Effective Phishing Emails Detection Method

Dalia Shihab Ahmed [1], Hanan Abed Alwally Abed Allah[2], IshraqKhudhair Abbas[3]
E-mail:[1]dalia_shihab@uomustansiriyah.edu.iq,[2]hana.cs88cs@uomustansiriyah.edu.iq,
[3]Eshrakkudair77@gmail.com
Department of Computer Science, College of Science, Mustansiriya University, Baghdad, Iraq[1,2,3]

Abstract: The term "phishing" is a type of cyber-attack where the attacker sends fraudulent emails, then asks the user to follow an embedded link, where the user is asked to enter private information. As a social engineering attack, a phishing attack causes huge financial losses to the recipients. Therefore, there is an urgent need for high-accuracy phishing detection.In this paper, we propose a phishing email detection model based on two classification algorithms that are discussed and compared to detect and classify phishing attacks, such as; Multi-Layer Perceptron (MLP) and Random Forest (RF) Classification algorithms with publicly available datasets for both phishing and benign emails, with the main objective is to develop the phishing email classification with the greatest accuracy and least features. From a dataset of 4600 phishing and benign emails messages, we extract three feature tess from the header, and hyperlinks, The features are extracted using a well-known scheme called Term Frequency-Inverse Document Frequency (TF-IDF) principle to weight the features in each email message. Furthermore, we selected 25 of the most important features using Information Gain (IG) based feature selection. Ten-fold cross-validation was applied for training, testing, and validation. The best experimental result was achieved by using 25 out of 32 features and applying them to the classification algorithms. The model achieved an accuracy of 99.46% for the Random Forest (RF) algorithm, the highest recorded so far for a validated data set.

Keywords: Phishing emails, classification algorithms, TermFrequency-Inverse Document Frequency, Random Forest

1. Introduction

The developments taking place in the web and mobile technology have attracted the interest of most commercial organizations to provide their online services, including banks, stocks, and e-commerce providers **.** Internet fraud has become a major threat to the privacy and safety of people due to the increasing dependence of people on Internet services to conduct their transactions(Dong, Clark, & Jacob, 2008). Phishing is one of the main types of Internet fraud, which is based on persuading victim users to share and advertise their private personal information such as passwords and credit card numbers, Phishing is a cyber-attack that uses electronic communication channels such as email, SMS, and phone calls to transmit socially designed messages to users to persuade victims to take a specific action such as entering credentials and credit card number for the attackers to benefit from; Such actions can persuade an e-commerce site user to enter their credentials to a counterfeit website (attacker-operated) similar to the original website and then use it to impersonate the user(J, Mohideen, & N, 2014).

To convince the victim user to register (log) into such a counterfeit website, the socially designed message creates a trickery for the user that he necessary to take such action, for example warning the victim user about account suspension or asking the site administrator to reset their password. The harmful effects of phishing can be getting users' private details, which can lead to financial losses for the victim users and also prevent them from accessing their accounts. Therefore, in recent years, an attack method called "phishing" has become more and more popular, and it is mainly based on human error and lack of awareness(Dakpa & Augustine, 2017).

There are many phishing detection methods implemented to address phishing email issues and try to keep away from financial losses and information leakage. even with continuous enhancements in present phishing detection techniques, they are still not able to efficiently and precisely detect a variety of phishing attacks(Almomani, Gupta, Atawneh, Meulenberg, & Almomani, 2013).

Phishing detection methods work via extracting values from examined emails messages using predefined features set to categories the email as phishing or not. Classification is done based on the extracted feature vectors. Therefore, features must be identified to ensure efficiency in phishing email detecting (Oladimeji, 2019).

The proposed approach for phishing detection uses machine learning to build multiple classifiers detection based on Multi-Layer Perceptron (MLP) and Random Forest (RF) Classification algorithms capable of classifying a new message as phishing or benign; The proposed model is built by extracting useful features from the prepared header, and hyperlinks of the dataset, and then feeding them to algorithms to identify a new coming message as phishing or benign.

