

DAFTAR PUSTAKA

- Adoma, A. F., Henry, N.-M., & Chen, W. (2020). Comparative Analyses of Bert, Roberta, Distilbert, and Xlnet for Text-Based Emotion Recognition. *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 117–121. <https://doi.org/10.1109/ICCWAMTIP51612.2020.9317379>
- Agarwal, M. (2019). An Overview of Natural Language Processing. *International Journal for Research in Applied Science and Engineering Technology*, 7(5), 2811–2813. <https://doi.org/10.22214/ijraset.2019.5462>
- Agung, F., Ayomi, J., & Dewi, K. E. (2023). ANALISIS EMOSI PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE MULTINOMIAL NAÏVE BAYES DAN SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE. *KOMPUTA : Jurnal Ilmiah Komputer dan Informatika*, 12(2). <https://www.statista.com>
- Altalhan, M., Algarni, A., & Alouane, M. T.-H. (2025). Imbalanced data problem in machine learning: A review. *IEEE access*, 13, 13686–13699.
- Antonakaki, D., Fragopoulou, P., & Ioannidis, S. (2021). A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. *Expert Systems with Applications*, 164. <https://doi.org/10.1016/j.eswa.2020.114006>
- Assunção, G. O., Izbicki, R., & Prates, M. O. (2023). *Is augmentation effective to improve prediction in imbalanced text datasets?* <http://arxiv.org/abs/2304.10283>
- Baihaqi, M. Y., Halawa, E., Syah, R. A. S., Nurrahma, A., & Wijaya, W. (2023). Emotion Classification in Indonesian Language: A CNN Approach with

Hyperband Tuning. *Jurnal Buana Informatika*, 14(02), 137–146.

<https://doi.org/10.24002/jbi.v14i02.7558>

Bushuiev, A., Bushuiev, R., Sedlar, J., Pluskal, T., Damborsky, J., Mazurenko, S., & Sivic, J. (2024). *Revealing data leakage in protein interaction benchmarks*.

<http://arxiv.org/abs/2404.10457>

Cahya, L. D., Luthfiarta, A., Krisna, J. I. T., Winarno, S., & Nugraha, A. (2024).

Improving Multi-label Classification Performance on Imbalanced Datasets Through SMOTE Technique and Data Augmentation Using IndoBERT Model.

Jurnal Nasional Teknologi dan Sistem Informasi, 9(3), 290–298.

<https://doi.org/10.25077/teknosi.v9i3.2023.290-298>

Cahyani, D. E., & Patasik, I. (2021). Performance comparison of tf-idf and word2vec models for emotion text classification. *Bulletin of Electrical Engineering and Informatics*, 10(5), 2780–2788.

<https://doi.org/10.11591/eei.v10i5.3157>

Cai, Z., Li, Z., Liu, Y., Guo, L., & Song, Y. (2025). *Ustnlp16 at SemEval-2025 Task 9: Improving Model Performance through Imbalance Handling and Focal Loss*.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Dalam *Journal of Artificial Intelligence Research* (Vol. 16).

Chiorrini, A., Diamantini, C., Mircoli, A., & Potena, D. (2021). Emotion and sentiment analysis of tweets using BERT. *EDBT/ICDT Workshops*.

<https://api.semanticscholar.org/CorpusID:233261276>

De Angeli, K., Gao, S., Danciu, I., Durbin, E. B., Wu, X. C., Stroup, A., Doherty, J., Schwartz, S., Wiggins, C., Damesyn, M., Coyle, L., Penberthy, L., Tourassi, G. D., & Yoon, H. J. (2022). Class imbalance in out-of-distribution datasets: Improving the robustness of the TextCNN for the classification of rare cancer types. *Journal of Biomedical Informatics*, 125. <https://doi.org/10.1016/j.jbi.2021.103957>

Dereich, S., & Jentzen, A. (2024). *Convergence rates for the Adam optimizer*. <http://arxiv.org/abs/2407.21078>

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3), 550–553. <https://doi.org/10.1037/0033-295X.99.3.550>

elnahas, mohammed, Hussein, M., & Keshk, A. (2021). Imbalanced Data Oversampling Technique Based on Convex Combination Method. *IJCI. International Journal of Computers and Information*, 0(0), 0–0. <https://doi.org/10.21608/ijci.2021.72508.1047>

Erfian Junianto, Mila Puspitasari, Salman Ilyas Zakaria, Toni Arifin, & Ignatius Wiseto Prasetyo Agung. (2024). Klasifikasi Emosi pada Teks Berbahasa Inggris Menggunakan Pendekatan Ensemble Bagging. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 13(4), 272–281. <https://doi.org/10.22146/jnteti.v13i4.14440>

Geetha, Dr. V, Gomathy, D. C. K., Vallab Yaratha Yagn, Mr. D. S. D., & Praneesh, S. (2023). THE ROLE OF NATURAL LANGUAGE PROCESSING.

INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, 07(11), 1–11. <https://doi.org/10.55041/IJSREM27094>

Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. Dalam *Expert Systems with Applications* (Vol. 73, hlm. 220–239). Elsevier Ltd. <https://doi.org/10.1016/j.eswa.2016.12.035>

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9), 1263–1284.

He, H., & Ma, Y. (2013). *Imbalanced learning: foundations, algorithms, and applications*.

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. Dalam *Science* (Vol. 349, Nomor 6245, hlm. 261–266). American Association for the Advancement of Science. <https://doi.org/10.1126/science.aaa8685>

Joeres, R., Blumenthal, D. B., & Kalinina, O. V. (2023). *DataSAIL: Data Splitting Against Information Leakage*. <https://doi.org/10.1101/2023.11.15.566305>

Khvatskii, G., Moniz, N., Doan, K. D., & Chawla, N. V. (2025). Class-aware contrastive optimization for imbalanced text classification. *Discover Data*, 3(1). <https://doi.org/10.1007/s44248-025-00064-0>

Kim, K. G. (2016). Book Review: Deep Learning. *Healthcare Informatics Research*, 22(4), 351. <https://doi.org/10.4258/hir.2016.22.4.351>

Koto, F., Lau, J. H., & Baldwin, T. (2021). *IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization*. <http://arxiv.org/abs/2109.04607>

- Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. Dalam *Progress in Artificial Intelligence* (Vol. 5, Nomor 4, hlm. 221–232). Springer Verlag. <https://doi.org/10.1007/s13748-016-0094-0>
- Kumar, P., Bhatnagar, R., Gaur, K., & Bhatnagar, A. (2021). Classification of Imbalanced Data: Review of Methods and Applications. *IOP Conference Series: Materials Science and Engineering*, 1099(1), 012077. <https://doi.org/10.1088/1757-899x/1099/1/012077>
- Kustiwa, A. F. A., Akbar, M. A., & Pinandito, A. (2024). Studi Perbandingan pada Metode CNN-LSTM dan LSTM dalam Mendeteksi Emosi pada Data Teks Berbahasa Indonesia pada Media Sosial Twitter. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 8(5).
- Labd, Z., & Housni, K. (2025). AFL-BERT : Enhancing Minority Class Detection in Multi-Label Text Classification with Adaptive Focal Loss and BERT. Dalam *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 16, Nomor 7). www.ijacsa.thesai.org
- Li, Z., & Shimada, K. (2024). Addressing Class Imbalance in Customer Review: Analysis Using Focal Loss and SVM with BERT. *2024 17th International Congress on Advanced Applied Informatics (IIAI-AAI-Winter)*, 59–64. <https://doi.org/10.1109/IIAI-AAI-Winter65925.2024.00020>
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). *Focal Loss for Dense Object Detection*. <http://arxiv.org/abs/1708.02002>
- Loshchilov, I., & Hutter, F. (2019). *Decoupled Weight Decay Regularization*. <http://arxiv.org/abs/1711.05101>

Madabushi, H. T., Kochkina, E., & Castelle, M. (2020). *Cost-Sensitive BERT for Generalisable Sentence Classification with Imbalanced Data*.
<http://arxiv.org/abs/2003.11563>

Masaling, N. A. P., Siswanto, R. R., & Girsang, A. S. (2024). Indonesian Tweet Emotion Detection Using IndoBERT. *2024 International Conference on Information Management and Technology (ICIMTech)*, 478–482.
<https://doi.org/10.1109/ICIMTech63123.2024.10780847>

Mohammad, S. M. (2015). Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. Dalam *Emotion Measurement*.

