

A hybrid machine learning approach for malicious website detection and accuracy enhancement

Ahmed Abu-Khadrah¹, Shayma Alkhamis², Ali Mohd Ali¹, Muath Jarrah³

¹Department of Electrical Engineering, College of Engineering Technology, Al-Balqa Applied University, Amman, Jordan

²College of Computing and Informatics, Saudi Electronic University, Riyadh, Saudi Arabia

³School of Computing, Skyline University College, Sharjah, United Arab Emirates

Article Info

Article history:

Received Oct 16, 2024

Revised Mar 19, 2025

Accepted Jul 1, 2025

Keywords:

Decision tree
Machine learning
Malicious URL
Random forest
XGBoost

ABSTRACT

Malicious URLs are web addresses purposely generated for a user's detriment. Some examples include phishing scams in which the victim is fooled into logging into a fake site or portals for downloading malware where any click on a link invites a hostile program to the user's device. The damage done to an individual's finances, confidential information, and even reputation due to malicious URLs makes it crucial to devise means of countering these threats. This can be achieved by creating an intelligent model that identifies suspicious characteristics common to these websites. The objective of this research is to design a novel hybrid machine learning algorithm-based model for detecting malicious websites. A random forest, decision tree, and extreme gradient boosting (XGBoost) are the three hybrid classification algorithms proposed for the study. Accuracy in detection will help prevent and reduce the effects of such websites. The accuracy rate in this research is 98.7%, precision is at 98.9%, and recall at 98.5%. With these results, it follows that the hybrid model is more effective than training any individual algorithm with the given dataset.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ahmed Abu-Khadrah

Department of Electrical Engineering, College of Engineering Technology, Al-Balqa Applied University
Amman, Jordan

Email: a.abukhadrah@bau.edu.jo

1. INTRODUCTION

A malicious URL is a link to an illegal website that attempts to trick users into visiting it. A single click on these malicious links can put the user in a dangerous situation, as these websites can install malware, cause confidential data leaks, or result in internet fraud. Attackers frequently alter one or more structural elements of a URL to deceive users into interacting with their malicious links [1]-[5]. Users may be redirected to unwanted websites, malicious websites, phishing websites, or pages where malware can be downloaded, allowing attackers to run codes on users' machines. Hackers typically use phishing to lure users into their malicious attacks. For example, they might create a malicious link that imitates a popular online shopping website to steal bank account information or impersonate a celebrity and share a malicious survey link to collect personal information. Additionally, attackers can send emails pretending to be from Microsoft, urging recipients to click a link to upgrade their Windows software. Once clicked, the link downloads malware that can crash the user's PC or steal sensitive information [6]-[10]. Nowadays, malicious URLs spread not only through email and websites but also through links found in online advertisements, status updates, tweets, and Facebook posts [11]. The impact of these malicious links extends beyond individuals; businesses can also be affected if one of their employees falls victim to these URLs. This could lead to the revelation of sensitive customer information, resulting in a loss of customer trust and reputation. For these

reasons, employing URL detection techniques is essential. Different types of URL detection methods rely on specific elements and features of the URL. Blacklisting, heuristic approaches, and machine learning have proven effective in detecting malicious URLs. Blacklisting is one of the most traditional methods, involving a database of previously detected malicious URLs. This method compares each newly visited URL with those in the blacklist. If there is a match, the URL is considered malicious, and the user is alerted; otherwise, it is considered benign [12]. However, this method has limitations. It cannot detect new malicious URLs that have not been previously identified, and attackers can evade detection by making minor alterations to the URL, such as changing the top-level domain, directory path, brand name, or query string [11]. Heuristic-based techniques rely on rules derived from past outcomes and learning to solve problems or facilitate learning. Although heuristic-based answers do not always lead to the best decisions, they come close. These techniques have proven successful in defending against zero-day attacks, with methods based on real phishing attack data. Despite showing a high rate of false positives, the outcomes are still better than blacklist-based approaches. Programs like Mozilla Firefox and Internet Explorer use heuristic methods to identify scams. Machine learning, a branch of artificial intelligence, uses specific algorithms to predict outcomes based on input data. It is one of the most effective methods for detecting malicious URLs. Machine learning techniques address large-scale binary classification problems by collecting and classifying features. In URL detection, these features can include lexical features, network-based features, and host-based features [13]. Several machine learning algorithms are used for detecting malicious URLs, including support vector machine (SVM), random forest, neural networks, Naïve Bayes, and extreme gradient boosting (XGBoost).

