REFERENCES

[1] A. S. Raamkumar, M. Erdt, H. Vijayakumar, E. Rasmussen, and Y.-L. Theng, "Understanding the Twitter usage of humanities and social sciences academic journals," *Proceedings of the Association for Information Science and Technology*, vol. 55, no. 1, pp. 430–439, 2018, https://doi.org/10.1002/pra2.2018.14505501047.

[2] A. Kumar and A. Jaiswal, "Systematic literature review of sentiment analysis on Twitter using soft computing techniques," *Concurr Comput*, vol. 32, p. e5107, 2019, https://doi.org/10.1002/cpe.5107.

[3] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artif Intell Rev*, vol. 53, no. 6, pp. 4335–4385, 2020, https://doi.org/10.1007/s10462-019-09794-5.

[4] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowl Inf Syst*, vol. 60, no. 2, pp. 617–663, 2019, https://doi.org/10.1007/s10115-018-1236-4.

[5] R. A. Yahya and E. B. Setiawan, "Feature Expansion with FastText on Topic Classification Using the Gradient Boosted Decision Tree on Twitter," in *2022 10th International Conference on Information and Communication Technology (ICoICT)*, pp. 322–327, 2022, https://doi.org/10.1109/ICoICT55009.2022.9914896

[6] L. Sheng and L. Xu, "Topic Classification Based on Improved Word Embedding," in *2017 14th Web Information Systems and Applications Conference (WISA)*, pp. 117–121, 2017, https://doi.org/10.1109/WISA.2017.44.

[7] A. Meddeb and L. Ben Romdhane, "Using Topic Modeling and Word Embedding for Topic Extraction in Twitter," *Procedia Comput Sci*, vol. 207, pp. 790–799, 2022, https://doi.org/10.1016/j.procs.2022.09.134.

[8] D. T. Maulidia and 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)*, pp. 87–92, 2022, https://doi.org/10.1109/ICoDSA55874.2022.9862907.

[9] A. Rafea and N. A. Gaballah, "Topic Detection Approaches in Identifying Topics and Events from Arabic Corpora," in *Procedia Computer Science*, pp. 270–277, 2018, https://doi.org/10.1016/j.procs.2018.10.492.

[10] K. Gligori´c, G. Epfl, A. Anderson, and R. West, "How Constraints Affect Content: The Case of Twitter's Switch from 140 to 280 Characters," in *Proceedings of the International AAAI Conference on Web and Social Media,* vol. 12, no. 1, 2018, https://doi.org/10.1609/icwsm.v12i1.15079.

[11] L. Parameswaran, B. K.R, and R. K, "A Text Classification Model Using Convolution Neural Network and Recurrent Neural Network," vol. 119, no. 7, 2018, http://www.acadpubl.eu/hub/.

[12] G. Carducci, G. Rizzo, D. Monti, E. Palumbo, and M. Morisio, "TwitPersonality: Computing Personality Traits From Tweets Using Word Embeddings and Supervised Learning," *Information (Switzerland)*, vol. 9, no. 5, 2018, https://doi.org/10.3390/info9050127.

[13] E. M. Dharma, F. Lumban Gaol, H. Leslie, H. S. Warnars, and B. Soewito, "The Accuracy Comparison Among Word2vec, Glove, And Fasttext Towards Convolution Neural Network (CNN) Text Classification," *J Theor Appl Inf Technol*, vol. 31, no. 2, 2022, http://www.jatit.org/volumes/Vol100No2/5Vol100No2.pdf.

[14] J. Claussen and C. Peukert, "Obtaining Data from the Internet: A Guide to Data Crawling in Management Research," *SSRN Electronic Journal*, 2019, https://doi.org/10.2139/ssrn.3403799.

[15] K. N. Alam *et al.*, "Deep Learning-Based Sentiment Analysis of COVID-19 Vaccination Responses from Twitter Data," *Comput Math Methods Med*, vol. 2021, p. 4321131, 2021, https://doi.org/10.1155/2021/4321131.

[16] F. Gargiulo, S. Silvestri, M. Ciampi, and G. De Pietro, "Deep neural network for hierarchical extreme multi-label text classification," *Appl Soft Comput*, vol. 79, pp. 125–138, 2019, https://doi.org/10.1016/j.asoc.2019.03.041.

