

DAFTAR PUSTAKA

- [1] T. Ko, H. D. Jeong, and J. Lee, "Natural Language Processing–Driven Similar Project Determination Using Project Scope Statements," *Journal of Management in Engineering*, vol. 39, no. 3, May 2023, doi: 10.1061/jmenea.meeng-5229.
- [2] S. Marwaha, R. C. Goyal, P. K. Malhotra, and A. Arora, "Decision Support System for research project duplication detection," *International Conference Hybrid Intelligent Systems*, pp. 518–524, Dec. 2012, doi: 10.1109/HIS.2012.6421388.
- [3] J. Motger de la Encarnación, C. Palomares Bonache, and J. Marco Gómez, "RESim - Automated Detection of Duplicated Requirements in Software Engineering Projects," *Requirements Engineering: Foundation for Software Quality*, pp. 1–6, Jan. 2020, [Online]. Available: <http://ceur-ws.org/Vol-2584/PT-paper1.pdf>
- [4] A. Modi, Y. S. Dhanjal, and A. Larhgotra, "Semantic similarity for text comparison between textual documents or sentences," Dec. 2023. doi: 10.1109/icses60034.2023.10465440.
- [5] A. Bernt and P. Gupta, "Semantic textual similarity system," Dec. 2019.
- [6] A. Bali, A. Bhagwat, A. Bhise, and S. Joshi, "Semantic similarity detection and analysis for text documents," Feb. 2024. doi: 10.1109/ic-etite58242.2024.10493834.
- [7] M. J. Procopio, G. Cobb, and D. Lonnon, "System and method for computation of document similarity," Jan. 2012. [Online]. Available: <https://patents.google.com/patent/US8886648B1/en>
- [8] X. Xiang, Z. Zhong, Y. Chen, and Z. Zhang, "Document similarity detection based on multi-feature semantic fusion and concept graph," in *Proc. IEEE Int. Conf. Electronic Technology, Communication and Information (ICETCI)*, pp. 469–472, May 2022. doi: 10.1109/icetci55101.2022.9832170.
- [9] I. Pfennigstorf, "Webscrapping with Symfony," doi: 10.59350/81wrn-nyh38.

- [10] M. I. Firdian, E. Darwiyanto, and M. Adrian, "Web scraping with HTML DOM method for website news API creation," *JlPI*, vol. 7, no. 4, pp. 1211–1219, 2022. doi: 10.29100/jipi.v7i4.3235.
- [11] H. Phan, *Building Application Powered by Web Scraping*, 2019. [Online]. Available: <https://www.theseus.fi/handle/10024/166489>
- [12] W. Ma, X. Chen, and W. Shang, "Advanced deep web crawler based on DOM," *Computational Sciences and Optimization*, pp. 605–609, Jun. 2012. doi: 10.1109/CSO.2012.138.
- [13] K. J. Millman and M. Aivazis, "Python for scientists and engineers," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 9–12, 2011. doi: 10.1109/MCSE.2011.36.
- [14] T. E. Oliphant, *A Guide to NumPy*, Trelgol Publishing, 2006.
- [15] W. McKinney, *Python for Data Analysis*, O'Reilly Media, 2017.
- [16] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*
- [17] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, O'Reilly Media, 2009.
- [18] M. Honnibal and I. Montani, "spaCy: Industrial-strength natural language processing in Python." [Online]. Available: <https://spacy.io/>
- [19] M. Grinberg, *Flask Web Development*, O'Reilly Media, 2014.
- [20] Y. Sankarasubramaniam, B. N. Ranganathan, and S. Ramani, "Processing of printed documents," Jan. 2009. [Online]. Available: <https://patents.google.com/patent/US8238665B2/en>
- [21] Y. Koda and T. Tanaka, "Document processing system," Aug. 1995. [Online]. Available: <https://patents.google.com/patent/US5749083A/en>
- [22] D. Wilson, "System and method for document processing," Mar. 2014. [Online]. Available: <https://patents.google.com/patent/US20150256493A1/en>
- [23] M. Anwar, *Digital Document Processing*, 2001. [Online]. <https://patents.google.com/patent/WO2001079984A1/en>

