

REFERENCES

- [1] International Telecommunication Union (ITU), “Internet Use 2025.” Accessed: Jan. 21, 2026. [Online]. Available: <https://www.itu.int/itu-d/reports/statistics/2025/10/15/ff25-internet-use/>
- [2] European Union Agency for Cybersecurity (ENISA), “Threat Landscape | ENISA.” Accessed: Jan. 21, 2026. [Online]. Available: <https://www.enisa.europa.eu/topics/cyber-threats/threat-landscape>
- [3] Anti-Phishing Working Group, “APWG PHISHING ACTIVITY TRENDS REPORT 1st Quarter 2025,” Jul. 2025.
- [4] Anti-Phishing Working Group, “APWG PHISHING ACTIVITY TRENDS REPORT 2nd Quarter 2025,” Aug. 2025.
- [5] Anti-Phishing Working Group, “APWG PHISHING ACTIVITY TRENDS REPORT 3rd Quarter 2025,” Dec. 2025. Accessed: Jul. 17, 2026. [Online]. Available: https://docs.apwg.org/reports/apwg_trends_report_q3_2025.pdf
- [6] “Microsoft Digital Defense Report 2025 | Microsoft.” Accessed: Jan. 21, 2026. [Online]. Available: <https://www.microsoft.com/en-us/corporate-responsibility/cybersecurity/microsoft-digital-defense-report-2025/>
- [7] Verizon Business, “2025 Data Breach Investigations Report,” 2025.
- [8] C. Opara, Y. Chen, and B. Wei, “Look before you leap: Detecting phishing web pages by exploiting raw URL and HTML characteristics,” *Expert Syst. Appl.*, vol. 236, Feb. 2024, doi: 10.1016/j.eswa.2023.121183.
- [9] A. Aljofey *et al.*, “An effective detection approach for phishing websites using URL and HTML features,” *Sci. Rep.*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-10841-5.
- [10] S. Das Gupta, K. T. Shahriar, H. Alqahtani, D. Alsalman, and I. H. Sarker, “Modeling Hybrid Feature-Based Phishing Websites Detection Using Machine Learning Techniques,” *Annals of Data Science*, vol. 11, no. 1, pp. 217–242, Feb. 2024, doi: 10.1007/s40745-022-00379-8.
- [11] N. Q. Do, A. Selamat, O. Krejcar, E. Herrera-Viedma, and H. Fujita, “Deep Learning for Phishing Detection: Taxonomy, Current Challenges and Future Directions,” 2022, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2022.3151903.

- [12] S. Bell and P. Komisarczuk, “An Analysis of Phishing Blacklists: Google Safe Browsing, OpenPhish, and PhishTank,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Feb. 2020. doi: 10.1145/3373017.3373020.
- [13] M. K. Prabakaran, P. Meenakshi Sundaram, and A. D. Chandrasekar, “An enhanced deep learning-based phishing detection mechanism to effectively identify malicious URLs using variational autoencoders,” *IET Inf. Secur.*, vol. 17, no. 3, pp. 423–440, May 2023, doi: 10.1049/ise2.12106.
- [14] A. Odeh, Q. A. Al-Haija, A. Aref, and A. A. Taleb, “Comparative Study of CatBoost, XGBoost, and LightGBM for Enhanced URL Phishing Detection: A Performance Assessment,” *Journal of Internet Services and Information Security*, vol. 13, no. 4, pp. 1–11, Nov. 2023, doi: 10.58346/JISIS.2023.I4.001.
- [15] D. M. Linh, H. D. Hung, H. M. Chau, Q. S. Vu, and T.-N. Tran, “Real-time phishing detection using deep learning methods by extensions,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, pp. 3021–3035, Jun. 2024, doi: 10.11591/ijece.v14i3.pp3021-3035.
- [16] M. N. Yeasmin, M. A. R. Refat, B. C. Singh, Z. Alom, Z. Aung, and M. Azim, “EnLeM: ensemble learning-based model to detect phishing websites,” *EURASIP J. Inf. Secur.*, vol. 2026, no. 1, Dec. 2025, doi: 10.1186/s13635-025-00219-1.
- [17] T. Berners-Lee, L. Masinter, and M. McCahill, “Uniform Resource Locators (URL),” Dec. 1994, doi: 10.17487/RFC1738.
- [18] J. Hawkins, “What’s in a Domain? Analysis of URL Features,” Academy and Industry Research Collaboration Center (AIRCC), Aug. 2023, pp. 85–93. doi: 10.5121/csit.2023.131409.
- [19] “HTTP vs HTTPS - Perbedaan Antara Protokol Transfer - AWS.” Accessed: Jan. 23, 2026. [Online]. Available: <https://aws.amazon.com/id/compare/the-difference-between-https-and-http/>
- [20] E. Rescorla, “The Transport Layer Security (TLS) Protocol Version 1.3,” Aug. 2018, doi: 10.17487/RFC8446.

