**Supplementary Material**

1. Abdollahi et al. (2024). An explainable artificial-intelligence-aided safety factor prediction of road embankments. Engineering Applications of Artificial Intelligence, 136(Part A), 108854. https://doi.org/10.1016/j.engappai.2024.108854
2. Abdulrauf Sharifai, G., & Zainol, Z. (2020). Feature Selection for High-Dimensional and Imbalanced Biomedical Data Based on Robust Correlation Based Redundancy and Binary Grasshopper Optimization Algorithm. Genes, 11(7), 717. https://doi.org/10.3390/genes11070717
3. Adeyeye et al. (2025). Integrating partial least square structural equation modelling and machine learning for causal exploration of environmental phenomena. Environmental Research, 274, 121358. Advance online publication. https://doi.org/10.1016/j.envres.2025.121358
4. Adler, A. I., & Painsky, A. (2022). Feature Importance in Gradient Boosting Trees with Cross-Validation Feature Selection. Entropy (Basel), 24(5), 687. https://doi.org/10.3390/e24050687
5. Aghaei et al. (2016). Applying a new quantitative global breast MRI feature analysis scheme to assess tumor response to chemotherapy. Journal of Magnetic Resonance Imaging, 44(5), 1099–1106. https://doi.org/10.1002/jmri.25276
6. Akshay et al. (2024). Machine Learning-Based Classification of Transcriptome Signatures of Non-Ulcerative Bladder Pain Syndrome. International Journal of Molecular Sciences, 25(3), 1568. https://doi.org/10.3390/ijms25031568
7. Alaimo et al. (2023). 2-step Gradient Boosting approach to selectivity bias correction in tax audit: an application to the VAT gap in Italy. Statistical Methods & Applications, 32, 237–270. https://doi.org/10.1007/s10260-022-00643-4
8. Alanis-Lobato et al. (2015). Highlighting nonlinear patterns in population genetics datasets. 1 Scientific Reports, 5, Article 8140. https://doi.org/10.1038/srep08140
9. Alipour et al. (2023). Enhanced instance space analysis for the maximum flow problem. European Journal of Operational Research, 304(2), 411–428. https://doi.org/10.1016/j.ejor.2022.04.012
10. Alirezanejad et al. (2020). Heuristic filter feature selection methods for medical datasets. Genomics, 112(2), 1173–1181. https://doi.org/10.1016/j.ygeno.2019.07.002
11. Altmann et al. (2010). Permutation importance: a corrected feature importance measure. Bioinformatics, 26(10), 1340–1347. https://doi.org/10.1093/bioinformatics/btq134
12. Amado et al. (2025). Robust estimation of heteroscedastic regression models: A brief overview and new proposals. Statistical Papers, 66, 65. https://doi.org/10.1007/s00362-025-01686-x
13. Ambroise, C., & McLachlan, G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. Proceedings of the National Academy of Sciences of the United States of America, 99(10), 6562–6566. https://doi.org/10.1073/pnas.102102699
14. Amornbunchornvej et al. (2021). Variable-lag Granger causality and transfer entropy for time series analysis. ACM Transactions on Knowledge Discovery from Data, 15(4), 67. https://doi.org/10.1145/3441452
15. Anandhi, P., & Nathiya, E. (2023). Application of linear regression with their advantages, disadvantages, assumption and limitations. International Journal of Statistics and Applied Mathematics, 8(6), 133-137. https://doi.org/10.22271/maths.2023.v8.i6b.1463
16. Aniceto et al. (2022). Exploring the Chemical Space of Urease Inhibitors to Extract Meaningful Trends and Drivers of Activity. Journal of Chemical Information and Modeling, 62(15), 3535–3550. https://doi.org/10.1021/acs.jcim.2c00150
17. Ardelean, E.-R., Portase, R. L., Potolea, R., & Dînșoreanu, M. (2025). A path-based distance computation for non-convexity with applications in clustering. Knowledge and Information Systems, 67, 1415–1453. https://doi.org/10.1007/s10115-024-02275-4
