

DAFTAR PUSTAKA

- [1] K. Wang, X. Liu, and Y. Han, "Exploring Goodreads reviews for book impact assessment," *J. Informetr.*, vol. 13, no. 3, pp. 874–886, Aug. 2019, doi: 10.1016/j.joi.2019.07.003.
- [2] T. Chen, P. Samaranayake, X. Cen, M. Qi, and Y.-C. Lan, "The Impact of Online Reviews on Consumers' Purchasing Decisions: Evidence From an Eye-Tracking Study," *Front. Psychol.*, vol. 13, Jun. 2022, doi: 10.3389/fpsyg.2022.865702.
- [3] H. Fürst and M. Kurfalı, "Behind the book reviews: Enduring hierarchies of occurrence and unpredictable sentiments," *Poetics*, vol. 113, p. 102063, Dec. 2025, doi: 10.1016/j.poetic.2025.102063.
- [4] C. Sich, "A Comparison of Traditional Book Reviews and Amazon.com Book Reviews of Fiction Using a Content Analysis Approach," *Evid. Based Libr. Inf. Pract.*, vol. 12, no. 1, pp. 85–96, Mar. 2017, doi: 10.18438/B8CW4N.
- [5] Y. Mao, Q. Liu, and Y. Zhang, "Sentiment analysis methods, applications, and challenges: A systematic literature review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 4, p. 102048, Apr. 2024, doi: 10.1016/j.jksuci.2024.102048.
- [6] B. Saberi and S. Saad, "Sentiment Analysis or Opinion Mining: A Review," vol. 7, no. 5, 2017.
- [7] M. Wan and J. McAuley, "Item recommendation on monotonic behavior chains," in *Proceedings of the 12th ACM Conference on Recommender Systems*, New York, NY, USA: ACM, Sep. 2018, pp. 86–94. doi: 10.1145/3240323.3240369.
- [8] A. Almjawel, S. Bayoumi, D. Alshehri, S. Alzahrani, and M. Alotaibi, "Sentiment Analysis and Visualization of Amazon Books' Reviews," in *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, IEEE, May 2019, pp. 1–6. doi: 10.1109/CAIS.2019.8769589.
- [9] D. Ikasari, "SENTIMENT ANALYSIS REVIEW NOVEL 'GOODREADS' BERBAHASA INDONESIA MENGGUNAKAN NAÏVE BAYES CLASSIFIER," in *Seminar Nasional Riset dan Inovasi Teknologi (SEMNAS RISTEK) 2021*, Jakarta, Jan. 2021, pp. 760–765.
- [10] R. A. Saputra and T. Fatimah, "Penerapan Algoritma Naive Bayes untuk Klasifikasi Buku Populer Berbasis Web," *6th Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, vol. 4, no. 2, pp. 449–458, Sep. 2025.
- [11] P. K. A. Duta, P. D. S. Pitnatri, and C. P. Loanata, "Big Data-based Sentiment Analysis on TripAdvisor Reviews Using Naïve Bayes Classification," *Media Wisata*, vol. 23, no. 1, pp. 92–107, May 2025, doi: 10.36276/mws.v23i1.864.
- [12] T. Kavzoglu and A. Teke, "Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost)," *Arab. J. Sci. Eng.*, vol. 47, no. 6, pp. 7367–7385, Jun. 2022, doi: 10.1007/s13369-022-06560-8.

