

DAFTAR PUSTAKA

- Aouragh, S., Gueddah, H. dan Yousfi, A., 2015, Adaptating the levenshtein distance to contextual spelling correction, *International Journal of Computer Science Applications* **12**: 127–133.
- Baldwin, T. dan Li, Y., (2015). An in-depth analysis of the effect of text normalization in social media, *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Association for Computational Linguistics, Denver, Colorado, pp. 420–429. <https://www.aclweb.org/anthology/N15-1045>
- Barik, A. M., Mahendra, R. dan Adriani, M., (2019). Normalization of indonesian-english code-mixed twitter data, *W-NUT@EMNLP*.
- Beckley, R., (2015). Bekli:a simple approach to twitter text normalization., *Proceedings of the Workshop on Noisy User-generated Text*, Association for Computational Linguistics, Beijing, China, pp. 82–86. <https://www.aclweb.org/anthology/W15-4312>
- Bojanowski, P., Grave, E., Joulin, A. dan Mikolov, T., 2017, *Transactions of the Association for Computational Linguistics* **5**: 135–146. https://doi.org/10.1162/tacl_a00051
- Caselles-Dupré, H., Lesaint, F. dan Royo-Letelier, J., 2018, Word2vec applied to recommendation: Hyperparameters matter, *CoRR* **abs/1804.04212**. <http://arxiv.org/abs/1804.04212>
- Chollet, F., 2017. *Deep Learning with Python*, Manning Publications.
- Clark, E. dan Araki, K., 2011, Text normalization in social media: Progress, problems and applications for a pre-processing system of casual english, *Procedia - Social and Behavioral Sciences* **27**: 2 – 11. Computational Linguistics and Related Fields. <http://www.sciencedirect.com/science/article/pii/S1877042811024049>
- Costas, R., van Honk, J. dan Franssen, T., 2017, Scholars on twitter: who and how many are they?

- Demaine, E., 2011, <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-006-introduction-to-algorithms-fall-2011/lecture-videos/lecture-8-hashing-with-chaining/> Hashing with chaining.
- Diez, P., (2018). Chapter 1 - introduction, in P. Diez (ed.), *Smart Wheelchairs and Brain-Computer Interfaces*, Academic Press, pp. 1 – 21. <http://www.sciencedirect.com/science/article/pii/B9780128128923000017>
- Dirkson, A., Verberne, S., Sarker, A. dan Kraaij, W., 2019, Data-driven lexical normalization for medical social media, *Multimodal Technologies and Interaction* **3**(3). <https://www.mdpi.com/2414-4088/3/3/60>
- Dreßler, K. dan Ngonga Ngomo, A.-C., 2015, On the efficient execution of bounded jaro-winkler distances.
- Fiona Maclean, Derek Jones², G. C.-L. H. H., 2013, Understanding twitter, *British Journal of Occupational Therapy* **76**: 295–298.
- Foster, J., Çetinoglu, Ö., Wagner, J., Roux, J. L., Hogan, S., Nivre, J., Hogan, D. dan van Genabith, J., (2011). hardtoparse: Pos tagging and parsing the twitterverse, *Analyzing Microtext*.
- Gupta, I. dan Joshi, N., (2017). Tweet normalization: A knowledge based approach, *2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS)*, pp. 157–162.
- Göker, S. dan Can, B., (2018). Neural text normalization for turkish social media, *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, pp. 161–166.
- Haldar, R. dan Mukhopadhyay, D., 2011, Levenshtein distance technique in dictionary lookup methods: An improved approach, *Computing Research Repository - CORR*.
- Hanafiah, N., Kevin, A., Sutanto, C., Fiona, Arifin, Y. dan Hartanto, J., 2017a, Text normalization algorithm on twitter in complaint category, *Procedia Computer Science* **116**: 20 – 26. Discovery and innovation of computer science technology in artificial intelligence era: The 2nd International Conference on Computer Science and Computational Intelligence (ICCSCI 2017). <http://www.sciencedirect.com/science/article/pii/S1877050917320410>

