URL Spam Detection Using Machine Learning Classifiers

Omar Almomani
Department of Networks and Cybersecurity, Al-Ahliyya Amman University, Amman, Jordan
o.almomani@ammanu.edu.jo

Adeeb Alsaaidah
Department of Networks and Cybersecurity, Al-Ahliyya Amman University, Amman, Jordan
a.alsaaidah@ammanu.edu.jo

Mosleh M. Abualhaj
Department of Networks and Cybersecurity, Al-Ahliyya Amman University, Amman, Jordan
m.abualhaj@ammanu.edu.jo

Mohammed Amin Almaiah
Department of Computer Science, King Abdullah the II IT School, The University of Jordan, Amman 11942, Jordan
m.almaiah@ju.edu.jo

Ammar Almomani
Department of Information Technology, Al-Huson University College, Al-Balqa Applied University, Irbid 19117, Jordan, Department of Computer Information Science, Higher Colleges of Technology, Sharjah, United Arab Emirates
ammarnav6@bau.edu.jo.

Shahzad Memon
Department of Computer Science and Digital Technologies ,Faculty of Architecture, Computing and Engineering, University of East London, UK
smemon@uel.ac.uk

***Abstract*— Cybersecurity has emerged as one of the most prevalent and significant challenges in recent years due to the advancement of technology. Among the most frequent and hazardous cybersecurity threats are spam URLs (Uniform Resource Locators), which are also one of the most popular methods for user fraud. Users are the victims of this attack, which also steals their data and infects their devices with harmful software. The detection of spam URLs has become very important in protecting the user. Therefore, this study aims to investigate the efficiency of machine learning classifiers in detecting spam URLs. The following machine learning classifiers were chosen: Random Forest, Decision Tree, and SVM. The evaluation was based on the ISCXURL2016 dataset, which is divided into three groups: All Features, BestFirst Features, and Infogain Features and evaluation matrices were the Accuracy, Precision, Sensitivity, and F-measure. The results obtained showed that Random Forest with All Features is superior to others with an accuracy of 99.75%, Precision of 99.74%, and Sensitivity of 99. 79%, and F-measure 99.76 %.**

Keywords: URL Spam, Random Forest, Decision Tree, SVM, ISCXURL2016

I. INTRODUCTION

In this era of digitized connectivity, the Internet acts as a medium for communication, commerce, and even dissemination of information to the maximum effect one can imagine . However, the phenomenal growth in the use of the Internet is also accompanied by increasing cyber threats [1] [2] [3] one of which is URL spam [4] [5] [6]. URL spam refers to causing link-baiting through deceptive web links for search engine ranking manipulation, distributing malicious software, or phishing purposes [7]. Such spam links are spread across various channels: emails, social messaging platforms, blogs, forums, and other unnoticed avenues that reach the targeted users. URL spam is also the greatest cause, besides being a major damage to the credibility of digital media, and to users' security and privacy [8]. Thus, URL spam detection and possible countermeasures have become a Subject of interest for researchers and practitioners from the cybersecurity domain[9] [10] [11].

Traditionally, URL spam detection relied on heuristic-based systems, blacklists, and manually made rules [12]. Existing methods do show effectiveness in certain scenarios, but their disadvantage lies in scalability and their drawback of quickly losing touch with new strategies employed by spammers. Some of the common techniques include domain obfuscation, URL shortening, and dynamic link generation that spammers would have used to bypass detection. These approaches make it less manageable for the rule-based system to detect spam links with a good degree of confidence. The use of blacklists, though useful, is contingent mostly on the definition and is not real-time in the detection of novel spam URLs. With the advent of more sophisticated spamming tactics, the need for more proactive, advanced detection mechanisms has also arisen.

