Cluster Computing (2022) 25:3819–3828
https://doi.org/10.1007/s10586-022-03604-4

An intelligent cyber security phishing detection system using deep learning techniques

Ala Mughaid[1] · Shadi AlZu'bi[2] · Adnan Hnaif[2] · Salah Taamneh[1] · Asma Alnajjar[1] · Esraa Abu Elsoud[1]

Received: 17 December 2021 / Revised: 20 April 2022 / Accepted: 22 April 2022 / Published online: 14 May 2022

Abstract

Recently, phishing attacks have become one of the most prominent social engineering attacks faced by public internet users, governments, and businesses. In response to this threat, this paper proposes to give a complete vision to what Machine learning is, what phishers are using to trick gullible users with different types of phishing attacks techniques and based on our survey that phishing emails is the most effective on the targeted sectors and users which we are going to compare as well. Therefore, more effective phishing detection technology is needed to curb the threat of phishing emails that are growing at an alarming rate in recent years, thus will discuss the techniques of mitigation of phishing by Machine learning algorithms and technical solutions that have been proposed to mitigate the problem of phishing and valuable awareness knowledge users should be aware to detect and prevent from being duped by phishing scams. In this work, we proposed a detection model using machine learning techniques by splitting the dataset to train the detection model and validating the results using the test data , to capture inherent characteristics of the email text, and other features to be classified as phishing or non-phishing using three different data sets, After making a comparison between them, we obtained that the most number of features used the most accurate and efficient results achieved. the best ML algorithm accuracy were 0.88, 1.00, and 0.97 consecutively for boosted decision tree on the applied data sets.

Keywords Cyber security · Phishing · Machine learning · Classifier · Algorithms

1 Introduction

Cybercrime refers to crimes that target computer or network. Computer crimes coated of a broad range of potentially criminal activities. Phishing is the most commonly used attack on social engineering. Through such attacks, the phisher tries to obtain confidential information from the user, with the purpose of using it fraudulently against user's [1, 2]. In today's digitized business world, more and more companies are taking advantage of the ever-evolving opportunities of cyberspace. Due to the growth of internet technology in our daily basis especially due to covid-19 impacts that forced all users to use more the internet in all sectors. Phishing is metaphorically similar to fishing in the water, but instead of trying to catch fish, attackers try to steal the user's personal information. Phishing websites look very similar to their corresponding legitimate websites to attract large numbers of Internet users. Recent developments in phishing detection have led to the growth of many new approaches based on visual similarity.

✉ Ala Mughaid
ala.mughaid@hu.edu.jo

Shadi AlZu'bi
smalzubi@zuj.edu.jo

Adnan Hnaif
adnan_hnaif@zuj.edu.jo

Salah Taamneh
Taamneh@hu.edu.jo

Asma Alnajjar
2070595@std.hu.edu.jo

Esraa Abu Elsoud
2070606@std.hu.edu.jo

[1] Department of Information Technology, Faculty of prince Al-Hussien bin Abdullah for IT, The Hashemite University, P.O. Box 330127, 13133 Zarqa, Jordan

[2] Faculty of Science and IT, Al-Zaytoonah University of Jordan, Amman, Jordan

Machine learning and modern AI techniques have been effeciently employed in several human life applications [3, 4], many previous researchers employed machine learning in security fields such as in [5–9]. Computer security attacks are classified into three types: physical attacks, synthetic attacks, and semantic attacks. Phishing is one of the semantic attack types [10]. In such attacks, the vulnerabilities of the users are targeted; for example, the way users interpret computer messages, because most users read information sources without verification and respond to their requests. Phishing is a type of social engineering attack often used to steal user data which is used to access important accounts and can result in identity theft and financial loss. It occurs when an attacker, posing as a trusted legitimate institution, dupes a victim through communication channels. The user is then lured into clicking a malicious link, which can cause the installation of malware, the freezing of the system as part of a ransomware attack, and revealing of sensitive information.

Phishers carry out their attacks by using E-mail"phishing" which is the most common channel for phishing and reverse social engineering attacks, Instant messaging "smseshing" are gaining popularity among social engineers as tools for phishing and reverse social engineering attacks, Telephone, Voice over IP "vishing" are common attack channels for social engineers to make their victim deliver sensitive information ,These are designed to lead consumers to counterfeit Web sites that trick recipients into divulging financial data such as usernames and passwords.

For example, according to PhishTank [11]: "Phishing is a fraudulent attempt, usually made through email, to steal your personal information" This definition restricts phishing attacks aimed at stealing personal information, which is not always the case. For example, a socially designed message could lure the victim into installing the message in the Browser malware. The software will transfer money to the attacker's bank account, whenever the victim logs in to perform his banking duties without having to steal the victim's personal information. Therefore, we believe that PhishTank's definition is not broad enough to cover the entire issue of fraud. Another definition is provided by Colin Whittaker et al. [12]: "We define a phishing page as any web page that, without permission, alleges to act on behalf of a third party with the intention of confusing viewers in to performing an action with which the viewer would only trust a true agent of the third party" ColinWhittaker et. Definition aims to be broader than PhishTank's definition in the sense that attacker'targets are no longer limited to stealing personal information from victims. On the other hand, the definition continues to limit phishing attacks to those acting on behalf of third parties, which is not always true. For example, phishing attacks can deliver socially designed messages to lure victims to websites that need to serve safe content, luring victims to install MITB malware. When the Work-group is loaded, it can record keystrokes to steal the victim's passwords. Note that the attacker in this scenario does not claim the identity of any third parties in the phishing process, only transmits messages with links that will lure victims to view videos or multimedia content. To address the limitations of the previous definitions above, we consider phishing attacks to be semantic attacks that use electronic communication channels to convey socially engineered messages. to convince the victim to take some action for the attacker's benefit.

