DAFTAR PUSTAKA

Angelov, D., 2020. Top2Vec: Distributed Representations of Topics. *CoRR.*

B.Dieng, A., J.R.Ruiz, F. & M.Blei, D., 2020. Topic Modeling in Embedding Spaces. *Transactions of the Association for Computational Linguistics,* Volume 8, pp. 439-453.

Bojanowski, P., Grave, E., Joulin, A. & Mikolov, T., 2017. Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics,* Volume 5, pp. 135-146.

Bouma, G., 2009. *Normalized (Pointwise) Mutual Information in Collocation Extraction.* s.l., s.n.

Boyd-Graber, J., Hu, Y. & Mimno, D., 2017. Applications of Topic Models. *Foundations and Trends in Information Retrieval,* Volume 11, pp. 143-296.

Budiarto, A. et al., 2021. Unsupervised News Topic Modelling with Doc2Vec and Spherical Clustering. *5th International Conference on Computer Science and Computational Intelligence 2020,* pp. 40-46.

Che, W. et al., 2018. *Towards Better {UD} Parsing: Deep Contextualized Word Embeddings, Ensemble, and Treebank Concatenation.* Brussels, Belgium, Association for Computational Linguistics.

Costa, G. & Ortale, R., 2020. *Document Clustering Meets Topic Modeling with Word Embeddings.* s.l., s.n.

Ethayarajh, K., 2019. *How Contextual are Contextualized Word Representations?Comparing the Geometry of BERT, ELMo, and GPT-2 Embeddings.* Hong Kong, China, Association for Computational Linguistics.

Fahlevvi, M. R. & A., 2022. Topic Modeling on Online News.Portal Using Latent. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems),* Volume 16, pp. 335-344.

Grootendorst, M., 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. Issue ArXiv abs/2203.05794 (2022).

Guangxu XUn, Y. L. W. X. Z., Gao, J. & Zhang, A., 2017. *A Correlated Topic Model Using Word Embeddings.* s.l., s.n.

Herwanto, G. B., 2018. DOCUMENT CLUSTERING DENGAN LATENT DIRICHLET ALLOCATION DAN HIERARICHAL CLUSTERING. *Jurnal Pseudocode,* Volume 5, pp. 29-37.

Ibrohim, M. O. & Budi, I., 2019. *Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter.* Florence, Italy, Association for Computational Linguistics.

Kherwa, P. & Bansal, P., 2019. Topic Modeling: A Comprehensive Review. *EAI Endorsed Transactions on Scalable Information Systems,* Volume 7.

Lau, J. H., Newman, D. & Baldwin, T., 2014. *Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality.* Gothenburg, Sweden, Association for Computational Linguistics.

Liu, Q., J.Kusner, M. & Blunsom, P., 2020. A Survey on Contextual Embeddings. *CoRR.*

Li, Y., Cai, J. & Wang, J., 2020. A Text Document Clustering Method Based on Weighted BERT Model. *IEEE 4th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC 2020,* pp. 1426-1430.

M.Alhawarat & M.Hegazi, 2016. Revisiting K-Means and Topic Modeling, a Comparison Study to Cluster Arabic Documents. *IEEE Access,* Volume 4.

M.Blei, D., Y.Ng, A. & I.Jordan, M., 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research 3,* pp. 993-1022.

Mahmood, A., 2013. Literature Survey on Topic Modeling.

Mikolov, T. et al., 2013. *Distributed Representations of Words and Phrasesand their Compositionality.* s.l., Curran Associates Inc.

Naseem, U., Razzak, I., Khan, S. K. & Prasad, M., 2020. A Comprehensive Survey on Word Representation Models: From Classical to State-Of-The-Art Word Representation Language Models. *ACM Transactions on Asian and Low-Resource Language Information Processing,* 20(5), pp. 1-35.

Nguyen, D. Q., Billingsley, R., Dau, L. & Johnson, M., 2018. Improving Topic Models with Latent Feature Word Representations. *Transactions of the Association for Computational Linguistics,* Volume 3.

Nurdin, A., Aji, B. A. S., Bustamin, A. & abidin, Z., 2020. Perbandingan kinerja word embedding Word2vec, Glove, dan FastText pada klasifikasi teks. *Jurnal TEKNOKOMPAK,* 14(2), pp. 74-79.