Machine learning (ML) has turned out to be quite beneficial for dealing with some of the inadequacies of URL spam detection through conventional approaches [13] [14]. It has enabled the analysis of large datasets automatedly [15] [16], the detection of complex patterns, and predictions on whether or not a particular URL is legitimate from the data-driven algorithms. ML-based approaches differ from static rule-based systems because they can learn from updated datasets to adapt to new spam techniques. In addition, feature types that can be processed by machine learning classifiers are lexical, host-based, and content-related features [17] [18] based upon which informed decisions can be made regarding whether a URL is spam or legitimate. Hence, machine learning appears to be a powerful tool that could help develop accurate, effective, and scalable URL spam detection systems.

Investigating the application of machine learning classifiers in URL spam detection is what this study aims. It will include a detailed study of the comparative performance of different algorithms in identifying spam URLs. This will mainly include testing out some supervised machine learning techniques such as decision trees [19] [20] [21], random forests [22] [23], and support vector machines [24] [25] [26] for their ability to classify links as legitimate or spam based on high-level features. The present study will systematically evaluate these classifiers toward establishing an efficient method in real-world implementations while providing insight into their practical usability in cybersecurity systems.

The presented study focuses on the evaluation of how well different classifiers can be evaluated based on some metrics such as accuracy, precision, recall, and F1-score [27] [28] [29]. Primarily the study compares algorithms with respect to strengths and limitations, so as to facilitate selection of best approaches for URL spam detection.

The paper's organization is as follows: Section 2 contains the literature review on URL spam detection using machine learning classifiers. Section 3 summarizes the methodology followed by the study, including data collection, feature engineering, and implementation of a machine learning classifier. A presentation of the experimental results with the implications on URL spam detection would be contained in Section 4. Finally, the whole work ends with Section 5, which offers a summary of the major findings as well as suggestions for future studies.

II. Related works

Detecting URL spam has been the main goal of researchers for some time now because of its role in the security of online platforms from malicious behaviors. Using machine-learning classifiers to identify spam URLs has been the focus of many studies conducted in this area. This section discusses the most recent studies.

A study by Mankar, Nikhilesh P., et al. [30] analyzes diverse ML classifiers for effective and efficient identification of malware URLs using the ISCX-URL-2016 dataset; the study evaluates seven machine learning models, including Decision Tree, Random Forest, AdaBoost, K-Nearest Neighbors, Support Vector Machine with Stochastic Gradient Descent, Extra Trees, and Naive Bayes. It has been proven from the results of this study that models that belong to the ensemble category, like random forest and extra trees, will give the very best performance, thereby outperforming all other models with over 91% accuracy.

Other study by Akar, Funda. [31], The eight machine learning algorithms are being compared using the Spam URLs Classification Dataset to test the algorithms for spam and non-spam URL detection. The algorithms employed in this research are logistic regression, decision tree, random forest, naive Bayes, K-nearest neighbor, XGBoost, AdaBoost, and gradient boosting. The results of the study revealed that the Random Forest Classifier outperformed other algorithms with a 94.16% accuracy rate, followed by Decision Tree, XGBoost, and Gradient Boosting.

Other Study by O. H. Odeh, A. Arram, and M. Njoum. [32], The present study has been conducted to compare the operational performance of the nine dissimilar machine learning algothims with respect to a Spam URLs classification dataset. Nine different machine learning techniques have found their implementations in this study; these are MLP, KNeighbours, Gradient Boosting, Decision Tree, Naive Bayes, Ada Boost, Random Forest, Bagging Classifier, and Stacking Classifier. As per the outcome, Bagging Classifier gained maximum accuracy of 96.52% followed by Stacking Classifier with 96.21%.

As well as, the study by M. Yıldırım. [33], In this research, the performance of various machine learning models was compared-analyzing the nine machine learning techniques in automatic spam URLs using Spam URLs classification dataset. These machine learning techniques include KNN, Gradient Boosting, Logistic Regression, XgBoost, Support Vector Machine (SVM), Naive Bayes, and Random Forest. The results of the study obtained showed that Random Forest classifier achieved the best accuracy of 93.77%, next by KNN, which achieved 91.39%, followed by Gradient Boosting with 84.58%. Naive Bayes classified it as the worst with an accuracy of 73.03%.

