



UNIVERSITAS  
GADJAH MADA

**GAN INDOBERT: Generative Adversarial Learning for Emotion Classification with Limited Labeled Data on Indonesian Tweets**

Norman Muhammad Mikhail, Dzikri Rahadian Fudholi, S.Kom., M.Comp.

Universitas Gadjah Mada, 2023 | Diunduh dari <http://etd.repository.ugm.ac.id/>

## REFERENCES

- Acheampong, F.A., Wenyu, C. and Nunoo-Mensah, H. (2020). Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, 2(7). doi:<https://doi.org/10.1002/eng2.12189>.
- Alswaidan, N. and Menai, M.E.B. (2020). A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, 62(8), pp.2937–2987. doi:<https://doi.org/10.1007/s10115-020-01449-0>.
- Arellano, L. J., Escalante, H. J., Villaseñor Pineda, L., Montes y Gómez, M., & Sanchez-Vega, F. (2022). *Overview of da-vincis at iberlef 2022: Detection of aggressive and violent incidents from social media in spanish*, Procesamiento del Lenguaje Natural. 2022, 69: 207-215. doi:<https://doi.org/10.26342/2022-69-18>.
- Auti, T., Sarkar, R., Stearns, B., Ojha, A.Kr., Paul, A., Comerford, M., Megaro, J., Mariano, J., Herard, V. and McCrae, J.P. (2022). *Towards Classification of Legal Pharmaceutical Text using GAN-BERT*. [online] ACLWeb. Available at: <https://aclanthology.org/2022.csrnlp-1.8> [Accessed 2 Nov. 2023].
- Birjali, M., Kasri, M. and Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, p.107134. doi:<https://doi.org/10.1016/j.knosys.2021.107134>.
- Breazzano, C., Croce, D. and Basili, R. (2021). MT-GAN-BERT: Multi-Task and Generative Adversarial Learning for Sustainable Language Processing.



Brock, A., Donahue, J. and Simonyan, K. (2018). *Large Scale GAN Training for High Fidelity Natural Image Synthesis*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1809.11096>.

Cho, W. and Choi, Y. (2022). LMGAN: Linguistically Informed Semi-Supervised GAN with Multiple Generators. *Sensors*, 22(22), p.8761. doi:<https://doi.org/10.3390/s22228761>.

Clark, K., Luong, M.-T., Le, Q.V. and Manning, C.D. (2020). ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators. *arXiv (Cornell University)*. doi:<https://doi.org/10.48550/arxiv.2003.10555>.

Croce, D., Castellucci, G. and Basili, R. (2020). *GAN-BERT: Generative Adversarial Learning for Robust Text Classification with a Bunch of Labeled Examples*. [online] ACLWeb. doi:<https://doi.org/10.18653/v1/2020.acl-main.191>.

Devlin, J., Chang, M., Lee, K. and Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv (Cornell University)*.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, [online] 6(3-4), pp.169–200. doi:<https://doi.org/10.1080/02699939208411068>.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4), pp.384–392. doi:<https://doi.org/10.1037/0003-066x.48.4.384>.

Engel, J., Agrawal, K.K., Chen, S., Gulrajani, I., Donahue, C. and Roberts, A. (2019). GANSynth: Adversarial Neural Audio Synthesis. *arXiv:1902.08710 [cs, eess, stat]*. [online] Available at: <https://arxiv.org/abs/1902.08710>.



Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). *Generative Adversarial Networks*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1406.2661>.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), pp.139–144. doi:<https://doi.org/10.1145/3422622>.

Han, Z., Wang, J., Fan, H., Wang, L. and Zhang, P. (2018). Unsupervised Generative Modeling Using Matrix Product States. 8(3). doi:<https://doi.org/10.1103/physrevx.8.031012>.

Hickman, L., Thapa, S., Tay, L., Cao, M. and Srinivasan, P. (2020). Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. *Organizational Research Methods*, 25(1), p.109442812097168. doi:<https://doi.org/10.1177/1094428120971683>.

M, H. and M.N, S. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, [online] 5(2), pp.01-11. doi:<https://doi.org/10.5121/ijdkp.2015.5201>.

Arfinda Ilmania, Abdurrahman, Cahyawijaya, S. and Ayu Purwarianti (2018). Aspect Detection and Sentiment Classification Using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis. doi:<https://doi.org/10.1109/ialp.2018.8629181>.

Izard, C.E. (1991). *The Psychology of Emotions*. Boston, MA: Springer US. doi:<https://doi.org/10.1007/978-1-4899-0615-1>.



