

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik Indonesia, “Statistik telekomunikasi indonesia 2023,” <https://www.bps.go.id/id/publication/2024/08/30/f4b846f397ea452bdc2178b3/statistik-telekomunikasi-indonesia-2023.html>, 2023, accessed: Mar. 27, 2025.
- [2] Reuters, “Digital news report 2024 | reuters institute for the study of journalism,” <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>, 2024.
- [3] A. Arrayhan, “Survei penetrasi internet indonesia 2024 asosiasi penyelenggara jasa internet indonesia,” <https://survei.apjii.or.id/survei>, 2024, accessed: Mar. 27, 2025.
- [4] H. Syaputra, D. P. Prahastiwi, R. Dewi, and N. Sari, “Statistik kriminal 2024,” <https://www.bps.go.id/id/publication/2024/12/12/13317138a55b2f7096589536/statistik-kriminal-2024.html>, 2024, accessed: Mar. 27, 2025.
- [5] M. R. Lubis, “Analisis kualitatif dampak hukum pidana terhadap masyarakat,” *Jurnal Ilmu Hukum, Humaniora dan Politik*, vol. 4, no. 4, May 2024.
- [6] Nugraha, “Kebijakan Kriminal Terhadap Kejahatan Begal Di Wilayah Hukum Polsek Medan Barat,” <https://repositori.uma.ac.id/handle/123456789/21979>, 2023.
- [7] D. Ulhaq, N. Suarna, and G. Dwilestari, “Klasifikasi berita kriminal menggunakan algoritma naïve bayes berbasis pso,” *Informatics for Educators and Professionals*, vol. 6, no. 2, pp. 12–21, 2022.
- [8] A. Widyanto, Y. Puspitarani, and A. Purnama, “Analisis pengelompokan data kriminalitas dan kejahatan di indonesia dari portal berita online menggunakan metode latent dirichlet allocation,” *Jurnal Ekonomi Manajemen Sistem Informasi*, vol. 6, no. 3, Jan. 2025.
- [9] J. Philip, B. VeerasesharReddy, M. Harshini, I. V. S. L. Haritha, S. Patil, and S. K. Shareef, “A comparative study of text classification using selective machine learning algorithms,” in *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, May 2023, pp. 482–484.
- [10] R. Dodda, A. Shahebaaz, M. M. Ali, and Z. A. Mohammed, “Performance analysis of various machine learning techniques for classifying topic categories from text news articles,” in *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Oct 2024, pp. 1126–1131.
- [11] T. Thaipisutikul, S. Tuarob, S. Pongpaichet, A. Amornvatcharapong, and T. K. Shih, “Automated classification of criminal and violent activities in thailand from online news articles,” in *13th International Conference on Knowledge and Smart Technology (KST)*, Jan 2021, pp. 170–175.
- [12] A. Rahmawati, A. Alamsyah, and A. Romadhony, “Hoax News Detection Analysis using IndoBERT Deep Learning Methodology,” in *IEEE Xplore*, Aug 2022, pp. 368–373.

- [13] N. Chandana, A. Sreelekha, K. Rasi, J. Sreeja, and P. L. Prassanna, "BCC NEWS Classification Comparison between Naïve Bayes, Support Vector Machine, Recurrent Neural Network," in *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Feb 2021, pp. 833–835.
- [14] Z. M. Mahdi, R. F. Istiqomah, N. Alfarelzi, S. Astuti, I. Asror, and R. Mayasari, "Text Classification Using NLP by Comparing LSTM and Machine Learning Method," in *2024 10th International Conference on Wireless and Telematics (ICWT)*, Jul 2024, pp. 1–7.
- [15] N. N. Qomariyah, T. Sun, and D. Kazakov, "NLP Analysis of COVID-19 Radiology Reports in Indonesian using IndoBERT," in *2022 4th International Conference on Biomedical Engineering (IBIOMED)*, Oct. 2022, pp. 65–70.
- [16] M. Kurniawan and B. Juarto, "Unlocking the Potential of IndoBERT for Classification of Indonesian Thesis," in *2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, Nov. 2024, pp. 785–791.
- [17] L. H. Suadaa, F. Ridho, A. K. Monika, and N. W. K. Projo, "Automatic Text Categorization to Standard Classification of Indonesian Business Fields (KBLI) 2020," in *2023 International Conference on Electrical Engineering and Informatics (ICEEI)*, Oct. 2023.
- [18] K. Akhil, T. B. Vamsi, S. Soujanya, K. Sai, M. Mounisha, and N. Srihitha, "Analyzing Unstructured Data: Natural Language Processing Frameworks, Challenges, Approaches, and Applications," in *2024 IEEE 4th International Conference on ICT in Business Industry & Government (ICTBIG)*, Dec 2024, pp. 1–9.
- [19] G. Jin, "Application Optimization of NLP System under Deep Learning Technology in Text Semantics and Text Classification," in *2022 International Conference on Education, Network and Information Technology (ICENIT)*, Sep 2022, pp. 279–283.
- [20] S. Das and D. Das, "Natural Language Processing (NLP) Techniques: Usability in Human-Computer Interactions," in *2024 6th International Conference on Natural Language Processing (ICNLP)*, vol. 9, Mar 2024, pp. 783–787.
- [21] L. Yang, "A Brief Introduction of the Text Classification Methods," <https://ieeexplore.ieee.org/document/9744845>, Feb 2022, accessed: Nov. 12, 2022.
- [22] S. Jayabharathi and V. Ilango, "A Brief Revolution of Evolution and Resurgence on Machine Learning," in *2021 Asian Conference on Innovation in Technology (ASI-ANCON)*, Aug 2021.
- [23] X. Gong, Q. Cheng, X. Wu, J. Wang, and Z. Feng, "Application of artificial intelligence technology in the electronic information age," in *2022 World Automation Congress (WAC)*, Oct 2022, pp. 128–132.
- [24] R. Farhat, Y. Mourali, M. Jemni, and H. Ezzedine, "An overview of Machine Learning Technologies and their use in E-learning," in *2020 International Multi-Conference on: "Organization of Knowledge and Advanced Technologies" (OCTA)*, Feb 2020.