18. Asensio et al. (2022). Predicting missing proteomics values using machine learning: Filling the gap using transcriptomics and other biological features. Computational and Structural Biotechnology Journal, 20, 2057–2069. https://doi.org/10.1016/j.csbj.2022.04.017
19. Attia et al. (2021). Deep neural networks learn by using human-selected electrocardiogram features and novel features. European Heart Journal - Digital Health, 2(3), 446–455. https://doi.org/10.1093/ehjdh/ztab060
20. Azim et al. (2022). CDSImpute: An ensemble similarity imputation method for single-cell RNA sequence dropouts. Computers in Biology and Medicine, 146, 105658. https://doi.org/10.1016/j.compbiomed.2022.105658
21. Banerjee et al. (2023). “Shortcuts” Causing Bias in Radiology Artificial Intelligence: Causes, Evaluation, and Mitigation. Journal of the American College of Radiology, 20(9), 842–851. https://doi.org/10.1016/j.jacr.2023.06.025
22. Bansal, S., & Singh, G. (2023). Multiple linear regression based analysis of weather data: Assumptions and limitations. In R. N. Shaw, M. Paprzycki, & A. Ghosh (Eds.), Advanced communication and intelligent systems. ICACIS 2023 (Vol. 1920, pp. 169-178). Springer, Cham. https://doi.org/10.1007/978-3-031-45121-8\_19
23. Barton-Henry et al. (2021). Decay radius of climate decision for solar panels in the city of Fresno, USA. Scientific Reports, 11(1), 8571. https://doi.org/10.1038/s41598-021-87714-w
24. Belz et al. (2017). Order Determination and Input Selection with Local Model Networks. IFAC-PapersOnLine, 50(1), 7327–7332. https://doi.org/10.1016/j.ifacol.2017.08.1475
25. Ben-Naim, A. (2023). Intermolecular Interactions, Correlations, and Mutual Information. In Information Theory and Selected Applications (pp. 19-38). Springer, Cham. https://doi.org/10.1007/978-3-031-21276-5\_2
26. Betz, J. L., & Sadler, J. R. (1981). Variants of a cloned synthetic lactose operator II. Chloramphenicol-resistant revertants retaining a lactose operator in the CAT gene of plasmid pBR325. Gene, 15(2-3), 187–200. https://doi.org/10.1016/0378-1119(81)90128-1
27. Bickel, D. R. (2008). Correcting the estimated level of differential expression for gene selection bias: application to a microarray study. Statistical Applications in Genetics and Molecular Biology, 7(1), Article10. https://doi.org/10.2202/1544-6115.1330
28. Bilodeau et al. (2024). Impossibility theorems for feature attribution. Proceedings of the National Academy of Sciences, 121(2), e2304406120. https://doi.org/10.1073/pnas.2304406120
29. Black et al. (2024). Towards machine learning-based quantitative hyperspectral image guidance for brain tumor resection. Communications Medicine, 4(1), Article 131. https://doi.org/10.1038/s43856-024-00562-3
30. Bobrowski, L. (1991). Design of piecewise linear classifiers from formal neurons by a basis exchange technique. Pattern Recognition, 24(9), 863–870. https://doi.org/10.1016/0031-3203(91)90005-P
31. Bobrowski, L., & Niemiro, W. (1984). A method of synthesis of linear discriminant function in the case of nonseparability. Pattern Recognition, 17(2), 205–210. https://doi.org/10.1016/0031-3203(84)90059-1
32. Bougioukos et al. (2010). An intensity-region driven multi-classifier scheme for improving the classification accuracy of proteomic MS-spectra. Computer Methods and Programs in Biomedicine, 99(2), 147–153. https://doi.org/10.1016/j.cmpb.2009.11.003
33. Brinkrolf et al. (2019). Differential privacy for learning vector quantization. Neurocomputing, 342, 125–136. https://doi.org/10.1016/j.neucom.2018.11.095
34. Burgard et al. (2024). Robustification of the k-means clustering problem and tailored decomposition methods: When more conservative means more accurate. Annals of Operations Research, 339, 1525–1568. https://doi.org/10.1007/s10479-022-04818-w
35. Bushra et al. (2024). AutoSCAN: Automatic detection of DBSCAN parameters and efficient clustering of data in overlapping density regions. PeerJ Computer Science, 10, e1921. https://doi.org/10.7717/peerj-cs.1921
