DAFTAR PUSTAKA

Afifah, Lutfia. 2020. "Evaluasi Model Machine Learning: Train/Test Split." *IlmudataPy* (blog). 20 Oktober 2020. https://ilmudatapy.com/evaluasi-model-machine-learning-dengan-train-test-split/.

Afinda, Angel Metanosa. 2024. "Overfitting vs. Underfitting, Apa Bedanya?" Dicoding Blog. 12 November 2024. https://www.dicoding.com/blog/overfitting-vs-underfitting-apa-bedanya/.

Airbyte. 2024. "What Are Word & Sentence Embedding? 5 Applications | Airbyte." 28 Agustus 2024. https://airbyte.com/data-engineering-resources/sentence-word-embeddings.

Algoritma Data Science School. 2022. "Pahami Data Dengan EDA (Exploratory Data Analysis)." *Algoritma* (blog). 1 Maret 2022. https://algorit.ma/blog/exploratory-data-analysis-2022/.

Anthony. 2023. "Perpustakaan Digital - Digilib ITB - Digital Library." 15 Agustus 2023. https://digilib.itb.ac.id/gdl/view/76390.

AWS. 2024. "Apa itu Visualisasi Data? - Penjelasan tentang Visualisasi Data - AWS." Amazon Web Services, Inc. 2024. https://aws.amazon.com/id/what-is/data-visualization/.

Barbella, Marcello, dan Genoveffa Tortora. 2022. "Rouge Metric Evaluation for Text Summarization Techniques." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.4120317.

Bopaiah, Yajna. 2023. "Unveiling the Power of ROUGE Metrics in NLP." Medium. 18 Agustus 2023. https://pub.aimind.so/unveiling-the-power-of-rouge-metrics-in-nlp-b6d3f96d3363.

BPJIID, Universitas Medan Area. 2024. "Hyperparameter: Memahami Dan Mengoptimalkan Kinerja Model Pembelajaran Mesin." 30 Agustus 2024. https://bpjiid.uma.ac.id/2024/08/30/hyperparameter-memahami-dan-mengoptimalkan-kinerja-model-pembelajaran-mesin/.

Brownlee, Jason. 2022. "Difference Between a Batch and an Epoch in a Neural Network." *MachineLearningMastery.Com* (blog). 9 Agustus 2022. https://www.machinelearningmastery.com/difference-between-a-batch-and-an-epoch/.

Casmir, Irene. 2023. "Regularization by Early Stopping." GeeksforGeeks. 21 September 2023. https://www.geeksforgeeks.org/regularization-by-early-stopping/.

Chung, Hyung Won, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, dkk. 2022. "Scaling Instruction-Finetuned Language Models." arXiv. http://arxiv.org/abs/2210.11416.

Coding, Rumah. 2024. "Sentiment Analysis Menggunakan Deep Learning - Bagian 2." 10 Juni 2024. https://rumahcoding.id/blog/sentiment-analysis-menggunakan-deep-learning-bagian-2/.

dewanpers. 2024. "SURVEI LANSKAP MEDIA PERS INDONESIA." 12 Juni 2024. https://dewanpers.or.id.

Eryana, Asep, Sriyanto, dan Firmansyah. 2023. "Prediksi Malaria Menggunakan Metode Pre-Trained Model Algoritma EfficientNet-B0 dan MobileNet-

V2." *Jurnal Ilmiah Komputasi* 22, no. 1 (Maret). https://doi.org/10.32409/jikstik.22.1.3332.

Fabbri, Alexander R., Wojciech Kryściński, Bryan McCann, Caiming Xiong, Richard Socher, dan Dragomir Radev. 2021. "SummEval: Re-Evaluating Summarization Evaluation." *Transactions of the Association for Computational Linguistics* 9, no. April (April): 391–409. https://doi.org/10.1162/tacl_a_00373.

Feby, Dita. 2023. "Konsep Model Machine Learning K-Means Clustering." 20 September 2023. https://dqlab.id/konsep-model-machine-learning-k-means-clustering.