- [25] Y. Wu and X. Zhang, "An Empirical Investigation into Complex Fake News Detection Using the Multinomial Naïve Bayes Framework," in *2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE)*, May 2024, pp. 943–946.
- [26] E. Hairani and I. Nurpalah, "Sentiment Analysis of Muslim Pro 'A1-Quran Adzan' on Google Play Using Multinomial Naïve Bayes," in *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, Oct 2024, pp. 1–5.
- [27] G. A. Muthulingam, K. Ponnuchamy, G. P. Kumar, J. Mahesh, G. N. H. B. Karthik, and J. V. Krishna, "A Comparative Analysis Between Support Vector Machine and Linear Regression for Improving Accuracy of Medal Tally in Asian Games," in *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, Jan 2025, pp. 1614–1618.
- [28] I. M. Fadhil and Y. Sibaroni, "Topic classification in indonesian-language tweets using fast-text feature expansion with support vector machine (svm)," in *2022 International Conference on Data Science and Its Applications (ICoDSA)*, Jul. 2022.
- [29] B. Pradito, S. D. Budiwati, E. Hernawati, and S. K. Sari, "Classification for human resource talent management using support vector machine model," in *2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*, Jul. 2024, pp. 282–287.
- [30] G. Naresh, J. Sreeja, G. Ramani, and G. Vishnu Teja, "Comparative study of classification algorithms on contrived news," in *2023 International Conference on Advanced Computing Communication Technologies (ICACCTech)*, 2023, pp. 478–483.
- [31] A. N, J. P, P. S, K. S, and S. S, "Online news article classification using machine learning approaches," in *2024 2nd International Conference on Disruptive Technologies (ICDT)*, 2024, pp. 1494–1496.
- [32] L. Shkurti and F. Kabashi, "Albanian news category predictor system using a multinomial naïve bayes and logistic regression algorithms," in *2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 2021, pp. 642–647.
- [33] T. Manvitha and K. S. Rekha, "Improved accuracy for prediction of leaf wetness using logistic regression algorithm compared with decision tree algorithm," in *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, 2023, pp. 1–5.
- [34] M. R. Camana Acosta, S. Ahmed, C. E. Garcia, and I. Koo, "Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks," *IEEE Access*, vol. 8, pp. 19 921–19 933, 2020.
- [35] S. Duman, E. Arik, S. Turnacigil, and M. A. Aktaş, "Estimating gold prices in financial markets using machine learning: A comparative analysis of random forest, k-nearest neighbor, and decision tree algorithms," in *2023 Innovations in Intelligent Systems and Applications Conference (ASYU)*, 2023, pp. 1–3.

