

DAFTAR PUSTAKA

- Allport, G. W. 1937. Introduction to personality (4th ed.). Longman
- Alsubhi, S. M., Alhothali, A. M., and Almansour, A. A., 2023, AraBig5: The Big-Five Personality Traits Prediction Using Machine Learning Algorithm on Arabic Tweets, IEEE vol 11, 112526-112534.
- Apriansyah, R., Winarko, E., 2021. *Klasifikasi Kepribadian Pengguna Twitter Berdasarkan Data Tweet Menggunakan Deep Learning, Thesis*. Diambil dari http://etd.repository.ugm.ac.id/home/detail_pencarian/197212
- Bird, S., Klein, E., and Loper, E., 2009, Natural Language Processing with Python.
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T, 2017, Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. https://doi.org/10.1162/tacl_a_00051
- Chaplin, J. P., 2006, *Kamus Lengkap Psikologi*. RajaGrafindo Persada. <https://doi.org/10.1139/b92-225>.
- Chen, T., & Guestrin, C., 2016, XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>.
- Chen, X., Wang, Y., & Smith, P., 2022, Cross-Cultural Personality Assessment using Translated Social Media Posts. *Journal of Cross-Cultural Psychology*, 53(6), 712-728.
- Cortes, C., & Vapnik, V., 1995, Support-vector networks. *Machine Learning*, 20(3),273–297. <https://doi.org/10.1007/BF00994018>
- DataReportal. (2025). Digital 2025: Indonesia — DataReportal – Global Digital Insights. We Are Social & Meltwater. Diakses dari <https://datareportal.com/reports/digital-2025-indonesia>
- Datta, A., Chakraborty, S., and Mukherjee, A., 2023, Personality Detection and Analysis using Twitter Data, arXiv:2309.05497v1.
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K., 2019, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.
- Digman, J. M., 1990, Personality structure: Emergence of the five-factor model.

Annual Review of Psychology, 41(1), 417-440.

Fadaee, M., Bisazza, A., & Monz, C., 2017, Data augmentation for low-resource neural machine translation. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 567–573.

Fieldman, R. S., 2015, *Essential of Understanding Psychology*.
<https://doi.org/10.1007/1-4020-4512-3>

Goldberg, L. R., 1993, The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26-34.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., 2003, A very brief measure of the Big-Five personality domains, *Journal of Research in Personality*.
[https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)

Hastie, T., Tibshirani, R., & Friedman, J., 2019, *The Elements of Statistical Learning*. Springer.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

<https://datareportal.com/reports/digital-2025-indonesia>

Ivcevic, Z., & Mayer, J. D., 2007, Creative Types and Personality: *Imagination, Cognition and Personality*. <https://doi.org/10.2190/0615-6262-g582-853u>

Johnson, J. A., 2014, Development of the IPIP-NEO-120. *Journal of Research in Personality*.

Jurafsky, D., & Martin, J. H., 2020, *Speech and Language Processing* (3rd ed.) Prentice Hall.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S., 2011, “Social Media? Get Serious! Understanding Functional Building Blocks of Social Media”, *Business Horizons*, Vol. 54, No.3, pp. 241-251.

Kobayashi, S., 2018, Contextual augmentation: Data augmentation by words with paradigmatic relations. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics*, 452–457.

Lee, J., Kim, S., & Park, J., 2023, Emotion Detection in Multiple Languages using Translated English Dataset. *International Journal of Computational Linguistics*, 15(2), 89-104.

Liu, Y., Liu, Z., Chua, T. S., & Sun, M., 2015, Topical word embeddings. In

Proceedings of the National Conference on Artificial Intelligence.

- Lucky, H., Nabiilah, G. Z., Jeremy, N. H., and Suhartono, D., 2022, A Three-Order Ensemble Model for User-level Big-Five Personality Prediction on Twitter Dataset, in *International Journal of Intelligent System and Application in Engineering*. *ijisae*, 11(2), pp. 283–292.
- Manning, C. D., 2020, Prabhakar Raghavan, Hinrich Schütze. *Introduction to Information Retrieval* (2nd ed). Cambridge University Press.
- Marston, W.M., 2013, *Emotions of Normal People*. Routledge.
- McCrae, R. R., and Costa, P. T., 1990, *Personality in Adulthood: A Five-Factor Theory Perspective*. Guilford Press.
- McCrae, R. R., and John, O. P., 1992, An introduction to the five-factor model and its applications, *Journal of personality*, vol. 60, no. 2, pp. 175–215.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J., 2013, Efficient Estimation of Word Representations in Vector Space. *arXiv preprint arXiv:1301.3781*.
- Myers, I. B., 1962, *The Myers-Briggs Type Indicator: Manual*. Consulting Psychologists Press.
- Nayak, S., Gupta, M., Suri, H., Kaushik, S., & Kumar, S., 2024, Personality Traits Prediction using DISC, 2024 3rd *International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pp. 1041-1046, <https://doi.org/10.1109/ICAAIC60222.2024.10575728>.
- Ong, V., Rahmanto, A. D., Williem, W., Jeremy, N. H., Suhartono, D., and Andangsari, E. W., 2021, Personality modelling of indonesian twitter users with XGBoost based on the five factor model, *International Journal of Intelligent Engineering and Systems*, vol. 14, pp. 248–261.
- Pennebaker, J. W., & King, L.A., 1999, Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6), 1296-1312. <https://doi.org/10.1037/0022-3514.77.6.1296>.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G., 2003, Psychological aspects of natural language use. *Annual Review of Psychology*, 54, 547–577. <https://doi.org/10.1146/annurev.psych.54.101601.145041>.
- Pham, T. H., Yin, J., Bui, D. T., & Hoang, T. M., 2022, BERT-Based Multi-Label Text Classification for Toxic Comment Detection with Class Imbalance. In *Proceedings of the International Conference on Machine Learning and*

