

## DAFTAR PUSTAKA

- Abdelrazek, A., Eid, Y., Gawish, E., Medhat, W., Hassan, A. (2023). "Topic modeling algorithms and applications: A survey". *Information Systems*. **112**:102131. <https://doi.org/10.1016/j.is.2022.102131>
- Aguilar, J., Salazar, C., Velasco, H., Monsalve-Pulido, J., Montoya, E. (2020). "Comparison and evaluation of different methods for the feature extraction from educational contents". *Computation*. **8(2)**:30. <https://doi.org/10.3390/COMPUTATION8020030>
- Akbarighatar, P., Pappas, I., Vassilakopoulou, P. (2023). "A sociotechnical perspective for responsible AI maturity models: Findings from a mixed-method literature review". *International Journal of Information Management Data Insights*. **3(2)**:100193. <https://doi.org/10.1016/j.ijime.2023.100193>
- Aleqabie, H. J., Sfoq, M. S., Albeer, R. A., Abd, E. H. (2024). "A Review Of Text Mining Techniques: Trends, and Applications In Various Domains". *Iraqi Journal for Computer Science and Mathematics*. **5(1)**:125-141. <https://doi.org/10.52866/ijcsm.2024.05.01.009>
- Awalina, A., Fawaid, J., Yunus Krisnabayu, R., Yudistira, N. (2021). "Indonesia's Fake News Detection using Transformer Network". *Proceedings of the 6th International Conference on Sustainable Information Engineering and Technology*. 247-251. <https://doi.org/10.1145/3479645.3479666>
- Birunda, S. S., Devi, R. K. (2021). "A review on word embedding techniques for text classification". *Innovative Data Communication Technologies and Application: Proceedings of ICIDCA 2020*. 267-281. [https://doi.org/10.1007/978-981-15-9651-3\\_23](https://doi.org/10.1007/978-981-15-9651-3_23)

- Chai, X., Xu, S., Li, S., Zhao, J. (2023). "The Process and Algorithm Analysis of Text Mining System Based on Artificial Intelligence". *Procedia Computer Science*. **228**:574-581. <https://doi.org/10.1016/j.procs.2023.11.066>
- Chen, W., Rabhi, F., Liao, W., Al-Qudah, I. (2023). "Leveraging State-of-the-Art Topic Modeling for News Impact Analysis on Financial Markets: A Comparative Study". *Electronics (Switzerland)*. **12(12)**. <https://doi.org/10.3390/electronics12122605>
- de Lima, B. C., Baracho, R. M. A., Mandl, T., Porto, P. B. (2023). "Reactions to science communication: discovering social network topics using word embeddings and semantic knowledge". *Social Network Analysis and Mining*. **13(1)**. <https://doi.org/10.1007/s13278-023-01125-5>
- Egger, R., Yu, J. (2022). "A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts". *Frontiers in Sociology*. **7**. <https://doi.org/10.3389/fsoc.2022.886498>
- Faizah, Lin, B. S. (2023). "Visualizing Change and Correlation of Topics With LDA and Agglomerative Clustering on COVID-19 Vaccine Tweets". *IEEE Access*. **11**:51647-51656. <https://doi.org/10.1109/ACCESS.2023.3278979>
- Fesseha, A., Xiong, S., Emiru, E. D., Diallo, M., Dahou, A. (2021). "Text classification based on convolutional neural networks and word embedding for low-resource languages: Tigrinya". *Information (Switzerland)*. **12(2)**:1-17. <https://doi.org/10.3390/info12020052>
- Gan, J., Qi, Y. (2021). "Selection of the optimal number of topics for LDA topic model—Taking patent policy analysis as an example". *Entropy*. **23(10)**:1301. <https://doi.org/10.3390/e23101301>
- Głowania, S., Kozak, J., Juszczuk, P. (2023). "Knowledge Discovery in Databases for a Football Match Result". *Electronics (Switzerland)*. **12(12)**:2712. <https://doi.org/10.3390/electronics12122712>

