REFERENCES

Albright, J. (2019). Forecasting the progression of Alzheimer's disease using neural networks and a novel preprocessing algorithm. Alzheimer's & Dementia: Translational Research & Clinical Interventions, 5(1), 483–491. https://doi.org/10.1016/j.trci.2019.07.001

Almuallim, H. (1996). An efficient algorithm for optimal pruning of decision trees. Artificial Intelligence, 83(2), 347–362. https://doi.org/10.1016/0004-3702(95)00060-7

Alzheimer's Association. (2023). 2023 Alzheimer's disease facts and figures. Alzheimer's & Dementia, 19, 1598-1695. https://doi.org/10.1002/alz.13016

An, N., Ding, H., Yang, J., Au, R., & Ang, T. F. A. (2020). Deep Ensemble Learning for Alzheimer's Disease Classification. Journal of Biomedical Informatics, 105, 103411. https://doi.org/10.1016/j.jbi.2020.103411

Bach, M., Werner, A., & Palt, M. (2019). The proposal of undersampling method for learning from imbalanced datasets. *Procedia Computer Science*, *159*, 125–134. https://doi.org/10.1016/j.procs.2019.09.167

Balamurugan, M., Nancy, A., & Vijaykumar, S. (2017). Alzheimer's disease diagnosis by using dimensionality reduction based on Knn classifier. Biomedical and Pharmacology Journal, 10(4), 1823–1830. https://doi.org/10.13005/bpj/1299

Bartz, E. (2023). Hyperparameter tuning for machine and Deep Learning with R. Springer Nature Singapore.

Breijyeh, Z., & Karaman, R. (2020). Comprehensive review on Alzheimer's disease: Causes and treatment. Molecules (Basel, Switzerland), 25(24), 5789. https://doi.org/10.3390/molecules25245789

Das, C., Sahoo, A. K., & Pradhan, C. (2022). Multicriteria recommender system using different approaches. Cognitive Big Data Intelligence with a Metaheuristic Approach, 259–277. https://doi.org/10.1016/b978-0-323-85117-6.00011-x

Devarriya, D., Gulati, C., Mansharamani, V., Sakalle, A., & Bhardwaj, A. (2020). Unbalanced breast cancer data classification using novel fitness functions in Genetic Programming. Expert Systems with Applications, 140, 112866. https://doi.org/10.1016/j.eswa.2019.112866

Elhassan, A. T., Aljourf, M., Al-Mohanna, F., & Shoukri, M. (2017). Classification of Imbalance Data using Tomek Link (T-Link) Combined with Random Under-sampling (RUS) as a Data Reduction Method. Global Journal of Technology and Optimization, S1, 111. https://doi.org/10.4172/2229-8711.S1:111

George, N. (2021). Practical data science with Python: Learn tools and techniques from hands-on examples to extract insights from data. Packt Publishing.

Géron, A. (2023). Hands-on machine learning with scikit-learn, Keras and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly.

Gustavsson, A., Norton, N., Fast, T., et al. (2023). Global estimates on the number of persons across the Alzheimer's disease continuum. Alzheimer's Dement, 19, 658–670. https://doi.org/10.1002/alz.12694

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. Electron Markets, 31, 685–695. https://doi.org/10.1007/s12525-021-00475-2

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Jolly, K. (2018). Machine learning with scikit-learn quick start guide: Classification, regression, and clustering techniques in Python. Packt.

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Deep learning and data sampling with imbalanced Big Data. *2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI)*. https://doi.org/10.1109/iri.2019.00038

Leevy, J. L., Hancock, J., Khoshgoftaar, T. M., & Abdollah Zadeh, A. (2023). Investigating the effectiveness of one-class and binary classification for fraud detection. *Journal of Big Data*, *10*(1). https://doi.org/10.1186/s40537-023-00825-1

Lin, M., Gong, P., Yang, T., Ye, J., Albin, R. L., & Dodge, H. H. (2018). Big Data Analytical Approaches to the NACC Dataset: Aiding Preclinical Trial Enrichment. Alzheimer Disease and Associated Disorders, 32(1), 18–27. https://doi.org/10.1097/WAD.0000000000000228

Liu, S., Masurkar, A. V., Rusinek, H., Chen, J., Zhang, B., Zhu, W., Fernandez-Granda, C., & Razavian, N. (2022). Generalizable deep learning model for early Alzheimer's disease detection from structural MRIs. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-20674-x

Lu, S.-C., Swisher, C. L., Chung, C., Jaffray, D., & Sidey-Gibbons, C. (2023). On the importance of interpretable machine learning predictions to inform clinical decision making in oncology. Frontiers in Oncology, 13. https://doi.org/10.3389/fonc.2023.1129380