Many researchers have employed different algorithms based on specific features extracted from URLs. Patgiri *et al.* [13] used random forests and SVM to test a dataset of benign and malicious URLs, utilizing lexical, host-based, and site popularity features extracted from Alexa.com. The highest accuracy achieved by random forest was 92%, while SVM reached only 89%. In another study [14], detection techniques based on lexical, host-based, and content-based features were used, employing SVM and random forest. URLs were collected from sources like Phishtank, Alexa, URLhaus, and Malicious_n_Non-Malicious URL. The results showed that random forest was more accurate and faster than SVM. The proposed solution in [15] aimed to use a multilayer perceptron (MLP) model to detect malicious URLs, focusing solely on lexical features to build a lightweight model and shorten feature extraction time. Using a dataset from GitHub, the model achieved an accuracy of 94.5%. Sahingoz *et al.* [16], around seven classification algorithms were tested on different features such as word vector, natural language processing (NLP) features, and hybrid features. The results showed that SVM, decision tree, and random forest achieved the highest accuracy with NLP features, with random forest reaching 97.98%. Manjeri *et al.* [17] tested the efficiency of random forest, decision trees, logistic regression, K-nearest neighbors (KNN), and SVM, with random forest achieving the highest accuracy at 96.58%. The paper on phishing detection based on machine learning and feature selection methods by [18] used feature selection to speed up model construction and improve performance on a unique phishing dataset. Random forest achieved the highest accuracy of 98.37% in this study, with a time of only 4.18 seconds to get the results. Other researchers who used SVM alone achieved an accuracy of 95.66% [19], relying on features such as lexical, host-based, and word vector to differentiate between phishing and legitimate URLs. A hybrid algorithm combining random forest and SVM improved accuracy from 95.6% to 96.8% [20]. Most related work has relied on single machine learning algorithms to classify URLs or compared different classifiers. While they achieved acceptable accuracy, it is not the best for critical tasks like detecting malicious websites, where a single click can have significant consequences. Improving a model's accuracy is necessary. Papers discussing hybrid algorithms have shown significant accuracy improvements. Relying on a single classifier may require longer training and validation times. According to [20], performance and accuracy increase with hybrid classifiers. Therefore, utilizing a hybrid classifier can improve reliability and accuracy in a shorter time. The aim of this study is to develop a new model for detecting malicious URLs using hybrid machine learning algorithms.

2. METHOD AND MATERIALS

This study aims to investigate the performance and accuracy of building a model that combines multiple classification algorithms to predict and detect phishing links. The classification process relies on extracting features from the URL and using them as input for the model, such as URL length, domain name, the existence of prefixes and suffixes, and any attribute that may indicate a phishing URL, using various tools. The approach allows for the development of a hybrid learning model comprised of various machine learning models, such as random forest, decision tree, and XGBoost. An advanced ensemble learning model will be designed to classify phishing websites. An outline of the suggested phishing prediction framework is provided in Figure 1.

After the dataset was cleaned by eliminating all the null and repeated values along with the outliers, the first step in the methodology is to face an accuracy related issue that stems from imbalanced data. An effective approach to overcome this imbalance is to use the synthetic minority oversampling technique known as (SMOTE). It is highly regarded as one of the best methods for imbalanced class problems. When the number of instances of the majority class is significantly greater than for the minority class, the model gets instruction bias because it does not learn enough features about the minority class. SMOTE addresses this problem by generating synthetic samples for both the instance of the minority class and its k nearest neighbors. Then, this process randomly picks a neighbor as a new instance on the line segment that connects both of the instances. The procedure continues until obtaining an adequate amount of samples for oversampling, resulting in an artificial balanced dataset with samples of the minority class [21].

The second step is to select the features that help identify phishing URLs, such as lexical features, URL length, and domain name. K-fold cross validation (KCV) will be used to divide the dataset into k subsets, with some subsets iteratively used for training the model and others for evaluating its performance. KCV is considered one of the most popular methods for classifier model selection and error estimation. The proposed method involves developing three ensemble classification algorithms during the modeling phase to handle the most important features used to identify malicious or benign websites.