- Koto, F., Rahimi, A., Lau, J.H. and 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*. doi:<https://doi.org/10.18653/v1/2020.coling-main.66>.
- Lang, P.J. (1995). The emotion probe. Studies of motivation and attention. *The American Psychologist*, [online] 50(5), pp.372–385. doi:<https://doi.org/10.1037/0003-066x.50.5.372>.
- Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C.H. and Kang, J. (2019). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, [online] 36(4). doi:<https://doi.org/10.1093/bioinformatics/btz682>.
- Li, X. and Roth, D. (2005). Learning question classifiers: the role of semantic information. *Natural Language Engineering*, 12(3), pp.229–249. doi:<https://doi.org/10.1017/s1351324905003955>.
- Lindsley, D. B. (1951). Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 473–516). Wiley.
- Maas, A.L., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y. and Potts, C. (2011). *Learning Word Vectors for Sentiment Analysis*. [online] ACLWeb. Available at: <https://aclanthology.org/P11-1015/>.
- Mirza, M. and Osindero, S. (2014). *Conditional Generative Adversarial Nets*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1411.1784>.
- Myszewski, J.J., Klossowski, E., Schroeder, K.M. and Schroeder, C.A. (2022). Utilization of sentiment analysis to assess and compare negative finding



reporting in veterinary and human literature. *Research in Veterinary Science*, 148, pp.27–32. doi:<https://doi.org/10.1016/j.rvsc.2022.04.010>.

Nash, J. (1953). Two-Person Cooperative Games. *Econometrica*, 21(1), p.128. doi:<https://doi.org/10.2307/1906951>.

Oord, A. van den, Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. and Kavukcuoglu, K. (2016). *WaveNet: A Generative Model for Raw Audio*. [online] Google Research. Available at: <https://research.google/pubs/pub45774/>.

Pang, B., Lee, L. and Vaithyanathan, S. (2002). Thumbs up? *Proceedings of the ACL-02 conference on Empirical methods in natural language processing - EMNLP '02*. [online] doi:<https://doi.org/10.3115/1118693.1118704>.

Pang, B. and Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, [online] 2(1–2), pp.1–135. doi:<https://doi.org/10.1561/1500000011>.

W Gerrod Parrott (2001). *Emotions in social psychology : essential readings*. Philadelphia: Psychology Press.

Purwarianti, A. and Crisdayanti, I.A.P.A. (2019). *Improving Bi-LSTM Performance for Indonesian Sentiment Analysis Using Paragraph Vector*. [online] arXiv.org. doi:<https://doi.org/10.1109/ICAICTA.2019.8904199>.

Pratama, Y., Roberto Tampubolon, A., Diantri Sianturi, L., Diana Manalu, R. and Frietz Pangaribuan, D. (2019). Implementation of Sentiment Analysis on Twitter Using Naïve Bayes Algorithm to Know the People Responses to Debate of DKI Jakarta Governor Election. *Journal of Physics: Conference Series*, 1175, p.012102. doi:<https://doi.org/10.1088/1742-6596/1175/1/012102>.



UNIVERSITAS  
GADJAH MADA

**GAN INDOBERT: Generative Adversarial Learning for Emotion Classification with Limited Labeled Data on Indonesian Tweets**

Norman Muhammad Mikhail, Dzikri Rahadian Fudholi, S.Kom., M.Comp.

Universitas Gadjah Mada, 2023 | Diunduh dari <http://etd.repository.ugm.ac.id/>

Riyadh, Md. and Shafiq, M.O. (2022). GAN-BElectra: Enhanced Multi-class Sentiment Analysis with Limited Labeled Data. *Applied Artificial Intelligence*, 36(1). doi:<https://doi.org/10.1080/08839514.2022.2083794>.

Riyadh, Md. and Shafiq, M.O. (2021). Towards Multi-class Sentiment Analysis With Limited Labeled Data. *2021 IEEE International Conference on Big Data (Big Data)*. doi:<https://doi.org/10.1109/bigdata52589.2021.9671692>.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X. and Chen, X. (2016). *Improved Techniques for Training GANs*. [online] Neural Information Processing Systems. Available at: [https://papers.nips.cc/paper\\_files/paper/2016/hash/8a3363abe792db2d8761d6403605aeb7-Abstract.html](https://papers.nips.cc/paper_files/paper/2016/hash/8a3363abe792db2d8761d6403605aeb7-Abstract.html) [Accessed 2 Nov. 2023].

Saputri, M.S., Mahendra, R. and Adriani, M. (2018). *Emotion Classification on Indonesian Twitter Dataset*. [online] IEEE Xplore. doi:<https://doi.org/10.1109/IALP.2018.8629262>.

Shaver, P., Schwartz, J., Kirson, D. and O'Connor, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of personality and social psychology*, [online] 52(6), pp.1061–86. doi:<https://doi.org/10.1037/0022-3514.52.6.1061>.

Shaver, P.R., Murdaya, U. and Fraley, R.C. (2001). Structure of the Indonesian Emotion Lexicon. *Asian Journal of Social Psychology*, 4(3), pp.201–224. doi:<https://doi.org/10.1111/1467-839x.00086>.



UNIVERSITAS  
GADJAH MADA

**GAN INDOBERT: Generative Adversarial Learning for Emotion Classification with Limited Labeled Data on Indonesian Tweets**

Norman Muhammad Mikhail, Dzikri Rahadian Fudholi, S.Kom., M.Comp.

