

DAFTAR PUSTAKA

- Agarwal, R. (2020) Understanding Transformers, the Data Science Way. Tersedia pada: <https://mlwhiz.com/blog/2020/09/20/transformers/> (Diakses: 7 Juli 2023).
- Albared, M., Omar, N., Aziz, M. J., Nazri, M. Z. A., 2010 Automatic part of speech tagging for arabic: An experiment using bigram hidden markov model, *International Conference on Rough Sets and Knowledge Technology*, hal.. 361–370, vol. 6401, doi: 10.1007/978-3-642-16248-0_52. Springer, Berlin, Heidelberg.
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Esesn, B. C. Van., Awwal, A. A. S., Asari, V. K., 2018, The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. *ArXiv Cornell University*, hal. 1-39, doi: 10.48550/arXiv.1803.01164.
- Bahasa, B. P. P., 2016, Kamus Besar Bahasa Indonesia Daring. <https://kbbi.kemdikbud.go.id/entri/ambigu>, diakses tanggal 08 Juli 2021.
- Bahasa, B. P. P., 2016, Kamus Besar Bahasa Indonesia Daring. <https://kbbi.kemdikbud.go.id/entri/homonim>, diakses tanggal 12 Januari 2023.
- Banten, K. B. P., 2017 Homonim dan Polisemi, <https://kantorbahasabanten.kemdikbud.go.id/?p=579>, diakses tanggal 13 Februari 2023.
- Candra, A., 2021, Pemodelan dan Implementasi BERT untuk Mendeteksi Kekerasan pada Teks Media Sosial Berbahasa Indonesia, *Skripsi*, Program Studi Sistem Informasi, Fakultas Teknik Informatika, Universitas Multimedia Nusantara, Tangerang.
- Chaer, A., 2009, *Sintaksis bahasa Indonesia*, Rineka Cipta, Jakarta.
- Cucus, A., Endra, R. Y. and Naralita, T., 2019, Chatter Bot Untuk Konsultasi Akademik Di Perguruan Tinggi, *Jurnal Sistem Informasi dan Telematika (Telekomunikasi, Multimedia, dan Informasi)*, hal. 20–25, vol. 10, No. 1.
- Devlin, J., Chang, M. W., Lee, K., Toutanova, K., 2019, BERT: Pre-training of deep bidirectional transformers for language understanding, *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, hal. 4171–4186.
- Dhumal, R. dan Kiwelekar, A., 2020, Deep Learning Techniques for Part of Speech Tagging by Natural Language Processing, *Proceedings of the Second International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020)*. IEEE, hal. 76–81.
- Dinakaramani, A., Rashel, F., Luthfi, A., Manurung, R., 2014, Designing an Indonesian part of speech tagset and manually tagged Indonesian corpus,

Proceedings of the International Conference on Asian Language Processing 2014, IALP 2014. IEEE, doi: 10.1109/IALP.2014.6973519, hal. 66–69.

- Espitia, F. A. dan Soto, L. R., 2020, *Novel Methods Based on Deep Learning Applied to Condition Monitoring in Smart Manufacturing Processes*, Martinez L. R., Rios, R. A. O., Prieto, M. D., *New Trends in the Use of Artificial Intelligence for the Industry 4.0.*, Intechopen, London, United Kingdom.
- Exxact (2021) *BERT Transformers – How Do They Work?* Tersedia pada: <https://www.exxactcorp.com/blog/Deep-Learning/how-do-bert-transformers-work> (Diakses: 5 Juli 2023).
- Fimoza, D., 2021, Analisis Sentimen Terhadap Film Indonesia dengan Pendekatan Bert. *Skripsi*, Departemen Ilmu Komputer, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara, Medan.
- Fosler-lussier, E., 1998, Markov Models and Hidden Markov Models: A Brief Tutorial, *ICSI Technical Report TR-98-041*, hal. 1-7.
- Fu, S., Lin, N., Zhu, G., Jiang., 2018, Towards Indonesian Part-of-Speech Tagging : Corpus and Models, *Proceedings of LREC 2018 Workshop on Belt and Road LRE*, 1, hal. 2-7.
- Ghosh, I., 2020, Ranked: The 100 Most Spoken Languages Around the World, Visual Capitalist, <https://www.visualcapitalist.com/100-most-spoken-languages/>, diakses tanggal 4 Desember 2022.
- Hadiyan, K. P., Bijaksana, M. A., Romadhony, A., 2021, Deteksi Penggunaan Kalimat Abusive Pada Teks Bahasa Indonesia Menggunakan Metode IndoBERT, *e-Proceeding of Engineering*, hal. 3028–3038, vol.8, no. 2.
- Halim, F., Liliana, L. and Gunadi, K., 2022, Ringkasan Ekstraktif Otomatis pada Berita Berbahasa Indonesia Menggunakan Metode BERT, *Jurnal Infra*, hal. 162-168, vol. 10, no.1.
- Han, J., Kamber, M., Pei, J., 2012, *Data Mining: Concept and Techniques*. Edisi ketiga, Elsevier, Amerika Serikat.
- Idris, A., 2018, Confusion Matrix.: <https://medium.com/@awabmohammedome/confusion-matrix-b504b8f8e1d1>, diakses tanggal 20 Juli 2020.
- Joshi, N., Darbari, H., Mathur, I., 2013, HMM Based POS Tagger for Hindi, *Third International Conference on Computer Science & Information Technology*, doi: 10.5121/csit.2013.3639, hal. 341–349.
- Kamayani, M., 2019, Perkembangan Part-of-Speech Tagger Bahasa Indonesia, *Jurnal Linguistik Komputasional (JLK)*, doi: 10.26418/jlk.v2i2.20, hal. 34-48, vol. 2, no. 2.
- Kikaben, 2021, Transformer's Encoder-Decoder: Let's Understand The Model Architectur. <https://kikaben.com/transformers-encoder-decoder/> diakses tanggal 5 November 2022.
- Koto, F., Rahimi, A., Lau, J. H., Baldwin, T., 2020, IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP',