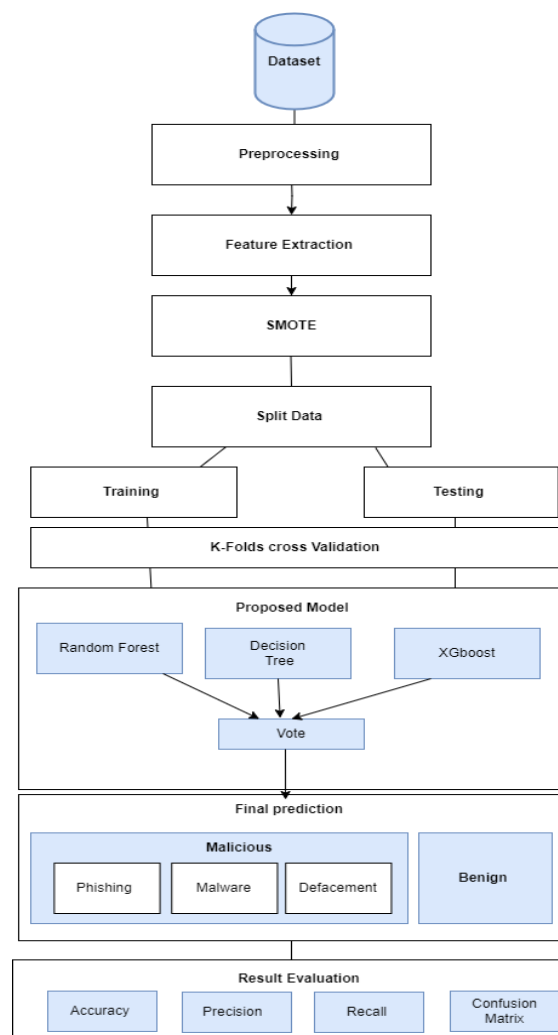


Figure 1. Overall methodology

The first algorithm to be used is random forest, which builds decision trees by training each tree on randomly chosen features and objects, a process known as the random subspace method. A prediction can then be made based on the outcome of each tree, often determined by a majority vote. This random technique reduces the model's error by decreasing the spread of predictions and improves classification quality while minimizing overfitting compared to algorithms such as the decision tree.

The second algorithm is the decision tree. This type of tree structure resembles a flowchart, where each internal node represents a test on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label. A tree can be “learned” by dividing the source set into subgroups based on an attribute value test, a process known as recursive partitioning. This operation is repeated on each derived subset until the split no longer improves the predictions or when the subset at a node has the same value for the target variable [16].

The third algorithm is XGBoost. Based on the gradient boosting decision tree (GBDT), XGBoost uses an ensemble learning boosting strategy to lower the classified error margin. The results of the XGBoost classification are improved by adjusting the weight of the data features that were incorrectly classified. The XGBoost technique is used to examine the reliability and accuracy of models for classifying dangerous URLs [22]. The final prediction will be determined by voting on the results of the three models, as all classifications operate in parallel. The most commonly used assessment metrics for phishing detection issues are classification accuracy, precision, and recall, which will be used in the evaluation phase to assess the overall performance of the suggested classification framework. The dataset used in this study is downloaded from the Kaggle website, comprising data collected from various sources such as the Canadian Institute for Cybersecurity, University of Brunswick, PhishTank, and the PhishStorm dataset. The dataset contains 651,191 URLs, including 428,103 benign or safe URLs, 96,457 defacement URLs, 94,111 phishing URLs, and 32,520 malware URLs. Since the dataset is raw, it is important to select lexical features for use in subsequent steps as input to train the model. Lexical features, known as the bag of words, describe the properties of the URLs, such as the presence of certain words or special characters that identify them as malicious [23]-[27].

3. RESULTS AND DISCUSSION

To evaluate the performance of the proposed model, several classification metrics have been used, including accuracy, precision, recall, and the confusion matrix. The dataset, available for download from Kaggle, has been collected from various sources, including the Canadian Institute for Cybersecurity, the University of Brunswick, PhishTank, and PhishStorm. It contains 651,191 URLs, of which 428,103 are benign or safe URLs, 96,457 are defacement URLs, 94,111 are phishing URLs, and 32,520 are malware URLs. After extracting the data, the correlation method is used to create a correlation matrix. This matrix illustrates the linear relationship between the features and the target, providing insights into the key characteristics of the model. The next crucial step before building our model is addressing imbalanced data using the SMOTE. SMOTE is one of the most common techniques for handling imbalanced data, where the majority class has significantly more instances than the minority class(es). The minority class’s low representation in the dataset can bias the model due to its inability to learn adequately about it. According to [21], SMOTE is the simplest method that involves duplicating examples from the minority class. While these examples don’t offer new information to the model, they can be used to create new data by synthesizing previously collected information. As shown in Figure 2, there is an imbalance in the data, with the number of benign URLs being significantly higher than the other types of malicious URLs.