According to APWG phishing attack trends reports [13, 14], the number of phishing attacks observed by APWG and its members grew through 2020, doubling over the course of the year. Phishing are spread via e-mail, SMS, instant messaging, social networking etc., but e-mail is a popular way to carry out this attack. The phishing email can lead to financial loss. Attacker always sending email tends to make user believe that they are communicating with trusted entity and deceive them into providing personal credentials in order to access service, such as credit card numbers, account login credential or identity information. In 2019, 293.6 billion emails were sent and received daily. This includes billions of promotional emails sent by merchants every day. While many email users believe that such content belongs in their spam folder, marketing emails are generally harmless if they are uncomfortable for the users. Spam messages accounted for 47.3 percent of e-mail traffic in September 2020 which caused serious economic losses and social problems. [www.statista.com] Spam e-mail It is almost impossible to think about email without considering the problem of spam. The world's most common variants of malicious spam include Trojan horses, spyware, and ransomware. There are many approaches that have been developed to deal with the spam problem [15].These days, three ways to mitigate such attacks stand out: Focus based on awareness, based on blacklists, and based on machine learning (ML). However, in the last days, Deep Learning (DL) has emerged as one of the most efficient techniques of machine learning [16].

In Sect. 1 of this paper, we applied machine learning on three different data sets where the first two datasets depend on multi features and the third one depends on text feature only. Section 2 we review the Related Work of classifiers used in detecting phishing emails, in Sect. 3 we mentioned the targeted victims in phishing. The methodology that has been followed to do this research has been introduced in Sect. 4. Section 5 presents the experiments for classifying Phishing Email Using Machine Learning, Finally, the work is concluded in Sect. 6.

2 Literature review

Ease of communicating with advent of email caused the problem of unsolicited bulk email, especially phishing attacks via emails. Various anti-phishing techniques have been developed to solve the problem of phishing attacks. This paper focuses on separating important emails from spam. One of the main factors for classification is how messages will be represented. Specifically, you need to decide which features to use and how to use those features when categorizing them. many researchers have employed AI in intelligent system, and many of them used the Deep learning in cybersecurity applications [17–20].

Fette et al. proposed an email filtering approach called PILFER which considered 10 features set including URL based and Script based features to detect phishing attacks [21]. By filtering phishing emails before they are read by users, it can reduce the percentage of users being fraudulent. Phishers can hide the URL and use tools like TinyUrl to make the URL appear valid. Phishers are becoming increasingly sophisticated in their approach and incorporating strategies to bypass existing anti-phishing tools.

Bhat et al. [22] came up with an approach which derives spam filter called Beaks. They classify emails into spam and no spam. Their pre-processing technique is designed to identify tag-of-spam words relevant to the dataset.

Kang Leng et al. proposed anti-phishing tools which depending on nine features derived from structure-based and behavior-based features. They are the sender of the domain's name, the words blacklisted in the subject and the content, the IP address in the URL, the dot in the URL, the symbol in the URL, the unique sender, the unique domain name, the hyperlink not. consistency and return path. All recommended features are selected based on the phishing technique commonly used by phishers, which achieve accuracy of 97.25 [23–25].

Teli [26] have showed a three-step system they designed for spam detection methods to classify each new coming emails according to the given algorithm as spam or legitimate email. Ram Basnet et al. [27] employed machine learning algorithms to detect a phishing attack by classifying phishing emails and legitimate emails. they have used sixteen features, the dataset that they used contains 4000 instances with a ratio of 0.75 legitimate emails and 0.25 phishing ones. they split the dataset into 0.50 training and the remaining as a test. the accuracy that they got was 97.99.

Moradpoor et al. [28] have used two datasets that contains 14,370 emails (benign/phishing) , in there detection and classification of phishing emails model based on neural network , the overall accuracies and inaccuracies come to 92.2%. Smadi et al. [29] The authors proposed a model to detect phishing emails by extracting 23 features, they compared different algorithms where random forest achieved the highest accuracy of 98.8%.

In this work, we applied machine learning techniques, to capture inherent characteristics of the email text and other features to be classified as phishing or non-phishing according to the selected data-sets.

