



DAFTAR PUSTAKA

Abualigah, L.M., Khader, A.T. and Al-Betar, M.A. (2016) 'Multi-objectives-based text clustering technique using K-Means algorithm', in *2016 7th International Conference on Computer Science and Information Technology (CSIT)*. IEEE.

Adriani, M. *dkk.* (2007) 'Stemming Indonesian: A Confix-Stripping Approach', in *Conferences in Research and Practice in Information Technology Series*. ACM Transactions on Asian Language Information Processing, pp. 307–314. Available at: <https://doi.org/10.1145/1316457.1316459>.

Alhawarat, M. and Hegazi, M. (2018) 'Revisiting K-Means and Topic Modeling, a Comparison Study to Cluster Arabic Documents', *IEEE Access*, 6, pp. 42740–42749. Available at: <https://doi.org/10.1109/ACCESS.2018.2852648>.

Alobed, M., Altrad, A.M.M. and Bakar, Z.B.A. (2021) 'A Comparative Analysis of Euclidean, Jaccard and Cosine Similarity Measure and Arabic Wordnet for Automated Arabic Essay Scoring', in *Proceedings - CAMP 2021: 2021 5th International Conference on Information Retrieval and Knowledge Management: Digital Technology for IR 4.0 and Beyond*. Institute of Electrical and Electronics Engineers Inc., pp. 70–74. Available at: <https://doi.org/10.1109/CAMP51653.2021.9498119>.

Alqaisi, R., Ghanem, W. and Qaroush, A. (2020) 'Extractive Multi-Document Arabic Text Summarization Using Evolutionary Multi-Objective Optimization With K-Medoids Clustering', *IEEE Access*, 8, pp. 228206–228224. Available at: <https://doi.org/10.1109/ACCESS.2020.3046494>.

Arthur, D. and Vassilvitskii, S. (2007) *K-Means++: The Advantages of Careful Seeding*.

Bingi, Y. and Yin, Y. (2023) 'An Analysis of Embedding Layers and Similarity Scores using Siamese Neural Networks'. Available at: <http://arxiv.org/abs/2401.00582>.

Dhir, S. *dkk.* (2024) 'News in Vectors: Optimizing News Analysis through Advanced Embeddings and Clustering', in *Proceedings - 2nd International Conference on Advancement in Computation and Computer Technologies, InCACCT 2024*. Institute of Electrical and Electronics Engineers Inc., pp. 40–45. Available at: <https://doi.org/10.1109/InCACCT61598.2024.10551154>.

Dhivya, R. and Shanmugapriya, N. (2023) 'An Efficient DBSCAN with Enhanced Agglomerative Clustering Algorithm', in *2023 4th International Conference on Electronics and Sustainable Communication Systems, ICESC 2023 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., pp. 1322–1327. Available at: <https://doi.org/10.1109/ICESC57686.2023.10193224>.

Duo, J., Zhang, P. and Hao, L. (2021) 'A K-Means Text clustering Algorithm Based on Subject Feature Vector', *Journal of Web Engineering*, 20(6), pp. 1935–1946. Available at: <https://doi.org/10.13052/jwe1540-9589.20612>.



Hasani, M.F. *dkk.* (2023) ‘Density Based Spatial Clustering of Applications with Noise and Sentence Bert Embedding for Indonesian Utterance Clustering’, *ICCoSITE 2023 - International Conference on Computer Science, Information Technology and Engineering: Digital Transformation Strategy in Facing the VUCA and TUNA Era*, pp. 386–391. Available at: <https://doi.org/10.1109/ICCoSITE57641.2023.10127683>.

Hu, W., Xu, D. and Niu, Z. (2021) ‘Improved K-Means Text clustering Algorithm Based on BERT and Density Peak’, in *2021 2nd Information Communication Technologies Conference, ICTC 2021*. Institute of Electrical and Electronics Engineers Inc., pp. 260–264. Available at: <https://doi.org/10.1109/ICTC51749.2021.9441505>.

Lahitani, A.R., Permanasari, A.E. and Setiawan, N.A. (2016) ‘Cosine similarity to determine similarity measure: Study case in online essay assessment’, *Proceedings of 2016 4th International Conference on Cyber and IT Service Management, CITSM 2016*, pp. 1–6. Available at: <https://doi.org/10.1109/CITSM.2016.7577578>.