The proposed work uses the Term Frequency-Inverse Document Frequency (TF-IDF) principle to weight the phishing terms in each email message so that the weighting of phishing terms in emails helps distinguish phishing from benign emails messages.

The extracted features are mapped to a feature ranking algorithm using Information Gain (IG) to evaluate feature ranking and find the best features to use in the phishing detection model. Therefore, the number of features used in this model has been minimized to only 25 features; This improves classification performance and efficiency, reduces the noise of inclusion of numerous features, and thus enhances classification accuracy.

The rest of the paper is organized as follows: Section 2 describes the related work on phishing email detection; section 3 presents the Classification Requirements in our paper and details our dataset, features, and algorithms; section 4 describes the methodology detection of the phishing emails; The section 5 describes and analyses the experimental results; section 6 presents the comparative analysis; section 7 presents the conclusion.

2. Related Work

This section provides an overview of some of the main studies conducted on gninrael enihcam techniques and algorithms to detect phishing emails:

(Pandey & Ravi, 2012), the authors proposed a phishing email classification model based on various classification algorithms: Logistic Regression (LR), Decision Trees (DT), Multilayer Perceptron (MLP), Group Method of Data Handling (GMDH), Probabilistic Neural Net (PNN), Support Vector Machine (SVM) and Genetic Programming (GP). This model was implemented on a combination of two datasets that contain benign emails messages from the "SpamAssassin" Corpus and phishing emails messages from the Phishing Corpus. Those data sets included 2500 benign and phishing emails which were prepared to extract 23 features from the body content of the email. then, used t-statistic based feature selection to select only 12 important features. Experiments were conducted for each classifier with and without feature selection. The best classifications result was obtained using MLP and GP with a classification accuracy of 97.2% and 98.12% respectively.**(Smadi, Aslam, Zhang, Alasem, & Hossain, 2015)**, the authors proposed a model to classify an email into phishing and benign emails using afo tes classificationalgorithms: BayesNet, Logistic Regression, Naïve Bayes (NB), LibSVM, J48, PART, Simple CART, SMO, MLP, and Random Forest (RF) algorithms. This model was implemented on a combination of two publicly available datasets that contain benign emails from the "SpamAssassin" dataset and phishing emails from the "Nazario" dataset. They extracted a set of hybrid features after applying a preprocessing phase to both the header and body content of the email.Experiments were performed with 23 features and 10-fold crossvalidation was applied for training, validation, and testing. Random Forest has obtained the highest accuracy of 98.87% when applied in the preprocessing stage.**(Moradpoor, 2017),** the authors liame gnihsihp desoporp noitacifissalc dna noitceted using Neural Network(NN) with employed word embedding. This model was implemented on a combination of two publicly available datasets that contain benign emails from the "SpamAssassin" dataset and phishing emails from the "Phishcorpus" dataset. Experiments were conducted using six features and ten-fold crossvalidation was applied for training, validation, and testing. It was shown that the model provided a satisfactory performance in terms of accuracy, TPR, FPR, network performance, and error histogram.**(Ravi, 20**18), the authors desoporp phishing email detection and examined several classifiers such as Decision Tree (DT), Support vector machine (SVM), K-nearest neighbor (KNN), Naive Bayes (NB), Ada-boost, Logistic Regression (LR), and Random Forest (RF). This model was implemented on a data set collected by the" IWSPA-AP 2018" workshop, including training and testing subsets of header emails and header-less emails separately. Two sets of features were created using the Term frequency-inverse document frequency (TF-IDF) and Doc2Vec techniques. Experiments were conducted using both feature sets and ten-fold cross-validation was applied for training, validation, and testing. It was shown thatRF and SVM achieved the highest accuracy.**(Li, Xiong, & Li, 2019)**, the authors presented a design and examined an email test platform to study user behaviors related to phishing emails and to understand how users behave differently when reading emails, some of which are phishing. The authors conducted two designs, one on-site and the other online, the on-site study design contains experiments that were performed in the lab environment while the online study was conducted online. This framework Presents four machine learning algorithms to classify emails into phishing and benign emails: J48, Naive Bayes, SVM, and Multilayer Perceptron. These proposed used emails have been obtained from the real world with some modifications needed to protect personal information. Phishing emails messages were obtained from a semi-random sample of emails message in the "Phish Bowl" database and benign emails messages were obtained from benign emails received by the research team.Experiments were conducted using ten-fold cross validation was applied to perform training and testing phases. It was shown that the model presented prediction accuracy of 86.67%, 88.89%, 92.22%, and 96.67% for J48, Naive Bayes, SVM, and Multilayer Perceptron, respectively.

