

Referensi

- [1] D. Setiawan, "Dampak Perkembangan Teknologi Informasi dan Komunikasi Terhadap Budaya," *JURNAL SIMBOLIKA: Research and Learning in Communication Study*, vol. 4, no. 1, p. 62, 2018, doi: 10.31289/simbollika.v4i1.1474.
- [2] E. B. Setiawan, D. H. Widyantoro, and K. Surendro, "Feature Expansion using Word Embedding for Tweet Topic Classification," *Proceeding of 2016 10th International Conference on Telecommunication Systems Services and Applications, TSSA 2016: Special Issue in Radar Technology*, 2017, doi: 10.1109/TSSA.2016.7871085.
- [3] M.-L. Kamberg, "Evan Williams, Biz Stone, Jack Dorsey, and Twitter," *The Rosen Publishing Group, Inc.*, 2012.
- [4] A. B. Boot, E. Tjong Kim Sang, K. Dijkstra, and R. A. Zwaan, "How character limit affects language usage in tweets," *Palgrave Communications*, vol. 5, no. 1, pp. 1–13, 2019, doi: 10.1057/s41599-019-0280-3.
- [5] M. A. Fauzi, R. F. N. Firmansyah, and T. Afrianto, "Improving sentiment analysis of short informal Indonesian product reviews using synonym based feature expansion," 2018, doi: 10.12928/TELKOMNIKA.v16i3.7751.
- [6] T. Kulesza, S. Amershi, R. Caruana, D. Fisher, and D. Charles, "Structured labeling to facilitate concept evolution in machine learning," in *Conference on Human Factors in Computing Systems - Proceedings*, 2014, pp. 3075–3084. doi: 10.1145/2556288.2557238.
- [7] J. H. Friedman, "999 REITZ LECTURE GREEDY FUNCTION APPROXIMATION: A GRADIENT BOOSTING MACHINE 1," 2001.
- [8] Raewyn. Peart, Kate. Mulcahy, Natasha. Garvan, and Environmental Defence Society (N.Z.), *Managing freshwater : an EDS guide*. Environmental Defence Society, 2010.
- [9] Z. Wen, B. He, K. Ramamohanarao, S. Lu, and J. Shi, "Efficient Gradient Boosted Decision Tree Training on GPUs." [Online]. Available: <https://github.com/zeyiwen/gbdt>
- [10] M. Chen, X. Jin, and D. Shen, "Short Text Classification Improved by Learning Multi-Granularity Topics." [Online]. Available: <http://www.wikipedia.org>
- [11] J. Eka Sembodo, E. Budi Setiawan, and Z. Abdurahman Baizal, "Data Crawling Otomatis pada Twitter," Sep. 2016, pp. 11–16. doi: 10.21108/indosc.2016.111.
- [12] Y. Sun, E. Lank, and M. Terry, "Label-And-learn: Visualizing the likelihood of machine learning classifier's success during data labeling," in *International Conference on Intelligent User Interfaces, Proceedings IUI*, Mar. 2017, pp. 523–534. doi: 10.1145/3025171.3025208.
- [13] G. Orellana, B. Arias, M. Orellana, V. Saquicela, F. Baculima, and N. Piedra, "A study on the impact of pre-processing techniques in Spanish and english text classification over short and large text documents," in *Proceedings - 3rd International Conference on Information Systems and Computer Science, INCISCOS 2018*, Dec. 2018, vol. 2018-December, pp. 277–283. doi: 10.1109/INCISCOS.2018.00047.
- [14] S. Gharatkar, A. Ingle, T. Naik, and A. Save, "Review preprocessing using data cleaning and stemming technique," *Proc. 2017 Int. Conf. Innov. Information, Embed. Commun. Syst. ICIIECS 2017*, vol. 2018-January, pp. 1–4, 2018, doi: 10.1109/ICIIECS.2017.8276011.
- [15] D. Gunawan, "khazanah informatika Jurnal Ilmu Komputer dan Informatika Evaluasi Performa Pemecahan Database dengan Metode Klasifikasi pada Data Preprocessing Data Mining," 2016.
- [16] J. Chen, C. Chen, and Y. Liang, "Optimized TF-IDF Algorithm with the Adaptive Weight of Position of Word," 2016.
- [17] E. B. Setiawan, D. H. Widyantoro, and K. Surendro, "Feature Expansion using Word Embedding for Tweet Topic Classification." [Online]. Available: <http://blog.twitter.com/2012/03/twitter-turns-six.html>
- [18] J. Lilleberg, Y. Zhu, and Y. Zhang, "Support vector machines and Word2vec for text classification with semantic features," *Proc. 2015 IEEE 14th Int. Conf. Cogn. Informatics Cogn. Comput. ICCI*CC 2015*, pp. 136–140, 2015, doi: 10.1109/ICCI-CC.2015.7259377.
- [19] M. Rapp, "BOOMER — An algorithm for learning gradient boosted multi-label classification rules," *Software Impacts*, vol. 10, p. 100137, Nov. 2021, doi: 10.1016/j.simpa.2021.100137.
- [20] A. K. Santra and C. J. Christy, "Genetic Algorithm and Confusion Matrix for Document Clustering," *Int. J. Comput. Sci. Issues*, vol. 9, no. 1, pp. 322–328, 2012.
- [21] C. Kim, V. Zhu, J. Obeid, and L. Lenert, "Natural language processing and machine learning algorithm to identify brain MRI reports with acute ischemic stroke," *PLoS ONE*, vol. 14, no. 2, Feb. 2019, doi: 10.1371/journal.pone.0212778.
- [22] Scikit-learn: Machine Learning in Python, Pedregosa et al., *JMLR* 12, pp. 2825–2830, 2011.
- [23] S. Bird and E. Loper, "NLTK: The Natural Language Toolkit," *Proc. 42nd Annu. Meet. Assoc. Comput. Linguist.*, pp. 1–4, 2004.
- [24] J. Asian, H. E. Williams, and S. M. M. Tahaghoghi, "Stemming Indonesian," *Conf. Res. Pract. Inf. Technol. Ser.*, vol. 38, no. 4, pp. 307–314, 2005, doi: 10.1145/1316457.1316459.
- [25] Gensim Python library, <https://radimrehurek.com/gensim/>, last accessed 2022/02/15