3 Targeted victims

In below sections, we summarize some of the identified characteristics of potential phishing victims based on previous studies:

A. Victim's Age: performed a role-play demographics and phishing susceptibility. They found that participants' age linearly predicts their susceptibility to phishing. Older one was less likely to fall prey for phishing, while younger users particularly between the age of 18–25 consistently more vulnerable to phishing attack [30–32].
B. Victim's Gender: Most studies showed that women are more likely to fall for phishing attack than men [30, 31, 33]. Jagatic et al. conducted a real phishing attack experiment on 1731 students from Indiana University their result showed that 77 percentage of female students fell for attacks while 0.65 of the male students fell for the same attack [33].
C. Victim's ant-phishing education: Kumaraguru et al. evaluated Phish Guru which is an embedded anti-phishing training system with 515 participants in a real-world study which analyzed response over a course of 28 days [32].
D. Victim's General Educational: a study reported by Kumaraguru et al, showed that users with computer science backgrounds performed slightly better than users with other backgrounds when being attacked by phishing.
E. Ant-phishing training delivery method: A study showed that users learned more effectively when the training material was presented after they fell victims for the phishing experiment [34, 35].
F. Victim's Personality: Table 3 summarizes all the factors we found to be correlated with susceptibility to phishing attacks.

Factors	High susceptible	Less susceptible
Age	18–24 years old or less	25 years old or more
Gender	Female	Male
Anti phishing training	No training	Anti-phishing trained

Factors	High susceptible	Less susceptible
Education	Humanities	Computer Since
Training delivery method	Non-embedded	Embedded
Personality	Agreeableness	Consciousness
Internet usage behaviour	E-commerce and online banking	E-mails and simple browser

Young people are accountable for their usage of their devices, but they should know how to protect their devices security. A study with 83 teenagers found that teenagers were poor at distinguishing between legitimate and phishing messages in an experimental task. participants exhibited riskier behavior while making decisions on unfamiliar messages. In a Cross-sectional study of 350 children aged 6 months to 4 years, results show most households had television (0.97), tablets (0.83), and smartphones (0.77). At age 4, half the children had their own television and their own mobile device. Almost all children (96.6) used mobile devices, and most started using before one-year-old. Parents gave children devices when doing house chores (0.70), to keep them calm (0.65), and at bedtime (0.29). At age 2, most children used a device daily. Most 3- and 4-year-old's used devices without monitoring, and one-third engaged in media multitasking [36]. In addition, children aged 5–15 years go online for a minimum of 8 h every week. Examples of common online activities for this age children include communicating through social media, watching YouTube videos, and playing games [37]. One of the primary digital risks that children need to be aware about is phishing, a common social engineering attack ranked as one of the most dangerous online risks for children. More than 1 million children (below 17 years of age) in 2017 in U.S. alone were victims of identity theft with estimated costs of 2.6 billion dollars Several efforts have been made towards designing mechanisms and training tools to help protect people against phishing.

With the prevalence and potential consequences of phishing, continuous efforts are made to improve the cybersecurity knowledge of citizens and to develop protections against phishing attacks. Researchers have explored technical solutions, awareness through cybersecurity educational games and training materials, the addition of cues in the user interface to aid in phish detection. As energetic users of social media websites, teens can regularly share a number of data while interacting and communicating with each other. This sharing may be risky The number of phishing attacks has expanded through the years. A recent survey on nearly 15,000 end-users from seven countries showed that 0.83 of the respondents had experienced a phishing attack in 2018 compared to 0.76 in 2017. Phishing severely impacts businesses; mid-sized companies pay an average of 1.6 million dollars to recover from a successful phishing attack where the consequences include malware infections, compromised accounts, and data loss. More than 1 million children in the U.S. under the age of 17 fell victim to identity theft in 2017 costing approximately 2.6 billion dollars.

Cain et al. made a study on people aged 18 to 55 years and observed that younger people have poor cyber security habits related to password management and phishing. Adults have been shown to have poor calibration between confidence and actual performance when it comes to identifying phishing messages which can then increase the likelihood that the attack is successful [38].

4 Methodology

Our methodology is categorized into the following phases: Datasets Collection, Datasets Preprocessing, Using machine learning classification techniques. We proposed models to classify emails as each model has been built with different functions based on the three datasets with different features. With high probability and filter out legitimate emails as little as possible. This promising result is superior to the existing detection methods and verifies the effectiveness of our models in detecting phishing emails. Experiments phases are presented in details in a separate subsection. Figure 1 illustrates the structure of the proposed detection model.

As The first most important step is to have the required dataset, we've used three datasets which has been taken from publicly available resources. The reason for using three datasets with different features is the high rate of changing phishing attack techniques which increases the difficulty of detecting and filtering phishing email attacks. In order to be able to classify the phishing emails and to identify how number of features will affect the efficiency

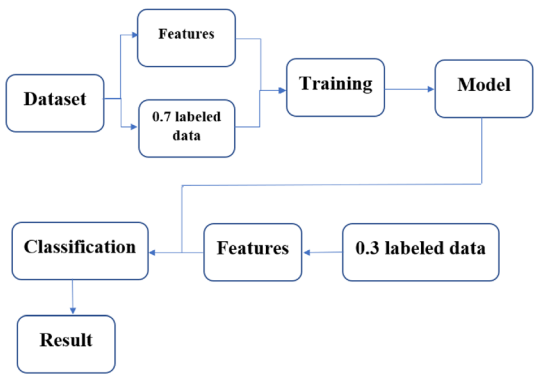


Fig. 1 Proposed detection model

of the model to detect the phishing emails. After collecting data, we reprocessed the datasets by removing the duplicated rows, removing missing values and balancing the instances to achieve the most accurate rate. After importing the processed dataset, we split it into 0.70 to train and 0.30 to test the model. In many cases, the selected ratio is 0.70 of the training set and 0.30 of the test. The idea is that more training data is a more professional as because it makes the classification model better, and more test data makes the error estimate more accurate.