Cybernetics (ICMLC), pages 234-241. IEEE.

- Prasanna, K. R., Bharati, M. G., and Gudivada, D. S., 2023, Ensemble Machine Learning Models in Predicting Personality Traits and Insights using Myers-Briggs Dataset, *2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*.
- Shah, C., Thacker, C., and Patel, Y., 2022, Multi-Label Personality Prediction on Twitter data using Machine Learning, *2022 International Mobile and Embedded Technology Conference (MECON)*.
- Shannon, C. E., 1948, A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Sidi, F., Hanafi, M., & Mahmood, A. N., 2012, Text Normalization in Social Media. *International Conference on Data Mining and Optimization*.
- Song, X., Salcianu, A., Song, Y., Dopson, D., & Zhou, D., 2021, Fast WordPiece Tokenization. *arXiv:2012.15524v3*.
- Sorower, M., 2010, A literature survey on algorithms for multi-label learning. *Oregon State University, Corvallis*, 1-25. Diambil dari <https://people.oregonstate.edu/~sorowem/pdf/Qual-Multilabel-Shahed-CompleteVersion.pdf>
- Sun, C., Qiu, X., Xu, Y., & Huang, X., 2019, How to fine-tune BERT for text classification? *China National Conference on Chinese Computational Linguistics*.
- Sun, X., Liu, P., Zhang, C., & Li, Y., 2020, CB-BERT: A Cost-Based BERT Framework for Imbalanced Text Classification. In *IEEE Access*, vol. 8, pages 143228-143238.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177/0261927X09351676>.
- Thapa, L., Pandey, A., Gupta, D., Deep, A., and Garg, R., 2024, A Framework for Personality Prediction for E-Recruitment Using Machine Learning Algorithms, *2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*.
- Utami, E., Hartanto, A. D., Adi, S., Oyong, I., and Raharjo, S., 2019, Profiling analysis of DISC personality traits based on Twitter posts in Bahasa Indonesia,

Journal of King Saud University - Computer and Information Sciences 34 (2022), 264-269.

- Utami, N. A., Maharania, W., and Atastina, I., 2021, Personality classification of facebook users according to Big-Five personality using svm (support vector machine) method, in *Procedia Computer Science*, vol. 196. Elsevier B.V., pp. 177–184.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I., 2017, Attention is all you need, in *31st Conference on Neural Information Processing Systems*.
- Wang, S., Cui, L., Liu, L., Li, Q., 2020, Personality Traits Prediction Based on Users Digital Footprints in Social Networks via Attention RNN, *2020 IEEE International Conference on Service Computing (SCC)*.
- Wei, J., and Zou, K., 2019, EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks.
- Wijaya, D. T., & Yeniterzi, R., 2011, Understanding Indonesian Microblogs: Part-of-speech tagging for Bahasa Indonesia Twitter Messages. *HTL-NAACL Workshop*
- Wilie, B., Winata, G. I., Cahyawijaya, S., Vincentio, A., Li, X., Kuncoro, A., and Fung, P., 2020, IndoBERT: A Pre-trained Language Model for Indonesian.
- Witten, I. H., Frank, E., & Hall, M. A., 2004, *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier.
- Zhang, M. L., & Zhou, Z. H., 2014, A review on multi-label learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 26(8), 1819-1837.
- Zhou, X., Chen, L., & Li, W., 2023, BERT-based Chinese Personality Recognition from Social Media. *Neural Computing and Applications*, 35(8), 6123-6138.
- Zumma, M. T., Munia, J. A., Halder, D., and Rahman, M. S., 2022, Personality Prediction from Twitter Dataset using Machine Learning, *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 54827.