

DAFTAR PUSTAKA

- [1] D. A. Taylor, "The development of interpersonal relationships: Social penetration processes," *J. Soc. Psychol.*, vol. 75, no. 1, pp. 79–90, 1968.
- [2] A. Carpenter and K. Greene, "Social Penetration Theory," *The International Encyclopedia of Interpersonal Communication*. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 1–4, 01-Dec-2015.
- [3] H. Dutton, K. L. Deane, and P. Bullen, "Exploring the benefits and risks of mentor self-disclosure: relationship quality and ethics in youth mentoring," *Kōtuitui N. Z. J. Soc. Sci. Online*, vol. 17, no. 1, pp. 116–133, 2022.
- [4] M. Skjuve, A. Følstad, K. I. Fostervold, and P. B. Brandtzaeg, "My chatbot companion - a study of human-chatbot relationships," *Int. J. Hum. Comput. Stud.*, vol. 149, no. 102601, p. 102601, 2021.
- [5] Z. Peng and Y. Wan, "Human vs. AI: What makes students prefer to self-disclosure to AI teaching assistant? The effect of psychological safety and relationship norms," *Educ. Inf. Technol. (Dordr.)*, vol. 30, no. 14, pp. 20851–20878, 2025.
- [6] L. L. Chung and J. Kang, "'I'm hurt too': The effect of a chatbot's reciprocal self-disclosures on users' painful experiences," *Arch. Des. Res.*, vol. 36, no. 4, pp. 67–85, 2023.
- [7] Y. Yamazaki, Y. Chiba, T. Nose, and A. Ito, "Construction and analysis of a multimodal chat-talk corpus for dialog systems considering interpersonal closeness," *LREC*, pp. 443–448, 2020.
- [8] H. Chen, J. Zhao, J. Shi, X. Wang, L. Shen, and Y. Sun, "Self-disclosure causality detection: Discover what cause the self-disclosure," in *2023 8th International Conference on Data Science in Cyberspace (DSC)*, 2023, pp. 298–305.
- [9] Sutranggono, A. N., Sarno, R., & Ghozali, I. (2024). Multi-class multi-level classification of mental health disorders based on textual data from social media. *Journal of Information and Communication Technology*, 23(1), 77-104.
- [10] P. Khodabakhshi, M. Mahootchi, and H. Mosadegh, "A novel hybrid clustering and classification framework for multi-level depression detection in social media texts," *Eng. Appl. Artif. Intell.*, vol. 160, no. 111952, p. 111952, 2025.
- [11] H. Kaur and S. Ghosal, "Multi-level depressive disorder detection incorporating supervised contrastive learning and GNN," *Procedia Comput. Sci.*, vol. 283, pp. 584–593, 2026.
- [12] M. Bader, M. Abdelwanis, M. Maalouf, and H. F. Jelinek, "Detecting depression severity using weighted random forest and oxidative stress biomarkers," *Sci. Rep.*, vol. 14, no. 1, p. 16328, 2024.
- [13] U. Naseem, A. G. Dunn, J. Kim, and M. Khushi, "Early identification of depression severity levels on reddit using ordinal classification," in *Proceedings of the ACM Web Conference 2022*, 2022.
- [14] M. Abuhassan *et al.*, "Lexicon-based graph attention for severity estimation of eating disorders on social media," *IEEE Trans. Affect. Comput.*, vol. 17, no. 1, pp. 602–615, 2026.
- [15] M. Mitsios *et al.*, "Improved text emotion prediction using combined valence and arousal ordinal classification," in *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*, 2024.
- [16] V. A. Wijayasekara, K. Martin, N. Wiratunga, S. Massie, and A. Wijekoon, "Few-shot essay grading: Weighted Prototypical Networks for ordinal text classification," in *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2025.