Liu, Y. and Rong, Y. (2020) ‘Staged text clustering algorithm based on K-Means and hierarchical agglomeration clustering’, in *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*. IEEE.

Liu, Y., Xu, Q. and Tang, Z. (2019) *Research on Text Classification Method Based on PTF-IDF and Cosine Similarity*. Available at: <https://www.kaggle.com/shineucc/bbc-news->.

Metre, V.A., Popat, S.K. and Deshmukh, P.B. (2017) ‘Hierarchical Document Clustering based on Cosine Similarity measure’, in. IEEE.

Onan, A. (2017) ‘A K-Medoids Based Clustering Scheme with an Application to Document Clustering’, in *(UBMK'17) International Conference on Computer Science and Engineering*. IEEE.

Pramudita, Y.D. *dkk.* (2022) ‘Automatic Text Summarization of Madura Tourism Articles Using TF-IDF and K-Medoids Clustering’, in *Proceeding - IEEE 8th Information Technology International Seminar, ITIS 2022*. Institute of Electrical and Electronics Engineers Inc., pp. 168–172. Available at: <https://doi.org/10.1109/ITIS57155.2022.10009983>.

Ristanti, Y., Wibawa, A.P. and Utomo, P. (2019) ‘Cosine Similarity for Title and Abstract of Economic Journal Classification’, in *2019 5th International Conference on Science in Information Technology (ICSITech) : ‘Embracing Industry 4.0 : Towards Innovation in Cyber Physical System’ : proceeding : October 23-24, 2019, Yogyakarta, Indonesia*. IEEE.

Saipech, P. and Seresangtakul, P. (2018) *Automatic Thai Subjective Examination using Cosine Similarity*. IEEE.

Salsabila, N.A. *dkk.* (2018) ‘Colloquial Indonesian Lexicon’, in *2018 International Conference on Asian Language Processing*. Institute of Electrical and Electronics Engineers, p. 327.

Schubert, E. and Rousseeuw, P.J. (2021) ‘Fast and eager K-Medoids clustering: O(k) runtime improvement of the PAM, CLARA, and CLARANS algorithms’, *Information Systems*, 101. Available at: <https://doi.org/10.1016/j.is.2021.101804>.

Setiawati, F.A. *dkk.* (2019) *Feature Selection using K-Medoids Algorithm for Categorization of Hadith Translation in English*. IEEE.



Shabrina, A.N., Afdal, M. and Monalisa, S. (2023) *Comparison Of K-Means, K-Medoids, and Fuzzy C-Means Algorithms for Clustering Drug User's Addiction Levels, Jurnal Sistem Cerdas.*

Si, X.L. *dkk.* (2022) 'An Online Dirichlet Model based on Sentence Embedding and DBSCAN for Noisy Short Text Stream Clustering', in *Proceedings of the International Joint Conference on Neural Networks*. Institute of Electrical and Electronics Engineers Inc. Available at: <https://doi.org/10.1109/IJCNN55064.2022.9892414>.

Tan, Y. (2018) 'An Improved KNN Text Classification Algorithm Based on K-Medoids and Rough Set', in *Proceedings - 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2018*. Institute of Electrical and Electronics Engineers Inc., pp. 109–113. Available at: <https://doi.org/10.1109/IHMSC.2018.00032>.

Ünver, M. (2023) 'Failure Mode Extraction via Vectorization and Clustering: A Case Study on Amazon Reviews', in *International Conference on Electrical, Computer and Energy Technologies, ICECET 2023*. Institute of Electrical and Electronics Engineers Inc. Available at: <https://doi.org/10.1109/ICECET58911.2023.10389603>.

Wang, X. *dkk.* (2020) 'Application of Improved DBSCAN Clustering Algorithm on Industrial Fault Text Data', in *IEEE International Conference on Industrial Informatics (INDIN)*. Institute of Electrical and Electronics Engineers Inc., pp. 461–468. Available at: <https://doi.org/10.1109/INDIN45582.2020.9442093>.

Zhang, X. *dkk.* (2024) 'mGTE: Generalized Long-Context Text Representation and Reranking Models for Multilingual Text Retrieval', in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*. Association for Computational Linguistics, pp. 1393–1412. Available at: <https://doi.org/10.18653/v1/2024.emnlp-industry.103>.