36. Cai et al. (2018). Concussion classification via deep learning using whole-brain white matter fiber strains. PLoS ONE, 13(5), e0197992. https://doi.org/10.1371/journal.pone.0197992
37. Caprihan et al. (2008). Application of principal component analysis to distinguish patients with schizophrenia from healthy controls based on fractional anisotropy measurements. NeuroImage, 42(2), 675-682. https://doi.org/10.1016/j.neuroimage.2008.04.255
38. Carletti et al. (2023). Interpretable Anomaly Detection with DIFFI: Depth-based feature importance of Isolation Forest. Engineering Applications of Artificial Intelligence, 119, 105730. https://doi.org/10.1016/j.engappai.2022.105730
39. Caserini, N. A., & Pagnottoni, P. (2022). Effective transfer entropy to measure information flows in credit markets. Statistical Methods & Applications, 31, 729–757. https://doi.org/10.1007/s10260-021-00614-1
40. Castaldi et al. (2011). An empirical assessment of validation practices for molecular classifiers. Briefings in Bioinformatics, 12(3), 189–202. https://doi.org/10.1093/bib/bbq073
41. Cechinel et al. (2024). Enhancing wastewater treatment efficiency through machine learning-driven effluent quality prediction: A plant-level analysis. Journal of Water Process Engineering, 58, 104758. https://doi.org/10.1016/j.jwpe.2023.104758
42. Cerqueti et al. (2024). Kendall correlations and radar charts to include goals for and goals against in soccer rankings. Computational Statistics. Advance online publication. https://doi.org/10.1007/s00180-024-01542-w
43. Chanderraj et al. (2022). The bacterial density of clinical rectal swabs is highly variable, correlates with sequencing contamination, and predicts patient risk of extraintestinal infection. Microbiome, 10(1), 2. https://doi.org/10.1186/s40168-021-01190-y
44. Chanderraj et al. (2020). Gut Microbiota Predict Enterococcus Expansion but Not Vancomycin-Resistant Enterococcus Acquisition. mSphere, 5(6), e00537-20. https://doi.org/10.1128/mSphere.00537-20
45. Chaudhari et al. (2020). DeepRMethylSite: a deep learning based approach for prediction of arginine methylation sites in proteins. Molecular Omics, 16(5), 448–454. https://doi.org/10.1039/d0mo00025f
46. Chen et al. (2023). Nonlinear Structural Equation Model Guided Gaussian Mixture Hierarchical Topic Modeling. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers) (pp. 10377-10390). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.578
47. Chen et al. (2023). Relationship between prediction accuracy and feature importance reliability: An empirical and theoretical study. NeuroImage, 274, 120115–120115. https://doi.org/10.1016/j.neuroimage.2023.120115
48. Chen et al. (2023). Linear, nonlinear, parametric, and nonparametric regression models for nonstationary flood frequency analysis. Journal of Hydrology, 616, 128772. https://doi.org/10.1016/j.jhydrol.2022.128772
49. Chen et al. (2022). Practical guide to using Kendall's τ in the context of forecasting critical transitions. Royal Society Open Science, 9(7), 211346. https://doi.org/10.1098/rsos.211346
50. Chen et al. (2024). Applying interpretable machine learning in computational biology—pitfalls, recommendations and opportunities for new developments. 2 Nature Methods, 21(8), 1454–1461. https://doi.org/10.1038/s41592-024-02359-7
51. Chen, Y. T., & Witten, D. M. (2023). Selective inference for k-means clustering. Journal of Machine Learning Research, 24, 152. https://doi.org/10.48550/arXiv.2203.15267
52. Chen et al. (2023). Human-Centered Design to Address Biases in Artificial Intelligence. Journal of Medical Internet Research, 25, e43251. https://doi.org/10.2196/43251
53. Cheng et al. (2024). GB-DBSCAN: A fast granular-ball based DBSCAN clustering algorithm. Information Sciences, 674, 120731. https://doi.org/10.1016/j.ins.2024.120731
