.9	98.08	98.69

Figure 12. Comparison of the three outcomes

D. Discussion

The findings from our experiments highlight the importance of feature selection and machine learning approaches for phishing and spam detection. In the first experiment, using URL-based features with the Support Vector Machine (SVM) model resulted in a low accuracy of 70.42%, suggesting that URL features alone may not be sufficient for detecting phishing attacks. This underscores the need for additional contextual features, such as text content or behavioural patterns, to improve performance.

In contrast, the second and third experiments, which focused on text-based features, demonstrated significantly better accuracy (98.83%) with the use of natural language processing (NLP). The models effectively handled imbalanced (6:1) and balanced (3:1) class distributions, showcasing the robustness and adaptability of NLP methods across different datasets. The consistency in accuracy across varying class distributions emphasizes the potential for deploying these models in real-world scenarios.

Overall, this study highlights the superiority of NLP-based text features over URL-based approaches for phishing and spam detection. Future work could explore hybrid models combining both URL and text features to further enhance accuracy, as well as real-time detection capabilities and evaluation on larger, more diverse datasets for broader applicability in dynamic cybersecurity environments.

VI. Conclusion and future work

This study evaluated machine learning algorithms for phishing and spam detection using three datasets. Random Forest achieved the highest accuracy (95.10%) for phishing detection, while Locally Deep SVM (98.83%) and Neural Networks (98.69%) excelled in spam classification. Natural language processing (NLP) techniques outperformed URL-based feature extraction, demonstrating robustness against class imbalances. These findings suggest that NLP-based approaches are highly effective for phishing and spam detection, while URL-based methods can be improved with additional contextual features. Future work should explore hybrid models integrating deep learning and ensemble methods for enhanced accuracy and real-time detection.

While this study provides important insights, there are several areas for future research. One promising direction is the development of hybrid models that combine deep learning with ensemble methods, which could improve both feature extraction and decision-making for more accurate detection systems. Additionally, enhancing URL-based models with contextual features, such as website metadata and user behaviour, could improve phishing detection accuracy. Future work should also focus on adapting models to address the evolving nature of phishing tactics, including the development of defences against adversarial attacks. Optimizing machine learning models for real-time applications, particularly in areas like email filtering and web security, is another key area of exploration, focusing on scalability and low-latency performance. Furthermore, the applicability of these models could be expanded to other platforms like social media and messaging systems. Lastly, ethical considerations around data privacy are crucial as machine learning models are deployed in security systems, and future research should prioritize privacy-preserving techniques to ensure effective and ethical use of these models in real-world applications.

Declarations: The authors confirm that ethical approval and consent to participate are not applicable to this study. All authors have agreed to the publication of this manuscript. The availability of supporting data has been discussed in Section 4. The authors declare no competing interests in relation to this work. Furthermore, no funds, grants, or financial support were received during the preparation of this manuscript.

Regarding authorship contributions, SM and MA were responsible for conceptualizing the main idea and developing the methodology of the model. ST and BA handled the implementation, data acquisition, and dataset processing. Each author has contributed equally to this research. All authors have thoroughly reviewed and approved the final version of the manuscript for publication.

References

[1] James, L. T., & Rajamani, S. (2020). Phishing URL Detection Through Top-Level Domain Analysis: A Descriptive Approach. Journal of Computer Security, 28(1), 150-167.

[2] Chiew, K. L., Tan, C. L., Wong, K., & Yong, K. S. C. (2019). A Survey of Phishing Attacks: Their Types, Vectors, and Technical Approaches. Expert Systems with Applications, 106, 1-20.

[3] Liu, D., & Wang, Z. (2019). Deep Learning Based Phishing URL Detection: A Structural Approach. IEEE Transactions on Dependable and Secure Computing, 17(2), 362-379.

[4] Ahmed, H., Traore, I., & Saad, S. (2019). Detecting Phishing Websites Using a Decision Tree Machine Learning Approach. Soft Computing, 23(5), 1721-1733.

[5] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. Expert Systems with Applications, 117, 345-357.

[6] Al-Zubi, S., Aqel, D., Lafi, M.: An intelligent system for blood donation process optimization-smart techniques for minimizing blood wastages. Clust. Comput. 2022, 1–11 (2022). https://doi.org/10.1007/s10586-022-03594-3

[7] Aqel, D., Al-Zubi, S., Mughaid, A., Jararweh, Y.: Extreme learning machine for plant diseases classification: a sustainable approach for smart agriculture. Clust. Comput. 2021, 1–14 (2021). https://doi.org/10.1007/s10586-021-03397-y

