

DAFTAR PUSTAKA

- Abubakar, H. D., & Umar, M. (2022). Sentiment Classification: Review of Text Vectorization Methods: Bag of Words, Tf-Idf, Word2vec and Doc2vec. *SLU Journal of Science and Technology*, 4(1 & 2), 27–33. <https://doi.org/10.56471/slujst.v4i.266>
- Fahim, A. R., Arman, M. S., Sultana, I., Tasnim, N., Rifat Ahmed, K., & Mahmud, I. (2024). PhishGuard: Leveraging NLP and Machine Learning for Email Phishing Detection. *International Conference on Big Data Analytics in Bioinformatics, DABCon 2024*, 1–6. <https://doi.org/10.1109/DABCon63472.2024.10919349>
- Tchakoucht, A. T., Elkari, B., Chaibi, Y., & Kousksou, T. (2024). Random Forest with Feature Selection and K-fold Cross Validation for Predicting The Electrical and Thermal Efficiencies of Air Based Photovoltaic-thermal Systems. *Energy Reports*, 12(July), 988–999. <https://doi.org/10.1016/j.egyr.2024.07.002>
- Al Tawil, A., Almazaydeh, L., Qawasmeh, D., Qawasmeh, B., Alshinwan, M., & Elleithy, K. (2024). Comparative Analysis of Machine Learning Algorithms for Email Phishing Detection Using TF-IDF, Word2Vec, and BERT. *Computers, Materials and Continua*, 81(2), 3395–3412. <https://doi.org/10.32604/cmc.2024.057279>
- Alhuzali, A., Alloqmani, A., Aljabri, M., & Alharbi, F. (2025). In-Depth Analysis of Phishing Email Detection: Evaluating the Performance of Machine Learning and Deep Learning Models Across Multiple Datasets. *Applied Sciences (Switzerland)*, 15(6). <https://doi.org/10.3390/app15063396>
- Almomani, A., Gupta, B. B., Atawneh, S., Meulenberg, A., & Almomani, E. (2013). A Survey of Phishing Email Filtering Techniques. *IEEE Communications Surveys and Tutorials*, 15(4), 2070–2090. <https://doi.org/10.1109/SURV.2013.030713.00020>
- Al-Subaiey, A., Al-Thani, M., Abdullah Alam, N., Antora, K. F., Khandakar, A., & Uz Zaman, S. A. (2024). Novel interpretable and robust web-based AI platform for phishing email detection. *Computers and Electrical Engineering*, 120(PA), 109625. <https://doi.org/10.1016/j.compeleceng.2024.109625>
- Amer, F. A. H. A., & Tasnim, H. K. A. (2024). Overview of Machine Classification Techniques. *BIO Web of Conferences*, 97, 00133. <https://doi.org/10.1051/bioconf/20249700133>
- Ashmore, R., Calinescu, R., & Paterson, C. (2021). *Assuring the Machine Learning Lifecycle: Desiderata, Method, and Challenges*. *ACM Computing Surveys*, 54(5), 1–39. <https://doi.org/10.1145/3453444>
- Awad, M., & Khanna, R. (2015). Support Vector Machines for Classification. In *Efficient learning machines* (pp. 39–66). Apress. https://doi.org/10.1007/978-1-4302-5990-9_3
- Bountakas, P., Koutroumpouchos, K., & Xenakis, C. (2021). A Comparison of Natural Language Processing and Machine Learning Methods for Phishing Email Detection. *ACM International Conference Proceeding Series*, (August 2021). <https://doi.org/10.1145/3465481.3469205>