54. Cheungpasitporn et al. (2024). Artificial intelligence and machine learning’s role in sepsis-associated acute kidney injury. Kidney Research and Clinical Practice, 43(4), 417–432. https://doi.org/10.23876/j.krcp.23.298
55. Chow et al. (2024). SHARK enables sensitive detection of evolutionary homologs and functional analogs in unalignable and disordered sequences. Proceedings of the National Academy of Sciences, 121(42), e2401622121. https://doi.org/10.1073/pnas.2401622121
56. Coronado et al. (2022). Transfer entropy Granger causality between news indices and stock markets in U.S. and Latin America during the COVID-19 pandemic. Entropy, 24(10), 1420. https://doi.org/10.3390/e24101420
57. Costes et al. (2024). Multi-omics data integration for the identification of biomarkers for bull fertility. PLoS ONE, 19(2), e0298623. https://doi.org/10.1371/journal.pone.0298623
58. Cristian et al. (2024). Diffusion on PCA-UMAP manifold: The impact of data structure preservation to denoise high-dimensional single-cell RNA sequencing data. Biology, 13(7), Article 512. https://doi.org/10.3390/biology13070512
59. Cummings et al. (2023). Development of a Novel Clinicomolecular Risk Index to Enhance Mortality Prediction and Immunological Stratification of Adults Hospitalized with Sepsis in Sub-Saharan Africa: A Pilot Study from Uganda. The American Journal of Tropical Medicine and Hygiene, 108(3), 619–626. https://doi.org/10.4269/ajtmh.22-0483
60. Cun, Y., & Fröhlich, H. (2012). Biomarker gene signature discovery integrating network knowledge. Biology, 1(1), 5–17. https://doi.org/10.3390/biology1010005
61. Dai, H., & Bao, Y. (2009). An inverse probability 3 weighted estimator for the bivariate distribution function under right censoring. Statistics & Probability Letters, 79(16), 1789–1797. https://doi.org/10.1016/j.spl.2009.05.010
62. Dankwa-Mullan, I. (2024). Health Equity and Ethical Considerations in Using Artificial Intelligence in Public Health and Medicine. Preventing Chronic Disease, 21, E64. https://doi.org/10.5888/pcd21.240245
63. Debray et al. (2015). Interstitial lung disease in anti-synthetase syndrome: Initial and follow-up CT findings. European Journal of Radiology, 84(3), 516–523. https://doi.org/10.1016/j.ejrad.2014.11.026
64. Demircioğlu, A. (2021). Measuring the bias of incorrect application of feature selection when using cross-validation in radiomics. Insights Imaging, 12(1), 172. https://doi.org/10.1186/s13244-021-01115-1
65. Deng et al. (2019). Tractography-based classification in distinguishing patients with first-episode schizophrenia from healthy individuals. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 88, 66–73. https://doi.org/10.1016/j.pnpbp.2018.06.010
66. Derrac et al. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. Swarm and Evolutionary Computation, 1(1), 3-18. https://doi.org/10.1016/j.swevo.2011.02.002
67. Dey et al. (2025). The proper application of logistic regression model in complex survey data: a systematic review. BMC Medical Research Methodology, 25, 15. https://doi.org/10.1186/s12874-024-02454-5
68. Dey, R., & Lee, S. (2019). Asymptotic properties of principal component analysis and shrinkage-bias adjustment under the generalized spiked population model. Journal of Multivariate Analysis, 173, 145-164. https://doi.org/10.1016/j.jmva.2019.02.007
69. Diaz et al. (2021). Sequence-based dynamic handwriting analysis for Parkinson’s disease detection with one-dimensional convolutions and BiGRUs. Expert Systems with Applications, 168, 114405. https://doi.org/10.1016/j.eswa.2020.114405
70. Ding et al. (2015). Research and development of advanced computing technologies. The Scientific World Journal, 2015, 239723. https://doi.org/10.1155/2015/239723 6
71. Domingue et al. (2024). Ubiquitous bias and false discovery due to model misspecification in analysis of statistical interactions: The role of the outcome's distribution and metric properties. Psychological Methods, 29(6), 1164–1179. https://doi.org/10.1037/met0000532