- [36] K. Liang and L. Zhang, "Application of decision tree random forest algorithm in the evolution of dual innovation in smes," in *2025 IEEE International Conference on Electronics, Energy Systems and Power Engineering (EESPE)*, 2025, pp. 895–899.
- [37] L. Wei, "Genetic algorithm optimization of concrete frame structure based on improved random forest," in *2023 International Conference on Electronics and Devices, Computational Science (ICEDCS)*, 2023, pp. 249–253.
- [38] E. Zorarpaci, "A comparative study of ensemble learning methods for turkish text categorization," in *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2024, pp. 1–6.
- [39] Y. Sun, Y. Li, Q. Zeng, and Y. Bian, "Application research of text classification based on random forest algorithm," in *2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, 2020, pp. 370–374.
- [40] M. Y. Khan, A. Qayoom, M. S. Nizami, M. S. Siddiqui, S. Wasi, and S. M. K.-R. Raza, "Automated prediction of good dictionary examples (gdex): A comprehensive experiment with distant supervision, machine learning, and word embedding-based deep learning techniques," *Complexity*, vol. 2021, pp. 1–18, Sep 2021.
- [41] T. Yang and S. Du, "An improved weighted knn algorithm about text classification based on spark framework," in *2022 IEEE 10th International Conference on Information, Communication and Networks (ICICN)*, 2022, pp. 655–661.
- [42] N. A. Kalcheva and M. N. Karova, "A comparison of machine learning classification algorithms and methods for english author's works and their translations into bulgarian," in *2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, 2022, pp. 1–4.
- [43] S. Xiong and Q. Gong, "Knn text classification algorithm based on chaotic particle swarm optimization," in *2023 3rd International Conference on Intelligent Communications and Computing (ICC)*, 2023, pp. 350–353.
- [44] I. Alexander, Y. Heryadi, Lukas, W. Suparta, and I. Sonata, "Classifying tweet sentiment analysis on airlines using random forest classifier, naïve bayes classifier and k-nearest neighbors classifier algorithm," in *2023 International Conference on Artificial Intelligence Robotics, Signal and Image Processing (AIROSiP)*, 2023, pp. 342–347.
- [45] M. M. Hijazi and Q. Althebyan, "A review study on utilizing deep learning in arabic text classification," in *2024 25th International Arab Conference on Information Technology (ACIT)*, Dec. 2024, pp. 1–3.
- [46] C. M. Suneera and J. Prakash, "Performance analysis of machine learning and deep learning models for text classification," in *2020 IEEE 17th India Council International Conference (INDICON)*, Dec. 2020.
- [47] L. Deping, W. Hongjuan, L. Mengyang, and L. Pei, "News text classification based on bidirectional encoder representation from transformers," *IEEE Xplore*, May 2021, accessed: May 03, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9545942>

- [48] G. Wu, “A transformer-based diagnostic prediction model for bipolar disorder,” in *2024 5th International Conference on Machine Learning and Computer Application (ICMLCA)*, Oct. 2024, pp. 419–424.
- [49] J.-W. Sun, J.-Q. Bao, and L.-P. Bu, “Text classification algorithm based on tf-idf and bert,” in *2022 11th International Conference of Information and Communication Technology (ICTech)*, Feb. 2022.
- [50] B. Zhu and W. Pan, “A text classification model based on bert and attention,” in *2023 4th International Conference on Computers and Artificial Intelligence Technology (CAIT)*, Dec. 2023, pp. 90–95.
- [51] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, “Indolem and indobert: A benchmark dataset and pre-trained language model for indonesian nlp,” *arXiv preprint arXiv:2011.00677*, Nov. 2020.
- [52] I. N. Khasanah and A. A. Krisnadhi, “Extreme multilabel text classification on indonesian tax court ruling using single channel cnn and indobert embedding,” in *2021 6th International Workshop on Big Data and Information Security (IWBIS)*, Oct. 2021, pp. 59–66.
- [53] A. Putra and Y. Sibaroni, “Disinformation detection on 2024 indonesia presidential election using indobert,” in *2023 International Conference on Data Science and Its Applications (ICoDSA)*, Aug. 2023.
- [54] H. N. Soe, L. M. Htut, H. Htun, N. Lynn, K. H. Saw, and K. K. Min, “Myanmar sports sub-news categorization using chi-square, bag of words, and cosine similarity combining with smooth tf-idf,” in *2023 Seminar on Information Computing and Processing (ICP)*, Nov. 2023, pp. 80–85.
- [55] S. Jain, S. K. Jain, and S. Vasal, “An effective tf-idf model to improve the text classification performance,” in *2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT)*, vol. 10, Apr. 2024, pp. 1–4.
- [56] D. Chicco, M. J. Warrens, and G. Jurman, “The matthews correlation coefficient (mcc) is more informative than cohen’s kappa and brier score in binary classification assessment,” *IEEE Access*, vol. 9, pp. 78 368–78 381, 2021.
- [57] W. Arshad, M. Ali, M. M. Ali, A. Javed, and S. Hussain, “Multi-class text classification: Model comparison and selection,” <https://ieeexplore.ieee.org/abstract/document/9514108>, Jun. 2021, accessed: Apr. 07, 2023.
- [58] A. Pramudita, R. Wijaya, S. Cokro, G. Z. Nabiilah, and N. Rojali, “A comparison of text classification methods: Naïve bayes and support vector machine for e-commerce item classification,” in *2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, Aug. 2023, pp. 413–418.