

**PENERAPAN *FINE-TUNING* INDOBERTA DALAM
KLASIFIKASI SENTIMEN BERITA DIGITAL
BERBAHASA INDONESIA**



TUGAS AKHIR

Diajukan untuk memenuhi syarat kelulusan pada Program Strata Satu (S1)

Desi Masdin Dama

15210007

**Program Studi Sains Data
Fakultas Teknologi Informasi
Universitas Nusa Mandiri**

Jakarta

2025

PERSEMBAHAN

Alhamdulillah Rabbil Alamin, segala puji syukur penulis panjatkan untuk Allah SWT, Sang Kreator Agung yang memberikan segala kekuatan, meskipun dihadapkan pada berbagai tantangan dan banyak keterbatasan sepanjang proses penulisan tugas akhir ini. Dengan ini, penulis mempersembahkan karya ini kepada:

1. Kedua orang tua penulis, Ibu Suhartini Mohamad yang paling saya banggakan dan Bapak Usman Dama yang paling hebat, karena selalu memberikan cinta lewat rasa percaya, mengasihi dengan usaha dan memeluk dengan doa anak semata wayang mereka.
2. Adik penulis, Alm. Reza Dama yang satu-satunya kekurangannya adalah tidak berumur panjang, namun menyaksikan dan menjaga saya tumbuh dari tempat terbaik, semoga disana cahaya kami abadi dengan lebih sehat dan kuat.
3. Bibi Jocelyn Balangger, Paman Rinto Umar dan Ilham Arif Umar yang selalu menjadi garda terdepan pendukung impian saya setelah Ibu dan Bapak.
4. Alm. Saiful Umar, paman yang telah berpulang saat menanti penulis pulang, terima kasih atas yakin yang dipinjamkan untuk semua langkah menuju hal baik.
5. Zaid Afif Rahman, awal inspirasi saya memupuk berani untuk menyelesaikan masa belajar dan pulang menjadi kemenangan bagi keluarga dan aksa.
6. Dorothea Salsabela F. Hun, saudara tidak sedarah yang selalu mendampingi setiap proses berdarah-darah.
7. Sabeum Willem Yanes beserta istri semata wayang beliau, Sabeum Maulyanas, pemberi rumah dan tempat berteduh pertama di Ibu Kota.
8. Kak Wahyuni, Kak Lastry, Kak Nadya, selaku pendengar dan pemberi tenang yang selalu bersedia berbagi senang, susah dan sesak.

9. Mami, Papi, Kak Putri Zahwa dan Kak Destika Fitryana yang selalu memastikan penulis masih menghirup udara yang sama dengan mereka.
10. Seluruh anggota UKM Taekwondo Nusa Mandiri yang telah memberikan banyak pembelajaran lewat proses dan turut bersuka cita atas selesainya studi.
11. Sanggar Tangkasi, rumah kedua yang mengajarkan bahwa tempat pulang tidak selalu dalam bentuk bangunan, tapi siapa yang berteduh dibalik bangunan itu.
12. Alm. Faradila Indama, teman baik penulis yang menjadi gelap dalam terang.
13. Teman-teman angkatan pertama, Kanita Salsabila Dwi Irmanti, Rianggi Silvi Anti dan Ihsan Aulia Rahman yang telah kebersamai proses, memberikan semangat, dukungan dan tawa selama masa studi.
14. Sari Puspita Ayu Ningsih, Elita Victoria Mandang dan Nur Khofifah yang selalu memberi waktu dan bahu untuk berkisah.
15. Nadin Amizah, terima kasih atas album “Untuk Dunia, Cinta, dan Kotornya”. Sal Priadi, terima kasih untuk “Amin Paling Serius”, “Gala Bunga Matahari” dan “Kita Usahakan Rumah Itu”.
16. Semua pihak yang tidak tercantum namanya namun telah memberikan dukungan dalam bentuk yang beragam.
17. Sunya, diri saya sendiri yang memilih tetap hidup meski menyerah terasa lebih mudah dan tetap melangkah, sesuai kata Dori. Terima kasih selalu menabur *lily* disepanjang perjalanan menuju pulang, sekarang mari senang dengan setangkai semanggi untuk tumbuh dengan merengkuh utuh.