The second phase can be called the training stage. Here the classifier model with the help of the inserted ML algorithm will be trained using the 0.70 dataset to manually classify entered data into a spam or legitimate emails. When the first and second phases are completed, it begins to classify emails according to the given algorithm as spam or legitimate email.

5 Experiments

We have applied the above-mentioned model Fig. 1 to the three selected datasets but because of the different features for each dataset we had to use different functions in each model for each dataset so we concluded with seven models as we used seven ML algorithms to compare them to obtain the highest accuracy.

5.1 Techniques used

Seven supervised classification algorithms were selected, to train and test the accuracy of phishing email detection with the grouped features. The reason behind selecting these algorithms are the different training strategy were used for discovering the rules and the mechanism of learning and testing. The below listed algorithms are considered as well-known algorithms:

1. Locally-deep support vector machine.
2. Support vector machine.
3. Boosted decision tree.
4. Logistic regression.
5. Averaged perceptron.
6. Neural network.
7. Decision forest .

5.2 The chosen phishing datasets

5.2.1 Phishing email collection

This utilized dataset was the first dataset we've used, its consists 5,25,754 instances with 8351 as phishing emails and 5,17,402 legitimate emails. Which mean that 98.4% from data are legitimate and the dataset is imbalance. So, to avoid over-fitting we have to reprocessing that data and investigate the balance between phishing and legitimate instances. This dataset consists of 22 features (Total Number of Characters C-Vocabulary Richness W/C - Account-Access-Bank-Credit-Click-Identity-Inconvenience-Information-Limited-Minutes-Password-Recently-Risk-Social-Security-Service-Suspended-Total number of Function words/W-Unique Words-Phishing Status). For that, we chose a random sample with 8351 phishing and 8400 legitimate instances. Then we split the data to 0.70 train and 0.30 test, as detailed in the following Table 5.2.1 :

	Train	Test
Legitimate email	5881	2520
Phishing email	5846	2506
Total	11,727	5026

5.2.2 Phishing legitimate full

We decided to use here another dataset with different features as this dataset consists of 10,000 instances with 5000 phishing emails and 5000 legitimate emails. The dataset has 50 features (id-NumDots-SubdomainLevel PathLevel-UrlLength-NumDash-NumDashInHostname-AtSymbol-TildeSymbol-NumUnderscoreNumPercent-NumQueryComponents- ... etc.) We split the data to 0.70 train and 0.30 test as shown in the flowing Table 5.2.2:

	Train	Test
Legitimate email	3498	1502
Phishing email	3502	1498
Total	7000	3000

5.2.3 Spam or not spam dataset

The third dataset consists of 2500 ham and 500 spam emails, all the numbers and URLs were converted to strings as NUMBER and URL respectively. This is the simplified spam and ham dataset. We split the data to 0.70 train and 0.30 test as shown in the flowing Table 5.2.3:

	Train	Test

	Train	Test
Spam	351	149
Ham	1749	751
Total	2100	900

5.3 Experiments results

Experiment 1: The first experiment depends on first dataset were mentioned earlier in Sect. 5.2.1. The dataset here consists of 22 features but it was imbalance as it contained 5,25,754 instances as legitimate emails were about 98.4%. So we had to reprocessed it before import it to our model to make sure of getting the best accurate result. So after reprocessing the dataset we applied our model on the chosen random sample with 8351 phishing and 8400 legitimate instances and splitting the data as 0.70 train and 0.30 test. The seven selected ML algorithms were imployed on the first processed dataset sample using AZURE ML Microsoft tools. The results are represented in the following table for the first experement:

Algorithm	Accuracy	Precision	Recall	F-Score
Locally-deep support vector machine	0.83546	0.91663	0.73703	0.81708
Support vector machine	0.81616	0.89749	0.71269	0.79448
Boosted decision tree	0.88818	0.89099	0.88388	0.88742
Logistic regression	0.81417	0.92349	0.68396	0.78588
Averaged perceptron	0.79586	0.88066	0.68316	0.76944
Neural network	0.80661	0.88895	0.69952	0.78294
Decision forest	0.86968	0.87485	0.86193	0.86834

The first experiment showed the lowest accuracy was averaged with “0.79586” and it is clear from Fig. 2 that Boosted Decision Tree gives us the best accuracy with “0.888181”.

Experiment 2: The second experiment were employed the second dataset mentioned in Sect. 5.2.2. The dataset used here completely different with more features as it has 50 features and its instances was 5000 phishing eamils and 5000 legitimate. So after balancing and processing the data we applied the dataset into our model using the selected seven ML algorithms, the results are represented in the following table:

Algorithm	Accuracy	Precision	Recall	F-Score
Locally-deep support vector machine	0.995667	0.996003	0.99534	0.995671
Support vector machine	0.997	0.997335	0.996671	0.997003
Boosted decision tree	1	1	1	1
Logistic regression	0.998	0.998667	0.997337	0.998001
Averaged perceptron	0.996667	0.997333	0.996005	0.996669
Neural network	0.995333	0.999329	0.991345	0.995321
Decision forest	0.999667	0.999335	1	0.999667

The second experiment results were increased clearly as shown from the table. The lowest accuracy we have got was Neural Network with “0.995333” and It is clear from the Fig. 3 that Boosted Decision Tree gives us the best accuracy.