A Study by A. K. Jilani and J. Sultana. [34], The present study assessed spam URL classification as an aspect of machine-learning practices such as Random Forests, Naive Bayes, and Support Vector Machine, acting to differentiate between true websites from phony for system safety through Spam URLs classification dataset. From the study, Random Forest implemented on 10-fold cross-validation has a high classification accuracy of 97% on the URL data and is followed by Support Vector Machine with a classification accuracy of 92% and then Naive Bayes with classification accuracy of 91%.

A study by Y. Kontsewaya, E. Antonov, and A. Artamonov. [35], This study analyzes six algorithms performance comparison within the broader scope of machine learning techniques for spam detection: Naive Bayes, K-Nearest Neighbors, SVM, and Logistic Regression, Decision Tree, Random Forest based on Spam filter dataset, and finally concludes that Logistic regression and Naive Bayes yield maximum accuracy levels reaching 99%.

A another study by A. Begum and S. Badugu. [36], Machine Learning for Detection of Malicious URLs: Recent Methods and Progressing Techniques. All such methods involve the use of several machine learning algorithms like Support Vector Machine, Random Tree, Random Forest, Naive Bayes, Logistic, J48 Bayes Net, and Logistic regression. The detection results reveal that above methods can attain higher accuracy in the detection of malicious URLs.

III. Proposed URL Spam Detection model

Fig 1 shows the proposed URL spam detection model flowchart. The proposed model goes through several stages.

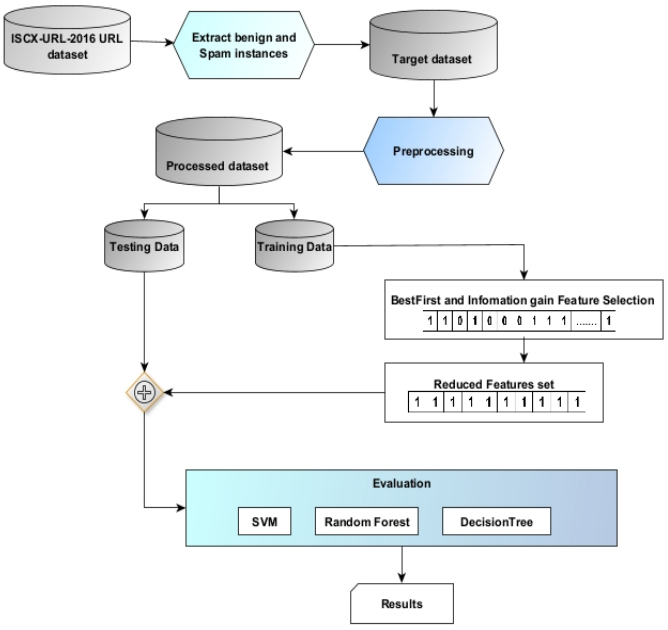

Fig.1: Proposed URL Spam Detection model

A. ISCX-URL-2016 URL Dataset

The ISCX-URL-2016 dataset [37] was made public for evaluation purposes of URL classification tasks from Canadian Institute for Cyber Security. The dataset contains 79 features in total, classified into four classes of URLs, Spam, Malware, Phishing and Benign.

B. Extract Benign and Spam Instances

The ISCX-URL-2016 dataset consists of four classes of URLs, i.e. spam, malware, phishing, defacement, and benign. This work focuses on spam detection, and therefore the data pertaining to attack spam URLs and benign URLs are extracted to form a new dataset called target dataset.

C. Preprocessing

Pre-processing is the act of giving initial raw data a suitable format for processing. Normally, the main steps of this process include: Cleaning - Removal of irrelevant or duplicated data; Normalization - Conversion of data into a common format (lowercase letters and inner special characters removed); Transformation - preparation of the raw data into a form suitable for a machine-learning model.