MOTTO

“Tidak ada berlian yang indah tanpa dipoles dan
tidak ada manusia yang hebat tanpa diproses”

-Desi 2021-

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Saya yang bertanda tangan di bawah ini:

Nama : Desi Masdin Dama
NIM : 15210007
Program Studi : Sains Data
Fakultas : Teknologi dan Informasi
Perguruan Tinggi : Universitas Nusa Mandiri

Dengan ini menyatakan bahwa Tugas Akhir yang telah saya buat dengan judul: **“Penerapan *Fine-Tuning* IndoBERTa Dalam Klasifikasi Sentimen Berita Digital Berbahasa Indonesia”**, adalah asli (orisinil) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun juga. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Tugas Akhir yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari **Universitas Nusa Mandiri** dicabut/dibatalkan.

Dibuat di: Depok
Pada Tanggal: 30 Juli 2025
Yang Menyatakan,



Desi Masdin Dama

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Saya yang bertanda tangan di bawah ini:

Nama : Desi Masdin Dama
NIM : 15210007
Program Studi : Sains Data
Fakultas : Teknologi dan Informasi
Perguruan Tinggi : Universitas Nusa Mandiri

Dengan ini menyetujui untuk memberikan ijin kepada pihak **Universitas Nusa Mandiri**, Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalti-Free Right*) atas karya ilmiah saya yang berjudul: **“Penerapan *Fine-Tuning* IndoBERTa Dalam Klasifikasi Sentimen Berita Digital Berbahasa Indonesia”**.

Dengan **Hak Bebas Royalti Non-Eksklusif** ini kepada pihak **Universitas Nusa Mandiri** berhak menyimpan, mengalih-media atau *format-kan*, mengelolaannya dalam pangkalan data (*database*), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak **Universitas Nusa Mandiri**, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di: Depok
Pada Tanggal: 30 Juli 2025
Yang Menyatakan,



Desi Masdin Dama

PERSETUJUAN DAN PENGESAHAN TUGAS AKHIR

Tugas Akhir ini diajukan oleh:

Nama : Desi Masdin Dama
NIM : 15210007
Program Studi : Sains Data
Fakultas : Teknologi Informasi
Jenjang : Sarjana (S1)
Judul Tugas Akhir : Penerapan Fine-Tuning IndoBERTa Dalam Klasifikasi Sentimen Berita Digital Berbahasa Indonesia

Telah dipertahankan pada periode 2025-1 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Sains Data (S.Si.D) pada Program Sarjana Program Studi Sains Data Fakultas Teknologi Informasi di Universitas Nusa Mandiri.

Jakarta, 14 Agustus 2025

PEMBIMBING TUGAS AKHIR

Dosen Pembimbing : Tati Mardiana, M.Kom.



Asisten Pembimbing : Riki Supriyadi, M.Kom.

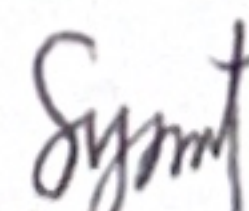


DEWAN PENGUJI

Penguji I : Nanang Ruhyana, M.Kom.



Penguji II : Syarah Seimahuira, M.Kom.



PEDOMAN PENGGUNAAN HAK CIPTA

Tugas Akhir yang berjudul **“Penerapan *Fine-Tuning* IndoBERTa Dalam Klasifikasi Sentimen Berita Digital Berbahasa Indonesia”** adalah karya tulis asli dari Desi Masdin Dama dan bukan merupakan hasil terbitan. Oleh karena hal tersebut, peredaran karya tulis ini hanya diperuntukkan di lingkungan akademik dan sepenuhnya memiliki hak cipta. Segala bentuk penggandaan, baik sebagian maupun seluruh isi dari karya tulis ini dilarang keras tanpa izin tertulis dari penulis.

