

Topic Classification Using the Long Short-Term Memory (LSTM) Method with FastText Feature Expansion on Twitter

Bella Adriani Putri
School of Computing
Telkom University
Bandung, Indonesia

bellaadrp@student.telkomuniversity.ac.id

Erwin Budi Setiawan
School of Computing
Telkom University
Bandung, Indonesia

erwinbudisetiawan@telkomuniversity.ac.id

REFERENCES

- [1] M. Marni, M. Paolanti, R. Pietrini, G. Pazzaglia, E. Frontoni, and P. Zingaretti, "Deep Learning Approaches for Fashion Knowledge Extraction from Social Media: A Review," *2022 10th International Conference on Information and Communication Technology (ICoICT)*, vol. 10. Institute of Electrical and Electronics Engineers Inc., pp. 1545–1576, 2022. doi: 10.1109/ACCESS.2021.3137893.
- [2] D. Antonakaki, P. Fragopoulou, and S. Ioannidis, "A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks," *Expert Syst Appl*, vol. 164, Feb. 2021, doi: 10.1016/j.eswa.2020.114006.
- [3] Naufal Adi Nugroho and Erwin Budi Setiawan, "Implementation Word2Vec for Feature Expansion in Twitter Sentiment Analysis," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 5, pp. 837–842, Oct. 2021, doi: 10.29207/resti.v5i5.3325.
- [4] E. B. Setiawan, D. H. Widyantoro, and K. Surendro, "Feature expansion using word embedding for tweet topic classification," in *Proceeding of 2016 10th International Conference on Telecommunication Systems Services and Applications, TSSA 2016: Special Issue in Radar Technology*, Institute of Electrical and Electronics Engineers Inc., Mar. 2017. doi: 10.1109/TSSA.2016.7871085.
- [5] S. Mukherjee, A. Malu, A. R. Balamurali, and P. Bhattacharyya, "TwiSent: A multistage system for analyzing sentiment in Twitter," in *ACM International Conference Proceeding Series*, 2012, pp. 2531–2534. doi: 10.1145/2396761.2398684.
- [6] M. A. Zingla, L. Chiraz, Y. Slimani, and C. Berrut, "Statistical and Semantic Approaches for Tweet Contextualization," *Procedia Comput Sci*, vol. 60, no. 1, pp. 498–507, Jan. 2015, doi: 10.1016/J.PROCS.2015.08.171.
- [7] M. A. Fauzi, R. F. N. Firmansyah, and T. Afrianto, "Improving sentiment analysis of short informal Indonesian product reviews using synonym based feature expansion," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 16, no. 3, pp. 1345–1350, Jun. 2018, doi: 10.12928/TELKOMNIKA.v16i3.7751.
- [8] R. A. Yahya and E. B. Setiawan, "Feature Expansion with FastText on Topic Classification Using the Gradient Boosted Decision Tree on Twitter," *2022 10th International Conference on Information and Communication Technology (ICoICT)*, 2022, doi: 10.1109/ICoICT55009.2022.9914896.
- [9] W. K. Sari, D. P. Rini, and R. F. Malik, "Text Classification Using Long Short-Term Memory With GloVe Features," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 5, no. 2, p. 85, Feb. 2020, doi: 10.26555/jiteki.v5i2.15021.
- [10] Nur Ghaniyanto Ramadhan, "Indonesian Online News Topics Classification using Word2Vec and K-Nearest Neighbor," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 6, pp. 1083–1089, Dec. 2021, doi: 10.29207/resti.v5i6.3547.
- [11] L. Jiang, C. Li, S. Wang, and L. Zhang, "Deep feature weighting for naive Bayes and its application to text classification," *Eng Appl Artif Intell*, vol. 52, pp. 26–39, Jun. 2016, doi: 10.1016/j.engappai.2016.02.002.
- [12] M. Goudjil, M. Koudil, M. Bedda, and N. Ghogali, "A Novel Active Learning Method Using SVM for Text Classification," *International Journal of Automation and Computing*, vol. 15, no. 3, pp. 290–298, Jun. 2018, doi: 10.1007/s11633-015-0912-z.
- [13] D. T. Maulidia E. Budi Setiawan, "Feature Expansion with Word2Vec for Topic Classification with Gradient Boosted Decision Tree on Twitter," in *2022 International Conference on Data Science and Its Applications (ICoDSA)*, 2022.
- [14] M. A. Riza and N. Charibaldi, "Emotion Detection in Twitter Social Media Using Long Short-Term Memory (LSTM) and Fast Text," *International Journal of Artificial Intelligence & Robotics (IJAIR)*, vol. 3, no. 1, pp. 15–26, May 2021, doi: 10.25139/ijair.v3i1.3827.
- [15] J. Eka Sembodo, E. Budi Setiawan, and Z. Abdurahman Baizal, "Data Crawling Otomatis pada Twitter," School of Computing, Telkom University, Sep. 2016, pp. 11–16. doi: 10.21108/indosc.2016.111.
- [16] 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*, Institute of Electrical and Electronics Engineers Inc., Dec. 2018, pp. 277–283. doi: 10.1109/INCISCOS.2018.00047.
- [17] Febiana Anistya and Erwin Budi Setiawan, "Hate Speech Detection on Twitter in Indonesia with Feature Expansion Using GloVe," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 6, pp. 1044–1051, Dec. 2021, doi: 10.29207/resti.v5i6.3521.
- [18] A. Alessa, M. Faezipour, and Z. Alhassan, "Text classification of flu-related tweets using FastText with sentiment and keyword features," in *Proceedings - 2018 IEEE International Conference on Healthcare Informatics, ICHI 2018*, Institute of Electrical and Electronics Engineers Inc., Jul. 2018, pp. 366–367. doi: 10.1109/ICHI.2018.00058.
- [19] A. Kumar and R. Rastogi nee Khemchandani, "Attentional recurrent neural networks for sentence classification," in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2019, pp. 549–559. doi: 10.1007/978-981-13-1966-2_49.