72. Dong et al. (2020). Predictive analysis methods for human microbiome data with application to Parkinson’s disease. PLoS ONE, 15(8), e0237779. https://doi.org/10.1371/journal.pone.0237779
73. Donmez et al. (2023). Road user behavior: Describing, inferring, predicting and beyond. Transportation Research Interdisciplinary Perspectives, 22, 100932. https://doi.org/10.1016/j.trip.2023.100932 7
74. Dukart et al. (2013). Meta-analysis based SVM classification enables accurate detection of Alzheimer’s disease across different clinical centers using FDG-PET and MRI. Psychiatry Research: Neuroimaging, 212(3), 230–236. https://doi.org/10.1016/j.pscychresns.2012.04.007
75. Dunne et al. (2023). Thresholding Gini variable importance with a single-trained random forest: An empirical Bayes approach. Computational and Structural Biotechnology Journal, 21, 4354-4360. https://doi.org/10.1016/j.csbj.2023.08.033
76. Dupeu, J. M. (1997). La consultation thérapeutique en pédo-psychiatrie: 10e journée annuelle de psychiatrie infantile. Journal de Pédiatrie et de Puériculture, 10(6), 349–355. https://doi.org/10.1016/S0987-7983(97)80099-0 8
77. Dyer, E. L., & Kording, K. (2023). Why the simplest explanation isn't always the best. Proceedings of the National Academy of Sciences, 120(52), Article e2319169120. https://doi.org/10.1073/pnas.2319169120
78. Eden et al. (2022). Nonparametric estimation of Spearman's rank correlation with bivariate survival data. Biometrics, 78(2), 421-434. https://doi.org/10.1111/biom.13453
79. Effrosynidis, D., & Arampatzis, A. (2021). An evaluation of feature selection methods for environmental data. Ecological Informatics, 61, 101224. https://doi.org/10.1016/j.ecoinf.2021.101224
80. Ehrlich et al. (2024). Partial information decomposition for continuous variables based on shared exclusions: Analytical formulation and estimation. Physical Review E, 110(1), 014115. https://doi.org/10.1103/PhysRevE.110.014115
81. Eichele et al. (2009). EEGIFT: A toolbox for group temporal ICA of event-related EEG. NeuroImage, 47(Supplement 1), S101. https://doi.org/10.1016/S1053-8119(09)70872-9
82. Eisele et al. (2024). Gene-expression memory-based prediction of cell lineages from scRNA-seq datasets. Nature Communications, 15(1), 2744. https://doi.org/10.1038/s41467-024-47158-y
83. Ekhlasi et al. (2023). Improving transfer entropy and partial transfer entropy for relative detection of effective connectivity strength between time series. Communications in Nonlinear Science and Numerical Simulation, 126, 107449. https://doi.org/10.1016/j.cnsns.2023.107449
84. El Allali et al. (2021). Machine learning applications in RNA modification sites prediction. Computational and Structural Biotechnology Journal, 19, 5510–5524. https://doi.org/10.1016/j.csbj.2021.09.025 9
85. Elhaik, E. (2022). Principal component analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. Scientific Reports, 12(1), Article 14683. https://doi.org/10.1038/s41598-022-14395-4
86. Erman, B. (2023). Mutual information analysis of mutation, nonlinearity, and triple interactions in proteins. Proteins, 91(1), 121–133. https://doi.org/10.1002/prot.26415
87. F., & Little, M. A. (2016). What to do when K-means clustering fails: A simple yet principled alternative algorithm. PLoS One, 11(9), e0162259. https://doi.org/10.1371/journal.pone.0162259
88. Fakhraei et al. (2014). Bias and stability of single variable classifiers for feature ranking and selection. Expert Systems with Applications, 41(15), 6945–6958. https://doi.org/10.1016/j.eswa.2014.05.007
89. Fallahpour et al. (2017). Using an ensemble classifier based on sequential floating forward selection for financial distress prediction problem. Journal of Retailing and Consumer Services, 34, 159–167. https://doi.org/10.1016/j.jretconser.2016.10.002 12