Mitra, S., Saha, S., & Hasanuzzaman, M. (2020). A multi-view deep neural network model for chemical-disease relation extraction from imbalanced datasets. IEEE Journal of Biomedical and Health Informatics, 24(11), 3315–3325. https://doi.org/10.1109/jbhi.2020.2983365

Mohammad, H., & N, S. M. (2015). A review on evaluation metrics for data classification evaluations. International Journal of Data Mining & Knowledge Management Process, 5(2), 01–11. https://doi.org/10.5121/ijdkp.2015.5201

Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine learning with oversampling and undersampling techniques: Overview study and experimental results. 2020 11th International Conference on Information and Communication Systems (ICICS). https://doi.org/10.1109/icics49469.2020.239556

Müller, A. C., & Guido, S. (2016). Introduction to machine learning with Python. O'Reilly & Associates Inc.

National Alzheimer's Coordinating Center. (2023). National Alzheimer's Coordinating Center. https://naccdata.org/

National Health Service UK. (2023). Causes Alzheimer's Disease. NHS choices. https://www.nhs.uk/conditions/alzheimers-disease/causes/

Pang, Y., Kukull, W., Sano, M., Albin, R. L., Shen, C., Zhou, J., & Dodge, H. H. (2023). Predicting progression from normal to MCI and from MCI to AD using clinical variables in the National Alzheimer's Coordinating Center Uniform Data Set Version 3: Application of machine learning models and a probability calculator. The Journal of Prevention of Alzheimer's Disease. https://doi.org/10.14283/jpad.2023.10

Petch, J., Di, S., & Nelson, W. (2022). Opening the black box: The promise and limitations of explainable machine learning in cardiology. Canadian Journal of Cardiology, 38(2), 204–213. https://doi.org/10.1016/j.cjca.2021.09.004

Reddy, G. T., Reddy, M. P., Lakshmanna, K., Kaluri, R., Rajput, D. S., Srivastava, G., & Baker, T. (2020). Analysis of dimensionality reduction techniques on Big Data. IEEE Access, 8, 54776–54788. https://doi.org/10.1109/access.2020.2980942

Silva, M. V., Loures, C. de, Alves, L. C., de Souza, L. C., Borges, K. B., & Carvalho, M. das. (2019). Alzheimer's disease: Risk factors and potentially protective measures. Journal of Biomedical Science, 26(1). https://doi.org/10.1186/s12929-019-0524-y

Shahabadi, M. S. E., Tabrizchi, H., Kuchaki Rafsanjani, M., Gupta, B. B., & Palmieri, F. (2021). A combination of clustering-based under-sampling with ensemble methods for solving imbalanced class problem in Intelligent Systems. *Technological Forecasting and Social Change*, *169*, 120796. https://doi.org/10.1016/j.techfore.2021.120796

Song, Y. Y., & Lu, Y. (2015). Decision tree methods: Applications for classification and prediction. Shanghai Archives of Psychiatry, 27(2), 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044

Wang, K., Tian, J., Zheng, C., Yang, H., Ren, J., Li, C., Han, Q., & Zhang, Y. (2021). Improving risk identification of adverse outcomes in chronic heart failure using SMOTE+ENN and machine learning. *Risk Management and Healthcare Policy*, *Volume 14*, 2453–2463. https://doi.org/10.2147/rmhp.s310295

Wang, Z., Wu, C., Zheng, K., Niu, X., & Wang, X. (2019). Smotetomek-based resampling for personality recognition. *IEEE Access*, *7*, 129678–129689. https://doi.org/10.1109/access.2019.2940061

Yuan, M., & Kennedy, K. M. (2023). Utility of environmental complexity as a predictor of Alzheimer's disease diagnosis: A big-data machine learning approach. The Journal of Prevention of Alzheimer's Disease. https://doi.org/10.14283/jpad.2023.18

Zhang, X.-X., Tian, Y., Wang, Z.-T., Ma, Y.-H., Tan, L., & Yu, J.-T. (2021). The epidemiology of Alzheimer's disease modifiable risk factors and prevention. The Journal of Prevention of Alzheimer's Disease, 1–9. https://doi.org/10.14283/jpad.2021.15

Zheng, M., Li, T., Zheng, X., Yu, Q., Chen, C., Zhou, D., Lv, C., & Yang, W. (2021). UFFDFR: Undersampling Framework with denoising, fuzzy C-means clustering, and representative sample selection for Imbalanced Data Classification. Information Sciences, 576, 658–680. https://doi.org/10.1016/j.ins.2021.07.053

Zheng, M., Wang, F., Hu, X., Miao, Y., Cao, H., & Tang, M. (2022). A method for analyzing the performance impact of imbalanced binary data on machine learning models. *Axioms*, *11*(11), 607. https://doi.org/10.3390/axioms11110607