Experiment 3: In this experiment, we chose the dataset mentioned in Sect. 5.2.3 with only text feature that contains 2500 ham and 500 spam. For doing so, We have built classifier model Using Python and TensorFlow/Keras neural network. Also, we used Tensorflow which is one of the most popular deep learning libraries to classify Email text. The accuracy was calculated using python model for text classifying (0.992%). For more efficiency and to compare more algorithm techniques we have built special model using AZURE ML Microsoft tools. We have built the model to be able to train and test the dataset text classification efficiently. Then after, we have applied the

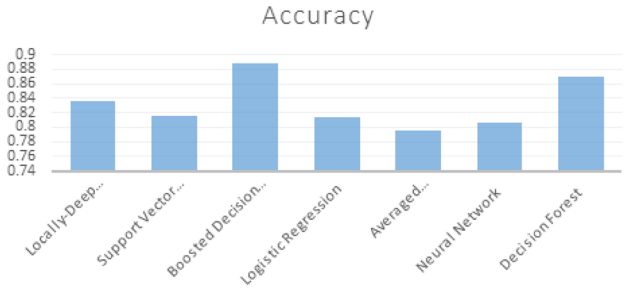


Fig. 2 Results for the first experiment

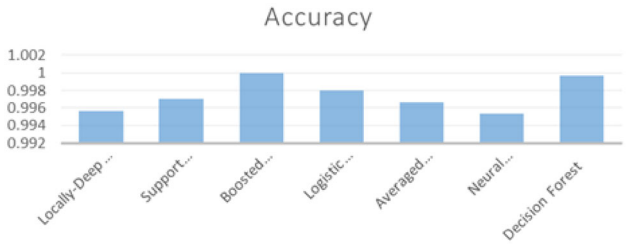


Fig. 3 Results for the second experiment

seven selected ML algorithms on the desined model, and the results are represented in the following table:

Algorithm	Accuracy	Precision	Recall	F-Score
Locally-deep support vector machine	0.968	0.900	0.906	0.903
Support vector machine	0.974	0.950	0.893	0.920
Boosted decision tree	0.972	0.931	0.899	0.915
Logistic regression	0.956	0.950	0.772	0.852
Averaged perceptron	0.977	0.944	0.913	0.928
Neural network	0.977	0.964	0.893	0.927
Decision forest	0.953	0.950	0.758	0.843

Clearly we can see here in the third experiment results that the lowest accuracy we've got was Decision Forest with"0.953" and it is clear from the Fig. 4 that Averaged Perception and Neural Network give us the best accuracy with "0.977".

5.4 Experiments results

Comparing the results of the previous three experiments, since in the first experiments we used dataset with 22 features, the second experiment we used dataset consists of 50 features while the third experiment used dataset with text feature only. We have end up with results summarised in the following Fig. 5.

The summarizing Fig. 5 shows that the best ML algorithm accuracy rates achieved was for boosted decision tree and Neural Network and the lowest algorithm was for Decision Forest.

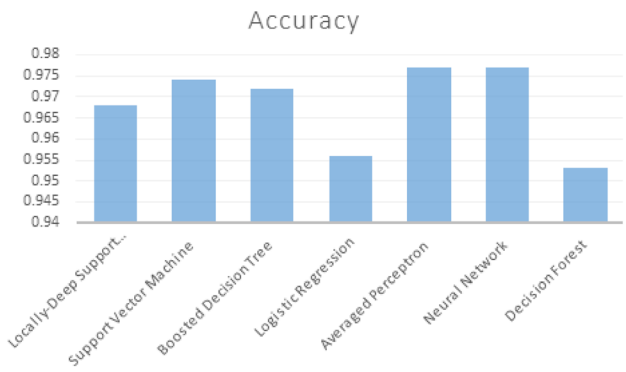


Fig. 4 Results for the third experiment

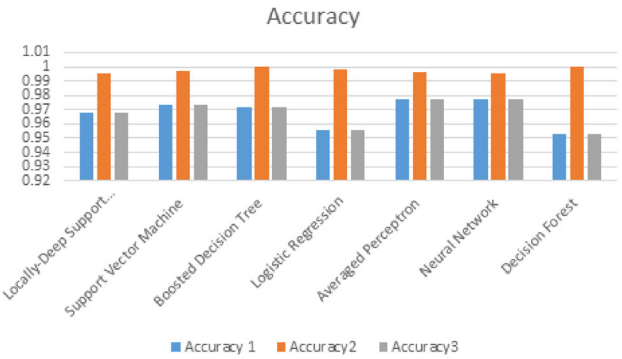


Fig. 5 Comparison between the three results

5.5 Discussion

Back to related work mentioned in Sect. 2, starting on comparing our results with Fette et al. [21] they used two non-spam, non-phishing datasets with 10 features on their proposed approach (PILFER). The overall accuracy that they have achieved on their approach was 99.5. However, we used in our first experiment dataset with 22 features and another dataset with 50 features on the second experiment, and as we saw the accuracy increased as we used more features. the difference between our results is that over-viewed the dataset and balance the phishing and non-phishing instances in each dataset which makes our result more accurate since they used an imbalanced dataset with 6950 non-phishing emails and 860 phishing emails which means that the percentage for their dataset is 0.80 non-phishing emails. On the other hand, we have made pre-processing for our used datasets. while the first dataset that we used (phishing email collection) was consists of 5,17,402 legitimate emails and 8351 phishing emails the balanced dataset that we preprocessed contains 8351 phishing and 8400 legitimate instances.