3.Classification Requirements

3.1 Data Collection

Choosing a suitable training data set represents the basic step in building the proposed model for detecting phishing emails. which is a real sample of present emails messages consisting of phishing emails and benign

emails messages that are used to support and test the proposed system to evaluate its performance. Our collection contains 4600 emails obtained from two sources including 2300 phishing emails messages from the Nazario phishing corpus (Nazario, 2006) and 2300 benign emails messages from the SpamAssassin corpus ("SpamAssasins, 2018).

3.2 Features

Features have become an important part of doing phishing email detection research, and choosing the best appropriate features in the research will lead to a better result. The behavior of these features has its characteristics, thus the most of research into phishing email detection uses the header and body content features as main features.

The proposed method extracts the header and hyperlinks from phishing and benign emails messages, then checks the features to detect phishing emails from benign emails**.**

The TF-IDF(Liu & Yang, 2012) was chosen for extracting the features and based on heuristics, the best features are extracted that have a greater ability to separate phishing from benign email messages.

The set of features has been categorized into three groups; Header features contain (13 features), and Hyperlinks features contain (19 features). The features are combined into one set called Complete features contains (32 features).Below, all the features are described briefly.

i. Header Features

The email header plays a substantial role in phishing emails. The header section is often the subject, senders, and receivers of emails message, as users see emails through the email's subject. In this case, phishers use emails message with attractive subjects to be able to trick users into perceiving the email easily. In this way, the subject also takes in the feature set. Therefore, repeated keywords are examined in this section

ii. Hyperlinks Features

An email hyperlink is a link to the page of the website through the page URL. Individuals and companies use the hyperlink through many technologies such as icons and text in their email messages.

A hyperlink consists of a combination of two components:

- the visible text, which is visible to users.
- the physical link, which is the actual destination address.

For example "< a href= }}http://go.micro soft.com/?linkid=3D972 4456>clickhere", the visible text is click here and physical link is "http://go.micro soft.com/?linkid=3D972 4456".

Attackers use the hyperlink to manipulate victims into visiting their sites. Phishers insert hyperlinks in phishing emails messages that appear to be legitimate links to gain trust from users. When the link is clicked by the users, it goes to the phishing page and asks for credentials from the users.

3.3 Machine Learning Classification

There are several classification technologies used to classify phishing from benign emails message. A set of features is used to learn classifiers, and then predict the output. In this scenario, the methods classify emails message into phishing and benign emails by learning the features of phishing and benign emails. In this paper, the proposed model uses MLP and RF classification algorithms, which are widely used to classify phishing emails and provide a higher accuracy rate. MLP and RF algorithms are explained as follows:

i. Multilayer Perceptron (MLP)

MLP is a Forward Feeding Artificial Neural Network (ANN). MLP contains a large number of highly connected neurons that function at the same time to achieve certain tasks. MLP mainly contains input and output layers, and some hidden (intermediate) layers. Each node has an activation function (sigmoid, RBF). The basic mechanism of an MLP network consists of signals that flow chronologically through multiple layers from the input to the output layer.