- Hanafiah, N., Kevin, A., Sutanto, C., Fiona, Arifin, Y. dan Hartanto, J., 2017b, Text normalization algorithm on twitter in complaint category, *Procedia Computer Science* **116**: 20 – 26. Discovery and innovation of computer science technology in artificial intelligence era: The 2nd International Conference on Computer Science and Computational Intelligence (ICCSCI 2017). <http://www.sciencedirect.com/science/article/pii/S1877050917320410>
- Jurafsky, D. dan Martin, J. H., 2008. *Speech and Language Processing*, Prentice Hall.
- Khurana, D., Koli, A., Khatter, K. dan Singh, S., 2017, Natural language processing: State of the art, current trends and challenges, *CoRR* **abs/1708.05148**. <http://arxiv.org/abs/1708.05148>
- Kusumawardani, R. P., Faizal, S. P. dan Atletiko, J., 2018, Context-sensitive normalization of social media text in bahasa indonesia based on neural word embeddings, *Procedia Computer Science* **144**: 105–117.
- Li, Y. dan Yang, T., (2017). *Word Embedding for Understanding Natural Language: A Survey*, Vol. 26.
- Liu, X., Zhang, S., Wei, F. dan Zhou, M., (2011). Recognizing named entities in tweets, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Association for Computational Linguistics, Portland, Oregon, USA, pp. 359–367. <https://www.aclweb.org/anthology/P11-1037>
- Manning, C., Raghavan, P. dan Schütze, H., 2010, Introduction to information retrieval, *Natural Language Engineering* **16**(1): 100–103.
- Mikolov, T., Chen, K., Corrado, G. dan Dean, J., 2013, Efficient estimation of word representations in vector space.
- Mistry, V., 2013, Critical care training: using twitter as a teaching tool, *British Journal of Nursing* **20**: 1292–1296.
- Naili, M., Chaibi, A. H. dan Ghezala, H. H. B., 2017, Comparative study of word embedding methods in topic segmentation, *Procedia Computer Science* **112**: 340–349.

- Ozer, Z., Ozer, I. dan Findik, O., 2018, Diacritic restoration of turkish tweets with word2vec, *Engineering Science and Technology, an International Journal* **21**(6): 1120 – 1127. <http://www.sciencedirect.com/science/article/pii/S2215098618308668>
- Powers, D., 2008, Evaluation: From precision, recall and f-factor to roc, informedness, markedness correlation, *Mach. Learn. Technol.* **2**.
- Rinaldi, E. dan Musdholifah, A., (2017). Fvec-svm for opinion mining on indonesian comments of youtube video, *2017 International Conference on Data and Software Engineering (ICoDSE)*, pp. 1–5.
- Rozovskaya, A., Habash, N., Eskander, R., Farra, N. dan Salloum, W., (2014). The Columbia system in the QALB-2014 shared task on Arabic error correction, *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP)*, Association for Computational Linguistics, Doha, Qatar, pp. 160–164. <https://www.aclweb.org/anthology/W14-3622>
- Sammut, C. dan Webb, G. I. (eds), (2010). *Accuracy*, Springer US, Boston, MA, pp. 9–10. https://doi.org/10.1007/978-0-387-30164-8_3
- Sarker, A., 2017, A customizable pipeline for social media text normalization, *Social Network Analysis and Mining* **7**.
- Setiadi, I., 2013, Damerau-levenshtein algorithm and bayes theorem for spell checker optimization.
- Turian, J., Ratnoff, L.-A. dan Bengio, Y., (2010). Word representations: A simple and general method for semi-supervised learning, *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Uppsala, Sweden, pp. 384–394. <https://www.aclweb.org/anthology/P10-1040>
- van der Goot, R., (2016). Normalizing social media texts by combining word embeddings and edit distances in a random forest regressor.
- Wang, B., Wang, A., Chen, F., Wang, Y. dan Kuo, C. J., 2019, Evaluating word embedding models: Methods and experimental results, *CoRR* **abs/1901.09785**. <http://arxiv.org/abs/1901.09785>

Watson, D., Zalmout, N. dan Habash, N., (2018). Utilizing character and word embeddings for text normalization with sequence-to-sequence models, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Brussels, Belgium, pp. 837–843. <https://www.aclweb.org/anthology/D18-1097>

Zaghouni, W., Mohit, B., Habash, N., Obeid, O., Tomeh, N., Rozovskaya, A., Farra, N., Alkuhlani, S. dan Oflazer, K., (2014). Large scale Arabic error annotation: Guidelines and framework, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, European Language Resources Association (ELRA), Reykjavik, Iceland. http://www.lrec-conf.org/proceedings/lrec2014/pdf/956_paper.pdf