[17] J. Hartmann, J. Huppertz, C. Schamp, and M. Heitmann, "Comparing automated text classification methods," *International Journal of Research in Marketing*, vol. 36, no. 1, pp. 20–38, 2019, https://doi.org/10.1016/j.ijresmar.2018.09.009.

[18] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B.-W. On, "Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM)," *IEEE Access*, vol. 8, pp. 156695–156706, 2020, https://doi.org/10.1109/ACCESS.2020.3019735.

[19] M. Rosid, A. Fitrani, I. Astutik, N. Mulloh, and H. Gozali, "Improving Text Preprocessing For Student Complaint Document Classification Using Sastrawi," *IOP Conf Ser Mater Sci Eng*, vol. 874, p. 12017, 2020, https://doi.org/10.1088/1757-899X/874/1/012017.

[20] M. Anandarajan, C. Hill, and T. Nolan, "Text Preprocessing," in *Practical Text Analytics*, pp. 45–59, 2019, https://doi.org/10.1007/978-3-319-95663-3_4 .

[21] J. Yao, "Automated Sentiment Analysis of Text Data with NLTK," *J Phys Conf Ser*, vol. 1187, no. 5, p. 52020, 2019, https://doi.org/10.1088/1742-6596/1187/5/052020.

[22] E. Setiawan, D. Widyantoro, and K. Surendro, "Measuring information credibility in social media using combination of user profile and message content dimensions," *International Journal of Electrical and Computer Engineering*, vol. 10, pp. 3537–3549, 2020, https://doi.org/10.11591/ijece.v10i4.pp3537-3549.

[23] D. J. Ladani and N. P. Desai, "Stopword Identification and Removal Techniques on TC and IR applications: A Survey," in *6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 466– 472, 2020, https://doi.org/10.1109/ICACCS48705.2020.9074166.

[24] D. Merlini and M. Rossini, "Text categorization with WEKA: A survey," *Machine Learning with Applications*, vol. 4, p. 100033, 2021, https://doi.org/10.1016/j.mlwa.2021.100033.