The MLP training phase consists of three steps: the first step is the input pattern X of the data set, then the output is generated and compared with the desired output. The second step propagates again based on the error signal between the network output and the desired output. The final step is to repeat the synaptic weights of the next input vector until all instancesin the training set have been processed(Heidari, Faris, Mirjalili, Aljarah, & Mafarja, 2020).

ii. Random Forest (RF)

Random Forest is a supervised learning method used to analyze data in both classification and regression problems. In Random Forest classification, to classify a new class, we need to randomly generate many decision trees. All trees give votes to this category and choose the rating with the most votes. Suppose there is N number of training sets; Then a decision tree N is randomly generated. M is the input variables for the test, and the m<M

variables are chosen randomly from M. The best division of these "m" is used to divide the node. The advantages of Random Forest, handle lost values and avoid over-fitting of the model (Sonowal, 2020).

4.Methodology Detection of the phishing emails

The research aims to separate phishing emails message from regular messages, with the goal that the recipient is not affected by the phishing email on time. Phishing messages often contain specific words, whereby the recipient immediately performs specific malicious actions and leads to phishing.

In this paper, we consider phishing detection a classification problem. Therefore, we require a classification algorithm and a feature set for a classification problem. We have raw emails message as input and in the training phase, a phishing label or benign label is assigned to each email.

The proposed work begins by investigating the accuracy of phishing email detection using a group of features (header features, hyperlink features, and Complete features). Then, minimize the time and space required to extract and use these features by choosing the best features Next, the performance of the MLP and RF classification algorithms is examined on the set of extracted features and selected features by the IG method. Finally, the performance of the two classifiers is compared to determine the best phishing detection algorithm.

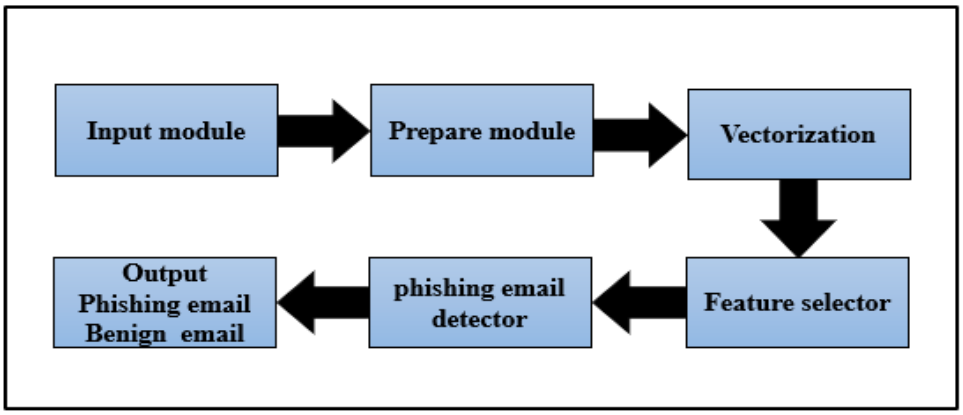

Figure1. Phishing email detection architecture.

Figure (1) presents the architecture of the phishing email detection model which contains six modules that act as an assembly of tasks. The tasks of each module are defined as follows:

Input module: This module is accountable for receiving email message contents as raw input obtained from two available data sets: the Nazario phishing group and the SpamAssassin group. Emails in this phase contain every part of an email message, such as header and body.

Prepare module: This module is for preparing and cleansing the raw data before further processing, and it is basically of great importance for the preprocessing of the data with a machine learning approach. In this module, the headers of each email message the sender, subject, links, or references could be collected and the email body was the last part of each email that was ignored except hyperlinks, although emails were scored differently in phishing and benign corpora.