Nurifan, F., Sarno, R. & Sungkono, K. R., 2019. Aspect Based Sentiment Analysis for Restaurant Reviews Using Hybrid ELMo-Wikipedia and Hybrid Expanded Opinion Lexicon-SentiCircle. *International Journal of Intelligent Engineering and Systems,* Volume 12, pp. 47-58.

Pennington, J., Socher, R. & D.Manning, C., 2014. *GloVe: Global Vectors for Word Representation.* Doha, Qatar, Association for Computational Linguistics.

Peters, M. E. et al., 2018. *Deep Contextualized Word Representations.* New Orleans, Louisiana, Association for Computational Linguistics.

Pilehvar, M. T. & Camacho-Collados, J., 2021. *Embeddings in Natural Language Processing: Theory and Advances in Vector Representations of Meaning.* 1 ed. s.l.:Springer Cham.

Rachman, F. F. & Pramana, S., 2020. Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter. *Indonesian of Health Information Management Journal,* 8(2), pp. 100-109.

Rahmanti, A. R. et al., 2022. Social media sentiment analysis to monitor the performance of vaccination coverage during the early phase of the national COVID-19 vaccine rollout. *Computer Methods and Programs in Biomedicine,* Volume 221.

Rashif, F., Nirvana, G. I. P., Noor, M. A. & Rakhmawati, N., 2021. Implementasi LDA untuk Pengelompokan Topik Cuitan. *Cogito Smart Journal,* Volume 7, pp. 170-181.

Reimers, N. et al., 2019. *Classification and Clustering of Arguments with Contextualized Word Embeddings.* Florence, Italy, Association for Computational Linguistics.

Roder, M., Both, A. & Hinneburg, A., 2015. *Exploring the Space of Topic Coherence Measures.* s.l., s.n.

Sia, S., Dalmia, A. & Mielke, S. J., 2020. *Tired of Topic Models? Clusters of Pretrained Word Embeddings Make for Fast and Good Topics too!.* s.l., Association for Computational Linguistic.

Singh, R. V. & Bhatia, M., 2011. Data Clustering with Modified K-means Algorithm. *IEEE-International Conference on Recent Trends in Information Technology (ICRTIT),* pp. 717-721.

Steinskog, A. O., Therkelsen, J. F. & Gambäck, B., 2017. *Twitter Topic Modeling by Tweet Aggregation.* Gothenburg,Sweden, Linköping University Electronic Press.

Thompson, L. & Mimno, D., 2020. Topic Modeling with Contextualized Word Representation Clusters. *CoRR.*

Walkowiak, T. & Gniewkowski, M., 2019. *Evaluation of vector embedding models in clustering of text documents.* Varna, Bulgaria, s.n.

Wang, A. & Zhang, J., 2020. Topic discovery method based on topic model combined with hierarchical clustering. *IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC 2020,* pp. 814-818.

Wang, B., Wang, A., Chen, F. & Wang, Y., 2019. Evaluating Word Embedding Models: Methods and Experimental Results. *APSIPA Transactions on Signal and Information Processing,* Volume 8.

Wijayanto, I. R., Cholissodin, I. & Sari, Y. A., 2021. Pengaruh Metode Word Embedding dalam Vector Space Model pada Pemerolehan Informasi Materi

IPA Siswa SMP. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer,* 5(3), pp. 950-959.

Yi, F., Jiang, B. & Wu, J., 2020. Topic Modeling for Short Texts via WordEmbedding and Document Correlation. *IEEE Access,* Volume 8, pp. 30692-30705.

Yin, H., Song, X., Yang, S. & Li, J., 2022. Sentiment analysis and topic modeling for COVID-19 vaccine discussions. *World Wide Web,* Volume 25, pp. 1067-1083.

Zhang, D. & Li, S., 2011. *Topic Detection Based on K-means.* Ningbo,China, s.n.

Zhang, P. et al., 2020. Combine Topic Modeling with Semantic Embedding: Embedding Enhanced Topic Model. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING,* 32(12), pp. 2322-2335.