[8] Srivastava, S., Singh, A.K.: Fraud detection in the distributed graph database. Clust. Comput. 2022, 1–23 (2022). https://doi. org/10.1007/s10586-022-03540-3
[9] Obeidat, I., Mughaid, A., Alzoubi, S.: A secure encrypted pro tocol for clients' handshaking in the same network. Int. J. Inter act. Mob. Technol. 13, 47–57 (2019)
[10] Mahmoud Khonji, Youssef Iraqi, Senior Member, IEEE, and Andrew Jones. (2013). Phishing Detection: A Literature Survey.
[11] Hong, J.: The state of phishing attacks. Commun. ACM 55(1), 74–81 (2012)
[12] Maqableh, M., Alia, M.: Evaluation online learning of under graduate students under lockdown amidst covid-19 pandemic: the online learning experience and students' satisfaction. Child Youth Serv. Rev. 128, 106160 (2021)
[13] C. Whittaker, B. Ryner, and M. Nazif, "Large-scale automatic classifi cation of phishing pages," in NDSS '10, 2010.
[14] X. Dong, J. Clark, and J. Jacob, "Modelling user-phishing interaction," in Human System Interactions, 2008 Conference on, may 2008, pp. 627–632.
[15] Li, J., Wang, Y., & Sheng, S. (2019). Feature-based email phishing detection using machine learning. Journal of Cybersecurity and Information Management, 3(2), 45-58.
[16] AlZu'bi, S., Al-Qatawneh, S., Alsmirat, M.: Transferable hmm trained matrices for accelerating statistical segmentation time. In: Proceedings of the 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), pp. 172–176. IEEE (2018)
[17] Garcia, M., Lopez, R., & Martinez, J. (2020). Content-based phishing detection using natural language processing techniques. International Journal of Information Security, 8(4), 321-336.
[18] AlKhatib, A.A., Sawalha, T., AlZu'bi, S.: Load balancing tech niques in software-defined cloud computing: an overview. In: Proceedings of the 2020 Seventh International Conference on Software Defined Systems (SDS), pp. 240–244. IEEE (2020).
[19] Zhang, Q., Liu, W., & Chen, H. (2021). URL analysis for phishing detection: A machine learning approach. Journal of Computer Security, 15(3), 189-204.
[20] Singh, S.; Singh, M.P.; Pandey, R. Phishing detection from URLs using deep learning approach. In Proceedings of the 2020 5th International Conference on Computing, Communication and Security (ICCCS), Patna, India, 14–16 October 2020.
[21] Wang, L., Zhang, H., & Chen, G. (2020). Behavioral analysis for email phishing detection using machine learning. IEEE Transactions on Dependable and Secure Computing, 17(3), 690-704.
[22] Bhat, V.H., Malkani, V.R., Shenoy, P.D., Venugopal, K., Patnaik, L.: Classification of email using beaks: behaviour and keyword stemming. In: Proceedings of the TENCON 2011-2011 IEEE Region 10 Conference, pp. 1139–1143. IEEE (2011)
[23] Abu-Nimeh, S., Nappa, D., & Wang, X. (2018). Deep learning for email phishing detection. IEEE Transactions on Information Forensics and Security, 13(5), 1255-1267.
[24] Teli, S.P., Biradar, S.K.: Effective email classification for spam and non-spam. Int. J. Adv. Res. Comput. Softw. Eng. 4, 2014 (2014)
[25] Atlas VPN. (2023). Phishing Statistics 2023. Retrieved from [https://atlasvpn.com/blog/phishing-statistics-2023] (https://atlasvpn.com/blog/phishing-statistics-2023)
[26] Verizon. (2022). 2022 Data Breach Investigations Report. Retrieved from [https://enterprise.verizon.com/resources/reports/dbir/] (https://enterprise.verizon.com/resources/reports/dbir/)
[27] Mughaid, A., AlZu'bi, S., Hnaif, A., Taamneh, S., Alnajjar, A., & Abu Elsoud, E. (2022). An intelligent cybersecurity phishing detection system using deep learning techniques.
[28] Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L.F., Downs, J.: Who falls for phish? a demographic analysis of phishing susceptibility and effectiveness of interventions. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp. 373–382 (2010)
[29] Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L.F., Hong, J.: Teaching johnny not to fall for phish. ACM Trans. Internet Technol. 10(2), 1–31 (2010)
[30] Kabali, H.K., Irigoyen, M.M., Nunez-Davis, R., Budacki, J.G., Mohanty, S.H., Leister, K.P., Bonner, R.L.: Exposure and use of mobile media devices by young children. Pediatrics 136(6), 1044–1050 (2015)

[31] kpea, O. (2020). Logistic Regressions Regularized by Elastic Net, SVM, Random Forest, and Artificial Neural Networks for Heart Disease Prediction. *The University of Texas at San Antonio Journal of Undergraduate Research and Scholarly Works*, 8, 1-12.
[32] Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
[33] Kotsiantis, S. B. (2013). Decision Trees: A Recent Overview. *Artificial Intelligence Review*, 39(4), 261-283.
[34] Belgiu, M., & Drăguţ, L. (2016). Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24-31.
[35] Wang, J., & Perez-Cruz, F. (2011). Training Support Vector Machines Using the SVM-Path. *Proceedings of the 2011 European Signal Processing Conference (EUSIPCO)*, 623-627.
[36] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436-444.
[37] Lone, A. N., Mustajab, S., & Alam, M. (2023). A comprehensive study on cybersecurity challenges and opportunities in the IoT world. Security and Privacy, 6(6), e318.
[38] Tiwari, S. (2023). Web Page Phishing Detection Dataset [Data set]. Kaggle. https://www.kaggle.com/datasets/shashwatwork/web-page-phishing-detection-dataset
[39] Qureshi, M. F. (2023). Email Spam Detection – 98% Accuracy [Notebook]. Kaggle. https://www.kaggle.com/code/mfaisalqureshi/email-spam-detection-98-accuracy/input
[40] Tung, N. V. (2023). Spam-Ham Emails Dataset [Data set]. Kaggle. https://www.kaggle.com/datasets/tungnv01/spam-ham-emails.