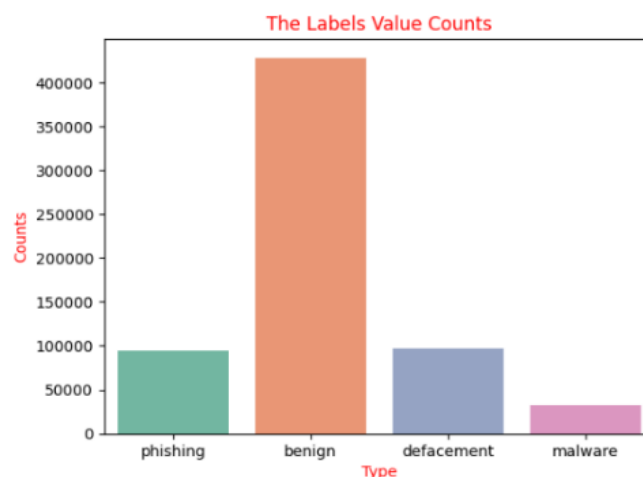


Figure 2. Count of URL in each label

To determine whether three ensemble classification models can effectively detect malicious websites, this study will investigate various machine learning techniques. The goal is to build an ensemble model that can classify websites as legitimate or malicious and assess the extent of their legitimacy or maliciousness. Identifying malicious websites is treated as a classification issue in data mining. The classification process relies on characteristics and attributes that distinguish malicious websites, such as spelling mistakes, special characters (e.g., @), prefixes, and suffixes. These attributes are extracted from the input websites using various technologies. To evaluate the performance of the model, the accuracy metric is used. Accuracy represents the proportion of correct predictions made by the model and is calculated using (1).

$$Accuracy = \frac{Ture\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad (1)$$

To analyze the model, precision and recall were used as evaluation metrics too. Precision measures the accuracy of positive claims. In other words, it is the ratio of true positive results to all positive results. Precision is defined by (2). This term assists in evaluating the correctness of predictions made automatically during the positive class estimation for the model under study.

$$Precision = \frac{Ture\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

In contrast, recall shows the ratio of detected positive cases to all actual positive cases in the dataset. It quantifies the ability of the model to find all relevant cases in the data. Recall is determined by (3). This term assists in evaluating the number of true positives found for all positives that were in the scope of the study or were actual positives. As a whole, both precision and recall serve to explain the model's success in detecting positives cases, negative predictions are measured by precision, whereas positive are measured by recall. The results of the evaluation metrics are displayed in Table 1 as an overall performance of the model with regard to the precision and recall metrics.

$$Recall = \frac{Ture\ Positive}{True\ Positive + False\ Negative} \quad (3)$$

Table 1. Result of evaluation metrics

Accuracy	Precision	Recall
0.987	0.989	0.985

The final metric used is the confusion matrix, which visualizes the model's performance by displaying the counts of true positives, false positives, true negatives, and false negatives. It allows the calculation of key metrics such as accuracy, misclassification rate, and precision. The confusion matrix also aids in deriving recall and specificity, providing a clear comparison of prediction outcomes. Figure 3 shows the confusion matrix for the proposed model, highlighting the accuracy of predictions for each class.

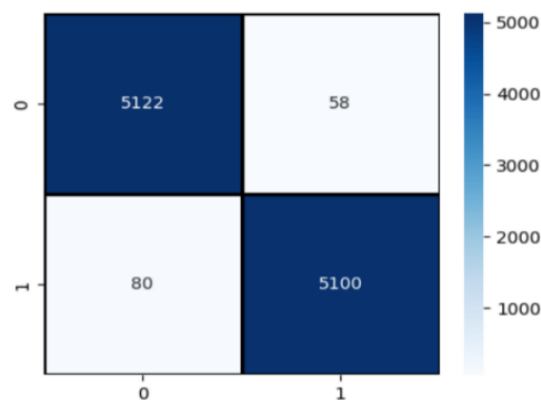


Figure 3. Confusion matrix of hybrid model

Table 2 illustrates the effectiveness of the proposed method, which outperforms other machine learning classifiers by achieving the highest accuracy of 98.7%. The combined use of random forest, decision tree, and XGBoost classifiers demonstrates superior performance in classifying phishing websites compared to individual classifiers. Additionally, Table 2 shows that other classifiers, such as random forest, SVM, and MLP, yielded lower accuracy results than the proposed ensemble model.