- [17] T. H. Chu, M. Sun, and L. Crystal Jiang, "Self-disclosure in social media and psychological well-being: A meta-analysis," *J. Soc. Pers. Relat.*, vol. 40, no. 2, pp. 576–599, 2023.
- [18] M. Choi and C. L. Toma, "An experiment on the effects of self-disclosure on perceived partner responsiveness and intimacy in zero-acquaintance relationships," *Commun. Stud.*, vol. 73, no. 3, pp. 297–313, 2022.
- [19] Y.-C. Lee, N. Yamashita, Y. Huang, and W. Fu, "'I hear you, I feel you': Encouraging deep self-disclosure through a chatbot," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.
- [20] E. Croes, M. L. Antheunis, and L. He, "You go first: The effects of self-disclosure reciprocity in human-chatbot interactions," in *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, 2023, pp. 1–4.
- [21] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," *Artif. Intell. Rev.*, vol. 26, no. 3, pp. 159–190, 2006.
- [22] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, 1999.
- [23] J. E. van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Mach. Learn.*, vol. 109, no. 2, pp. 373–440, 2020.
- [24] H. Zhang, "The Optimality of Naive Bayes," in *Proceedings of 17th International Florida Artificial Intelligence Research Society Conference, Menlo Park*, 12-14 May 2004, pp. 562–567.
- [25] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [26] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [27] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [28] G. Tsoumakas and I. Katakis, "Multi-Label Classification: An Overview," *Int. J. Data Warehous. Min.*, vol. 3, no. 3, pp. 1–13, 2007.
- [29] P. A. Gutierrez, M. Perez-Ortiz, J. Sanchez-Monedero, F. Fernandez-Navarro, and C. Hervas-Martinez, "Ordinal regression methods: Survey and experimental study," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 1, pp. 127–146, 2016.
- [30] S. Robertson, "Understanding inverse document frequency: on theoretical arguments for IDF," *J. Doc.*, vol. 60, no. 5, pp. 503–520, 2004.
- [31] M. Jabardi, "Support Vector Machines: Theory, algorithms, and applications," *Infocommunications J.*, vol. 17, no. 1, pp. 66–75, 2025.
- [32] A. Vaswani *et al.*, "Attention is all you need," in *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 6000–6010.
- [33] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North*, 2019.
- [34] N. Patwardhan, S. Marrone, and C. Sansone, "Transformers in the real world: A survey on NLP applications," *Information (Basel)*, vol. 14, no. 4, p. 242, 2023.
- [35] T. Lin, Y. Wang, X. Liu, and X. Qiu, "A survey of transformers," *AI Open*, vol. 3, pp. 111–132, 2022.
- [36] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," in *Proceedings of the 28th International Conference on Computational Linguistics*, 2020, pp. 757–770.

- [37] M. A. Nur, N. Umar, Z. Feng, and H. Gani, "Evaluation of IndoBERT and Roberta: Performance of Indonesian language transformer models in sentiment classification," *jiko*, vol. 8, no. 2, pp. 121–127, 2025.
- [38] Z. Qin, D. Kim, and T. Gedeon, "Rethinking softmax with cross-entropy: Neural network classifier as mutual information estimator," *arXiv preprint*, 2019.
- [39] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-Margin Softmax Loss for Convolutional Neural Networks," in *The 33rd International Conference on International Conference on Machine Learning*, 2016, pp. 507–516.
- [40] S. Sun *et al.*, "Sparse-softmax: A simpler and faster alternative softmax transformation," *arXiv preprint*, 2021.
- [41] W. Cao, V. Mirjalili, and S. Raschka, "Rank consistent ordinal regression for neural networks with application to age estimation," *Pattern Recognit. Lett.*, vol. 140, pp. 325–331, 2020.
- [42] J. Terven, D.-M. Cordova-Esparza, J.-A. Romero-González, A. Ramírez-Pedraza, and E. A. Chávez-Urbiola, "A comprehensive survey of loss functions and metrics in deep learning," *Artif. Intell. Rev.*, vol. 58, no. 7, 2025.
- [43] A. Mao, M. Mohri, and Y. Zhong, "Cross-Entropy Loss Functions: Theoretical Analysis and Applications," in *Proceedings of the 40th International Conference on Machine Learning*, 2023, pp. 23803–23828.
- [44] C. Li, K. Liu, and S. Liu, "A survey of loss functions in deep learning," *Mathematics*, vol. 13, no. 15, p. 2417, 2025.
- [45] J. Cohen, "Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit," *Psychol. Bull.*, vol. 70, no. 4, pp. 213–220, 1968.
- [46] J. de la Torre, D. Puig, and A. Valls, "Weighted kappa loss function for multi-class classification of ordinal data in deep learning," *Pattern Recognit. Lett.*, vol. 105, pp. 144–154, 2018.
- [47] S. Haas and E. Hüllermeier, "Uncertainty quantification in ordinal classification: A comparison of measures," *Int. J. Approx. Reason.*, vol. 186, no. 109479, p. 109479, 2025.
- [48] H. Xu, H. He, Y. Zhang, L. Ma, and J. Li, "A comparative study of loss functions for road segmentation in remotely sensed road datasets," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 116, no. 103159, p. 103159, 2023.
- [49] M. Kohler and S. Langer, "Statistical theory for image classification using deep convolutional neural network with cross-entropy loss under the hierarchical max-pooling model," *J. Stat. Plan. Inference*, vol. 234, no. 106188, p. 106188, 2025.
- [50] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.
- [51] A. Tharwat, "Classification assessment methods," *Appl. Comput. Inform.*, vol. 17, no. 1, pp. 168–192, 2021.
- [52] 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.
- [53] J. S. Cardoso and R. Sousa, "Measuring the performance of ordinal classification," *Intern. J. Pattern Recognit. Artif. Intell.*, vol. 25, no. 08, pp. 1173–1195, 2011.
- [54] T. Sakai, "Evaluating evaluation measures for ordinal classification and ordinal quantification," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2021, pp. 2759–2769.