- Grootendorst, M. (2022). "BERTopic: Neural topic modeling with a class-based TF-IDF procedure". *arXiv preprint*. **arXiv:2203.05794**. <https://doi.org/10.48550/arXiv.2203.05794>
- Hananto, V. R. (2023). "Implementation of Dynamic Topic Modeling to Discover Topic Evolution on Customer Reviews". *Jurnal Online Informatika*. **8(2)**:147-157. <https://doi.org/10.15575/join.v8i2.963>
- Hutama, L. B., Suhartono, D. (2022). "Indonesian Hoax News Classification with Multilingual Transformer Model and BERTopic". *Informatika (Slovenia)*. **46(8)**:81-90. <https://doi.org/10.31449/inf.v46i8.4336>
- Jansevskis, M., Osis, K. (2023). "Knowledge Discovery Frameworks and Characteristics". *Baltic Journal of Modern Computing*. **11(4)**:686-702. <https://doi.org/10.22364/bjmc.2023.11.4.08>
- Jung, Y. J., Kim, Y. (2023). "Research trends of sustainability and marketing research, 2010–2020: Topic modeling analysis". *Heliyon*. **9(3)**. <https://doi.org/10.1016/j.heliyon.2023.e14208>
- KARADAĞ, T. (2023). "Can we identify the similarity of courses in computer science?". *Sigma: Journal of Engineering & Natural Sciences/Mühendislik ve Fen Bilimleri Dergisi*. **41(4)**. <https://doi.org/10.14744/sigma.2023.00089>
- Khadija, M. A., Nurharjadmo, W. (2024). "Enhancing Indonesian customer complaint analysis: LDA topic modelling with BERT embeddings". *Sinergi (Indonesia)*. **28(1)**:153-162. <https://doi.org/10.22441/sinergi.2024.1.015>
- Khairunnisa, S., Adiwijaya, A., Faraby, S. Al. (2021). "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)". *JURNAL MEDIA INFORMATIKA BUDIDARMA*. **5(2)**:406. <https://doi.org/10.30865/mib.v5i2.2835>
- Khurana, D., Koli, A., Khatter, K., Singh, S. (2023). "Natural language processing: state of the art, current trends and challenges". *Multimedia Tools and Applications*. **82(3)**:3713-3744. <https://doi.org/10.1007/s11042-022-13428-4>

- Kukushkin, K., Ryabov, Y., Borovkov, A. (2022). "Digital Twins: A Systematic Literature Review Based on Data Analysis and Topic Modeling". *Data*. **7(12)**:173. <https://doi.org/10.3390/data7120173>
- Kumar, A. N., Raj, R. K., Aly, S. G., Anderson, M. D., Becker, B. A., Blumenthal, R. L., Eaton, E., Epstein, S. L., Goldweber, M., Jalote, P., Lea, D., Oudshoorn, M., Pias, M., Reiser, S., Servin, C., Simha, R., Winters, T., Xiang, Q. (2024). "Computer Science Curricula 2023". *Association for Computing Machinery*. <https://doi.org/10.1145/3664191>
- Lee, H., Lee, S. H., Lee, K. R., Kim, J. H. (2023). "ESG Discourse Analysis Through BERTopic: Comparing News Articles and Academic Papers". *Computers, Materials and Continua*. **75(3)**:6023-6037. <https://doi.org/10.32604/cmc.2023.039104>
- Lee, S., Kang, M. (2022). "Trend Analysis of Physical Activity Measurement Research Using Text Mining in Big Data Analytics". *International Journal of Physical Activity and Health*. **1(1)**:5. <https://doi.org/10.18122/ijpah1.1.5.boisestate>
- Lu, C., Zhu, L., Xie, Y., Xu, W., Zhao, Y., Cao, Y. (2024). "Analysis of Hot Topics and Evolution of Research in World-class Agricultural Universities Based on BERTopic". *Applied Mathematics and Nonlinear Sciences*. **9(1)**. <https://doi.org/10.2478/amns-2024-0327>
- Mendonça, M., Figueira, Á. (2024). "Topic Extraction: BERTopic's Insight into the 117th Congress's Twittersverse". *Informatics*. **11(1)**. <https://doi.org/10.3390/informatics11010008>
- Mulia, F. R., Gupta, A. K. (2022). "A Review Paper on Dimensionality Reduction Techniques". *Journal of Pharmaceutical Negative Results*. **13(SO3)**:1263-1272. <https://doi.org/10.47750/pnr.2022.13.s03.198>
- Naseem, U., Razzak, I., Eklund, P. W. (2021). "A survey of pre-processing techniques to improve short-text quality: a case study on hate speech detection on twitter". *Multimedia Tools and Applications*, **80(28–29)**, 35239–35266. <https://doi.org/10.1007/s11042-020-10082-6>