Fu, Xue-Yong, Md Tahmid Rahman Laskar, Elena Khasanova, Cheng Chen, dan Shashi Bhushan TN. 2024. "Tiny Titans: Can Smaller Large Language Models Punch Above Their Weight in the Real World for Meeting Summarization?" arXiv. http://arxiv.org/abs/2402.00841.

Fuadi, Mukhlish, Adhi Dharma Wibawa, dan Surya Sumpeno. 2023. "idT5: Indonesian Version of Multilingual T5 Transformer." https://arxiv.org/pdf/2302.00856.

Geeksforgeeks. 2023. "Learning Rate in Neural Network." GeeksforGeeks. 31 Juli 2023. https://www.geeksforgeeks.org/impact-of-learning-rate-on-a-model/.

geeksforgeeks. 2024. "Dropout vs Weight Decay." GeeksforGeeks. 10 Februari 2024. https://www.geeksforgeeks.org/dropout-vs-weight-decay/.

Gupta, Akanksha. 2020. "Understanding Text Summarization Using K-Means Clustering." *Medium* (blog). 29 September 2020. https://medium.com/@akankshagupta371/understanding-text-summarization-using-k-means-clustering-6487d5d37255.

Gupta, Dishashree. 2017. "Transfer Learning and the Art of Using Pre-Trained Models in Deep Learning." *Analytics Vidhya* (blog). 1 Juni 2017. https://www.analyticsvidhya.com/blog/2017/06/transfer-learning-the-art-of-fine-tuning-a-pre-trained-model/.

Hadwirianto, Muhammad Raihan, Faqih Hamami, dan Oktariani Nurul Pratiwi. 2024. "Extractive Text Summarization Terhadap Artikel Berita Indonesia Berbasis Machine Learning," Agustus (Agustus). https://openlibrary.telkomuniversity.ac.id/pustaka/200930/extractive-text-summarization-terhadap-artikel-berita-indonesia-berbasis-machine-learning.html.

Hakim, RB Fajriya. 2024. "Word Embedding." *Medium* (blog). 13 April 2024. https://medium.com/@986110101/word-embedding-2e969566fe11.

Halim, Franky, Kartika Gunadi, dan Liliania Liliana. 2022. "Ringkasan Ekstraktif Otomatis pada Berita Berbahasa Indonesia Menggunakan Metode BERT." https://publication.petra.ac.id/index.php/teknik-informatika/article/download/12036/10570.

Hans, Rizal. 2024. "Konsep Learning Rate pada Algoritma Machine Learning." 2024. https://dqlab.id/konsep-learning-rate-pada-algoritma-machine-learning.