D. Processed Dataset

The preprocessed data is divided into the training data used to train the machine-learning model and the test data to evaluate the efficiency of the model after training.

E. Feature Selection (BestFirst and Information Gain)

Dimensional reduction through feature selection to enhance model performance has been a vital step in machine learning. The ISCXURL2016 dataset has two feature selection methods, BestFirst Features, and Infogain Features. 79 features are present, where BestFirst reduces it to 6 features, and Infogain reduces it to 5 features.

F. Machine Learning Classifiers

The processed data is trained on various machine learning models, and in this study, three classifiers are enlisted: Support Vector Machine (SVM): very optimal classifier finding hyperplane giving best separation among analytically disparate classes. Random Forests: an ensemble learning approach that uses several possible decision trees to enable prediction accuracy and avoid overfitting. Decision Tree: Very simple model partitions its data using features for prediction.

IV. RESULTS AND DISCUSSION

The selected ML classifiers for detecting spam URL attacks on the ISCXURL2016 dataset are SVM, random forest , and decision tree. The experiments are carried out using spider Python 5.5.1 on a GHz i7 CPU with 6.0 GB RAM. The output of three classifiers is then evaluated using a variety of evaluation matrices, including precision, accuracy, F-measure, and Sensitivity. The obtained results are shown in Table.1

TABLE I. RESULTS

		Random Forest	Decision Tree	SVM
All Features	Accuracy	99.75	99.61	99.19
	F-measure	99.76	99.64	99.25
	Sensitivity	99.79	99.74	98.84
	Precision	99.74	99.53	99.65
BestFirst	Accuracy	98.69	98.71	98.39
	F-measure	98.78	98.80	98.49
	Sensitivity	98.84	98.89	97.94
	Precision	98.72	98.72	99.05
Infogain	Accuracy	99.54	99.56	97.86
	F-measure	99.57	99.59	97.98
	Sensitivity	99.61	99.66	96.91
	Precision	99.53	99.53	99.08

