REFERENCES

[1] Worldometer. World population clock: 7.9 billion people (2021). [Online] Available at: https://www.worldometers.info/world-population/ [Accessed 16 June 2022]

[2] Badan Pusat Statistik Indonesia. Sensus penduduk 2020. [Online] Available at: https://www.bps.go.id/indicator/12/1886/1/jumlah-penduduk-hasil-proyeksi-menurut-provinsi-dan-jenis-kelamin.html [Accessed 16 June 2022]

[3] Kementrian Energi dan Sumber Daya Mineral. Rencana umum ketena- galistrikan nasional 2019-2038. [Online] Available at: https://gatrik.esdm.go.id/assets/uploads/download_index/files/0e997-rukn-2019-2038.pdf [Accessed 16 June 2022]

[4] PT. PLN (Persero). Laporan statistik - pt pln (persero). [Online] Available at: https://web.pln.co.id/stakeholder/laporan-statistik [Accessed 16 June 2022]

[5] Neo Wu, Bradley Green, Xue Ben, and Shawn O'Banion. 2020. Deep transformer models for time series forecasting: The influenza prevalence case. arXiv preprint arXiv:2001.08317.

[6] Albert Zeyer, Parnia Bahar, Kazuki Irie, Ralf Schlu¨ter, and Hermann Ney. 2019. A comparison of transformer and lstm encoder decoder models for asr. In 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), pages 8–15. IEEE.

[7] Hongfei Xu, Qiuhui Liu, Deyi Xiong, and Josef van Genabith. 2020. Transformer with depth-wise lstm. arXiv preprint arXiv:2007.06257.

[8] Jay Heizer and Banry Rander. 2001. Operations management sidh edition. New Jersey: Prentice Hall.

[9] Rob J Hyndman and George Athanasopoulos. 2018. Forecasting: principles and practice. OTexts.

[10] Musaed Alhussein, Khursheed Aurangzeb, and Syed Irtaza Haider. 2020. Hybrid cnn-lstm model for short-term individual household load fore- casting. Ieee Access, 8:180544–180557.

[11] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. Advances in neural information processing systems, 30.

[12] N. Parmar, A. Vaswani, J. Uszkoreit, Łukasz Kaiser, N. Shazeer, A. Ku, and D. Tran, "Image transformer," 2018.

[13] D. Povey, H. Hadian, P. Ghahremani, K. Li, and S. Khudanpur, "A time-restricted self-attention layer for asr," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018, pp. 5874–5878.

[14] P. J. Liu, M. Saleh, E. Pot, B. Goodrich, R. Sepassi, L. Kaiser, and N. Shazeer, "Generating wikipedia by summarizing long sequences," 2018.

[15] C.-Z. A. Huang, A. Vaswani, J. Uszkoreit, N. Shazeer, I. Simon, C. Hawthorne, A. M. Dai, M. D. Hoffman, M. Dinculescu, and D. Eck, "Music transformer," 2018.

[16] Andemariam Mebrahtu Tesfagergis. 2021. Transformer networks for short- term forecasting of electricity prosumption.

[17] Zhijun Zhao, Shaw Chen Liu, Run Liu, Ziyin Zhang, Yanzi Li, Huisi Mo, and Yanxing Wu. 2021. Contribution of climate/meteorology to winter haze pollution in the fenwei plain, china. International Journal of Climatology, 41(10):4987–5002.

[18] Davide Chicco, Matthijs J Warrens, and Giuseppe Jurman. 2021. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. PeerJ Computer Science, 7:e623.

[19] Tianfeng Chai and Roland R Draxler. 2014. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. Geoscientific model development, 7(3):1247–1250.

[20] Hair, Joseph F. Multivariate Data Analysis: A Global Perspective. 7th ed. Upper Saddle River: Prentice Hall, 2009. Print.