- Cahyani, D. E., & Patasik, I. (2021). Performance Comparison of Tf-Idf and Word2Vec models for Emotion Text classification. *Bulletin of Electrical Engineering and Informatics*, 10(5), 2780–2788. <https://doi.org/10.11591/eei.v10i5.3157>
- Cahyani, E. F., Ikhsan, A. N., & Astrida, D. N. (2026). *Event-Based Detection of Provocative Political Discourse on Indonesian Twitter: A Comparative Study of SVM and IndoBERT*. 8(1), 530–548. <https://doi.org/10.63158/journalisi.v8i1.1409>
- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges and Trends. *Neurocomputing*, 408, 189–215. <https://doi.org/10.1016/j.neucom.2019.10.118>
- Çınar, A. (2024). Multi-Class Classification with the Gaussian Naive Bayes Algorithm. *Journal of Data Applications*, 0(2), 1–13. <https://doi.org/10.26650/joda.1389471>
- Fares, H., Kilani, J., Fagroud, F. E., Toumi, H., Lakrami, F., Baddi, Y., & Aknin, N. (2024). Machine Learning Approach for Email Phishing Detection. *Procedia Computer Science*, 251, 746–751. <https://doi.org/10.1016/j.procs.2024.11.179>
- Firman, Tukiyyat, & Wiharjo, S. (2025). *Phishing Email Classification Approach Using Machine Learning Algorithms - A Literature Review*. *Data and Metadata*, 3(3), 135–145. <https://doi.org/10.61978/data.v3i3.692>
- Golazad, S. Z., Mohammadi, A., Rashidi, A., & Ilbeigi, M. (2024). Data preprocessing for Construction Machine Learning (ML), Deep Learning (DL), and Reinforcement Learning (RL) Models. *Automation in Construction*, 168, 105844. <https://doi.org/10.1016/j.autcon.2024.105844>
- Gorritz, J. M., Clemente, R. M., Segovia, F., Ramirez, J., Ortiz, A., & Suckling, J. (2024). *Is K-Fold Cross Validation The Best Model Selection Method for Machine Learning?* *arXiv*. <https://doi.org/10.48550/arXiv.2401.16407>
- Ismail, M., Hassan, N., & Saleh Bafjaish, S. (2020). Comparative Analysis of Naive Bayesian Techniques in Health-Related for Classification Task. *Journal of Soft Computing and Data Mining*, 1(2), 1–10. <https://doi.org/10.30880/jscdm.2020.01.02.001>
- Kamel, H., Abdulah, D., & Al-Tuwaijari, J. M. (2019). Cancer Classification Using Gaussian Naive Bayes Algorithm. *Proceedings of the 5th International Engineering Conference, IEC 2019*, 165–170. <https://doi.org/10.1109/IEC47844.2019.8950650>
- Li, D., Chen, *, Wang, & Lun. (2024). Phishing Attacks: Detection and Prevention Techniques. *Journal of Industrial Engineering and Applied Science Journal Home: jieas.suaspress.org*, 2(4). <https://doi.org/10.5281/zenodo.12789572>
- Ma, T. M., Yamamori, K., & Thida, A. (2020). A Comparative Approach to Naïve Bayes Classifier and Support Vector Machine for Email Spam Classification. *2020 IEEE 9th Global Conference on Consumer Electronics, GCCE 2020*, 324–326. <https://doi.org/10.1109/GCCE50665.2020.9291921>

- Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data Pre-processing and Data Augmentation Techniques. *Global Transitions Proceedings*, 3(1), 91–99. <https://doi.org/10.1016/j.gltp.2022.04.020>
- Martins, P., Cardoso, F., Váz, P., Silva, J., & Abbasi, M. (2025). Performance and Scalability of Data Cleaning and Preprocessing Tools: A Benchmark on Large Real-World Datasets. *Data*, 10(5), 1–22. <https://doi.org/10.3390/data10050068>
- Mbadiwe, O. N., Nwokonkwo, O. C., Ikerionwu, C. O., Otuonye, A. I., Etus, C., & Okloegbo, C. A. (2024). *Email Anti-Phishing: Machine Learning Models and Evaluation Overview*. <https://doi.org/10.5281/zenodo.13327914>
- Meléndez, R., Ptaszynski, M., & Masui, F. (2024). Comparative Investigation of Traditional Machine-Learning Models and Transformer Models for Phishing Email Detection. *Electronics (Switzerland)*, 13(24). <https://doi.org/10.3390/electronics13244877>
- Mohammed, C. N., & Ahmed, A. M. (2024). A Semantic-based Model with A Hybrid Feature Engineering Process for Accurate Spam Detection. *Journal of Electrical Systems and Information Technology*, 11(1), 26. <https://doi.org/10.1186/s43067-024-00151-3>
- Mumuni, A., & Mumuni, F. (2025). Automated Data Processing and Feature Engineering for Deep Learning and Big Data Applications: A survey. *Journal of Information and Intelligence*, 3(2), 113–153. <https://doi.org/10.1016/j.jiixd.2024.01.002>
- Octaviani, N. L., Hari Rachmawanto, E., Sari, C. A., & Rosal Ignatius Moses Setiadi, D. (2020). Comparison of Multinomial Naïve Bayes Classifier, Support Vector Machine, and Recurrent Neural Network to Classify Email Spams. *Proceedings - 2020 International Seminar on Application for Technology of Information and Communication: IT Challenges for Sustainability, Scalability, and Security in the Age of Digital Disruption, iSemantic 2020*, 17–21. <https://doi.org/10.1109/iSemantic50169.2020.9234296>
- Rafat, K. F., Xin, Q., Javed, A. R., Jalil, Z., & Ahmad, R. Z. (2022). Evading Obscure Communication From Spam Emails. *Mathematical Biosciences and Engineering*, 19(2), 1926–1943. <https://doi.org/10.3934/mbe.2022091>
- Salloum, S., Gaber, T., Vadera, S., & Shaalan, K. (2022). A Systematic Literature Review on Phishing Email Detection Using Natural Language Processing Techniques. Dalam *IEEE Access* (Vol. 10, hlm. 65703–65727). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2022.3183083>
- Sharma, J. (2024). Review of Machine Learning Algorithms and Their Applications. *International Journal of Innovations in Science Engineering And Management*, 113–119. <https://doi.org/10.69968/ijisem.2024v3i2113-119>
- Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019). Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification. *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, 593–596. <https://doi.org/10.1109/ICACTM.2019.8776800>

- Somesha, M., & Pais, A. R. (2022). Classification of Phishing Email Using Word Embedding and Machine Learning Techniques. *Journal of Cyber Security and Mobility*, 11(3), 279–320. <https://doi.org/10.13052/jcsm2245-1439.1131>
- Talabani, H. S., & Jumaa, I. H. (2024). A Review of Various Machine Learning Techniques and its Application on IoT and Cloud Computing. *Tikrit Journal of Pure Science*, 29(1), 185–195. <https://doi.org/10.25130/tjps.v29i1.1618>
- Tausif, S., Alam, M., Mustajab, S., & Alam, B. (2024). Dynamic Phishing Detection in Cybersecurity Method Using Machine and Deep Learning Approaches. *Journal of Electrical Systems*, 20(3), 9508–9523. <https://hdl.handle.net/20.500.14536/5923>
- Teruel, Z. C., García-vallejo, D., Pintado, C. N., & Abascal, J. D. (2026). Cross-Validated Neural Network Models for Fretting Fatigue Life Prediction Across Multiple Materials and Contact Geometries. *Tribology International*, 112058. <https://doi.org/10.1016/j.triboint.2026.112058>
- Thakur, K., Ali, M. L., Obaidat, M. A., & Kamruzzaman, A. (2023). A Systematic Review on Deep-Learning-Based Phishing Email Detection. Dalam *Electronics (Switzerland)* (Vol. 12, Nomor 21). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/electronics12214545>
- Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based Feature Selection Introduction and Review. *Journal of Biomedical Informatics*, 85(July), 189–203. <https://doi.org/10.1016/j.jbi.2018.07.014>
- Verma, P., Goyal, A., & Gigras, Y. (2020). Email Phishing: Text Classification Using Natural Language Processing. *Computer Science and Information Technologies*, 1(1). <https://doi.org/10.11591/csit.v1i1.p1-12>
- Wilk-Jakubowski, J. L., Pawlik, L., Wilk-Jakubowski, G., & Sikora, A. (2025). Machine Learning and Neural Networks for Phishing Detection: A Systematic Review (2017–2024). *Electronics (Switzerland)*, 14(18), 1–65. <https://doi.org/10.3390/electronics14183744>