Penggunaan referensi diperkenankan untuk dicatat, tetapi pengutipan atau peringkasan isi hanya dapat dilakukan dengan persetujuan penulis serta harus mencantumkan sumber secara ilmiah sesuai kaidah penulisan yang berlaku.

Untuk keperluan perizinan atau penggunaan lebih lanjut, dapat menghubungi penulis melalui informasi berikut:

Nama : Desi Masdin Dama

Alamat : Kel. Pondok Cina, Kec. Beji, Kota Depok, Jawa Barat

No.Telp : 0882-4595-6254

Email : desimdama1234@gmail.com

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT atas segala limpahan rahmat, karunia, serta kekuatan yang telah diberikan sehingga penulis mampu menyelesaikan tugas akhir yang berjudul **“Penerapan *Fine-Tuning* IndoBERTa Dalam Klasifikasi Sentimen Berita Digital Berbahasa Indonesia”** ini dengan baik dan lancar, sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sains Data di Universitas Nusa Mandiri.

Penyusunan Tugas Akhir ini tentu tidak terlepas dari dukungan, bimbingan, serta doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Rektor Universitas Nusa Mandiri
2. Wakil Rektor I Bidang Akademik Universitas Nusa Mandiri
3. Wakil Rektor II Bidang Non Akademik Universitas Nusa Mandiri
4. Dekan Fakultas Teknologi Informasi Universitas Nusa Mandiri
5. Kepala Program Studi Sains Data sekaligus Dosen Pembimbing Pertama, Ibu Tati Mardiana, M.Kom yang telah banyak memberikan bantuan kepada penulis dalam aspek teknis dan substansi isi dalam pengembangan sistem serta penajaman analisis.
6. Bapak Riki Supriyadi, M.Kom selaku Dosen Pembimbing Kedua, yang secara konsisten memberikan arahan mengenai kesesuaian format penulisan dengan mengacu pada pedoman akademik yang berlaku serta memastikan seluruh elemen penulisan sesuai standar dalam penyempurnaan dari tugas akhir penulis.
7. Seluruh Dosen dan Staf Universitas Nusa Mandiri, atas ilmu dan pengalaman yang telah diberikan selama masa perkuliahan.

8. Ibu Suhartini Mohamad dan Bapak Usman Dama atas segala doa, cinta dan dukungan tanpa batas yang menjadi sumber semangat penulis.
9. Ni Kadek Sri Wahyuni Pradnyawati, teman penulis yang turut serta menjadi rekan yang dapat diandalkan dalam pencarian dan pertimbangan referensi yang relevan selama penyusunan tugas akhir ini.
10. Moh. Alfarizqy selaku mentor selama magang, yang telah memberikan banyak pembelajaran berharga terkait pemahaman model, pengenalan lebih mendalam terkait *Data Science* serta mendelegasikan tugas dengan kepercayaan penuh kepada penulis, dalam melakukan perancangan fungsi dan melakukan analisis.
11. Pihak-pihak lainnya yang tidak dapat disebutkan satu per satu, namun turut berperan dalam mendukung terselesaikannya tugas akhir ini.

Penulis menyadari bahwa Tugas Akhir ini masih memiliki kekurangan, baik dari segi isi maupun teknis penulisan. Oleh karena itu, penulis sangat terbuka terhadap segala kritik dan saran yang membangun demi perbaikan di masa mendatang. Lebih dari pada itu, penulis berharap karya ini dapat memberikan manfaat bagi pembaca serta menjadi kontribusi kecil dalam pengembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen dan penerapan *Large Language Model* (LLM) di Indonesia.