- [21] T. Berners-Lee and D. Connolly, “Hypertext Markup Language - 2.0,” Nov. 1995, doi: 10.17487/RFC1866.
- [22] “HTML Standard.” Accessed: Jan. 23, 2026. [Online]. Available: <https://html.spec.whatwg.org/multipage/>
- [23] R. Tabarés, “HTML5 and the evolution of HTML; tracing the origins of digital platforms,” *Technol. Soc.*, vol. 65, May 2021, doi: 10.1016/j.techsoc.2021.101529.
- [24] “The Web Within the Web - IEEE Spectrum.” Accessed: Jan. 23, 2026. [Online]. Available: <https://spectrum.ieee.org/the-web-within-the-web>
- [25] I. Gur *et al.*, “Understanding HTML with Large Language Models,” *Findings of the Association for Computational Linguistics: EMNLP 2023*, p. 2803, 2023, doi: 10.18653/v1/2023.findings-emnlp.185.
- [26] “Browser Extensions: How to Vet and Install Safely | Information Security Office.” Accessed: Jan. 24, 2026. [Online]. Available: <https://security.berkeley.edu/education-awareness/browser-extensions-how-vet-and-install-safely>
- [27] “Everything to Know about Browser Extensions | Microsoft Edge.” Accessed: Jan. 24, 2026. [Online]. Available: <https://www.microsoft.com/en-us/edge/learning-center/everything-to-know-about-browser-extensions?form=MA13I2>
- [28] N. S. Gurjar, S. R. Sudheendra, C. S. Kumar, and K. S. Krishnaveni, “WebSecAsst - A Machine Learning based Chrome Extension,” in *Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021*, Institute of Electrical and Electronics Engineers Inc., Jul. 2021, pp. 1631–1635. doi: 10.1109/ICCES51350.2021.9488953.
- [29] L. Tang and Q. H. Mahmoud, “A Deep Learning-Based Framework for Phishing Website Detection,” *IEEE Access*, vol. 10, pp. 1509–1521, 2022, doi: 10.1109/ACCESS.2021.3137636.
- [30] Ltd. Huawei Technologies Co., *Artificial Intelligence Technology*. Springer Nature, 2022. doi: 10.1007/978-981-19-2879-6.

- [31] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *J. Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00444-8.
- [32] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
- [33] H. Bourlard and S. H. Kabil, “Autoencoders reloaded,” *Biol. Cybern.*, vol. 116, no. 4, pp. 389–406, Aug. 2022, doi: 10.1007/s00422-022-00937-6.
- [34] H. Torabi, S. L. Mirtaheri, and S. Greco, “Practical autoencoder based anomaly detection by using vector reconstruction error,” *Cybersecurity*, vol. 6, no. 1, Dec. 2023, doi: 10.1186/s42400-022-00134-9.
- [35] J. Qian, Z. Song, Y. Yao, Z. Zhu, and X. Zhang, “A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes,” Dec. 15, 2022, *Elsevier B.V.* doi: 10.1016/j.chemolab.2022.104711.
- [36] K. Berahmand, F. Daneshfar, E. S. Salehi, Y. Li, and Y. Xu, “Autoencoders and their applications in machine learning: a survey,” *Artif. Intell. Rev.*, vol. 57, no. 2, Feb. 2024, doi: 10.1007/s10462-023-10662-6.
- [37] Z. Arif Ali, Z. H. Abduljabbar, H. A. Tahir, A. Bibo Sallow, and S. M. Almufti, “eXtreme Gradient Boosting Algorithm with Machine Learning: a Review,” *Academic Journal of Nawroz University*, vol. 12, no. 2, pp. 320–334, May 2023, doi: 10.25007/ajnu.v12n2a1612.
- [38] Y. Xie and S. P. Tummala, “Machine Learning for Sensor Analytics: A Comprehensive Review and Benchmark of Boosting Algorithms in Healthcare, Environmental, and Energy Applications,” Dec. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/s25237294.
- [39] Z. G. Fang, S. Q. Yang, C. X. Lv, S. Y. An, and W. Wu, “Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: a time-series study,” *BMJ Open*, vol. 12, no. 7, Jul. 2022, doi: 10.1136/bmjopen-2021-056685.

- [40] M. Misiek and T. Hyla, “XGBoost-based URL phishing detection method with cross-dataset validation,” *IEEE Access*, 2026, doi: 10.1109/ACCESS.2026.3672690.
- [41] Microsoft Corporation, *LightGBM Release 4.6.0.99*. 2026. Accessed: May 01, 2026. [Online]. Available: <https://lightgbm.readthedocs.io/>
- [42] L. Li, “LightGBM integration with modified data balancing and whale optimization algorithm for rock mass classification,” *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-73742-9.
- [43] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” *Advances in neural information processing systems 31*, 2018.
- [44] K. Ileri, “Comparative analysis of CatBoost, LightGBM, XGBoost, RF, and DT methods optimised with PSO to estimate the number of k-barriers for intrusion detection in wireless sensor networks,” *International Journal of Machine Learning and Cybernetics*, vol. 16, no. 9, pp. 6937–6956, Sep. 2025, doi: 10.1007/s13042-025-02654-5.
- [45] D. Boldini, F. Grisoni, D. Kuhn, L. Friedrich, and S. A. Sieber, “Practical guidelines for the use of gradient boosting for molecular property prediction,” *J. Cheminform.*, vol. 15, no. 1, Dec. 2023, doi: 10.1186/s13321-023-00743-7.
- [46] M. T. Syamkalla, S. Khomsah, and Y. S. R. Nur, “Implementasi Algoritma Catboost Dan Shapley Additive Explanations (SHAP) Dalam Memprediksi Popularitas Game Indie Pada Platform Steam,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, pp. 777–786, Aug. 2024, doi: 10.25126/jtiik.1148503.
- [47] Q. Ding *et al.*, “Revolutionizing oncology care: pioneering AI models to foresee pneumonia-related mortality,” *Front. Oncol.*, vol. 15, 2025, doi: 10.3389/fonc.2025.1520512.
- [48] L. Gao, L. Zhang, J. Chen, L. Peng, L. Guo, and L. Yang, “Autoencoder-CatBoost Model for Accurate Hyperspectral Quality Assessment of Yunling Snowflake Beef,” *IEEE Access*, vol. 12, pp. 184701–184713, 2024, doi: 10.1109/ACCESS.2024.3510035.