Table 2. Comparison of the accuracy, precision, and recall

Author	Used approach	Accuracy	Precision	Recall
Patgiri <i>et al.</i> [13]	Random forest	92	NA	NA
Do Xuan <i>et al.</i> [14]	Random forest	96.28	91.44	94.42
Ahmed and Jameel [15]	MLP	94.51	95.0	94.1
Sahingoz <i>et al.</i> [16]	Random forest	97.98	NA	NA
Manjeri <i>et al.</i> [17]	Random forest	96.58	98.34	94.76
Almseidin <i>et al.</i> [18]	Random forest	98.3	NA	NA
Rashid <i>et al.</i> [19]	SVM	95.66	NA	NA
Kulkarni and Brown [20]	Used hybrid algorithm which combine the random with SVM	96.8	96.8	NA
Proposed model	Used hybrid algorithm which combine the random forest, decision tree and XGBoost	98.7	98.9	98.5

In addition to that, the newly suggested model that integrates ensemble learning using random forest, decision tree, and XGBoost classifiers performs better compared to the existing classifiers with respect to accuracy, precision, and recall. The novel model records the highest accuracy of 98.7%. This is a great improvement from the other researches which ranges from 92% to 98.3% [13]-[20]. In addition, the novel method outperformed SVM [19] which has an accuracy of 95.66%. and MLP [15] with lower accuracy of 94.51%. The hybrid approach used in [20] which combines random forest with SVM also comes short to the novel method with 96.8% accuracy. This model also gives better results in precision (98.9%) and recall (98.5%) which makes it more robust in phishing website classification. These results validate the claim on the usefulness of employing multiple classifiers in one model, outperforming the traditional models, indeed come to conclusion that the model proposed is better than the existing ones.

4. CONCLUSION

The rapid escalation of phishing attacks and their sophisticated evasion techniques have rendered traditional blacklist methods increasingly ineffective. As cybercriminals adapt their strategies to launch short-lived, rapidly evolving campaigns, the limitations of blacklisting become more pronounced. This scenario highlights the need for more advanced approaches in the reliable detection of phishing URLs. Machine learning has proven to be an incredibly useful tool in overcoming this challenge to the extent of being able to flag and tell apart dits and legitimate URLs, even those which are recently generated and have not yet been blacklisted. Unlike traditional techniques, machine learning algorithms are able to evolve with new threats, providing a more effective, dynamic, and phishing resistant defense. In this study, a hybrid model is introduced which is a fusion of three powerful machine learning algorithms, random forest, decision tree, and XGBoost. This approach has shown significant improvement in detection accuracy as high as 98.7% which no longer individual classifiers can achieve. The hybrid model improves not only accuracy but also the model's strength against an array of phishing attacks, making it more robust. The development and its verification constitutes the main claim of the research which certainly is a breakthrough in phishing detection a hybrid model. There are countless approaches in developing solutions and the methods fails to address this gap, but this model does. The solution is reliable and effective for ideal.

ACKNOWLEDGEMENTS

Thanks to ALLAH who helped us in this work and thanks to the Al-Balqa Applied University and Saudi Electronic University for sponsoring this work.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ahmed Abu-Khadrah	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓	✓	
Shayma Alkhamis	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓		✓
Ali Mohd Ali	✓		✓	✓			✓	✓	✓	✓	✓			✓
Muath Jarrah			✓	✓	✓		✓		✓	✓				✓

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author AA on request.