Depok, 30 Juli 2025
Penulis



Desi Masdin Dama

ABSTRAK

Setiap harinya, terdapat ribuan artikel berita dipublikasikan oleh berbagai portal berita nasional, yang masing-masing memiliki gaya bahasa, pendekatan jurnalistik, serta *framing* (pembingkaiian informasi) yang beragam. Pendekatan manual yang dilakukan untuk membaca dan mengkaji ribuan berita tentu memerlukan waktu, tenaga dan sumber daya yang besar. Realitas tersebut menimbulkan kebutuhan terhadap pendekatan otomatis yang mampu mengidentifikasi dan mengklasifikasikan sentimen secara cepat dan akurat, yang tidak hanya berguna untuk mendeteksi kecenderungan opini publik, tetapi juga membantu dalam mengidentifikasi potensi bias atau kecenderungan opini publik terhadap isu tertentu. Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk menguji secara sistematis performa model klasifikasi sentimen pada berita digital. Dengan memanfaatkan teknologi kecerdasan buatan, khususnya *Natural Language Processing* (NLP), penelitian ini menerapkan model IndoBERTa yang telah melalui proses *fine-tuning* untuk melakukan klasifikasi sentimen pada berita digital dengan mengacu pada kerangka kerja CRISP-DM yang terdiri dari enam tahapan. Berdasarkan hasil evaluasi, model menunjukkan performa tinggi dengan *Accuracy* dan *Cohen's Kappa* sebesar 98%, yang berarti bahwa distribusi dari klasifikasi sentimen menunjukkan bahwa model mampu menangkap konteks topik berita. Berdasarkan hal tersebut, untuk mendukung penerapannya, model ini diintegrasikan ke dalam *pipeline* berbasis Streamlit yang mengotomatisasi proses mulai dari pengumpulan berita, klasifikasi sentimen, hingga visualisasi dan ekspor hasil analisis.

Kata Kunci: Berita Digital, Analisis Sentimen, NLP, IndoBERTa, *Fine-Tuning*

UNIVERSITAS
NUSA MANDIRI

ABSTRACT

Every day, thousands of news articles are published by various national news portals, each employing distinct linguistic styles, journalistic approaches and information framing techniques. A manual approach to reading and analyzing such a vast volume of news requires considerable time, effort, and resources. This condition underscores the necessity for an automated method capable of rapidly and accurately identifying and classifying sentiment—an approach that not only assists in detecting public opinion trends but also in uncovering potential biases or public inclinations toward specific issues. This research adopts a quantitative approach through experimental methods to systematically assess the performance of a sentiment classification model applied to digital news content. By utilizing artificial intelligence technologies, particularly Natural Language Processing (NLP), this study implements the IndoBERTa model—refined through a fine-tuning process—to classify sentiment in digital news articles, based on the CRISP-DM framework comprising six stages. Evaluation results indicate that the model exhibits excellent performance, achieving 98% in both Accuracy and Cohen’s Kappa scores, which signifies that the sentiment classification outcomes successfully capture the contextual nuances of the news topics. To support its practical implementation, the model has been integrated into a Streamlit-based pipeline that automates the entire process, encompassing news collection, sentiment classification, visualization, and analysis exportation.