In relative to reference [22], they used a text dataset consists of 1897 spam, 3900 ham, and 250 'hard' ham. while we used in text dataset in the third experiment 2500 ham emails and 500 spam emails. In their work, they have identified Random Forests as the classifier to detect spam mails from ham mails, and the average accuracy that they have is 98.3 using WEKA the open-source software. While we used in our experiments seven algorithms mentioned in section 6, the highest accuracy that we got is 97.7 using AZURE when applying the Neural Network algorithm.

Islam et al. [8] They have proposed a multi-stage classification technique using three popular learning algorithms as NB, SVM and AdaBoost. The dataset they used public datasets PUA, It has been shown that the accuracy of their proposed system (97.05). Compared to our work we have used three datasets with different features and seven

algorithms the best ML algorithm accuracy rates achieved was for boosted decision tree and Neural Network.

Study	Study methodology	Our methodology
Fette et al. [5]	They used two non-spam, non-phishing datasets with 10 features on their proposed approach the overall accuracy they have achieved was 99.5	We used three datasets with different features , our result more accurate since they've used an imbalanced dataset
Bhat et al. [6]	They used a text dataset consists of 1897 spam, 3900 ham, and 250 'hard' ham The result identified Random Forests as the classifier with 98.3 accuracy	We used in the third experiment text dataset with 2500 ham emails and 500 spam emails, In our experiments we used seven algorithms, highest accuracy that we got is Neural Network algorithm with 97.7 accuracy
Islam et al. [8]	They used public data sets PUA and used three classification algorithms as NB, SVM and AdaBoost, It has been shown that the accuracy of their proposed system (97.05)	We used three datasets with different features and seven algorithms the best ML algorithm accuracy rates achieved was for boosted decision tree and Neural Network

6 Conclusion

Phishing emails have come to be a common problem in the latest years. Phishing email attacks are intelligently crafted social engineering email attacks in which victims are conned by email to provide important information and then directly sent it to the phisher.

Young users are more likely to fall for phishing attacks. Furthermore, users with agreeable personality trait are likely to be lured by phishing scam more than other users. Women are more likely to provide their personal and financial details to phishing emails and websites. This causal relationship between gender and social engineering is influenced by the internet usage behavior. So the detection of that type of email is necessary.

There are numerous techniques for detecting phishing emails. However, there are a few limitations like accuracy is low. The content material may be the same as legitimate email so cannot be detected, the detection rate is not high.

This work employed machine learning techniques to achieve better results, and to capture inherent characteristics of the email text and other features to classify emails as phishing or non-phishing. This research have come up with a better accuracy of phishing email detection. Which evaluated based on three supervised datasets, and comparison between these classifiers were conducted.

Finally, comparison of the results was obtained using different algorithms. The noted results that using an algorithm based on multi feature of (50) gave us the highest accuracy, and less features of (20) the accuracy was high enough but this result is not effective enough to detect phishing emails. The limitation of this work was finding the predefined dataset.

7 Future work

In Future Work, we noted that Feature selection techniques need more improvement to cope with the continuous development of new techniques by the phishers over the time. Therefore, we recommend developing a new automated tool in order to extract new features from new raw emails to improve the accuracy of detecting phishing email and to cope with the expanding with phisher techniques.

Author Contributions All Three Authors worked in an equivalent load at all stages to produce this research.

Funding This work was supported by the Hashemite University and AL Zaytoonah University of Jordan.

Data availability The data set used in the work will be available upon request.

Declarations

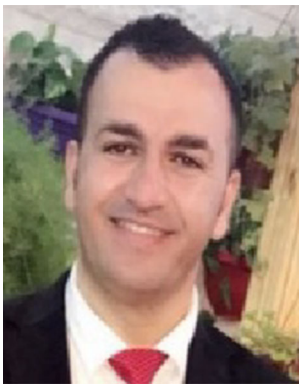
Conflict of interest The authors have not disclosed any competing interests.

Informed consent I have read and I understand the journal information and have agreed to all mentioned terms and conditions.