Hugging Face. 2020a. “AutoModels — transformers 3.0.2 documentation.” 2020. https://huggingface.co/transformers/v3.0.2/model_doc/auto.html.
———. 2020b. “Glossary — transformers 3.5.0 documentation.” 2020. https://huggingface.co/transformers/v3.5.1/glossary.html.
———. 2024a. “Padding and truncation.” 2024. https://huggingface.co/docs/transformers/pad_truncation.
———. 2024b. “SEACrowd/indosum · Datasets at Hugging Face.” 2024. https://huggingface.co/datasets/SEACrowd/indosum.
IBM. 2024. “Apa itu peringkasan teks? | IBM.” 6 November 2024. https://www.ibm.com/id-id/topics/text-summarization.
Iskandar, Muhammad Iqbal. 2024. “Natural Language Processing: Cara AI Memahami Bahasa Manusia.” Phintraco Technology. 26 April 2024. https://phintraco-tech.com/blog/natural-language-processing/.
Jeremy, Albert. 2024. “Understanding AI Models: How Size Affects Their Ability to Handle Text - A Case Study from Llama 3.1 and FLAN-T5 Models.” 18 September 2024. https://www.linkedin.com/pulse/understanding-ai-models-how-size-affects-ability-handle-albert-jeremy-zabcc.
Kashyap, Piyush. 2024. “Understanding Dropout in Deep Learning: A Guide to Reducing Overfitting.” *Medium* (blog). 30 Oktober 2024. https://medium.com/@piyushkashyap045/understanding-dropout-in-deep-learning-a-guide-to-reducing-overfitting-26cbb68d5575.
Kementerian Komunikasi dan Informatika. 2024. “Pengguna Internet Meningkat, Kominfo Galang Kolaborasi Tingkatkan Kualitas Layanan.” 31 Januari 2024. https://www.kominfo.go.id/content/detail/54481/siaran-pers-no-80hmkominfo012024-tentang-pengguna-internet-meningkat-kominfo-galang-kolaborasi-tingkatkan-kualitas-layanan/0/siaran_pers.
Khasanah, Aini Nur, dan Mardhiya Hayaty. 2023. “ABSTRACTIVE-BASED AUTOMATIC TEXT SUMMARIZATION ON INDONESIAN NEWS USING GPT-2.” *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)* 10, no. 1 (Desember): 9–18. https://doi.org/10.33330/jurteksi.v10i1.2492.
Kirmani, Mahira, Gagandeep Kaur, dan Mudasir Mohd. 2024. “Analysis of Abstractive and Extractive Summarization Methods.” *International Journal of Emerging Technologies in Learning (iJET)* 19, no. 01 (Januari): 86–96. https://doi.org/10.3991/ijet.v19i01.46079.
Kolena. 2024. “BERTScore - Testing with Kolena.” 15 Juli 2024. https://docs.kolena.com/metrics/bertscore/.
Krishna, Kundan, Prakhar Gupta, Sanjana Ramprasad, Byron C. Wallace, Jeffrey P. Bigham, dan Zachary C. Lipton. 2023. “USB: A Unified Summarization Benchmark Across Tasks and Domains.” arXiv. http://arxiv.org/abs/2305.14296.
Lin, Chin-Yew. 2004. “ROUGE: A Package for Automatic Evaluation of Summaries.” Dalam *Text Summarization Branches Out*, 74–81. Barcelona, Spain: Association for Computational Linguistics. https://aclanthology.org/W04-1013.
Liu, Yang, dan Mirella Lapata. 2019. “Text Summarization with Pretrained Encoders.” arXiv. http://arxiv.org/abs/1908.08345.

Lumbansiantar, Sarima, Saruni Dwiasnata, dan Nenden Siti Fatonah. 2023. "Penerapan Metode Cosine Similarity Dalam Mendeteksi Plagiarisme Pada Jurnal | Lumbansiantar | Format : Jurnal Ilmiah Teknik Informatika." 2023. https://publikasi.mercubuana.ac.id/index.php/format/article/view/18437/pdf.

Mahajaya, Nyoman Sarasuartha, Putu Desiana Wulaning Ayu, dan Roy Rudolf Huizen. 2024. "Pengaruh Optimizer Adam, AdamW, SGD, Dan LAMB Terhadap Model Vision Transformer Pada Klasifikasi Penyakit Paru-Paru." *Seminar Hasil Penelitian Informatika Dan Komputer (SPINTER) | Institut Teknologi Dan Bisnis STIKOM Bali*, Juni (Juni), 818–23.

Mandia, Nyoman. 2017. "View of Kohesi dan Koherensi sebagai Dasar Pembentukan Wacana yang Utuh." 2 Juli 2017. https://ojs.pnb.ac.id/index.php/SOSHUM/article/view/531/492.

Mashudi, Irsyad, Sofyan Arief, Deasy Sandhya, dan Triana Fatmawati. 2023. "(PDF) JURNAL MEDIA INFORMATIKA BUDIDARMA Klasterisasi Jawaban Uraian Mahasiswa Menggunakan TF-IDF Dan K-Means Untuk Membantu Koreksi Ujian." *ResearchGate*, Oktober (Oktober). https://doi.org/10.30865/mib.v7i4.6688.