- Ng, Q. X., Lee, D. Y. X., Yau, C. E., Lim, Y. L., Liew, T. M. (2023). “Public perception on “healthy ageing” in the past decade: An unsupervised machine learning of 63,809 Twitter posts”. *Heliyon*. **9(2)**. <https://doi.org/10.1016/j.heliyon.2023.e13118>
- Niroomand, K., Saady, N. M. C., Bazan, C., Zendehboudi, S., Soares, A., Albayati, T. M. (2023). “Smart investigation of artificial intelligence in renewable energy system technologies by natural language processing: Insightful pattern for decision-makers”. *Engineering Applications of Artificial Intelligence*. **126**. <https://doi.org/10.1016/j.engappai.2023.106848>
- Ogunleye, B., Maswera, T., Hirsch, L., Gaudoin, J., Brunsdon, T. (2023). “Comparison of Topic Modelling Approaches in the Banking Context”. *Applied Sciences*. **13(2)**. <https://doi.org/10.3390/app13020797>
- Palacios, C. A., Reyes-Suárez, J. A., Bearzotti, L. A., Leiva, V., Marchant, C. (2021). “Knowledge discovery for higher education student retention based on data mining: Machine learning algorithms and case study in chile”. *Entropy*. **23(4)**:485. <https://doi.org/10.3390/e23040485>
- Prakash, R., Anoop, V. S., Asharaf, S. (2022). “Blockchain technology for cybersecurity: A text mining literature analysis”. *International Journal of Information Management Data Insights*. **2(2)**:100112. <https://doi.org/10.1016/j.jjime.2022.100112>
- Qin, H., Zeng, J., Ma, X. (2021). “Trend Analysis of Research Direction in Computer Science Based on Microsoft Academic Graph”. *ACM International Conference Proceeding Series*. **PartF**:168982. <https://doi.org/10.1145/3448734.3450470>
- Samsir, Saragih, R. S., Subagio, S., Aditya, R., Watrianthos, R. (2023). “BERTopic Modeling of Natural Language Processing Abstracts: Thematic Structure and Trajectory”. *JURNAL MEDIA INFORMATIKA BUDIDARMA*. **7(3)**:1514–1520. <https://doi.org/10.30865/mib.v7i3.6426>
- Scopus. (n.d.). Scopus. Retrieved August 21, 2024, from <https://www.scopus.com/>

- Sharma, C., Batra, I., Sharma, S., Malik, A., Sanwar Hosen, A. S. M., Ra, I. H. (2022). “Predicting Trends and Research Patterns of Smart Cities: A Semi-Automatic Review Using Latent Dirichlet Allocation (LDA)”. *IEEE Access*. **10**:121080–121095. <https://doi.org/10.1109/ACCESS.2022.3214310>
- Shu, X., Ye, Y. (2023). “Knowledge Discovery: Methods from data mining and machine learning”. *Social Science Research*. **110**:102817. <https://doi.org/10.1016/j.ssresearch.2022.102817>
- Siino, M., Tinnirello, I., La Cascia, M. (2024). “Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers”. *Information Systems*. **121**:102342. <https://doi.org/10.1016/j.is.2023.102342>
- Stewart, G., Al-Khassaweneh, M. (2022). “An Implementation of the HDBSCAN\* Clustering Algorithm”. *Applied Sciences (Switzerland)*. **12(5)**:2405. <https://doi.org/10.3390/app12052405>
- Takacs, V., O’Brien, C. D. (2023). “Trends and gaps in biodiversity and ecosystem services research: A text mining approach”. *Ambio*. **52(1)**:81–94. <https://doi.org/10.1007/s13280-022-01776-2>
- Wang, C., Nulty, P., Lillis, D. (2020). “A Comparative Study on Word Embeddings in Deep Learning for Text Classification”. *ACM International Conference Proceeding Series*. 37–46. <https://doi.org/10.1145/3443279.3443304>
- Wang, Y., Bashar, M. A., Chandramohan, M., Nayak, R. (2023). “Exploring topic models to discern cyber threats on Twitter: A case study on Log4Shell”. *Intelligent Systems with Applications*. **20**:200280. <https://doi.org/10.1016/j.iswa.2023.200280>
- Widyaningsih, T. W., Dewi, M. A., Andrianingsih, A. (2021). “Analisis Bibliometrik untuk Memetakan Tren Penelitian Covid-19 dalam Topik Ilmu Komputer”. *Techno. Com*. **20(3)**:440–454. <https://doi.org/10.33633/tc.v20i3.4593>

Yu, D., Fang, A., Xu, Z. (2023). "Topic research in fuzzy domain: Based on LDA topic modelling". *Information Sciences*. **648**:119600. <https://doi.org/10.1016/j.ins.2023.119600>

Yukselen, M., Mutlu, A., Karagoz, P. (2022). "Predicting the Trending Research Topics by Deep Neural Network based Content Analysis". *IEEE Access*. **10**:90887-90902. <https://doi.org/10.1109/ACCESS.2022.3202654>