Keyword: *Digital News, Sentiment Analysis, NLP, IndoBERTa, Fine-Tuning*



UNIVERSITAS
NUSA MANDIRI

DAFTAR PUSTAKA

- [1] S. Kemp, "Digital 2024: Global Overview Report," Data Reportal. Accessed: Jun. 03, 2025. [Online]. Available: <https://datareportal.com/reports/digital-2024-global-overview-report>
- [2] R. Chandra, B. Zhu, Q. Fang, and E. Shinjikashvili, "Large language models for sentiment analysis of newspaper articles during COVID-19: The Guardian," *Applied Soft Computing Journal*, vol. 171, Jan. 2025, doi: 10.1016/j.asoc.2025.112743.
- [3] M. Alam, A. Iana, A. Grote, K. Ludwig, P. Müller, and H. Paulheim, "Towards Analyzing the Bias of News Recommender Systems Using Sentiment and Stance Detection," *WWW 2022 - Companion Proceedings of the Web Conference 2022*, pp. 448–457, Mar. 2022, doi: 10.1145/3487553.3524674.
- [4] Zulham, F. A. Lubis, D. Priyono, Fauzan, S. Julina, and D. A. Deryansyah, "Framing Media dalam Berita Kontroversial: Studi Kasus pada Kasus-Kasus Politik atau Sosial," *Jurnal Review Pendidikan dan Pengajaran*, vol. 7, no. 3, 2024, Accessed: May 06, 2025. [Online]. Available: <http://journal.universitaspahlawan.ac.id/index.php/jrpp>
- [5] A. R. Hanum *et al.*, "Analisis Kinerja Algoritma Klasifikasi Teks BERT dalam Mendeteksi Berita Hoaks," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 36, pp. 537–546, Jun. 2024, doi: 10.25126/jtiik938093.
- [6] K. Kirtac and G. Germano, "Sentiment trading with large language models," *Financ Res Lett*, vol. 62, Mar. 2024, doi: 10.1016/j.frl.2024.105227.
- [7] Syahrudin, F. A. Artanto, A. R. Maulana, and Filsafat, "Metode Support Vector Machine (SVM) dan Lexicon-Based dalam Analisis Sentiment Ulasan Pengguna Aplikasi Wink," *JUMINTAL: Jurnal Manajemen Informatika dan Bisnis Digital*, vol. 4, no. 1, pp. 59–73, May 2025, doi: 10.55123/jumintal.v4i1.5236.
- [8] F. J. Rodrigo-Ginés, J. Carrillo-de-Albornoz, and L. Plaza, "A systematic review on media bias detection: What is media bias, how it is expressed, and how to detect it," *Expert Syst Appl*, vol. 237, 2024, doi: 10.1016/j.eswa.2023.121641.
- [9] A. Mahendra and S. Styawati, "Implementasi Lowk-Rank Adaptation of Large Langauge Model (LoRA) Untuk Effisiensi Large Language Model," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 4, pp. 1881–1890, Nov. 2024, doi: 10.29100/jipi.v9i4.5519.
- [10] M. E. Syahputra, A. Putera Kemala, and D. Ramdhan, "Clickbait Detection in Indonesia Headline News Using Indobert and Roberta," *JURNAL RISET INFORMATIKA*, vol. 5, no. 3, Jun. 2023, doi: 10.34288/jri.v5i3.237.

