Impact Factor: 6.252

Phishing Website Detection using Machine Learning

Indubu Nutana[1], J. Tagore Babu[2], J. Jahnavi[3], J. Yogendra[4], J. Siva Tanay Akash[5]
GMR Institute of Technology, Rajam, Andhra Pradesh, India
20341A0571@gmrit.edu.in[1], 20341A0572@gmrit.edu.in[2], 20341A0575@gmrit.edu.in[3],
20341A0574@gmrit.edu.in[4], 20341A0573@gmrit.edu.in[5]

Abstract: *Phishing websites today pose serious risks because of their almost invisible risk. In these attacks, fraudsters trick users into handing over their login information or other sensitive information through a login form that mimics the target website and submits the information to an evil server. In recent years, numerous anti-phishing strategies have been created, using various resources such as the URL and HTML code from trustworthy and fraudulent index websites. These methods have significant restrictions when predicting authentic login websites because they frequently lack login forms to specify the appropriate class that was used to train the suggested model. The URL, HTML, and web technology properties are used to detect phishing websites in actual situations. In this work the phishing website datasets will be used with machine learning algorithms like Random Forest, Logistic regression, KNN and SVM to test the best method, so the crawl trusted websites are taken to align with phishing perspective. The study involves comparison of the four algorithms and finding efficient algorithm in the basis of accuracy.*

Keywords: Phishing, Cybercrime, Detection, Phishing Dataset, Machine Learning

REFERENCES

[1]. Sánchez-Paniagua, M., Fidalgo, E., Alegre, E., & Alaiz-Rodríguez, R. (2022). Phishing websites detection using a novel multipurpose dataset and web technologies features. Expert Systems with Applications, 207, 118010.

[2]. Hannousse, A., & Yahiouche, S. (2021). Towards benchmark datasets for machine learning based website phishing detection: An experimental study. Engineering Applications of Artificial Intelligence, 104, 104347.

[3]. Shahrivari, V., Darabi, M. M., & Izadi, M. (2020). Phishing detection using machine learning techniques. arXiv preprint arXiv:2009.11116.

[4]. Ramana, A. V., Rao, K. L., & Rao, R. S. (2021). Stop-Phish: an intelligent phishing detection method using feature selection ensemble. Social Network Analysis and Mining, 11(1), 1-9.

[5]. Rao, R. S., Vaishnavi, T., & Pais, A. R. (2020). CatchPhish: detection of phishing websites by inspecting URLs. Journal of Ambient Intelligence and Humanized Computing, 11(2), 813-825.

[6]. Fitzpatrick, B., Liang, X., & Straub, J. (2021). Fake news and phishing detection using a machine learning trained expert system. arXiv preprint arXiv:2108.08264.

[7]. Christou, O., Pitropakis, N., Papadopoulos, P., McKeown, S., & Buchanan, W. J. (2020). Phishing url detection through top-level domain analysis: A descriptive approach. arXiv preprint arXiv:2005.06599.

[8]. Tupsamudre, H., Jain, S., & Lodha, S. (2021). PhishMatch: A Layered Approach for Effective Detection of Phishing URLs. arXiv preprint arXiv:2112.02226.

[9]. Lakshmanarao, A., Rao, P. S. P., & Krishna, M. B. (2021, March). Phishing website detection using novel machine learning fusion approach. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1164-1169). IEEE.

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

[11]. Abuzuraiq, A., Alkasassbeh, M., & Almseidin, M. (2020, April). Intelligent methods for accurately detecting phishing websites. In 2020 19th International Conference on Information and Communication Systems (ICICS) (pp. 085-090). IEEE.

[12]. He, S., Li, B., Peng, H., Xin, J., & Zhang, E. (2021). An effective cost-sensitive XGBoost method for malicious URLs detection in imbalanced dataset. IEEE Access, 9, 93089-93096.

[13]. Mao, J., Bian, J., Tian, W., Zhu, S., Wei, T., Li, A., & Liang, Z. (2019). Detecting phishing websites via aggregation analysis of page layouts. Procedia Computer Science, 129, 224-230.

[14]. Aljofey, A., Jiang, Q., Rasool, A., Chen, H., Liu, W., Qu, Q., & Wang, Y. (2022). An effective detection approach for phishing websites using URL and HTML features. Scientific Reports, 12(1), 1-19.

[15]. Wood, T., Basto-Fernandes, V., Boiten, E., & Yevseyeva, I. (2022). Systematic Literature Review: Anti-Phishing Defences and Their Application to Before-the-click Phishing Email Detection. arXiv preprint arXiv:2204.13054.