REFERENCES

- [1] A. K. Jain and B. B. Gupta, "A survey of phishing attack techniques, defence mechanisms and open research challenges," *Enterprise Information Systems*, vol. 16, no. 4, pp. 527–565, Mar. 2021, doi: 10.1080/17517575.2021.1896786.
- [2] E. Nowroozi, Abhishek, M. Mohammadi, and M. Conti, "An adversarial attack analysis on malicious advertisement URL detection framework," *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 1332–1344, Jun. 2023, doi: 10.1109/TNSM.2022.3225217.
- [3] A. S. Rafsanjani, N. B. Kamaruddin, M. Behjati, S. Aslam, A. Sarfaraz, and A. Amphawan, "Enhancing malicious URL detection: a novel framework leveraging priority coefficient and feature evaluation," *IEEE Access*, vol. 12, pp. 85001–85026, 2024, doi: 10.1109/ACCESS.2024.3412331.
- [4] S. S. Nair, "Securing against advanced cyber threats: a comprehensive guide to phishing, XSS, and SQL injection defense," *Journal of Computer Science and Technology Studies*, vol. 6, no. 1, pp. 76–93, Jan. 2024, doi: 10.32996/jcsts.2024.6.1.9.
- [5] M. Al-Khateeb, M. R. Al-Mousa, A. S. Al-Sherideh, D. Almajali, M. Asassfeh, and H. Khafajeh, "Awareness model for minimizing the effects of social engineering attacks in web applications," *International Journal of Data and Network Science*, vol. 7, no. 2, pp. 791–800, 2023, doi: 10.5267/j.ijdns.2023.1.010.
- [6] Z. Abdin, "Empowering the hydrogen economy: the transformative potential of blockchain technology," *Renewable and Sustainable Energy Reviews*, vol. 200, p. 114572, Aug. 2024, doi: 10.1016/j.rser.2024.114572.
- [7] V. Rishiwal, U. Agarwal, A. Alotaibi, S. Tanwar, P. Yadav, and M. Yadav, "Exploring secure V2X communication networks for human-centric security and privacy in smart cities," *IEEE Access*, vol. 12, pp. 138763–138788, 2024, doi: 10.1109/ACCESS.2024.3467002.
- [8] K. Haritha, S. S. Vellela, R. D. L. R. Vuyyuru, N. Malathi, and L. Dalavai, "Distributed blockchain-SDN models for robust data security in cloud-integrated IoT networks," in *2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS)*, Dec. 2024, pp. 623–629, doi: 10.1109/ICACRS62842.2024.10841584.
- [9] S. Ali, Q. Li, and A. Yousafzai, "Blockchain and federated learning-based intrusion detection approaches for edge-enabled industrial IoT networks: a survey," *Ad Hoc Networks*, vol. 152, p. 103320, Jan. 2024, doi: 10.1016/j.adhoc.2023.103320.
- [10] H. Su, S. Dong, N. Wang, and T. Zhang, "An efficient privacy-preserving authentication scheme that mitigates TA dependency in VANETs," *Vehicular Communications*, vol. 45, p. 100727, Feb. 2024, doi: 10.1016/j.vehcom.2024.100727.
- [11] B. B. Gupta, K. Yadav, I. Razzak, K. Psannis, A. Castiglione, and X. Chang, "A novel approach for phishing URLs detection using lexical based machine learning in a real-time environment," *Computer Communications*, vol. 175, pp. 47–57, Jul. 2021, doi: 10.1016/j.comcom.2021.04.023.
- [12] M. Ferreira, "Malicious URL detection using machine learning," in *Proceedings of the Digital Privacy and Security Conference 2019*, 2019, doi: 10.11228/dpsc.01.01.
- [13] R. Patgiri, H. Katari, R. Kumar, and D. Sharma, "Empirical study on malicious URL detection using machine learning," in *Distributed Computing and Internet Technology*, Springer International Publishing, 2018, pp. 380–388.
- [14] C. Do Xuan, H. Dinh, and T. Victor, "Malicious URL detection based on machine learning," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 1, 2020, doi: 10.14569/ijacsa.2020.0110119.
- [15] W. Ahmed and N. G. M. Jameel, "Malicious URL detection using decision tree-based lexical features selection and multilayer perceptron model," *UHD Journal of Science and Technology*, vol. 6, no. 2, pp. 105–116, Nov. 2022, doi: 10.21928/uhdjst.v6n2y2022.pp105-116.
- [16] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from URLs," *Expert Systems with Applications*, vol. 117, pp. 345–357, Mar. 2019, doi: 10.1016/j.eswa.2018.09.029.
- [17] A. S. Manjeri, R. Kaushik, M. N. V. Ajay, and P. C. Nair, "A machine learning approach for detecting malicious websites using URL features," in *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, Jun. 2019, pp. 555–561, doi: 10.1109/iceca.2019.8821879.
- [18] M. Almseidin, A. Abu Zuraiq, M. Al-kasassbeh, and N. Alniami, "Phishing detection based on machine learning and feature selection methods," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 13, no. 12, p. 171, Dec. 2019, doi: 10.3991/ijim.v13i12.11411.