- [11] C. H. Lin and U. Nuha, "Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy," *J Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00782-9.
- [12] D. Dwi Saputra, R. Fahlapi, A. Yadi Kuntoro, and T. Asra, "Analisis Sentimen Terhadap Twitter Direktorat Jenderal Bea dan Cukai Menggunakan komparasi Algoritma Naïve Bayes dan Support Vector Machine," *J-INTECH (Journal of Information and Technology)*, pp. 285–296, 2024.
- [13] A. S. Aribowo and S. Khomsah, "Implementation Of Text Mining For Emotion Detection Using The Lexicon Method (Case Study: Tweets About Covid-19) Implementasi Text Mining Untuk Deteksi Emosi Menggunakan Metode Leksikon (Studi Kasus: Twit Tentang Covid-19)," *Jurnal Informatika dan Teknologi Informasi*, vol. 18, no. 1, pp. 49–60, Feb. 2021, doi: 10.31515/telematika.v18i1.4341.
- [14] S. H. Caryarini, F. Adibha, M. K. Luthfi, A. R. Arfianti, and P. K. Nisa, "Konglomerasi Media (Koran Kompas) ke Berita Digital Terhadap Masyarakat," *Jurnal Ilmiah Research and Development Student*, vol. 2, no. 2, pp. 103–116, May 2024, doi: 10.59024/jis.v2i2.757.
- [15] G. J. Wiladi and M. D. Afrianti, "Pengaruh Literasi Media Digital Terhadap Tindakan Penyebaran Berita Palsu Pada Mahasiswa Universitas Bhayangkara," *Jurnal Ilmiah Wahana Pendidikan*, vol. 10, no. 21, pp. 352–360, Nov. 2024, doi: 10.5281/zenodo.14405369.
- [16] J. Lantowa and R. Idul, "Representasi Aksi Sosial dalam Konstruksi Ideologi Media Berita Digital Terkait Kebijakan Pemerintah Selama Pandemi," *Ranah: Jurnal Kajian Bahasa*, vol. 12, no. 1, pp. 87–100, Jun. 2023, doi: 10.26499/rnh.v12i1.5269.
- [17] O. Mailani, I. Nuraeni, S. A. Syakila, and J. Lazuardi, "Bahasa Sebagai Alat Komunikasi Dalam Kehidupan Manusia," *KAMPRET Journal*, vol. 1, no. 2, pp. 1–10, Jan. 2022, [Online]. Available: www.plus62.isha.or.id/index.php/kampret
- [18] Y. Salim and M. Hasnawi, "Konversi Bahasa Indonesia ke Perintah Data Manipulation Language pada Structured Query Language menggunakan Natural Language Processing," *Buletin Sistem Informasi dan Teknologi Islam*, vol. 3, no. 3, pp. 181–187, Aug. 2022, Accessed: May 22, 2025. [Online]. Available: <https://garuda.kemdikbud.go.id/documents/detail/3106930>
- [19] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, B. MIT Press, 2016.
- [20] R. Danar, D. M. Kom, M. M. Kom, A. Bahtiar, M. Kom, and I. Ali, *Dasar Dasar Natural Language Processing (NLP)*, 1st ed. Minhaj Pustaka, 2024.
- [21] F. H. Rachman, *Komputasi Bahasa Alami*. Malang: Media Nusa Creative (MNC Publishing), 2020. Accessed: May 22, 2025. [Online]. Available: <https://play.google.com/store/books/details?id=Gn5JEAAQBAJ&rdid=book-Gn5JEAAQBAJ&rdot=1>

- [22] R. Aulianita, A. M. B. Aji, and Y. E. Achyani, "TEXT MINING MENGGUNAKAN NAIVE BAYES BERBASIS PARTICLE SWARM OPTIMIZATION UNTUK SENTIMENT RESTAURANT," *Jurnal Teknik Informatika Musirawas* Rizki Aulianita, vol. 6, no. 1, pp. 21–29, Jun. 2021, doi: 10.32767/jutim.v6i1.1300.
- [23] M. S. 'Afif, M. Muzakir, and Moh. I. A. G. Al Awalaen, "TEXT MINING UNTUK MENGGKLASIFIKASI JUDUL BERITA ONLINE STUDI KASUS RADAR BANJARMASIN MENGGUNAKAN METODE NAÏVE BAYES," *Kumpulan jurnaL Ilmu Komputer (KLIK)*, vol. 08, no. 2, Jun. 2021, doi: 10.20527/klik.v8i2.389.
- [24] R. S. Lutfiyani and N. Retnowati, "IMPLEMENTASI PENDETEKSIAN SPAM EMAIL MENGGUNAKAN METODE TEXT MINING DENGAN ALGORITMA NAÏVE BAYES DAN DECISION TREE J48," *Jurnal Komputer dan Informatika*, vol. 9, no. 2, pp. 244–252, Oct. 2021, doi: 10.35508/jicon.v9i2.5304.
- [25] M. A. Palomino and F. Aider, "Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis," *Applied Sciences (Switzerland)*, vol. 12, no. 17, Sep. 2022, doi: 10.3390/app12178765.
- [26] S. M. Weiss, N. Indurkha, and T. Zhang, *Fundamentals of Predictive Text Mining*. Springer Publishing Company, Incorporated, 2010.
- [27] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [28] S. Han, M. Wang, J. Zhang, D. Li, and J. Duan, "A Review of Large Language Models: Fundamental Architectures, Key Technological Evolutions, Interdisciplinary Technologies Integration, Optimization and Compression Techniques, Applications, and Challenges," *Electronics (Switzerland)*, vol. 13, no. 24, Dec. 2024, doi: 10.3390/electronics13245040.
- [29] M. A. AlAfnan, "Large Language Models as Computational Linguistics Tools: A Comparative Analysis of ChatGPT and Google Machine Translations," *Journal of Artificial Intelligence and Technology*, vol. 5, pp. 20–32, Jun. 2024, doi: 10.37965/jait.2024.0549.
- [30] H. Yang, S. Li, and T. Gonçalves, "Enhancing Biomedical Question Answering with Large Language Models," *Information (Switzerland)*, vol. 15, Aug. 2024, doi: 10.3390/info15080494.
- [31] E. León-Sandoval, M. Zareei, L. I. Barbosa-Santillán, and L. E. Falcón Morales, "Measuring the Impact of Language Models in Sentiment Analysis for Mexico's COVID-19 Pandemic," *Electronics (Switzerland)*, vol. 11, no. 16, Aug. 2022, doi: 10.3390/electronics11162483.

