



Personality Detection On Twitter User With RoBERTa

Rianda Khusuma^{1*}, Warih Maharani², Prati Hutari Gani³

School of Computing, Informatics, Telkom University, Bandung, Indonesia

Email: ^{1*}riandakhusuma@student.telkomuniversity.ac.id, ²wmaharani@telkomuniversity.ac.id,

³pratihutarigani@telkomuniversity.ac.id

Correspondence Author Email: riandakhusuma@student.telkomuniversity.ac.id

REFERENCES

- [1] D. J. Holman and D. J. Hughes, "Transactions between Big-5 personality traits and job characteristics across 20 years," *J Occup Organ Psychol*, vol. 94, no. 3, pp. 762–788, Sep. 2021, doi: 10.1111/joop.12332.
- [2] D. T. Alidemi and F. Fejza, "Theories Of Personality: A Literature Review," *International Journal of Progressive Sciences and Technologies (IJPSAT)*, vol. 25, no. 2, pp. 194–200, 2021, [Online]. Available: <http://ijpsat.ijsht-journals.org>
- [3] K. Simon, "DIGITAL 2022: INDONESIA," Feb. 15, 2022. <https://datareportal.com/reports/digital-2022-indonesia> (accessed Jan. 22, 2023).
- [4] H. J. Kawekas, "Application of Social Media Twitter as a Strategy for Government's Transparency: Study on #Kemala Jateng Program," *Forum Ilmu Sosial*, vol. 47, no. 1, pp. 1–7, 2020, doi: 10.15294/fis.v47i1.23424.
- [5] N. Hutagalung, "Klasifikasi Tipe Kepribadian Pengguna Sosial Media Berdasarkan Teori BIG Five Menggunakan K-Nearest Neighbor," Skripsi Sarjana, Universitas Sumatera Utara, Medan, 2018.
- [6] W. Bleidorn and C. James, "Using Machine Learning to Advance Personality Assessment and Theory," *Personality and Social Psychology Review*, vol. 23, no. 2, pp. 190–203, 2019, doi: 10.1177/1088868318772990.
- [7] Md. T. Zumma, J. A. Munia, D. Halder, and Md. S. Rahman, "Personality Prediction from Twitter Dataset using Machine Learning," in *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2022, pp. 1–5. doi: 10.1109/ICCCNT54827.2022.9984495.
- [8] Y. Liu *et al.*, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," *CoRR*, vol. abs/1907.11692, 2019, [Online]. Available: <http://arxiv.org/abs/1907.11692>
- [9] M. Hercog, P. Jaroński, J. Kolanowski, P. Mieczynski, D. Wiśniewski, and J. Potoniec, "Sarcastic RoBERTa: A RoBERTa-Based Deep Neural Network Detecting Sarcasm on Twitter," in *Big Data Analytics and Knowledge Discovery*, 2022, pp. 46–52.
- [10] H. Jiang, X. Zhang, and J. D. Choi, "Automatic Text-based Personality Recognition on Monologues and Multiparty Dialogues Using Attentive Networks and Contextual Embeddings," *CoRR*, vol. abs/1911.09304, 2019, [Online]. Available: <http://arxiv.org/abs/1911.09304>
- [11] H. Christian, D. Suhartono, A. Chowanda, and K. Z. Zamli, "Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging," *J Big Data*, vol. 8, no. 1, p. 68, 2021, doi: 10.1186/s40537-021-00459-1.
- [12] D. Lu, "Masked Reasoner at SemEval-2020 Task 4: Fine-Tuning RoBERTa for Commonsense Reasoning," in *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, Dec. 2020, pp. 411–414. doi: 10.18653/v1/2020.semeval-1.49.
- [13] M. A. Ayub, K. Ahmad, K. Ahmad, N. Ahmad, and A. I. Al-Fuqaha, "NLP Techniques for Water Quality Analysis in Social Media Content," *CoRR*, vol. abs/2112.11441, 2021, [Online]. Available: <https://arxiv.org/abs/2112.11441>
- [14] A. F. Adoma, N.-M. Henry, and W. Chen, "Comparative Analyses of Bert, Roberta, Distilbert, and Xlnet for Text-Based Emotion Recognition," in *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 2020, pp. 117–121. doi: 10.1109/ICCWAMTIP51612.2020.9317379.
- [15] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *CoRR*, vol. abs/1810.04805, 2018, [Online]. Available: <http://arxiv.org/abs/1810.04805>
- [16] D. Dhimi, "Understanding BERT Word Embeddings," *Medium*, Jul. 05, 2020. <https://medium.com/@dhardtidhami/understanding-bert-word-embeddings-7dc4d2ea54ca> (accessed Jan. 22, 2023).
- [17] "flax-community/indonesian-roberta-base," *Huggingface.co*, Dec. 02, 2022. <https://huggingface.co/flax-community/indonesian-roberta-base> (accessed Jan. 22, 2023).
- [18] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.
- [19] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit*, vol. 91, pp. 216–231, 2019, doi: <https://doi.org/10.1016/j.patcog.2019.02.023>.
- [20] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, 2020, doi: 10.1186/s12864-019-6413-7.