# DAFTAR PUSTAKA

Abou Omar, K. B. (2018). XGBoost and LGBM for Porto Seguro's Kaggle challenge: A comparison. *Preprint Semester Project*, 1–20.

Al-Jawaldeh, A., & Abbass, M. M. S. (2022). Unhealthy Dietary Habits and Obesity: The Major Risk Factors Beyond Non-Communicable Diseases in the Eastern Mediterranean Region. *Frontiers in Nutrition*, *9*(March). https://doi.org/10.3389/fnut.2022.817808

Bray, G. A. (2004). Medical consequences of obesity. *Journal of Clinical Endocrinology and Metabolism*, *89*(6), 2583–2589. https://doi.org/10.1210/jc.2004-0535

Brown, J. D. (2001). Biserial Correlation Coefficients. *JLT Testing & Evlution SIG Newsletter*, *5*(3), 13–17.

Chandak, G. (2017). *Binary Classification*. https://medium.com/learning-machine-learning-by-gc-nit/binary-classification-35590076ca69

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*(Sept. 28), 321–357. https://doi.org/10.1613/jair.953

Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2020). A survey on ensemble learning. *Frontiers of Computer Science*, *14*(2), 241–258. https://doi.org/10.1007/s11704-019-8208-z

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Garreta, R., & Moncecchi, G. (2021). *Learning scikit-learn : Machine Learning in Python*.

Ha, J., Kambe, M., & Pe, J. (2011). Data Mining: Concepts and Techniques. In *Data Mining: Concepts and Techniques*. https://doi.org/10.1016/C2009-0-61819-5

ImbLearn. (2024). *Imbalanced Learn Documentation 0.12.3*. https://imbalanced-learn.org/stable/

Islahulhaq, Wibowo, W., & Ratih, I. D. (2021). Classification of non-performing financing using logistic regression and synthetic minority over-sampling technique-nominal continuous (SMOTE-NC). *International Journal of Advances in Soft Computing and Its Applications*, *13*(3), 115–128. https://doi.org/10.15849/ijasca.211128.09

Jindal, K., Baliyan, N., & Rana, P. S. (2018). Obesity prediction using ensemble machine learning approaches. In *Advances in Intelligent Systems and Computing* (Vol. 708). Springer Singapore. https://doi.org/10.1007/978-981-10-8636-6_37

Jupyter. (2015). *The Jupyter Notebook*. https://jupyter-notebook.readthedocs.io/en/stable/notebook.html

Koyejo, O., Natarajan, N., Ravikumar, P., & Dhillon, I. S. (2015). Consistent multilabel classification. *Advances in Neural Information Processing Systems*, *2015-Janua*,

3321–3329.

Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). *Multi-label Classifier Performance Evaluation with Confusion Matrix*. 01–14. https://doi.org/10.5121/csit.2020.100801

Machado, M. R., Karray, S., & De Sousa, I. T. (2019). LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry. *14th International Conference on Computer Science and Education, ICCSE 2019*, *Nips*, 1111–1116. https://doi.org/10.1109/ICCSE.2019.8845529

Marsland, S. (2011). *Esmeble Method Foundation* (pp. 1–232).

Matplotlib. (2007). *Matplotlib 3.9.0 documentation*. https://matplotlib.org/stable/index.html

Microsoft. (2016). *LightGBM Documentation*. https://lightgbm.readthedocs.io

Mondal, P. K., Foysal, K. H., Norman, B. A., & Gittner, L. S. (2023). Predicting Childhood Obesity Based on Single and Multiple Well-Child Visit Data Using Machine Learning Classifiers. *Sensors*, *23*(2). https://doi.org/10.3390/s23020759

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, *7*(DEC). https://doi.org/10.3389/fnbot.2013.00021

Nosrati, M. (2011). Python: An appropriate language for real world programming. *World Applied Programming*, *1*(2), 110–117. www.waprogramming.com