- [13] T. Chen and C. Guestrin, "XGBoost," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [14] Y. Zhang, J. Liu, and W. Shen, "A Review of Ensemble Learning Algorithms Used in Remote Sensing Applications," *Applied Sciences*, vol. 12, no. 17, p. 8654, Aug. 2022, doi: 10.3390/app12178654.
- [15] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992, doi: 10.1016/S0893-6080(05)80023-1.
- [16] A. Gaius, R. W. Mwangi, and A. Ngunyi, "A Stacking-Based Ensemble Approach with Embeddings from Language Models for Depression Detection from Social Media Text," *Journal of Data Analysis and Information Processing*, vol. 11, no. 04, pp. 420–453, 2023, doi: 10.4236/jdaip.2023.114022.
- [17] A. McCallum and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," 1998.
- [18] C. D. Manning, Prabhakar. Raghavan, and Hinrich. Schütze, *Introduction to information retrieval*. Cambridge University Press, 2009.
- [19] M. Veziroğlu, E. Veziroğlu, and İ. Ö. Bucak, "Performance Comparison between Naive Bayes and Machine Learning Algorithms for News Classification," in *Bayesian Inference - Recent Trends*, IntechOpen, 2024. doi: 10.5772/intechopen.1002778.
- [20] V. Chaurasia and S. Pal, "Stacking-Based Ensemble Framework and Feature Selection Technique for the Detection of Breast Cancer," *SN Comput. Sci.*, vol. 2, no. 2, p. 67, Apr. 2021, doi: 10.1007/s42979-021-00465-3.
- [21] M. Chen *et al.*, "Automatic text classification of drug-induced liver injury using document-term matrix and XGBoost," *Front. Artif. Intell.*, vol. 7, Jun. 2024, doi: 10.3389/frai.2024.1401810.
- [22] G. Rohini, T. Mariprath, S. Al Bustanji, and I. Zaitsev, "A stacked Gradient Boosting–XGBoost ensemble with ridge meta-learner for accurate short-term solar PV power forecasting in smart grids," *Sci. Rep.*, vol. 16, no. 1, p. 16871, Apr. 2026, doi: 10.1038/s41598-026-47042-3.
- [23] T. G. Dietterich, "Ensemble Methods in Machine Learning," 2000, pp. 1–15. doi: 10.1007/3-540-45014-9_1.
- [24] K. Nigam, J. Lafferty, and A. McCallum, "Using Maximum Entropy for Text Classification."
- [25] F. E. Kavabilla, T. Widiari, and B. Warsito, "ANALISIS SENTIMEN PADA ULASAN APLIKASI INVESTASI ONLINE AJAIB PADA GOOGLE PLAY MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN MAXIMUM ENTROPY," *Jurnal Gaussian*, vol. 11, no. 4, pp. 542–553, Feb. 2023, doi: 10.14710/j.gauss.11.4.542-553.
- [26] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992, doi: 10.1016/S0893-6080(05)80023-1.
- [27] K. C. K. and P. B. D. Karuna C Khobragade and Pankaj B Dhumane, "REVIEW OF DATA PRE-PROCESSING TECHNIQUES IN DATA MINING," *International Journal of Researches in Biosciences and Agriculture Technology*, 2022, doi: 10.29369/ijrbat.2022.010.3.0016.

- [28] K. C. K. and P. B. D. Karuna C Khobragade and Pankaj B Dhumane, "REVIEW OF DATA PRE-PROCESSING TECHNIQUES IN DATA MINING," *International Journal of Researches in Biosciences and Agriculture Technology*, 2022, doi: 10.29369/ijrbat.2022.010.3.0016.
- [29] Salvador. García, J. Luengo, and Francisco. Herrera, *Data Preprocessing in Data Mining*. Springer International Publishing : Imprint : Springer, 2015.
- [30] M. IŞIK and H. DAĞ, "The impact of text preprocessing on the prediction of review ratings," *TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES*, vol. 28, no. 3, pp. 1405–1421, May 2020, doi: 10.3906/elk-1907-46.
- [31] M. Anandarajan, C. Hill, and T. Nolan, "Text Preprocessing," 2019, pp. 45–59. doi: 10.1007/978-3-319-95663-3_4.
- [32] L. Hickman, S. Thapa, L. Tay, M. Cao, and P. Srinivasan, "Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations," *Organ. Res. Methods*, vol. 25, no. 1, pp. 114–146, Jan. 2022, doi: 10.1177/1094428120971683.
- [33] S. E. . Robertson and Hugo. Zaragoza, *The probabilistic relevance framework : BM25 and beyond*. Now Publishers, 2009.
- [34] C. D. . Manning, Prabhakar. Raghavan, and Hinrich. Schütze, *Introduction to information retrieval*. Cambridge University Press, 2009.
- [35] M. Mendoza, "A new term-weighting scheme for naïve Bayes text categorization," *International Journal of Web Information Systems*, vol. 8, no. 1, pp. 55–72, Mar. 2012, doi: 10.1108/17440081211222591.
- [36] Z. Zhang, "An improved BM25 algorithm for clinical decision support in Precision Medicine based on co-word analysis and Cuckoo Search," *BMC Med. Inform. Decis. Mak.*, vol. 21, no. 1, p. 81, Dec. 2021, doi: 10.1186/s12911-021-01454-5.
- [37] P. Lepagnol, S. Ghannay, T. Gerald, C. Servan, and S. Rosset, "Leveraging Information Retrieval to Enhance Spoken Language Understanding Prompts in Few-Shot Learning," Jun. 2025, [Online]. Available: <http://arxiv.org/abs/2506.03035>
- [38] A. I. Kadhim, "Term Weighting for Feature Extraction on Twitter: A Comparison Between BM25 and TF-IDF," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, IEEE, Apr. 2019, pp. 124–128. doi: 10.1109/ICOASE.2019.8723825.
- [39] S. Kanwal, S. Nawaz, M. K. Malik, and Z. Nawaz, "A Review of Text-Based Recommendation Systems," *IEEE Access*, vol. 9, pp. 31638–31661, 2021, doi: 10.1109/ACCESS.2021.3059312.
- [40] M. Isnani, G. N. Elwirehardja, and B. Pardamean, "Sentiment Analysis for TikTok Review Using VADER Sentiment and SVM Model," *Procedia Comput. Sci.*, vol. 227, pp. 168–175, 2023, doi: 10.1016/j.procs.2023.10.514.
- [41] D. C. Youvan, "Understanding Sentiment Analysis with VADER: A Comprehensive Overview and Application," 2024.
- [42] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, May 2014, doi: 10.1609/icwsm.v8i1.14550.