- [32] R. M. T. Rahardiansyah, S. R. Perdana, and T. N. Fatyanosa, "Analisis Teknik Embedding Model NV-Embed pada Large Language Models Berbasis Retrieval Augmented Generation," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 2, Feb. 2025, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [33] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.
- [34] V. Chakkarwar, S. Tamane, and A. Thombre, "A Review on BERT and Its Implementation in Various NLP Tasks," in *Advances in Computer Science Research*, Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022), May 2023, pp. 112–121. doi: 10.2991/978-94-6463-136-4_12.
- [35] Y. Liu *et al.*, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," Jul. 2019, doi: <https://doi.org/10.48550/arXiv.1907.11692> Focustolearnmore.
- [36] X. You *et al.*, "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Science China Information Sciences*, vol. 64, no. 1, pp. 1–74, Jan. 2021, doi: 10.1007/s11432-020-2955-6.
- [37] A. Rahmawati, A. Alamsyah, and A. Romadhony, "Hoax News Detection Analysis using IndoBERT Deep Learning Methodology," in *5th International Conference on Information and Communications Technology (ICOIACT)*, Institute of Electrical and Electronics Engineers, Aug. 2022, pp. 368–373. doi: 10.1109/ICoICT55009.2022.9914902.
- [38] A. Rogers, O. Kovaleva, and A. Rumshisky, "A Primer in BERTology: What We Know About How BERT Works," in *Transactions of the Association for Computational Linguistics*, M. Johnson, B. Roark, and A. Nenkova, Eds., Cambridge, MA: MIT Press, 2020, pp. 842–866. doi: 10.1162/tac1_a_00349.
- [39] E. M. Pusung and I. N. Dewi, "Optimasi RoBERTa dengan Hyperparameter Tuning untuk Deteksi Emosi berbasis Teks," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 10, no. 3, pp. 240–248, Feb. 2025, doi: 10.25077/TEKNOSI.v10i3.2024.240-248.
- [40] A. Bonfigli, L. Bacco, M. Merone, and F. Dell'Orletta, "From pre-training to fine-tuning: An in-depth analysis of Large Language Models in the biomedical domain," *Artif Intell Med*, vol. 157, Oct. 2024, doi: 10.1016/j.artmed.2024.103003.
- [41] A. V Ganesan, M. Matero, A. Reddy Ravula, H. Vu, and H. A. Schwartz, "Empirical Evaluation of Pre-trained Transformers for Human-Level NLP: The Role of Sample Size and Dimensionality," in *Proceedings of the 2021*

Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2021), Association for Computational Linguistics (ACL), Jun. 2021, pp. 4515–4532. doi: 10.18653/v1/2021.naacl-main.357.