NumPy. (2008). *NumPy Documentations*. https://numpy.org/doc/stable/

Nuttall, F. Q. (2015). Body mass index: Obesity, BMI, and health: A critical review. *Nutrition Today*, *50*(3), 117–128. https://doi.org/10.1097/NT.0000000000000092

Nyuytiymbiy, K. (2020). *Parameters and Hyperparameters in Machine Learning and Deep Learning*. https://towardsdatascience.com/parameters-and-hyperparameters-aa609601a9ac

Pandas. (2024). *Pandas Documentation*. https://pandas.pydata.org/docs/

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Phelps, N. H., Singleton, R. K., & Zhou, B. (2023). Worldwide trends in the burden of underweight and obesity from 1990 to 2022: a pooled analysis of 3663 population-representative studies with 222 million school-aged children, adolescents and adults. *The Lancet*, *6736*(23), 1–24. https://doi.org/10.1016/S0140-6736(23)02750-2

Powers, D. M. W. (2020). *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*. 37–63. http://arxiv.org/abs/2010.16061

Prasad, A. (2021). *Regression Trees | Decision Tree for Regression | Machine Learning*. https://medium.com/analytics-vidhya/regression-trees-decision-tree-for-regression-

machine-learning-e4d7525d8047

Rufo, D. D., Debelee, T. G., Ibenthal, A., & Negera, W. G. (2021). Diagnosis of diabetes mellitus using gradient boosting machine (Lightgbm). *Diagnostics*, *11*(9), 1–14. https://doi.org/10.3390/diagnostics11091714

Ryan Hoens, T., & Chawla, N. V. (2013). Imbalanced datasets: From sampling to classifiers. *Imbalanced Learning: Foundations, Algorithms, and Applications*, 43–59. https://doi.org/10.1002/9781118646106.ch3

Sasaki, Y. (2007). The truth of the F-measure. *Teach Tutor Mater*, 1–5. http://www.cs.odu.edu/~mukka/cs795sum09dm/Lecturenotes/Day3/F-measure-YS-26Oct07.pdf

Seaborn. (2022). *An introduction to seaborn — seaborn 0.12.1 documentation*. https://seaborn.pydata.org/tutorial/introduction

Soofi, A. A., & Awan, A. (2017). Classification Techniques in Machine Learning: Applications and Issues. *Journal of Basic & Applied Sciences*, *13*, 459–465.

Teoh, T. T., & Rong, Z. (2022). *Python for Data Analysis*. https://doi.org/10.1007/978-981-16-8615-3_7

Thamrin, S. A., Arsyad, D. S., Kuswanto, H., Lawi, A., & Nasir, S. (2021). Predicting Obesity in Adults Using Machine Learning Techniques: An Analysis of Indonesian Basic Health Research 2018. *Frontiers in Nutrition*, *8*(June). https://doi.org/10.3389/fnut.2021.669155

Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*, *17*(1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003

Uslu, Ç. (2022). *What is Kaggle?* https://www.datacamp.com/blog/what-is-kaggle

Visual Studio Code. (2024). *Visual Studio Code Documentation - Overview*. https://code.visualstudio.com/docs

WHO. (2010). *A Healthy Lifestyle - WHO Recommendations*. https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle---who-recommendations

WHO. (2024a). *Obesity*. https://www.who.int/health-topics/obesity

WHO. (2024b). *Obesity and overweight*. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Yadav, D. (2019). *Categorical encoding using Label-Encoding and One-Hot-Encoder*. https://towardsdatascience.com/categorical-encoding-using-label-encoding-and-one-hot-encoder-911ef77fb5bd

Zhang, D., & Gong, Y. (2020). The Comparison of LightGBM and XGBoost Coupling Factor Analysis and Prediagnosis of Acute Liver Failure. *IEEE Access*, *8*, 220990–221003. https://doi.org/10.1109/ACCESS.2020.3042848