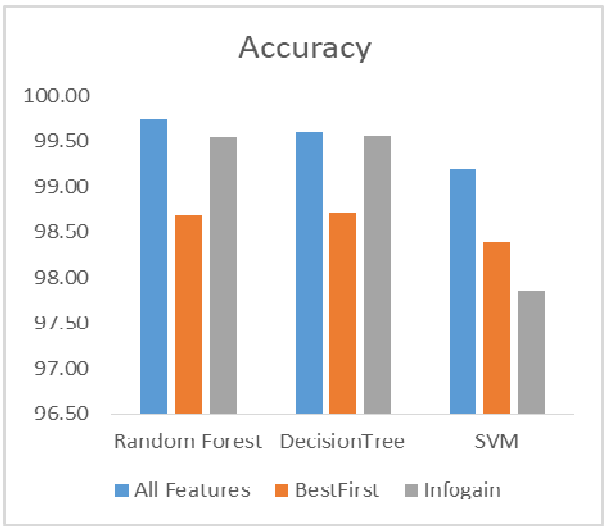

Fig 2. Accuracy

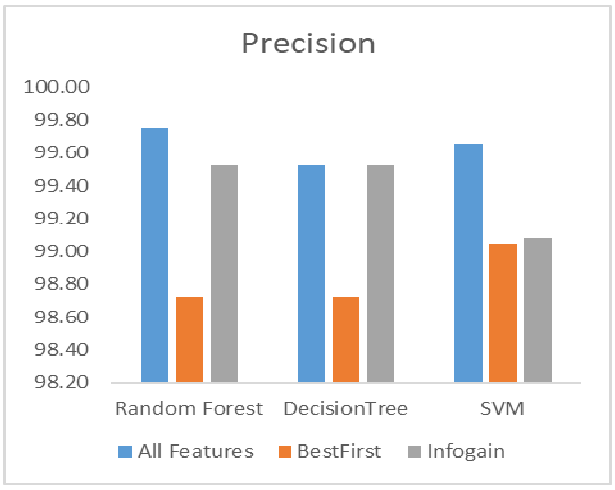

Fig 3. Precision

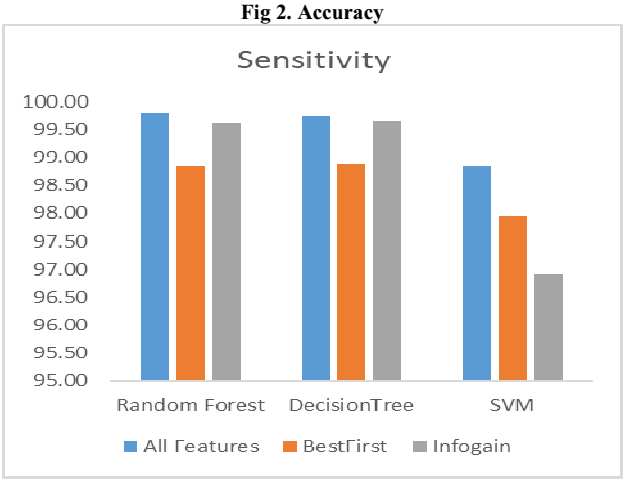

Fig 4. Sensitivity

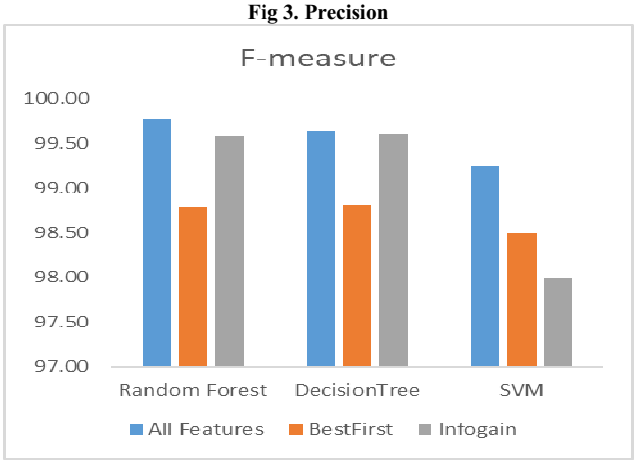

Fig 5. F-measure

Fig 2 shows the accuracy of three different ML classifiers Random Forest, Decision Tree, and SVM, across All Features and the other feature selection methods: BestFirst, and Infogain. The results revealed that random forest with All features Achieves the highest accuracy (99.75%) as compared to other classifiers and feature selection methods. Fig 3 shows the precision of three different ML classifiers Random Forest, Decision Tree, and SVM concerning All Features, BestFirst, and Infogain. The obtained results demonstrate that the random forest has a robust precision across different feature selection methods compared to other classifiers. Fig 4 shows the sensitivity of the three tested ML classifiers across different feature selection methods. The obtained results demonstrate that the random forest classifier consistently shows strong sensitivity. Fig 5 shows the F-measure of the three ML classifiers across different feature selection methods. Regarding the F-measure, the random forest with all features is the highest with 99.76% as compared to other classifiers and feature selection methods. Finally, Fig 6 shows the obtained results of all tested classifiers concerning all features and different feature selection methods.