- [42] D. Purnamasari *et al.*, *Pengantar Metode Analisis Sentimen*. Depok: Gunadarma Penerbit, 2023.
- [43] I. R. Yunita, W. Maulana Baihaqi, A. Shafira, T. Damayanti, and L. Akhaerunnisa, “Analisis Performa Algoritma Klasifikasi pada Sentimen Ulasan Pengguna terhadap Aplikasi Muamalat DIN Analysis of Classification Algorithm Performance on User Review Sentiment of the Muamalat DIN Application,” *Cogito Smart Journal*, vol. 9, no. 2, Dec. 2023.
- [44] F. K. Ihtada, R. Alfianita, and O. Q. Aziz, “Aspect-based Multilabel Classification of E-commerce Reviews Using Fine-tuned IndoBERT,” *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, Jan. 2025, doi: 10.22219/kinetik.v10i1.2088.
- [45] S. Saadah, M. K. Auditama, A. A. Fattahila, I. F. Amorokhman, A. Aditsania, and A. A. Rohmawati, “Implementation of BERT, IndoBERT, and CNN-LSTM in Classifying Public Opinion about COVID-19 Vaccine in Indonesia,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 2, pp. 648–655, Aug. 2022, doi: 10.29207/resti.v6i4.4215.
- [46] Y. A. Singgalen, “Performance Analysis of IndoBERT for Sentiment Classification in Indonesian Hotel Review Data,” *Journal of Information System Research*, vol. 6, no. 2, pp. 976–986, 2025, doi: 10.47065/josh.v6i2.6505.
- [47] A. S. Talaat, “Sentiment analysis classification system using hybrid BERT models,” *J Big Data*, vol. 10, no. 1, 2023, doi: 10.1186/s40537-023-00781-w.
- [48] Aripin, S. A. Santoso, and H. Haryanto, “Mengoptimalkan Akurasi pada Klasifikasi Emosi Majemuk Berdasarkan Semantik Kalimat Menggunakan XLM-RoBERTa,” *JURNAL NASIONAL TEKNIK ELEKTRO DAN TEKNOLOGI INFORMASI*, vol. 12, Feb. 2023, Accessed: May 20, 2025. [Online]. Available: <https://jurnal.ugm.ac.id/v3/JNTETI/article/view/6084/2279>
- [49] A. Bello, S. C. Ng, and M. F. Leung, “A BERT Framework to Sentiment Analysis of Tweets,” *Sensors (MDPI)*, vol. 23, no. 1, Jan. 2023, doi: 10.3390/s23010506.
- [50] Rina, “Memahami Confusion Matrix: Accuracy, Precision, Recall, Specificity, dan F1-Score untuk Evaluasi Model Klasifikasi | by Rina | Medium.” Accessed: Jul. 07, 2025. [Online]. Available: <https://esairina.medium.com/memahami-confusion-matrix-accuracy-precision-recall-specificity-dan-f1-score-610d4f0db7cf>
- [51] A. A. Chamid, Widowati, and R. Kusumaningrum, “Labeling Consistency Test of Multi-Label Data for Aspect and Sentiment Classification Using the Cohen

Kappa Method,” *Ingenierie des Systemes d’Information*, vol. 29, no. 1, pp. 161–167, Feb. 2024, doi: 10.18280/isi.290118.

- [52] Effendi and R. Noviana, “Perancangan Web Sistem Analisis Sentimen Media Sosial Twitter Dengan Metode Valence Aware Dictionary And Sentimen Reasoner (Vader) Menggunakan PHP & MySQL pada Pemerintah Kota Bekasi,” *Jurnal Ilmiah Komputasi*, vol. 20, no. 1, Mar. 2021, doi: 10.32409/jikstik.20.1.369.

- [53] A. Rahmadian, “Public Sentiment Towards Mandatory Halal Certification: A Large Language Model (LLM) Approach,” *Jurnal Ekonomi Industri Halal*, vol. 4, no. 2, pp. 1–15, 2024, doi: <https://doi.org/10.15575/likuid.v4i2.35185>.



UNIVERSITAS
NUSA MANDIRI