[25] A. Kadhim, "An Evaluation of Preprocessing Techniques for Text Classification," *International Journal of Computer Science and Information Security,* vol. 16, no. 6, pp. 22–32, 2018, https://www.slideshare.net/IJCSISResearchPublic/an-evaluation-of-preprocessing-techniques-for-text-classification.
[26] A. I. Kadhim, "Term Weighting for Feature Extraction on Twitter: A Comparison Between BM25 and TF-IDF," in *International Conference on Advanced Science and Engineering (ICOASE)*, pp. 124–128, 2019, https://doi.org/10.1109/ICOASE.2019.8723825.
[27] P. M. Prihatini, I. K. Suryawan, and I. N. Mandia, "Feature extraction for document text using Latent Dirichlet Allocation," *J Phys Conf Ser*, vol. 953, no. 1, p. 12047, 2018, https://doi.org/10.1088/1742-6596/953/1/012047.
[28] N. Nasser, L. Karim, A. El Ouadrhiri, A. Ali, and N. Khan, "n-Gram based language processing using Twitter dataset to identify COVID-19 patients," *Sustain Cities Soc*, vol. 72, p. 103048, 2021, https://doi.org/10.1016/j.scs.2021.103048.
[29] S. W. Kim and J. M. Gil, "Research paper classification systems based on TF-IDF and LDA schemes," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, 2019, https://doi.org/10.1186/s13673-019-0192-7.
[30] A. I. Kadhim, "Survey on supervised machine learning techniques for automatic text classification," *Artif Intell Rev*, vol. 52, no. 1, pp. 273–292, 2019, https://doi.org/10.1007/s10462-018-09677-1.
[31] R. Dzisevič and D. Šešok, "Text Classification using Different Feature Extraction Approaches," in *Open Conference of Electrical, Electronic and Information Sciences (eStream)*, pp. 1–4, 2019, https://doi.org/10.1109/eStream.2019.8732167.
[32] M. Birjali, M. Kasri, and A. B-Hssane, "A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends," *Knowl Based Syst*, vol. 226, p. 107134, 2021, https://doi.org/10.1016/j.knosys.2021.107134.
[33] C. I. Eke, A. Norman, L. Shuib, F. B. Fatokun, and I. Omame, "The Significance of Global Vectors Representation in Sarcasm Analysis," in *International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS)*, pp. 1–7, 2020, https://doi.org/10.1109/ICMCECS47690.2020.246997.
[34] K. Kowsari *et al.*, "Text Classification Algorithms: A Survey," *Information (Switzerland)*, vol. 10, 2019, https://doi.org/10.3390/info10040150.
[35] D. Alsaleh and S. Larabi-Marie-Sainte, "Arabic Text Classification Using Convolutional Neural Network and Genetic Algorithms," *IEEE Access*, vol. 9, pp. 91670–91685, 2021, https://doi.org/10.1109/ACCESS.2021.3091376.
[36] S. V Georgakopoulos, S. K. Tasoulis, A. G. Vrahatis, and V. P. Plagianakos, "Convolutional Neural Networks for Toxic Comment Classification," *Proceedings of the 10th Hellenic Conference on Artificial Intelligence*, pp. 1-6, 2018, https://doi.org/10.1145/3200947.3208069.
[37] N. I. Widiastuti, "Convolution Neural Network for Text Mining and Natural Language Processing," *IOP Conf Ser Mater Sci Eng*, vol. 662, no. 5, p. 52010, 2019, https://doi.org/10.1088/1757-899X/662/5/052010.
[38] E. Tutubalina, Z. Miftakhutdinov, S. Nikolenko, and V. Malykh, "Medical Concept Normalization in Social Media Posts with Recurrent Neural Networks," *J Biomed Inform*, vol. 84, 2018, https://doi.org/10.1016/j.jbi.2018.06.006.
[39] E. C. Nisa and Y. D. Kuan, "Comparative assessment to predict and forecast water-cooled chiller power consumption using machine learning and deep learning algorithms," *Sustainability (Switzerland)*, vol. 13, no. 2, pp. 1–18, 2021, https://doi.org/10.3390/su13020744.
[40] M. Azizjon, A. Jumabek, and W. Kim, "1D CNN based network intrusion detection with normalization on imbalanced data," in *International Conference on Artificial Intelligence in Information and Communication*, pp. 218–224, 2020, https://doi.org/10.1109/ICAIIC48513.2020.9064976.
[41] H. Mohaouchane, A. Mourhir, and N. S. Nikolov, "Detecting Offensive Language on Arabic Social Media Using Deep Learning," in *Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 466–471, 2019, https://doi.org/10.1109/SNAMS.2019.8931839.
[42] G. K. Soon, C. K. On, N. M. Rusli, T. S. Fun, R. Alfred, and T. T. Guan, "Comparison Of Simple Feedforward Neural Network, Recurrent Neural Network And Ensemble Neural Networks In Phishing Detection," *J Phys Conf Ser*, vol. 1502, no. 1, p. 012033, 2020, https://doi.org/10.1088/1742-6596/1502/1/012033.
[43] E. Singh, N. Kuzhagaliyeva, and S. M. Sarathy, "Chapter 9 - Using deep learning to diagnose preignition in turbocharged spark-ignited engines," in *Artificial Intelligence and Data Driven Optimization of Internal Combustion Engines*, pp. 213–237, 2022, https://doi.org/10.1016/B978-0-323-88457-0.00005-9.
[44] G. Tanaka *et al.*, "Recent Advances In Physical Reservoir Computing: A Review," *Neural Networks*, vol. 115, pp. 100–123, 2019, https://doi.org/10.1016/j.neunet.2019.03.005.
[45] C. Du, L. Huang, C. Du, and L. Huang, "Text Classification Research with Attention-based Recurrent Neural Networks," *International Journal of Computers Communications & Control,* vo. 13, no.1, pp. 50-61, 2018, https://doi.org/10.15837/ijccc.2018.1.3142.
[46] H. Jelodar, Y. Wang, R. Orji, and S. Huang, "Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach," *IEEE J Biomed Health Inform*, vol. 24, no. 10, pp. 2733–2742, 2020, https://doi.org/10.1109/JBHI.2020.3001216.

[47] X. Liu, Y. Su and B. Xu, "The Application of Graph Neural Network in Natural Language Processing and Computer Vision," *3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, pp. 708714, 2021, https://doi.org/10.1109/MLBDBI54094.2021.00140.

[48] S. V. Bhoir, "An Efficient Fake News Detector," in *International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1-9, 2020, https://doi.org/10.1109/ICCCI48352.2020.9104177.

R. Hamad, L. Yang, W. L. Woo, and B. Wei, "Joint Learning of Temporal Models to Handle Imbalanced Data for Human Activity Recognition," *Applied Sciences*, vol. 10, p. 5293, 2020, https://doi.org/10.3390/app10155293.