- [55] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006.
- [56] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014.
- [57] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Clim. Res.*, vol. 30, pp. 79–82, 2005.
- [58] J. S. Armstrong and F. Collopy, "Error measures for generalizing about forecasting methods: Empirical comparisons," *Int. J. Forecast.*, vol. 8, no. 1, pp. 69–80, 1992.
- [59] D. R. Legates and G. J. McCabe Jr, "Evaluating the use of 'goodness-of-fit' Measures in hydrologic and hydroclimatic model validation," *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, 1999.
- [60] A. Doewes, N. A. Kurdhi, and A. Saxena, "Evaluating Quadratic Weighted Kappa as the standard performance metric for automated Essay Scoring." Zenodo, 2023.
- [61] A. Galdran, "Performance metrics for probabilistic ordinal classifiers," in *Lecture Notes in Computer Science*, Cham: Springer Nature Switzerland, 2023, pp. 357–366.
- [62] M. Li, Q. Gao, and T. Yu, "Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters," *BMC Cancer*, vol. 23, no. 1, p. 799, 2023.
- [63] J. Hauke and T. Kossowski, "Comparison of values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of data," *Quaest. Geogr.*, vol. 30, no. 2, pp. 87–93, 2011.
- [64] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson Correlation Coefficient," in *Noise Reduction in Speech Processing*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 1–4.
- [65] M. M. Mukaka, "Statistics corner: A guide to appropriate use of correlation coefficient in medical research," *Malawi Med. J.*, vol. 24, no. 3, pp. 69–71, 2012.
- [66] S. Couch, Z. Kazan, K. Shi, A. Bray, and A. Groce, "Differentially Private Nonparametric Hypothesis Testing," *Proc. ACM SIGSAC Conf. Computer and Communications Security (CCS)*, 2019.
- [67] X. Liu, Y. Zhou, Y. Tang, J. Qian, and Y. Zhou, "Human-in-the-loop Online Just-in-Time Software Defect Prediction," *IEEE Transactions on Software Engineering*, 2023.
- [68] S. Kwak, "Are only p-values less than 0.05 significant? A p-value greater than 0.05 is Also significant!," *J. Lipid Atheroscler.*, vol. 12, no. 2, pp. 89–95, 2023.
- [69] A. A. Chamid, Widowati, and R. Kusumaningrum, "Labeling consistency test of multi-label data for aspect and sentiment classification using the Cohen kappa method," *Ing. Syst. D Inf.*, vol. 29, no. 1, pp. 161–167, 2024.
- [70] L. Chen and H. Suominen, "An approach to the frugal use of human annotators to scale up auto-coding for text classification tasks," *ALTA*, pp. 12–21, 2021.
- [71] N. P. Sastra, Linawati, G. Sukadarmika, I. P. G. H. Suputra, N. M. Ariwilani, and N. L. I. D. Swandi, "Dataset on cognitive distortions for text classification in Indonesian language," *Data Brief*, vol. 61, no. 111836, p. 111836, 2025.