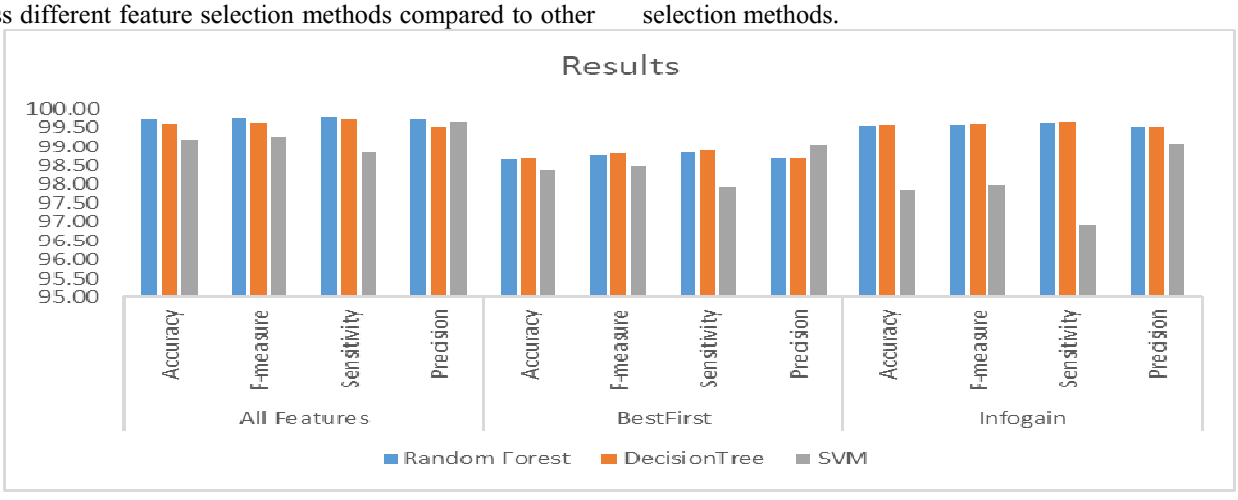

Fig 6. Results obtained

V. CONCLUSION

This paper proposed a model for detecting URL Spam Attacks based on Random Forest, Decision Tree, and SVM ML classifiers and feature selection methods. The ISCXURL2016 dataset was used to test the proposed model, and the model was evaluated using accuracy, precision, sensitivity, and F-measure. The results of the experiment demonstrated that the Random Forest is the most efficient classifier among other examined classifiers to detect URL spam attacks with all features. Its accuracy of 99.75%, Precision of 99.74%, and Sensitivity of 99. 79%, and F-measure 99.76 %. The obtained results also prove that the BestFirst, and Infogain feature selection methods did not improve classification results, so this led to the use of another method for feature selection. In the future direction of the research, modern feature selection methods using bio-inspired metaheuristic algorithms will be tested as well as different types of URL attacks such as malware, phishing, and defacement will be investigated.

REFERENCES

[1] Y. Baddi, M. A. Almaiah, O. Almomani, and Y. Maleh, *The Art of Cyber Defense: From Risk Assessment to Threat Intelligence*. CRC Press, 2024.

[2] A. Sholiyi, J. A. Alzubi, O. A. Alzubi, O. Almomani, and T. O'Farrell, "Near capacity irregular turbo code," *Indian Journal of Science and Technology,* vol. 8, no. 23, 2016.

[3] M. M. Abualhaj, M. Al-Zyoud, A. Alsaaidah, A. Abu-Shareha, and S. Al-Khatib, "Enhancing Malware Detection through Self-Union Feature Selection Using Firefly Algorithm with Random Forest Classification," *International Journal of Intelligent Engineering Systems,* vol. 17, no. 4, 2024.

[4] İ. Yurtseven, S. Bagriyanik, and S. Ayvaz, "A review of spam detection in social media," in *2021 6th International Conference on Computer Science and Engineering (UBMK)*, 2021, pp. 383-388: IEEE.

[5] M. Abualhaj, "Enhancing Spam Detection Using Hybrid of Harris Hawks and Firefly Optimization Algorithms," *Journal of Soft Computing Data Mining,* vol. 5, no. 2, pp. 161-174, 2024.

[6] M. M. Abualhaj, M. O. Hiari, A. Alsaaidah, M. Al-Zyoud, and S. Al-Khatib, "Spam Feature Selection Using Firefly Metaheuristic Algorithm," *Journal of Applied Data Sciences,* vol. 5, no. 4, pp. 1692-1700, 2024.

[7] S. Kaddoura, G. Chandrasekaran, D. E. Popescu, and J. H. Duraisamy, "A systematic literature review on spam content detection and classification," *PeerJ Computer Science,* vol. 8, p. e830, 2022.