- Proceedings of the 28th International Conference on Computational Linguistics*. Barcelona, Spain (Online): International Committee on Computational Linguistics, doi: 10.18653/v1/2020.coling-main.66, hal. 757–770.
- Lestari, A., 2016, Ambiguitas dalam Judul Berita pada Surat Kabar Harian Kompas Dan Pos Kota, *Skripsi*, Program Sarjana Sastra Indonesia, Universitas Negeri Jakarta, Jakarta.
- Lubis, P. H. H., 1993, *Jengala Bahasa Indonesia*, Angkasa, Bandung.
- Marliana, N. L. (2004) *Kamus Polisemi Bahasa Indonesia*. Yogyakarta: Penerbit Andi.
- Martinez, A. R., 2010, Natural language processing, *Wiley Interdisciplinary Reviews: Computational Statistics*. University of Strathclyde, doi: 10.1002/wics.76hal. 352–357, vol 2, no. 3.
- Muljono, Afini, U., Supriyanto, C., 2017, Morphology Analysis for Hidden Markov Model based Indonesian Part-of-Speech Tagger, *1st International Conference on Informatics and Computational Sciences (ICICoS)*, hal. 237–240.
- Mulyanto, A., Nurhuda, Y. A., Wiyanto, N., 2017, Penyelesaian Kata Ambigu Pada Proses POS Tagging Menggunakan Algoritma Hidden Markov Model (HMM), *Prosiding Seminar Nasional Metode Kuantitatif*, hal. 347–358, vol. 1, no.1.
- Muslich, M., 2009, *Tata Bentuk Bahasa Indonesia*, Bumi Aksara, Jakarta.
- Muth'mainnah, R. L., 2014, Pembangkit Grafik Berbasis Bahasa Alami pada Database Management System (DBMS), *Skripsi*, Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, UIN Maulana Malik Ibrahim, Malang.
- Narein T, A., 2021, Embeddings in BERT, OpenGenus IQ, <https://iq.opengenus.org/embeddings-in-bert/>, diakses tanggal 09 September 2022.
- Ngafenan, M. (1989) *Kamus Homonim Bahasa Indonesia*. Semarang: Dahara Press.
- Nivre, J., Marneffe, M., Ginter, F., Hajie, J., Manning, C. D., Pyysalo, S., Schuster, S., Tyers, F., Zeman, D., 2020, Universal Dependencies v2: An Evergrowing Multilingual Treebank Collection, *Proceedings of The 12th Language Resources and Evaluation Conference*. Marseille, Perancis, hal. 4034–4043.
- Nugroho, K. S., Sukmadewa, A. Y., DW, H. W., Bachtiar, F. A., Yudistira, N., 2021, BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews, *CoRR*, abs/2107.0, doi: 10.48550/arXiv.2107.06802.
- Otter, D. W., Medina, J. R., Kalita, J. K., 2021, A Survey of the Usages of Deep Learning for Natural Language Processing, *IEEE Transactions on Neural Networks and Learning Systems*, hal. 604–624, vol. 32, no. 2.
- Parera, J. D., 2009, *Dasar-dasar Analisis Sintaksis*, Erlangga, Yogyakarta.