References

1. Al-Masalha, H., Hnaif, A.A., Kanan, T.: Cyber-crime effect on Jordanian society. Int. J. Adv. Soft Comput. Appl. **12**(3), 123–139 (2020)
2. Saini, B., Srivastava, S., Bajpai, A.: Deep CNN model for nanotoxicity classification using microscopic images. Int. J. Adv. Soft Comput. Appl. **12**(2), 22 (2020)
3. 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
4. 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

5. 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
6. Kim, D., Kim, Y.-H., Shin, D., Shin, D.: Fast attack detection system using log analysis and attack tree generation. Clust. Comput. **22**(1), 1827–1835 (2019)
7. Aldabbas, H., Amin, R.: A novel mechanism to handle address spoofing attacks in sdn based iot. Clust. Comput. **24**(4), 3011–3026 (2021)
8. Abusukhon, A., AlZu'bi, S.: New direction of cryptography: a review on text-to-image encryption algorithms based on rgb color value. In: Proceedings of the 2020 Seventh International Conference on Software Defined Systems (SDS), pp. 235–239. IEEE (2020)
9. Obeidat, I., Mughaid, A., Alzoubi, S.: A secure encrypted protocol for clients' handshaking in the same network. Int. J. Interact. Mob. Technol. **13**, 47–57 (2019)
10. Salahdine, F., Kaabouch, N.: Social engineering attacks: a survey. Future Internet **11**(4), 89 (2019)
11. Khonji, M., Iraqi, Y., Jones, A.: Phishing detection: a literature survey. IEEE Commun. Surv. Tutor. **15**(4), 2091–2121 (2013)
12. Whittaker, C., Ryner, B., Nazif, M.: Large-scale automatic classification of phishing pages. In: Proceedings of the Network and Distributed System Security Symposium (2010)
13. Hong, J.: The state of phishing attacks. Commun. ACM **55**(1), 74–81 (2012)
14. Maqableh, M., Alia, M.: Evaluation online learning of undergraduate students under lockdown amidst covid-19 pandemic: the online learning experience and students' satisfaction. Child Youth Serv. Rev. **128**, 106160 (2021)
15. Zhao, W., Zhu, Y.: An email classification scheme based on decision-theoretic rough set theory and analysis of email security. In: Proceedings of the TENCON 2005-2005 IEEE Region 10 Conference, pp. 1–6. IEEE (2005)
16. Vinayakumar, R., Soman, K., Poornachandran, P., Akarsh, S., Elhoseny, M.: Deep learning framework for cyber threat situational awareness based on email and url data analysis. In: Hassanien, A.E., Elhoseny, M. (eds.) Cybersecurity and Secure Information Systems, pp. 87–124. Springer, New York (2019)
17. 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)
18. Al-Zubi, S., Hawashin, B., Mughaid, A., Baker, T.: Efficient 3d medical image segmentation algorithm over a secured multimedia network. Multimed. Tools Appl. **80**(11), 16887–16905 (2021)
19. AlZu'bi, S., Jararweh, Y.: Data fusion in autonomous vehicles research, literature tracing from imaginary idea to smart surrounding community. In: Proceedings of the 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC), pp. 306–311. IEEE (2020)
20. AlKhatib, A.A., Sawalha, T., AlZu'bi, S.: Load balancing techniques 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)
21. Fette, I., Sadeh, N., Tomasic, A.: Learning to detect phishing emails. In: Proceedings of the 16th international conference on World Wide Web, pp. 649–656 (2007)
22. Bhat, V.H., Malkani, V.R., Shenoy, P.D., Venugopal, K., Patnaik, L.: Classification of email using beaks: behavior and keyword stemming. In: Proceedings of the TENCON 2011-2011 IEEE Region 10 Conference, pp. 1139–1143. IEEE (2011)
23. Form, L.M., Chiew, K.L., Tiong, W.K.: Phishing email detection technique by using hybrid features. In: Proceedings of the 2015 9th International Conference on IT in Asia (CITA), pp. 1–5. IEEE (2015)
24. Elbes, M., Alrawashdeh, T., Almaita, E., AlZu'bi, S., Jararweh, Y.: A platform for power management based on indoor localization in smart buildings using long short-term neural networks''. Trans. Emerg. Telecommun. Technol. **33**, e3867 (2020)
25. AlZu'bi, S., Shehab, M.A., Al-Ayyoub, M., Benkhelifa, E., Jararweh, Y.: Parallel implementation of fcm-based volume segmentation of 3d images. In: Proceedings of the IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), **2016**, pp. 1–6. IEEE (2016)
26. Teli, S.P., Biradar, S.K.: Effective email classification for spam and non-spam. Int. J. Adv. Res. Comput. Softw. Eng. **4**, 2014 (2014)
27. Basnet, R., Mukkamala, S., Sung, A.H.: Detection of phishing attacks: a machine learning approach. In: Proceedings of the Soft computing applications in industry, pp. 373–383. Springer (2008)
28. Moradpoor, N., Clavie, B., Buchanan, B.: Employing machine learning techniques for detection and classification of phishing emails. Comput. Conf. **2017**, 149–156 (2017)
29. Smadi, S., Aslam, N., Zhang, L., Alasem, R., Hossain, M.A.: Detection of phishing emails using data mining algorithms. In: Proceedings of the 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), pp. 1–8. IEEE (2015)
30. 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)
31. Jagatic, T.N., Johnson, N.A., Jakobsson, M., Menczer, F.: Social phishing. Commun. ACM **50**(10), 94–100 (2007)
32. 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)
33. Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor,L., Hong, J., Blair, M.A., Pham, T.: School of phish: a real-world evaluation of anti-phishing training. In: Proceedings of the 5th Symposium on Usable Privacy and Security, pp. 1–12 (2009)
34. Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L.F., Hong, J.: Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In Proceedings of the Anti-phishing Working Groups 2nd Annual eCrime Researchers Summit, pp. 70–81 (2007)
35. Parrish, J.L., Jr., Bailey, J.L., Courtney, J.F.: A Personality Based Model for Determining Susceptibility to Phishing Attacks, pp. 285–296. University of Arkansas, Little Rock (2009)
36. 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)
37. Nikken, P., Schols, M.: How and why parents guide the media use of young children. J. Child Fam. Stud. **24**(11), 3423–3435 (2015)
38. Nicholson, J., Javed, Y., Dixon, M., Coventry, L., Ajayi, O.D., Anderson, P.: Investigating teenagers ability to detect phishing messages. In: Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroS &PW). IEEE **2020**, pp. 140–149 (2020)