Universitas Gadjah Mada, 2023 | Diunduh dari <http://etd.repository.ugm.ac.id/>

Shayesteh, S. and Inkpen, D. (2022). Generative Adversarial learning with Negative Data Augmentation for Semi-supervised Text Classification. *The International FLAIRS Conference Proceedings*, 35.

doi:<https://doi.org/10.32473/flairs.v35i.130722>.

Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A. and Potts, C. (2013). *Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank*. [online] ACLWeb. Available at: <https://aclanthology.org/D13-1170/>.

Sinha, A., Ayush, K., Song, J., Uzkent, B., Jin, H. and Ermon, S. (2021). *Negative Data Augmentation*. [online] arXiv.org. doi:<https://doi.org/10.48550/arXiv.2102.05113>.

Ta, H.T., Rahman, A.B.S., Najjar, L. and Gelbukh, A. (2022). GAN-BERT: 2022 Iberian Languages Evaluation Forum, IberLEF 2022. *CEUR Workshop Proceedings*, [online] 3202. Available at: <https://ipn.elsevierpure.com/en/publications/gan-bert-adversarial-learning-for-detection-of-aggressive-and-vio> [Accessed 2 Nov. 2023].

Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, [online] 37(2), pp.267–307. doi:[https://doi.org/10.1162/coli\\_a\\_00049](https://doi.org/10.1162/coli_a_00049).



Tanvir, R., Shawon, M.T.R., Mehedi, M.H.K., Mahtab, M.M. and Rasel, A.A. (2022). A GAN-BERT Based Approach for Bengali Text Classification with a Few Labeled Examples. *easychair.org*. [online] Available at: <https://easychair.org/publications/preprint/hp4Q> [Accessed 2 Nov. 2023].

Triawati, C. (2009). *Metode Pembobotan Statistical Concept Based untuk Klastering dan Kategorisasi Dokumen Berbahasa Indonesia Statistical Concept Based Weighting Method for Indonesian Language Document Clustering and Categorisation*. [online] *openlibrary.telkomuniversity.ac.id*. Universitas Telkom. Available at: <https://openlibrary.telkomuniversity.ac.id/pustaka/94970/metode-pembobtan-statistical-concept-based-untuk-klastering-dan-kategorisasi-dokumen-berbahasa-indonesia-statistical-concept-based-weighting-method-for-indonesian-language-document-clustering-and-categorisation.html> [Accessed 2 Nov. 2023].

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I. (2017). *Attention is All you Need*. [online] Neural Information Processing Systems. Available at: [https://papers.nips.cc/paper\\_files/paper/2017/hash/3f5ee243547dee91fb053c1c4a845aa-Abstract.html](https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fb053c1c4a845aa-Abstract.html).

Wen, J., Kozak, M., Yang, S. and Liu, F. (2020). COVID-19: potential effects on Chinese citizens' lifestyle and travel. *Tourism Review*, 76(1). doi:<https://doi.org/10.1108/tr-03-2020-0110>.

Williams, A., Nangia, N. and Bowman, S.R. (2018). A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference. *arXiv:1704.05426 [cs]*. [online] Available at: <https://arxiv.org/abs/1704.05426>.



UNIVERSITAS  
GADJAH MADA

**GAN INDOBERT: Generative Adversarial Learning for Emotion Classification with Limited Labeled Data on Indonesian Tweets**

Norman Muhammad Mikhail, Dzikri Rahadian Fudholi, S.Kom., M.Comp.

Universitas Gadjah Mada, 2023 | Diunduh dari <http://etd.repository.ugm.ac.id/>

Wilie, B., Vincentio, K., Winata, G.I., Cahyawijaya, S., Li, X., Lim, Z.Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S. and Purwarianti, A. (2020).

*IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding.* [online] ACLWeb. Available at: <https://aclanthology.org/2020.aacl-main.85/>.

Yi, X., Walia, E. and Babyn, P. (2019). Generative adversarial network in medical imaging: A review. *Medical Image Analysis*, 58, p.101552. doi:<https://doi.org/10.1016/j.media.2019.101552>.

Zhang, X., Zhao, J. and LeCun, Y. (2015). *Character-level Convolutional Networks for Text Classification*. [online] Neural Information Processing Systems. Available at: [https://papers.nips.cc/paper\\_files/paper/2015/hash/250cf8b51c773f3f8dc8b4be867a9a02-Abstract.html](https://papers.nips.cc/paper_files/paper/2015/hash/250cf8b51c773f3f8dc8b4be867a9a02-Abstract.html).

Zhu, X. and Goldberg, A.B. (2009). *Introduction to Semi-Supervised Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning*. Cham: Springer International Publishing. doi:<https://doi.org/10.1007/978-3-031-01548-9>.

Zhuang, Q., Zhu, L. and Lin, S. (2022). A GAN-Bert Based Fault Diagnosis Model for CBTC Data Communication Systems Using Edge-to-edge Collaboration Training. *ICC 2022 - IEEE International Conference on Communications*. doi:<https://doi.org/10.1109/icc45855.2022.9838770>.