[8] A. K. Jain, S. R. Sahoo, and J. Kaubiyal, "Online social networks security and privacy: comprehensive review and analysis," *Complex Intelligent Systems,* vol. 7, no. 5, pp. 2157-2177, 2021.

[9] S. Rao, A. K. Verma, and T. Bhatia, "A review on social spam detection: Challenges, open issues, and future directions," *Expert Systems with Applications,* vol. 186, p. 115742, 2021.

[10] M. A. Almaiah, L. M. Saqr, L. A. Al-Rawwash, L. A. Altellawi, R. Al-Ali, and O. Almomani, "Classification of Cybersecurity Threats, Vulnerabilities and Countermeasures in Database Systems," *Computers, Materials & Continua* vol. 81, no. 2, 2024.

[11] O. Almomani, Alsaaidah, A., Abu-Shareha, A. A., Alza qebah, A., Almaiah, M. A., & Shambour, Q., "Enhance URL Defacement Attack Detection Using Particle Swarm Optimization and Machine Learning," *Journal of Computational and Cognitive Engineering,* vol. 4, no. 1, 2025.

[12] C. M. R. da Silva, E. L. Feitosa, and V. C. Garcia, "Heuristic-based strategy for Phishing prediction: A survey of URL-based approach," *Computers Security,* vol. 88, p. 101613, 2020.

[13] T. Tabassum, M. M. Alam, M. S. Ejaz, and M. K. Hasan, "A Review on Malicious URLs Detection Using Machine Learning Methods," *Journal of Engineering Research Reports,* vol. 25, no. 12, pp. 76-88, 2023.

[14] A. Almaiah and O. Almomani, "An investigator digital forensics frequencies particle swarm optimization for dectection and classification of apt attack in fog computing enviroment (idf-fpso)," *J. Theor. Appl. Inf. Technol,* vol. 15, p. 98, 2020.

[15] O. Almomani, "A Hybrid Model Using Bio-Inspired Metaheuristic Algorithms for Network Intrusion Detection System," *Computers, Materials & Continua,* vol. 68, no. 1, 2021.

[16] O. Almomani, "A feature selection model for network intrusion detection system based on PSO, GWO, FFA and GA algorithms," *Symmetry,* vol. 12, no. 6, p. 1046, 2020.

[17] T. A. Al-Amiedy, M. Anbar, B. Belaton, A. A. Bahashwan, and M. M. Abualhaj, "Towards a Lightweight Detection System Leveraging Ranking Techniques with Wrapper Feature Selection Algorithm for Selective Forwarding Attacks in Low power and Lossy Networks of IoTs," in *2024 4th International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, 2024, pp. 1-17: IEEE.

[18] A. H. Mohammad, T. Alwada'n, O. Almomani, S. Smadi, and N. ElOmari, "Bio-inspired hybrid feature selection model for intrusion detection," *Computers, Materials & Continua,* vol. 73, no. 1, pp. 133-150, 2022.

[19] M. Abualhaj *et al.*, "A fine-tuning of decision tree classifier for ransomware detection based on memory data," *International Journal of Data Network Science,* vol. 8, no. 2, pp. 733-742, 2024.

[20] M. M. Abualhaj and S. N. Al-Khatib, "Using decision tree classifier to detect Trojan Horse based on memory data," *TELKOMNIKA,* vol. 22, no. 2, pp. 393-400, 2024.

[21] A. H. Mohammad, O. Al-Momani, and T. Alwada'n, "Arabic text categorization using k-nearest neighbour, Decision Trees (C4. 5) and Rocchio classifier: a comparative study," *International Journal of Current Engineering Technology,* vol. 6, no. 2, pp. 477-482, 2016.

[22] L. Al-Dabbas and A. A. Abu-Shareha, "Early Detection of Female Type-2 Diabetes using Machine Learning and Oversampling Techniques," *Journal of Applied Data Sciences,* vol. 5, no. 3, pp. 1237-1245, 2024.