MYSCALE. 2024. "Enhancing NLP with Cosine Similarity: Text Summarization Insights." 6 Juni 2024. https://myscale.com/blog/transforming-nlp-text-summarization-cosine-similarity/.

NLP Group, Stanza. 2020. "Tokenization & Sentence Segmentation." Stanza. 2020. https://stanfordnlp.github.io/stanza/tokenize.html.

Nugroho, Kuncahyo Setyo. 2020. "Dasar Text Preprocessing Dengan Python." *Medium* (blog). 4 Juni 2020. https://ksnugroho.medium.com/dasar-text-preprocessing-dengan-python-a4fa52608ffe.

Papers With Code. 2024. "Papers with Code - Flan-T5 Explained." 2024. https://paperswithcode.com/method/flan-t5.

Petrus, Johannes, Ermatita Ermatita, Sukemi Sukemi, dan Erwin Erwin. 2023. "An Adaptable Sentence Segmentation Based on Indonesian Rules." *IAES International Journal of Artificial Intelligence (IJ-AI)* 12, no. 3 (September): 1491–99. https://doi.org/10.11591/ijai.v12.i3.pp1491-1499.

Poclitari, Radu. 2024. "5 Best NLP Algorithms to Get Document Similarity." 15 Februari 2024. https://www.index.dev/blog/best-nlp-algorithms-to-get-document-similarity.

ProjectPro. 2024. "17 Best Open-Source LLMs Data Scientists Must Know in 2023." ProjectPro. 21 Maret 2024. https://www.projectpro.io/article/open-source-llm-models/879.

Putra, Kharisma Dharma. 2024. "Apa Itu Random State Di Machine Learning Dan Cara Kerjanya." *Bengkel TI* (blog). 25 Maret 2024. https://www.bengkelti.com/blog/apa-itu-random-state-di-machine-learning-dan-cara-kerjanya/.

Putri, Roza Yuniza, Widodo, dan Hamidillah Ajie. 2023. "PERBANDINGAN PERINGKASAN MULTI DOKUMEN ILMIAH BERBAHASA INDONESIA MENGGUNAKAN METODE K-MEANS DAN K-

NEAREST NEIGHBORS (K-NN)." *PINTER : Jurnal Pendidikan Teknik Informatika Dan Komputer* 7, no. 2 (Desember): 19–27. https://doi.org/10.21009/pinter.7.2.3.

Python.org. 2024. "Re — Regular Expression Operations." Python Documentation. 2024. https://docs.python.org/3/library/re.html.

Qi, Peng, Yuhao Zhang, Yuhui Zhang, Jason Bolton, dan Christopher D. Manning. 2020. "Stanza: A Python Natural Language Processing Toolkit for Many Human Languages." arXiv. https://doi.org/10.48550/arXiv.2003.07082.

Raffel, Colin, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, dan Peter J. Liu. 2023. "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer." arXiv. http://arxiv.org/abs/1910.10683.

Ranggianto, Narandha Arya. 2023. "Peringkasan Teks dengan Pendekatan Abstraktif dan Ekstraktif Pada Ulasan Perjalanan Berbahasa Indonesia." Masters, Institut Teknologi Sepuluh Nopember. https://repository.its.ac.id/101729/.

RevoU. 2024. "Batch Processing." RevoU. 2024. https://www.revou.co/kosakata/batch-processing.

Rully Widiastutik, Lukman Zaman P. C. S. W, dan Joan Santoso. 2019. "Peringkasan Teks Ekstraktif pada Dokumen Tunggal Menggunakan Metode Restricted Boltzmann Machine." *Journal of Intelligent System and Computation* 1, no. 2 (Desember): 58–64. https://doi.org/10.52985/insyst.v1i2.84.

Sahagian, Graham. 2020. "What Is Random State 42?" *Medium* (blog). 24 Desember 2020. https://grsahagian.medium.com/what-is-random-state-42-d803402ee76b.