- Pisceldo, F., Adriani, M., Manurung, R., 2009, Probabilistic Part of Speech Tagging for Bahasa Indonesia, *Proceedings of the 3rd International MALINDO Workshop, colocated event ACL-IJCNLP*, (January 2009), hal 1-6.
- Pratama, R. A., Suryani, A. A., Maharani, W., 2020, Part of Speech Tagging for Javanese Ngoko Language with Hidden Markov Model, *J. Cosine Journal*, hal. 84–91, vol.4, no.1.
- Putri, D. U. K., dan Pratomo, D. N., 2022, Clickbait Detection of Indonesian News Headlines using Fine-Tune Bidirectional Encoder Representations from Transformers (BERT), *Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, doi: 10.25139/inform.v7i2.4686hal. 162–168, vol.7, no.2.
- Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., 2018, Improving Language Understanding by Generative Pre-Training, *preprint: in progress*, hal. 1-12.
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu, P., 2020, Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer, *The Journal of Machine Learning Research*, hal. 5485-5551, vol. 21, no. 1.
- Ramadhanti, F., Wibisono, Y., Sukanto, R. A., 2019, Analisis Morfologi untuk Menangani Out-of-Vocabulary Words pada Part-of-Speech Tagger Bahasa Indonesia Menggunakan Hidden Markov Model, *Jurnal Linguistik Komputasional (JLK)*, doi: 10.26418/jlk.v2i1.13, hal. 6-12, vol. 2, no.1.
- Ratnasari, C. I., Kusumadewi, S., Rosita, L., 2014, Model Natural Language Processing untuk Perumusan Keluhan Pasien, *Seminar Nasional Informatika Medis (SNIMed) V*, hal. 11–18.
- Rich, L. and Knight, K., 1991, *Artificial Intelligence, Second edition*. McGraw-Hill Education, New York.
- Saadah, S., Auditama, K. M., Fattahila, A. A., Amorokhman, F. I., Aditsania, A., Rohmawati, A. A., 2022, Implementation of BERT, IndoBERT, and CNN-LSTM in Classifying Public Opinion about COVID-19 Vaccine in Indonesia, *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, doi: 10.29207/resti.v6i4.4215, hal. 648–655, vol. 6, no. 4.
- Samosir, F. V. P., Toba, H., Ayub, M., 2022, BESKlus : BERT Extractive Summarization with K-Means Clustering in Scientific Paper, *Jurnal Teknik Informatika dan Sistem Informasi*, doi: 10.28932/jutisi.v8i1.4474, hal. 202-217, vol. 8, no. 1.
- Sariwening, S. L., 2020, IndoBERT: Transformer-based Model for Indonesian Language Understanding, *Tesis*, Departemen Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada, Yogyakarta.
- Shaikh, R., 2018, Gentle Start to Natural Language Processing using Python. <https://towardsdatascience.com/gentle-start-to-natural-languageprocessing-using-python-6e46c07addf3> diakses tanggal 24 November 2022.

- Suyanto, D., Ramadhani, K. N., Mandala, S., 2019, *Deep Learning Modernisasi Machine Learning untuk Big Data*, Informatika, Bandung.
- Torfi, A., Shirvani, R. A., Keneshloo, Y., Tavaf, N., Fox, E. A., 2020, Natural Language Processing Advancements By Deep Learning: A Survey. *Glasgow: University of Strathclyde*. doi: 10.48550/ arXiv.2003.01200, hal. 1-23.
- Trismanto, 2018, Ambiguitas dalam bahasa indonesia, *Bangun Rekaprima*, hal. 42-48, vol. 4, no. 1.
- Trivusi, 2022, Mengenal BERT: Model NLP Canggih Buatan Google, <https://www.trivusi.web.id/2022/04/apa-itu-bert.html>, diakses tanggal 07 Oktober 2022.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I., 2017, Attention is all you need, *31st Conference on Neural Information Processing Systems (NIPS 2017)*, hal. 5999–6009.
- Wicaksono, A. F. dan Purwarianti, A., 2010, HMM Based Part-of-Speech Tagger for Bahasa Indonesia, *Proceedings of 4th International MALINDO (Malay and Indonesian Language) Workshop*, ha. 1-7.
- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., Purwarianti, A., 2020, IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*. Suzhou, China: Association for Computational Linguistics, hal. 843–857.
- Wojowasito, S., 1976, *Pengantar Sintaksis Indonesia*, Sintha Dharma, Bandung.
- Wulandari, D. W., Adikara, P. P., Adinugroho, S., 2018, Named Entity Recognition (NER) Pada Dokumen Biologi Menggunakan Rule Based dan Naïve Bayes Classifier, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, hal. 4555–4563, vol. 2, no. 11.
- Yazid, A. S. dan Fatwanto, A., 2018, Penentuan Kelas Kata Pada Part of Speech Tagging Kata Ambigu Bahasa Indonesia, *JISKA (Jurnal Informatika Sunan Kalijaga)*, doi: 10.14421/jiska.2018.23-05, hal. 157–166, vol 2, no. 3.
- Yuwana, R. S., Suryawati, E., Pardede, H. F., 2018, On Empirical Evaluation of Deep Architectures for Indonesian POS Tagging Problem, *2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*. Tangerang: IEEE, hal. 204–208.
- Zhang, Y., 2018, A Better Autoencoder for Image: Convolutional Autoencoder, *ANU Bio-inspired Computing conference*. Canberra, hal. 1–7.