[23] M. Madi, F. Jarghon, Y. Fazea, O. Almomani, and A. Saaidah, "Comparative analysis of classification techniques for network fault management," *Turkish Journal of Electrical Engineering Computer Sciences,* vol. 28, no. 3, pp. 1442-1457, 2020.

[24] A. Alsaaidah, O. Almomani, A. A. Abu-Shareha, M. M. Abualhaj, and A. Achuthan, "ARP Spoofing Attack Detection Model in IoT Network using Machine Learning: Complexity vs. Accuracy," *Journal of Applied Data Sciences,* vol. 5, no. 4, pp. 1850-1860, 2024.

[25] O. Almomani, A. Alsaaidah, A. A. A. Shareha, A. Alzaqebah, and M. Almomani, "Performance Evaluation of Machine Learning Classifiers for Predicting Denial-of-Service Attack in Internet of Things," *International Journal of Advanced Computer Science Applications,* vol. 15, no. 1, 2024.

[26] M. A. Almaiah *et al.*, "Performance investigation of principal component analysis for intrusion detection system using different support vector machine kernels," *Electronics,* vol. 11, no. 21, p. 3571, 2022.

[27] O. Almomani, M. A. Almaiah, M. Madi, A. Alsaaidah, M. A. Almomani, and S. Smadi, "Reconnaissance attack detection via boosting machine learning classifiers," in *AIP Conference Proceedings*, 2023, vol. 2979, no. 1: AIP Publishing.

[28] O. Almomani, M. A. Almaiah, A. Alsaaidah, S. Smadi, A. H. Mohammad, and A. Althunibat, "Machine learning classifiers for network intrusion detection system: comparative study," in *2021 International Conference on Information Technology (ICIT)*, 2021, pp. 440-445: IEEE.

[29] A. A. Abu-Shareha, H. Qutaishat, and A. Al-Khayat, "A Framework for Diabetes Detection Using Machine Learning and Data Preprocessing," *Journal of Applied Data Sciences,* vol. 5, no. 4, pp. 1654-1667, 2024.

[30] N. P. Mankar, P. E. Sakunde, S. Zurange, A. Date, V. Borate, and Y. K. Mali, "Comparative Evaluation of Machine Learning Models for Malicious URL Detection," in *2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon)*, 2024, pp. 1-7: IEEE.

[31] F. J. J. o. S. R.-A. Akar, "Data correlation matrix-based spam URL detection using machine learning algorithms," *Journal of Scientific Reports-A* no. 056, pp. 56-69, 2024.

[32] O. H. Odeh, A. Arram, and M. Njoum, "Classification of Spam URLs Using Machine Learning Approaches," *arXiv preprint arXiv:.05953,* 2023.

[33] M. Yıldırım, "Using and Comparing Machine Learning Techniques for Automatic Detection of Spam Website URLs," *NATURENGS : MTU Journal of Engineering and Natural Sciences,* vol. 3, no. 1, pp. 33-41, 2022.

[34] A. K. Jilani and J. Sultana, "A Random Forest Based Approach to Classify Spam URLs Data," in *2022 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*, 2022, pp. 268-272: IEEE.

[35] Y. Kontsewaya, E. Antonov, and A. Artamonov, "Evaluating the effectiveness of machine learning methods for spam detection," *Procedia Computer Science,* vol. 190, pp. 479-486, 2021.

[36] A. Begum and S. Badugu, "A study of malicious url detection using machine learning and heuristic approaches," in *International Conference on E-Business and Telecommunications*, 2019, pp. 587-597: Springer.

[37] M. S. I. Mamun, M. A. Rathore, A. H. Lashkari, N. Stakhanova, and A. A. Ghorbani, "Detecting Malicious URLs Using Lexical Analysis," in *Network and System Security, Springer International Publishing*, Cham, 2016, pp. 467-482: Springer International Publishing.