Samosir, Feliks Victor Parningotan, Hapnes Toba, dan Mewati Ayub. 2022. "BESKlus : BERT Extractive Summarization with K-Means Clustering in Scientific Paper." *Jurnal Teknik Informatika dan Sistem Informasi* 8, no. 1 (April). https://doi.org/10.28932/jutisi.v8i1.4474.

Sari, Yunita Maulidia, dan Nenden Siti Fatonah. 2021. "Peringkasan Teks Otomatis pada Modul Pembelajaran Berbahasa Indonesia Menggunakan Metode Cross Latent Semantic Analysis (CLSA)." *Jurnal Edukasi dan Penelitian Informatika (JEPIN)* 7, no. 2 (Agustus): 153. https://doi.org/10.26418/jp.v7i2.47768.

SBert.Net. 2024. "Training Overview — Sentence Transformers documentation." 2024. https://www.sbert.net/docs/sentence_transformer/training_overview.html.

Shivanandhan, Manish M. 2024. "Understanding AutoTokenizer in Huggingface Transformers." 19 Juni 2024. https://www.linkedin.com/pulse/understanding-autotokenizer-huggingface-transformers-m-shivanandhan-zfbsc.

shubham.singh. 2019. "NLP Essentials: Removing Stopwords and Performing Text Normalization Using NLTK and spaCy in Python." *Analytics Vidhya* (blog). 21 Agustus 2019.

https://www.analyticsvidhya.com/blog/2019/08/how-to-remove-stopwords-text-normalization-nltk-spacy-gensim-python/.

Simanjuntak, Anugerah, Rosni Lumbantoruan, Kartika Sianipar, Rut Gultom, Mario Simaremare, Samuel Situmeang, dan Erwin Panggabean. 2023. "Studi dan Analisis Hyperparameter Tuning IndoBERT Dalam Pendeteksian Berita Palsu." 7 November 2023. https://www.scribd.com/document/814794865/8532-Article-Text-56641-1-10-20240228.

sukardi. 2022. "Materi, Soal, Dan Pembahasan - Ruang Vektor Umum — Mathcyber1997." 17 April 2022. https://mathcyber1997.com/materi-soal-dan-pembahasan-ruang-vektor-umum/.

Tensorflow. 2024. "Transfer Learning and Fine-Tuning | TensorFlow Core." TensorFlow. 16 Agustus 2024. https://www.tensorflow.org/tutorials/images/transfer_learning.

Thony. 2023. "thonyyy/pegasus_indonesian_base-finetune · Hugging Face." 18 April 2023. https://huggingface.co/thonyyy/pegasus_indonesian_base-finetune.

Trivusi. 2023. "Beam Search: Pencarian Optimal pada Ruang Pencarian yang Besar." 3 April 2023. https://www.trivusi.web.id/2023/03/algoritma-beam-search.html.

Tsirmpas, Dimitrios, Ioannis Gkionis, Georgios Th Papadopoulos, dan Ioannis Mademlis. 2024. "Neural Natural Language Processing for Long Texts: A Survey on Classification and Summarization." *Engineering Applications of Artificial Intelligence* 133, no. Juli (Juli): 108231. https://doi.org/10.1016/j.engappai.2024.108231.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, dan Illia Polosukhin. 2023. "Attention Is All You Need." arXiv. https://doi.org/10.48550/arXiv.1706.03762.

Yassin, Ahmed. 2024. "Adam vs. AdamW: Understanding Weight Decay and Its Impact on Model Performance." *Medium* (blog). 8 November 2024. https://yassin01.medium.com/adam-vs-adamw-understanding-weight-decay-and-its-impact-on-model-performance-b7414f0af8a1.

Zhang, Tianyi, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, dan Yoav Artzi. 2020. "BERTScore: Evaluating Text Generation with BERT." arXiv. https://doi.org/10.48550/arXiv.1904.09675.