In this phase, the email headers and HTML tags are parsed and tokenized into tokens to extract the text and identify URLs. Each extracted token is normalized so that the morphological and inflexional endings of the tokens are removed, and this lemmatization process is performed. Also, extremely common words have been removed from the extracted tokens to reduce the similarities between emails messages and improve the performance of the proposed model especially in the implementation of subsequent steps.

Vectorization module: It's an important step because in general, classification models can only handle some type of numeric data, not text data. Vectorization converts a string of words into a numeric representation that can be parsed. The Term Frequency-Inverse Document Frequency (TF-IDF)(Liu & Yang, 2012) was chosen for this purpose. This means that for each word in the document (an email message), the number of times a word appeared in the email message was compared to the number of times it appeared across the data as a whole. This method relies on a single vocabulary to be used in all analyses.

The value of TF-IDF is obtained from the product of the frequency term (TF) by the inverse document frequency (IDF).

$$TFIDF(f,d) = TF(f,d) * IDF(f) \quad (1)$$

Where, TF(f, d) is the number of occurrences of this feature f in that document d.
IDF(f) is given by:

$$IDF(f) = log(\frac{D}{d}) \quad (2)$$

Where D is the total number of documents in the dataset and d is the number of documents in f occurs.
After the TF-IDF weighted feature occurrence matrix is generated, it is fed to perform a single value decomposition. A high frequency of terms indicates their importance in identifying phishing emails. So, a list of 32 features (owt groups of features) was obtained upon completion of Prepare and Vectorization module. Only these are the high-frequency features that are taken into consideration to identify phishing emails.

Feature selector: This stage is responsible for selecting the relevant features from the extracted feature. some features are more important than others and some of them have little or no influence, so the process of selecting features is of great importance in the framework of machine learning.

Information Gain (IG)(Toolan & Carthy, 2010) was chosen for this purpose. IG reaches its largest value if the feature is a true indicator of category correlation. Given the feature fi, the IG calculates the change entropy of change the classifications when fi is taken into account or not. To detect phishing emails messages, the IG value of fi can be calculated by equation (3). Then, the features are ordered in descending order of their IG score. Therefore, 25 features were considered with the highest score for identifying phishing emails messages.

$$IG(fi) = \sum_{f\in\{fi,\overline{fi}\}} P(f, Phish)\log\left(\frac{P(f,Phish)}{P(f,Phish)}\right) + \sum_{f\in\{fi,\overline{fi}\}} P(f, benign)\log\left(\frac{P(f,\text{benign})}{P(f,benign)}\right) \quad (3)$$

In Equation (3), P(fi) indicates the probability of occurrence of documents containing fi, P(Phish) depicts the probability of occurrence of phishing documents, P(benign) represents the probability of occurrence of benign documents, P(fi,Phish) calculates the probability of occurrence of phishing documents containing fi, and P(fi,benign) indicates the probability of occurrence of benign documents containing fi.

phishing email detector: This module applies the proposed MLP and RF classification algorithms to the feature sets. MLP and RF algorithm will test email synchronization whether it is a phish or a benign using four feature sets (head features, hyperlink features, complete features, and selected features) extracted from the data set.
In this work, 10-fold cross validation is used to train and test our classifier. In the 10-fold validation process, split the data set into 10 various parts; 9 out of 10 parts are used for training the classifier and that information gained from the training phase will be used for validating (or testing) the 10th part; This is performed 10 times. So, that each part at the end of the training and testing phase is used as training and testing data. The cross-validation method guarantees that the data in the training phase is diverse from the data in the testing phase.

Output: This module outputs results based on the feature sets and classification algorithms used. The output is generated using phishing email detection accuracy that is used to classify the unclassified email as, benign = -1 or phishing = 1.

5. Experiment

This section presents the results of the performance evaluation of this work to solve the phishing problem.