- [24] H. Saeidnia, "Gemini is multimodal," Social Science Research Network. doi: 10.2139/ssrn.4703935.
- [25] A. Nazarius, F. Saputra, N. N. Kamala Sari, and V. H. Pranatawijaya, "Penerapan Gemini AI dalam pembuatan deskripsi produk e-commerce," JATI, vol. 8, no. 3, pp. 3721–3725, 2024. doi: 10.36040/jati.v8i3.9670.
- [26] M. Qudisa and M. M. Fraz, Optimizing Article Screening and Information Extraction: A Hybrid Approach with GeminiAI and Vector Database, 2024, pp. 1–8. doi: 10.1109/icodt262145.2024.10740208.
- [27] R. M. Handoko et al., "Implementasi Gemini AI dalam pengembangan aplikasi e-commerce mobile generate deskripsi produk," Protekinfo, vol. 11, no. 1, pp. 21–25, 2024. doi: 10.30656/protekinfo.v11i1.8595.
- [28] D. Archana et al., "Image text detection and documentation using OCR," pp. 410–414, Jun. 2024. doi: 10.1109/icsseecc61126.2024.10649443.
- [29] P. Sisodia, "Optical character recognition development using Python," J. Inform. Electr. Electron. Eng. (JIEEE), vol. 4, no. 3, pp. 1–13, Jan. 2023. doi: 10.54060/jieee.2023.75.
- [30] K. Joshi, "Study of Tesseract OCR," vol. 1, no. 2, pp. 41–50, Mar. 2024. doi: 10.69974/glskalp.01.02.54.
- [31] C. A. Lestari et al., "Penggunaan kamus singkatan kata Bahasa Indonesia sehari-hari dalam pembangkitan fitur teks," J. Informatika Univ. Pamulang, vol. 8, no. 2, pp. 121–128, Jun. 2023. doi: 10.32493/informatika.v8i2.29306.
- [32] I. A. Rahma and L. H. Suadaa, "Penerapan text augmentation untuk mengatasi data yang tidak seimbang pada klasifikasi teks berbahasa Indonesia," J. Teknol. Informasi dan Ilmu Komputer, 2023. doi: 10.25126/jtiik.1067325.
- [33] D. H. Putra and N. A. A. N. Syam, "Automasi pembangkit lirik lagu dalam Bahasa Indonesia: implementasi GPT-Neo dalam pemrosesan bahasa kreatif," Syntax, vol. 5, no. 2, pp. 612–616, 2024. doi: 10.46576/syntax.v5i2.5542.
- [34] N. N. Widyastuti, A. Bijaksana, and I. L. Sardi, "Analisis Word2Vec untuk perhitungan kesamaan semantik antar kata," vol. 5, no. 3, Dec. 2018.
- [35] —, "Improving MiniBERT's semantic performance with semantic-rich sentence embeddings."

- [36] S. Zhao, Q. Gu, J. Zhang, G. Song, and Q. Lu, "SMGC-SBERT: A multi-feature fusion Chinese short text similarity computation model based on optimised SBERT," pp. 3091–3096, 2024. doi: 10.1109/smc54092.2024.10831117.
- [37] H. Bekamiri, D. S. Hain, and R. Jurowetzki, "PatentSBERTa: A deep NLP based hybrid model for patent distance and classification using augmented SBERT," arXiv: Learning, Mar. 2021. [Online]. Available: <https://ui.adsabs.harvard.edu/abs/2021arXiv210311933B/abstract>
- [38] Z. Ali et al., Fine-tuned training method for semantic text similarity measurement using SBERT, Bi-LSTM and attention network, 2023. doi: 10.1109/mvipit60427.2023.00028.
- [39] M. Fateen and T. Mine, Using similarity learning with SBERT to optimize teacher report embeddings for academic performance prediction, pp. 720–726, 2023. doi: 10.1007/978-3-031-36336-8_111.
- [40] M. S. Rahman et al., "Efficient medical image retrieval using DenseNet and FAISS for BIRADS classification," 2024. doi: 10.48550/arxiv.2411.01473.
- [41] E. Öztürk and A. Mesut, "Performance analysis of chroma, qdrant, and FAISS databases," Meždunarodna Naučna Konferenciâ Uniteh, 2024. doi: 10.70456/tbrn3643.
- [42] A. Ronacher, Flask documentation, 2024, [Online]. Retrieved from flask.palletsprojects.com
- [43] Pallets Projects, Werkzeug documentation, [Online]. Retrieved from <https://werkzeug.palletsprojects.com>
- [44] T. Otwell, Laravel: Up and Running, O'Reilly Media, 2019.
- [45] Laravel.com, Laravel documentation, 2024, [Online]. Retrieved from <https://laravel.com/docs>
- [46] Laravel.com, Eloquent ORM. [Online]. Retrieved from <https://laravel.com/docs/eloquent>
- [47] Laravel.com, Blade Templates. [Online]. Retrieved from <https://laravel.com/docs>
- [48] F. Zaninotto, "PHP framework benchmark," PHPFrameworks.com, 2024.
- [49] DigitalOcean Tutorials, "RESTful APIs with Laravel." [Online]. Retrieved from <https://www.digitalocean.com/community/tutorials>

- [50] MySQL.com, MySQL documentation, 2024. [Online]. Retrieved from <https://dev.mysql.com/doc/>
- [51] DigitalOcean, "MySQL performance optimization." [Online]. Retrieved from <https://www.digitalocean.com/community/tutorials>
- [52] P. DuBois, MySQL Cookbook, O'Reilly Media, 2013.
- [53] MySQL.com, "Replication in MySQL."
- [54] R. Copeland, Essential SQLAlchemy, 2008.
- [55] J. W. Krogh, MySQL Connector/Python Revealed. doi: 10.1007/978-1-4842-3694-9.
- [56] A. Lukaszewski, MySQL for Python, 2010.
- [57] L. Hakim, "Discover Idea: Aplikasi Berbasis AI untuk Membantu Mahasiswa Menemukan Judul PA", I. Azimi, Ed. Bandung: Universitas Telkom, D3 Rekayasa Perangkat Lunak Aplikasi, 2024.