DAFTAR PUSTAKA

Ahmad, M. W., Mourshed, M. and Rezgui, Y. (2017) 'Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption', *Energy and Buildings*. Elsevier B.V., 147, pp. 77–89. doi: 10.1016/j.enbuild.2017.04.038.

Ahmad, M. W., Reynolds, J. and Rezgui, Y. (2018) 'Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees', *Journal of Cleaner Production*. Elsevier Ltd, 203, pp. 810–821. doi: 10.1016/j.jclepro.2018.08.207.

Aljuaid, T. and Sasi, S. (2017) 'Proper imputation techniques for missing values in data sets', *Proceedings of the 2016 International Conference on Data Science and Engineering, ICDSE 2016*. doi: 10.1109/ICDSE.2016.7823957.

Andreas Chandra (2017) *PERBEDAAN SUPERVISED AND UNSUPERVISED LEARNING*. Available at: https://datascience.or.id/article/Perbedaan-Supervised-and-Unsupervised-Learning-5a8fa6e6.

Appbrain (2019) *Number of Android apps on Google Play*. Available at: https://www.appbrain.com/stats/number-of-android-apps.

Batista, G. and Monard, M. C. (2002) 'A Study of K-Nearest Neighbour as an Imputation Method', *HIS'02: 2nd International Conference on Hybrid Intelligent Systems*, pp. 251–260. Available at: http://conteudo.icmc.usp.br/pessoas/gbatista/files/his2002.pdf.

Bindu, K. H., BhanuJyothi, K. and Suryanarayana, D. (2017) 'A Comparative Study of Random Forest & K – Nearest Neighbors on HAR dataset Using Caret', *International Journal of Innovative Research in Technology*, 3(9), pp. 6–9. Available at: http://ijirt.org/Article?manuscript=144228.

Bishop, C. M. (2006) *Pattern recognition and machine learning*. springer.

Breiman, L. (2001) 'Random Forests', *Machine Learning*, 45(1), pp. 5–32. doi: 10.1023/A:1010933404324.

Burkov, A. (2019) 'The Hundred-Page Machine Learning Book-Andriy Burkov', *Expert Systems*, 5(2), pp. 132–150. doi: 10.1111/j.1468-0394.1988.tb00341.x.

Chand, N. *et al.* (2016) 'A comparative analysis of SVM and its stacking with other classification algorithm for intrusion detection', *Proceedings - 2016 International Conference on Advances in Computing, Communication and Automation, ICACCA 2016*. doi: 10.1109/ICACCA.2016.7578859.

Chen, W. *et al.* (2017) 'A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility', *Catena*. Elsevier B.V., 151, pp. 147–160. doi: 10.1016/j.catena.2016.11.032.

Dong, Y. and Peng, C. Y. J. (2013) 'Principled missing data methods for researchers', *SpringerPlus*, 2(1), pp. 1–17. doi: 10.1186/2193-1801-2-222.

Effendi, J. and M Jorgi, R. (2018) 'Analisis Cluster Aplikasi pada Google play Store dengan Menggunakan Metode K-Mean', 4(1), pp. 978–979.

google (2012a) *Introducing Google Play: All your entertainment, anywhere you go*, *googleblog*. Available at: https://googleblog.blogspot.com/2012/03/introducing-google-play-all-your.html.

google (2012b) *Introducing Google Play: All your entertainment, anywhere you go*, *googleblog*.

Hartmann-Boyce, J. *et al.* (2017) 'Insights From Google Play Store User Reviews for the Development of Weight Loss Apps: Mixed-Method Analysis', *JMIR mHealth and uHealth*, 5(12), p. e203. doi: 10.2196/mhealth.8791.

He, Y. and Pi, D. C. (2016) 'Improving KNN method based on reduced relational grade for microarray missing values imputation', *IAENG International Journal of Computer Science*, 43(3), pp. 356–362.

Hengshu Zhu *et al.* (2014) 'Popularity Modeling for Mobile Apps: A Sequential Approach', *IEEE Transactions on Cybernetics*, 45(7), pp. 1303–1314. doi: 10.1109/tcyb.2014.2349954.

Islam, M. J. *et al.* (2010) 'Investigating the Performance of Naive-Bayes Classifiers and K-Nearest Neighbor Classifiers Security and Trust for IoT View project Hybrid memristor-CMOS based linear feedback shift register design View project Investigating the Performance of Naïve-Bayes', *Article in Journal of Convergence Information Technology*, 5(2), pp. 1541–1546. doi: 10.1109/ICCIT.2007.148.

Jadhav, A., Pramod, D. and Ramanathan, K. (2019) 'Comparison of Performance of Data Imputation Methods for Numeric Dataset', *Applied Artificial Intelligence*. Taylor & Francis, 33(10), pp. 913–933. doi: 10.1080/08839514.2019.1637138.

Kandhasamy, J. P. and Balamurali, S. (2014) 'Performance analysis of classifier models to predict diabetes mellitus', *Procedia Computer Science*. Elsevier Masson SAS, 47(C), pp. 45–51. doi: 10.1016/j.procs.2015.03.182.

Kulkarni, V. Y. and Sinha, P. K. (2014) 'Effective Learning and Classification using Random Forest Algorithm', *International Journal of Engineering and Innovative Technolgy*, 3(11), pp. 267–273.

Larose, D. T. (2005) *Discovering knowledge in data*.

Minakshi, Vohra, R. and Gimpy (2014) 'Missing Value Imputation in Multi Attribute Data Set', *International Journal of Computer Science and Information Technologies*, 5(4), pp. 5315–5321.

Rodriguez-Galiano, V. *et al.* (2015) 'Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines', *Ore Geology Reviews*. Elsevier B.V., 71, pp. 804–818. doi: 10.1016/j.oregeorev.2015.01.001.

Shen, S., Lu, X. and Hu, Z. (2017) 'Towards Release Strategy Optimization for Apps in Google Play'. Available at: http://arxiv.org/abs/1707.06022.

Socialresearchmethods.net (2019) *Dummy Variables*.

statlab.id (2018) *Data Hilang (Missing Value)*. Available at: https://statlab.id/data-hilang-missing-value/.

Susanti, Martha, S. and Sulistianingsih, E. (2018) 'K Nearest Neighbor dalam Imputasi Missing Data', *Bimaster*, 07(1), pp. 1–18.

Wael M. Khedr and Ahmed M. Elshewey (2013) 'Pattern Classification for Incomplete Data Using PPCA and KNN', 4(8), pp. 628–632.

Zhu, H. *et al.* (2014) 'Mobile App Recommendations with Security and Privacy Awareness Categories and Subject Descriptors', *Proc. of the 20th ACM SIGKDD international conference on Knowledge Discovery and Data mining (KDD)*, pp. 951–960. doi: 10.1145/2623330.2623705.