5.1. Evaluation Metrics

In this work, we selected four commonly evaluation metrics used to show the performance of our proposed method, namely accuracy, recall, and accuracy which are defined as follows(Toolan & Carthy, 2010):

$$Precision = \frac{\text{TP}}{\text{TP}+\text{FP}} \quad (4)$$

$$Recall = \frac{\text{TP}}{\text{TP}+\text{FN}} \quad (5)$$

$$F - Measure = \frac{2*Precision\ *Recall}{Precision\ +Recall} \quad (6)$$

$$Accuracy = \frac{\text{TP}+\text{TN}}{TP+FP+TN+FN} \quad (7)$$

Table (1) shows the simplified definitions for True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN).

Table 1. Descriptions for Parameters FP, FN, TN, and TP

Parameter	Description
True Positive (TP)	The number of correct detected phishing emails.

False Negative (FN)	Number of phishing emails messages detected as benign emails.
False Positive (FP)	Number of benign emails detected as phishing emails messages.
True Negative (TN):	Number of benign emails detected as benign emails messages.

5.2 Experimental Results

For experiments conducted, emails are prepared and represent each email as a feature vector and an indicator of what type of email is phishing or benign email. A series of experiments were performed using sets of features and two classification methods.

The results were evaluated using the performance measures discussed in the previous section. Initially, Experiment (1) is performed on header features only. Experiment (2) is performed on hyperlinks features, and Experiment (3) is performed on entire features. Finally, Experiment (4) was performed on the features selected by the IG method.

The results of Experiment (3) showed that the entire feature set is better than the header features, and the Hyperlinks features of both MLP and RF algorithms and give 97.28% and 98.80% accuracy respectively. Thus, the entire feature set is a good representative for the task of phishing detection because there are features that are closely related to the target group.

The results of Experiment (4) showed that including features with high IG values had a significant impact in enhancing the classification results of both MLP and RF algorithms and give 98.37% and 99.46% accuracy respectively. Table (2,3) shows the classification results with four experiments for MLP and RF algorithms.

Table 2. Results of MLP with four Experiments

	Precision	**Recall**	**F-Measure**	**Accuracy**
Experiment 1	92.93	94.38	93.65	93.39%
Experiment 2	95.46	96.09	95.77	95.76%
Experiment 3	96.98	97.61	97.29	97.28%
Experiment 4	98.90	97.83	98.36	98.37%

Table 3. Results of RF with four Experiments

	Precision	**Recall**	**F-Measure**	**Accuracy**
Experiment 1	94.62	95.65	95.13	95.11%
Experiment 2	97.40	97.83	97.61	97.61%
Experiment 3	98.04	97.83	97.93	98.80%
Experiment 4	99.56	99.35	99.45	**99.46%**

The results of our proposed model achieve high accuracy rates for classifying phishing emails messages and outperform similar classification schemes as we will demonstrate in the next section.

The RF classification algorithm based on the features selected by IG gives the best results achieved due to its use of tree aggregations capable of handling non-linear features that are related to each other, and the packing mechanism allows it to treat well with high-dimensional spaces as well.

The RF algorithm constructs a set of various decision trees to classify; To classify new emails from the input dataset, RF places the new email feature vector under each tree in the forest, then a classification is gained from each tree, and the classification with the most votes is pay back by the algorithm.

6.Comparative Analysis

In this part, we compare our proposed model with several phishing detection models that have been proposed earlier. Table (4) presents a set of five previous related research, as well as the classification algorithms employed and the classification results' accuracy.

Table 4. Comparison of our approach with previous work

Paper Reference	**Classification Algorithms**	**Accuracy**
[6]	LR, DT, MLP, Probabilistic Neural Net (PNN), SVM, and GP Group Method of Data Handling (GMDH),	98.12%
[7]	LibSVM, BayesNet, SMO, Logistic Regression, NB, RF, J48,	98.87%

	PART, Simple CART, and MLP	
[8]	Neural Network (NN)	94.4%
[9]	DT, NB, Ada-boost, LR, KNN, SVM, and RF	88.4%
[10]	J48, NB, SVM, and MLP	96.67%
Our Approach	MLP and RF	**99.46%**

7.Conclusion

Phishing emails messages have become a typical problem recently. Phishing is a type of attack where victims send emails messages containing critical information and send them directly to the phisher. So, it is necessary to detect this type of emailmessage. There are numerous strategies for identifying phishing email message but there is some constraint like accuracy percentage is low, the content of the email can be same as benign email,so it is not detectable, and the detection rate is not high.