Nugroho, K. S., Bachtiar, F. A., & Mahmudy, W. F. (2022). Detecting Emotion in Indonesian Tweets: A Term-Weighting Scheme Study. *Journal of Information Systems Engineering and Business Intelligence*, 8(1), 61–70.
<https://doi.org/10.20473/jisebi.8.1.61-70>

Nurfauzan, A., & Maharani, W. (2021). *Klasifikasi Emosi Pada Pengguna Twitter Menggunakan Metode Klasifikasi Decision Tree*.

Olusegun, R., Oladunni, T., Audu, H., Houkpati, Y. A. O., & Bengesi, S. (2023). Text Mining and Emotion Classification on Monkeypox Twitter Dataset: A Deep Learning-Natural Language Processing (NLP) Approach. *IEEE Access*, 11, 49882–49894. <https://doi.org/10.1109/ACCESS.2023.3277868>

Pang, F., Chen, X., Li, L., Xu, X., & Xing, Z. (2024). F_MixBERT: Sentiment Analysis Model using Focal Loss for Imbalanced E-commerce Reviews. *KSII Transactions on Internet and Information Systems*, 18(2), 263–283.
<https://doi.org/10.3837/TIIS.2024.02.001>

Parrott, W. G. (2001). *Emotions in social psychology: Essential readings*. psychology press.

Parvin, T., & Hoque, M. M. (2021). An Ensemble Technique to Classify Multi-Class Textual Emotion. *Procedia Computer Science*, 193, 72–81. <https://doi.org/10.1016/j.procs.2021.10.008>

Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. *Science China Technological Sciences*, 63(10), 1872–1897. <https://doi.org/10.1007/s11431-020-1647-3>

Rahmat, A. F., & Rafi, M. (2022). Social Media Network Analysis on Twitter Users Network to the Pension Plan Policy. *Communicare : Journal of Communication Studies*, 8(1), 62. <https://doi.org/10.37535/101009120225>

Ramakrishnan, S., & Dhinesh Babu, L. D. (2025). Improving Multi-Label Emotion Classification on Imbalanced Social Media Data With BERT and Clipped Asymmetric Loss. *IEEE Access*, 13, 60589–60601. <https://doi.org/10.1109/ACCESS.2025.3557091>

Riccosan, Saputra, K. E., Pratama, G. D., & Chowanda, A. (2022). Emotion dataset from Indonesian public opinion. *Data in Brief*, 43. <https://doi.org/10.1016/j.dib.2022.108465>

Sailunaz, K., & Alhadj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, 36, 101003. <https://doi.org/10.1016/j.jocs.2019.05.009>

Salih, M. I., Mohammed, S. M., Ibrahim, A. K., Ahmed, O. M., & Haji, L. M. (2025). Fine-Tuning BERT for Automated News Classification. *Engineering*,

Technology and Applied Science Research, 15(3), 22953–22959.

<https://doi.org/10.48084/etasr.10625>

Shaver, P., Schwartz, J., Kirson, D., & O’connor, G. (1987). *ATTITUDES AND SOCIAL COGNITION Emotion Knowledge: Further Exploration of a Prototype Approach*.

Subramanian, S., Rahimi, A., Baldwin, T., Cohn, T., & Frermann, L. (2021). Fairness-aware Class Imbalanced Learning. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2045–2051. <https://doi.org/10.18653/v1/2021.emnlp-main.155>

Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). *How to Fine-Tune BERT for Text Classification?* <http://arxiv.org/abs/1905.05583>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. <http://arxiv.org/abs/1706.03762>

William, S., Kenny, & Chowanda, A. (2024). EMOTION RECOGNITION INDONESIAN LANGUAGE FROM TWITTER USING INDOBERT AND BI-LSTM. *Communications in Mathematical Biology and Neuroscience*, 2024. <https://doi.org/10.28919/cmbn/7858>

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, L., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., ... Dean, J. (2016). *Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*. <http://arxiv.org/abs/1609.08144>

Yeung, A. W. K., Kletecka-Pulker, M., Eibensteiner, F., Plunger, P.,
Völkl-Kernstock, S., Willschke, H., & Atanasov, A. G. (2021).
Implications of Twitter in Health-Related Research: A Landscape
Analysis of the Scientific Literature. *Frontiers in Public Health*, 9.
<https://doi.org/10.3389/fpubh.2021.654481>