- [19] J. Rashid, T. Mahmood, M. W. Nisar, and T. Nazir, "Phishing detection using machine learning technique," Nov. 2020, doi: 10.1109/smart-tech49988.2020.00026.
- [20] A. Kulkarni and L. L. Brown., "Phishing websites detection using machine learning," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 7, 2019, doi: 10.14569/ijacsa.2019.0100702.
- [21] A. J. Mohammed, M. M. Hassan, and D. H. Kadir, "Improving classification performance for a novel imbalanced medical dataset using SMOTE method," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 3, pp. 3161–3172, Jun. 2020, doi: 10.30534/ijacse/2020/104932020.
- [22] Y.-C. Chen, Y.-W. Ma, and J.-L. Chen, "Intelligent malicious URL detection with feature analysis," in *2020 IEEE Symposium on Computers and Communications (ISCC)*, Jul. 2020, pp. 1–5, doi: 10.1109/iscc50000.2020.9219637.
- [23] Z. Wang, J. Yao, M. Xu, M. Jiang, and J. Su, "Transformer-based network with temporal depthwise convolutions for sEMG recognition," *Pattern Recognition*, vol. 145, p. 109967, Jan. 2024, doi: 10.1016/j.patcog.2023.109967.
- [24] A. S. Dhanjal and W. Singh, "A comprehensive survey on automatic speech recognition using neural networks," *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 23367–23412, Aug. 2023, doi: 10.1007/s11042-023-16438-y.
- [25] V. Kukartsev, K. Kravtsov, O. Stefanenko, N. Podanyov, and A. Bezvorotnykh, "Using machine learning techniques to simulate network intrusion detection," in *2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)*, May 2024, pp. 1–4, doi: 10.1109/ISCS61804.2024.10581097.
- [26] K. U. Sarker, F. Yunus, and A. Deraman, "Penetration taxonomy: a systematic review on the penetration process, framework, standards, tools, and scoring methods," *Sustainability*, vol. 15, no. 13, p. 10471, Jul. 2023, doi: 10.3390/su151310471.
- [27] K. U. Sarker *et al.*, "A ranking learning model by K-means clustering technique for web scraped movie data," *Computers*, vol. 11, no. 11, p. 158, Nov. 2022, doi: 10.3390/computers11110158.

BIOGRAPHIES OF AUTHORS






Ahmed Abu-Khadrah    was born in United Arab Emirates in 1981. He received Bachelor of engineering in computer engineering from Alblqa Applied University in 2003. He received the master's degree in electronic engineering (computer engineering) from Universiti Teknikal Malaysia Melaka (UTeM) in 2013. He received a Ph.D. in computer engineering and communications from Universiti Teknikal Malaysia Melaka (UTeM) in 2017. He is currently Faculty member at Department of Electrical Engineering, College of Engineering Technology, Al-Balqa Applied University, Amman. His research interests in wireless network protocols, networking, communications, wireless mathematical model, also multimedia service over the networks. He can be contacted at email: a.abukhadrah@bau.edu.jo.






Shayma Alkhamis    was born in Saudi Arabia 1992; She received Bachelor degree in computer information system from Imam Abdulrahman bin Faisal University in 2016. She received the Master's degree in cyber security from Saudi Electronic University. Her research interests in cyber security methods and tools, malicious website detection and machine learning. She can be contacted at email: g210006087@seu.edu.sa.



Ali Mohd Ali    was born in 1982 in Jordan. Mutah University awarded him a Bachelor of Engineering in Computer Engineering in 2005. In 2013, he received a Master's degree in computer and communication engineering from Universiti Kebangsaan Malaysia (UKM). In 2021, he received a Ph.D. in computer and communications engineering from the University of Huddersfield in the United Kingdom. He is currently Faculty member at Department of Electrical Engineering, College of Engineering Technology, Al-Balqa Applied University, Amman. His primary research interests are in the analysis of communication system reliability using complex modelling techniques, as well as approaches to WLAN optimization. He can be contacted at email: a_90ali@hotmail.com.



Muath Jarrah    received his master and Ph.D. degrees in computer science from the Technical University of Malaysia, Melaka, in 2014 and 2018, respectively. He specializes in artificial intelligence and software engineering. His research interests include industrial computing, artificial intelligence, data science, machine learning, and modeling and optimization algorithms. He has been working at different universities at different countries. He can be contacted at email: aljarrahmuath@gmail.com.