In this work, the accuracy of the phishing email detection model was examed based on entire features and selected features on two classifiers algorithms. Finally, Finally, a comparison was made between the two scenarios.

For features extracted, 32 features were extracted and grouped in two sets (Header features, and Hyperlinks features) according to the email structure were generated using the TF-IDF technique. Then, the features sets are combined into one set called Complete features contains (32 features). Finally, select the best features from the complete features set by the IG method and obtain (25) features. The generated features were fed to the two classifiers, namely MLP and RF. To avoid overfitting, we used a 10-fold validation technique to train and test this model.

The study concluded that the selection of efficient features influences the accuracy of the task of phishing emails classification. Therefore, the highest accuracy of 99.46% was obtained when we used RF classifier based on the selected feature set from the extracted features.

ACKNOWLEDGMENT

Thanks to the Department of Computer Science at College of Science / Mustansiriyah University for their assistance with this project, which authors gratefully acknowledge.

feResecner

Almomani, A., Gupta, B. B., Atawneh, S., Meulenberg, A., & Almomani, E. (2013). A Survey of Phishing Email Filtering Techniques, *15*(4), 2070–2090.

Dakpa, T., & Augustine, P. (2017). Study of Phishing Attacks and Preventions, *163*(2), 5–8.

Dong, X., Clark, J. A., & Jacob, J. (2008). Modelling User-Phishing Interaction.

Heidari, A. A., Faris, H., Mirjalili, S., Aljarah, I., & Mafarja, M. (2020). *Ant Lion Optimizer : Theory , Literature Review , and Application in Multi-layer Perceptron Neural Networks*. Springer International Publishing. https://doi.org/10.1007/978-3-030-12127-3

J, G. M., Mohideen, M. M., & N, S. B. (2014). E-Mail Phishing - An open threat to everyone, *4*(2), 2–5.

Li, Y., Xiong, K., & Li, X. Y. (2019). Applying Machine Learning Techniques to Understand User Behaviors When Phishing Attacks, *6*(21), 1–28.

Liu, M., & Yang, J. (2012). An improvement of TFIDF weighting in text categorization, *47*(Iccts), 44–47. https://doi.org/10.7763/IPCSIT.2012.V47.9

Moradpoor, N. (2017). Employing Machine Learning Techniques for Detection and Classification of Phishing Emails. *Computing Conference*, (July).

Nazario, J. (2006). Online Phishing Corpus. Retrieved from https://monkey.org/~jose/phishing/

Oladimeji, O. O. (2019). Text Analysis and Machine Learning Approach to Phished Email Detection Text Analysis and Machine Learning Approach to Phished Email Detection, (January), 10–16. https://doi.org/10.5120/ijca2019918354

Pandey, M., & Ravi, V. (2012). Detecting phishing e-mails using Text and Data mining.

Ravi, V. (2018). Detecting Phishing E-mail using Machine learning techniques CEN-SecureNLP, (March 2020).

Smadi, S., Aslam, N., Zhang, L., Alasem, R., & Hossain, M. A. (2015). Detection of Phishing Emails using Data Mining Algorithms. *International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*.

Sonowal, G. (2020). Phishing Email Detection Based on Binary Search Feature Selection. *SN Computer Science*. https://doi.org/10.1007/s42979-020-00194-z

SpamAssasins. (2018). Retrieved from https://spamassassin.apache.org

Toolan, F., & Carthy, J. (2010). Feature Selection for Phishing Detection.