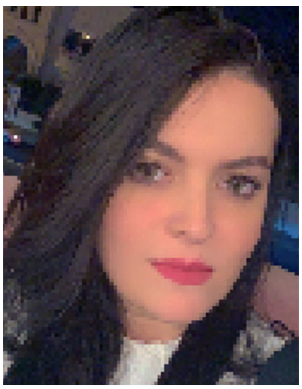


Alaa Mughaid was born in Irbid, Jordan, in 1984. He received the BSC degree in Computer Science from Jordan University of Science and Technology (JUST), Jordan, in 2006, and the MSC in Engineering degree in engineering Computer Network from the Western Sydney University, Sydney, Australia, in 2010. Dr. Mughaid received the Ph.D. degree in Computer Science from Newcastle University-Sydney, Australia, in 2018. In 2018, Dr. Mughaid joined the Department of Computer Science, The Hashemite University, as an assistant professor, Zarqa, Jordan. Dr. Mughaid current research interests include but not limited to Cyber Security, Cloud Computing, Network Security, Artificial Intelligence, Virtual reality, Data ining. He is working voluntarily in many social services.

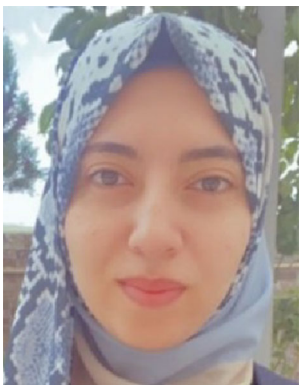
Shadi AlZu'bi was born in Irbid, Jordan, in 1984. He received the B.E. degree in computer and networks engineering from Jordan University of Science and Technology (JUST), Jordan, in 2007, and the M. Eng. degree in engineering Projects Management from the University of Technology, Sydney (UTS), Sydney, Australia, in 2008, and the PhD. degree in Computer engineering from Brunel University - London, The United Kingdom, in 2011. In 2011, he joined the Department of Computer Science, Jordan University of Science and Technology (JUST), as a Lecturer, and in 2012 became an assistant professor in the department of computer science in Zaytoonah University of Jordan, Amman, Jordan. Since September 2014, he has been with the Department of Multimedia Systems in Zaytoonah University of Jordan, Amman, Jordan till now. His current research interests include image processing, Medical imaging, 3D volume processing, medical imaging, computation acceleration using parallel processing and image compression and archiving. Dr. AlZu'bi is a Fellow of the Jordan Engineers association. He was a recipient of Vice-Chancellor Prize of Brunel University in 2009.

Adnan Hnaif is an associate professor at the computer science department, Faculty of Science and information technology, Al Zaytoonah University of Jordan. Dr. Hnaif received his PhD degree in Computer Science from University Sains Malaysia – National Advanced IPv6 Centre and Excellence (NAV6) in 2010. He received his MSc degree of Computer Science from department of Computer Science in 2003, and obtained his Bachelor degree of Computer Science from the department of Computer Science, in 1999/2000. His researches focus on the computer networks and communications, wireless sensor networks, network security, parallel processing, and algorithms.

Salah Taamneh is currently an Assistant Professor at the Department of Computer Science and Applications, Hashemite University, Zarqa, Jordan. He received the B.S. degree in computer science from Jordan University of Science and Technology, Irbid, Jordan, in 2005, the M.S. degree in computer science from Prairie View A &M University, Prairie View, Texas, in 2011 and the Ph.D. degree in computer science from University of Houston, Houston, Texas, USA, in 2016. He. His current research interests include parallel and distributed computing, machine learning and human- computer interaction.



Asma Alnajjar was born in UAE, Abu Dhabi, in 1992. She received the BSC degree in Software Engineering from AL Zaytoonah University of Jordan, in 2012, Currently she is studying MSC in Cyber Security at Hashemite University in Jordan. She has a successful career as an Application Support at Department of Finance in UAE/Abu Dhabi. Asma's current research interests include Cyber security, Machine Learning, Big Data, Mobile Network.



Esraa Abu Elsoud was born in Amman, Jordan, in 1991. She received the BSC degree in Electrical Engineering from Hashemite University, Jordan, in 2013, and now studying MSC in Cyber Security at Hashemite University too. Eng. Esraa's current research interests include Cyber Security, Machine Learning, Big Data and Mobile Network.