

DAFTAR PUSTAKA

- [1] P. Shah and A. Patel, "A comprehensive study and detailed review on hate speech classification: A systematic analysis," *2021 International Conference on Advances in Computing, Communication, and Control (ICAC3)*, pp. 1–6, 2021.
- [2] N. S. Mullah and W. Zainon, "Advances in machine learning algorithms for hate speech detection in social media: A review," *IEEE Access*, vol. 9, pp. 88 364–88 376, 2021.
- [3] G. I. Wijaya, "Pengembangan modul u-tapis berbasis naive bayes classifier untuk deteksi hujatan sebagai aspek spesifik harassment dalam klasifikasi teks," Universitas Multimedia Nusantara Repository, 2024. [Online]. Available: <https://kc.umn.ac.id/id/eprint/36022/>
- [4] P. Tendy, "U-Tapis: Deteksi dan Koreksi Kesalahan Penulisan Peluluhan Kata Menggunakan Algoritma Damerau-Levenshtein Distance," MBKM thesis, Universitas Multimedia Nusantara, jan 2025, mBKM thesis. [Online]. Available: <https://kc.umn.ac.id/id/eprint/36026/>
- [5] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," in *European Conference on Machine Learning*. Springer, 1998, pp. 137–142.
- [6] Z. Waseem and D. Hovy, "Hateful symbols or hateful people? predictive features for hate speech detection on twitter," *Proceedings of NAACL-HLT*, 2016.
- [7] P. Fortuna and S. Nunes, "A survey on automatic detection of hate speech in text," *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, pp. 1–30, 2018.
- [8] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [9] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18–28, 1998.
- [10] M. Lindsay and J. Krysik, "Online harassment among college students," *Information, Communication Society*, vol. 15, pp. 703–719, 2012.
- [11] S. Kim, A. Razi, A. Alsoubai, P. J. Wisniewski, and M. De Choudhury, "Assessing the impact of online harassment on youth mental health in private networked spaces," *ICWSM*, pp. 826–838, 2024.
- [12] L. Blackwell, T. Chen, S. Schoenebeck, and C. Lampe, "When online harassment is perceived as justified," *ICWSM*, pp. 22–31, 2018.

- [13] A. Gandhi, P. Ahir, K. Adhvaryu, P. Shah, R. Lohiya, E. Cambria, S. Poria, and A. Hussain, "Hate speech detection: A comprehensive review of recent works," *Expert Systems*, vol. 41, 2024.
- [14] A. Miran and A. Abdulazeez, "Detection of hate-speech tweets based on deep learning: A review," *JISA (Jurnal Informatika dan Sains)*, 2023.
- [15] K. Kowsari, K. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. E. Brown, "Text classification algorithms: A survey," *Information*, vol. 10, no. 4, p. 150, 2019.
- [16] R. Snow, B. O'Connor, D. Jurafsky, and A. Y. Ng, "Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks," in *EMNLP*, 2008, pp. 254–263.
- [17] M. Sabou, K. Bontcheva, and A. Scharl, "Corpus annotation through crowdsourcing: Towards best practice guidelines," *LREC*, 2014.
- [18] R. V. Krejcie and D. W. Morgan, "Determining sample size for research activities," *Educational and psychological measurement*, vol. 30, no. 3, pp. 607–610, 1970.
- [19] X. Zhang, "What is the kernel trick? why is it important?" <https://medium.com/@zxr.nju/what-is-the-kernel-trick-why-is-it-important-98a98db0961d>, 2020, accessed: 23 Juli 2025.
- [20] J. D. M. Rennie, L. Shih, J. Teevan, and D. R. Karger, "Tackling the poor assumptions of naive bayes text classifiers," in *Proceedings of the 20th International Conference on Machine Learning (ICML)*. Washington, DC, USA: AAAI Press, 2003, pp. 616–623.
- [21] GeeksforGeeks. (2025, Feb.) Understanding tf-idf (term frequency–inverse document frequency). Last updated on February 7, 2025; accessed July 15, 2025. [Online]. Available: <https://www.geeksforgeeks.org/machine-learning/understanding-tf-idf-term-frequency-inverse-document-frequency/>
- [22] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *Proceedings of Workshop at ICLR*, 2013. [Online]. Available: <https://arxiv.org/abs/1301.3781>
- [23] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008. [Online]. Available: <https://nlp.stanford.edu/IR-book/>
- [24] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," in *Transactions of the Association for Computational Linguistics (TACL)*, vol. 5, 2017, pp. 135–146.

- [25] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.
- [26] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.
- [27] C. J. van Rijsbergen, *Information Retrieval*, 2nd ed. London: Butterworth-Heinemann, 1979.
- [28] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2019, pp. 2623–2631.
- [29] X. Zhang, J. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," in *Advances in neural information processing systems*, 2015, pp. 649–657.
- [30] M. O. Ibrohim and I. Budi, "Multi-label hate speech and abusive language detection in Indonesian twitter," in *Proceedings of the Third Workshop on Abusive Language Online*. Florence, Italy: Association for Computational Linguistics, Aug. 2019, pp. 46–57. [Online]. Available: <https://www.aclweb.org/anthology/W19-3506>
- [31] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," in *Journal of Machine Learning Research*, vol. 13, no. Feb, 2012, pp. 281–305.