- [43] S. M. Mohammad, "A Practical Guide to Sentiment Annotation: Challenges and Solutions," in *Proceedings of NAACL-HLT*, 2016 Association for Computational Linguistics, 2016, pp. 174–179.
- [44] J. Guerreiro and P. Rita, "How to predict explicit recommendations in online reviews using text mining and sentiment analysis," *Journal of Hospitality and Tourism Management*, vol. 43, pp. 269–272, Jun. 2020, doi: 10.1016/j.jhtm.2019.07.001.
- [45] Haibo He and E. A. Garcia, "Learning from Imbalanced Data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009, doi: 10.1109/TKDE.2008.239.
- [46] C. Elkan, "The Foundations of Cost-Sensitive Learning," in *International Joint Conference on Artificial Intelligence (IJCAI'01)*, 2001.
- [47] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, Oct. 2018, doi: 10.1016/j.neunet.2018.07.011.
- [48] D. J. Hand and K. Yu, "Idiot's Bayes: Not So Stupid after All?," *Int. Stat. Rev.*, vol. 69, no. 3, p. 385, Dec. 2001, doi: 10.2307/1403452.
- [49] Mitchell Tom M., "Ensemble Methods Foundations and Algorithms," *AI Magazine*, vol. 18, pp. 11–20, 1997. [Online]. Available: <https://www.routledge.com/Chapman->
- [50] T. Hastie, R. Tibshirani, and J. Friedman, *Springer Series in Statistics The Elements of Statistical Learning Data Mining, Inference, and Prediction*. 2009.
- [51] Dan. Jurafsky and J. H. . Martin, *Speech and language processing*. Pearson Education, 2014.
- [52] Z.-H. Zhou, *Ensemble Methods*. Boca Raton: Chapman and Hall/CRC, 2025. doi: 10.1201/9781003587774.
- [53] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sens. Environ.*, vol. 62, no. 1, pp. 77–89, Oct. 1997, doi: 10.1016/S0034-4257(97)00083-7.
- [54] D. M. W. Powers, "Evaluation Evaluation a Monte Carlo study," Apr. 2015.
- [55] J. Opitz, "A Closer Look at Classification Evaluation Metrics and a Critical Reflection of Common Evaluation Practice," *Trans. Assoc. Comput. Linguist.*, vol. 12, pp. 820–836, Jun. 2024, doi: 10.1162/tacl_a_00675.
- [56] B. Sousa, F. Durão, M. Duarte, and R. Alves, "Goodreads Books and Reviews."
- [57] B. Sousa, "Encouraging Connections: A Genre Analysis of Goodreads Comments Amal Foufa Writer's Statement," *Writing for University and Beyond: A Journal of First Year Writing at UTM*, vol. 3, pp. 105–111, 20213.
- [58] S. M. Mudambi and D. Schuff, "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.Com1," *MIS Quarterly*, vol. 34, no. 1, pp. 185–200, Mar. 2010, doi: 10.2307/20721420.
- [59] G. Rao, W. Huang, Z. Feng, and Q. Cong, "LSTM with sentence representations for document-level sentiment classification," *Neurocomputing*, vol. 308, pp. 49–57, Sep. 2018, doi: 10.1016/j.neucom.2018.04.045.