90. Felippe et al. (2024). Network mutual information measures for graph similarity. Communications Physics, 7, 335. https://doi.org/10.1038/s42005-024-01830-3
91. Fisher et al. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. Journal of Machine Learning Research, 20, 177. https://doi.org/10.48550/arXiv.1801.01489
92. Fox, J., & Monette, G. (1992). Generalized Collinearity Diagnostics. Journal of the American Statistical Association, 87(417), 178–183. https://doi.org/10.1080/01621459.1992.10475190
93. Frénay et al. (2014). Estimating mutual information for feature selection in the presence of label noise. Computational Statistics & Data Analysis, 71, 832–848. https://doi.org/10.1016/j.csda.2013.05.001
94. Fujita et al. (2009). Comparing Pearson, Spearman, and Hoeffding's D measure for gene expression association analysis. Journal of Bioinformatics and Computational Biology, 7(4), 663–684. https://doi.org/10.1142/s0219720009004230
95. Galligan et al. (2013). Greedy feature selection for glycan chromatography data with the generalized Dirichlet distribution. BMC Bioinformatics, 14, 155. https://doi.org/10.1186/1471-2105-14-155
96. Gandhudi et al. (2024). Explainable hybrid quantum neural networks for analyzing the influence of tweets on stock price prediction. Computers and Electrical Engineering, 118(Part A), 109302. https://doi.org/10.1016/j.compeleceng.2024.109302
97. Ge et al. (2019). Study Progress of Radiomics With Machine Learning for Precision Medicine in Bladder Cancer Management. Frontiers in Oncology, 9, 1296. https://doi.org/10.3389/fonc.2019.01296
98. Ge et al. (2016). McTwo: a two-step feature selection algorithm based on maximal information coefficient. BMC Bioinformatics, 17, 142. https://doi.org/10.1186/s12859-016-0990-0
99. Gibson, J. D. (2025). Entropy and Mutual Information. In Information Theoretic Principles for Agent Learning (pp. 15-33). Springer, Cham. https://doi.org/10.1007/978-3-031-65388-9\_2
100. Gnambs, T. (2023). A brief note on the standard error of the Pearson correlation. Collabra: Psychology, 9(1). https://doi.org/10.1525/collabra.87615
101. Goldfarb et al. (2022). High-Risk Drinkers Engage Distinct Stress-Predictive Brain Networks. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 7(8), 805–813. https://doi.org/10.1016/j.bpsc.2022.02.010
102. Grabowski, E., & Kuo, J. (2023). Comparing K-means and OPTICS clustering algorithms for identifying vowel categories. Proceedings of the Linguistic Society of America, 8(1), 5488. https://doi.org/10.3765/plsa.v8i1.5488
103. Grover et al. (2019). Using supervised learning to select audit targets in performance-based financing in health: An example from Zambia. 1 PLoS ONE, 14(1), e0211262. 2 https://doi.org/10.1371/journal.pone.0211262
104. Gu et al. (2024). Machine learning and neuroimaging: Understanding the human brain in health and disease. In Neuroimaging Methods and Applications, Computational and Network Modeling of Neuroimaging Data (pp. 261–285). Academic Press. https://doi.org/10.1016/B978-0-443-13480-7.00010-7
105. Guan et al. (2009). Ovarian cancer detection from metabolomic liquid chromatography/mass spectrometry data by support 3 vector machines. BMC Bioinformatics, 10, 259. https://doi.org/10.1186/1471-2105-10-259
106. Guidotti et al. (2015). Visual Learning Induces Changes in Resting-State fMRI Multivariate Pattern of 4 Information. The Journal of Neuroscience, 35(27), 9786–9798. https://doi.org/10.1523/JNEUROSCI.3920-14.2015
107. Guyader et al. (2018). Groupwise image registration based on a total correlation dissimilarity measure for quantitative MRI and dynamic imaging data. Scientific Reports, 8, 13112. https://doi.org/10.1038/s41598-018-31474-7
108. Guyatt et al. (1984). Determining causation—A case study: Adrenocorticosteroids and osteoporosis: Should the fear of inducing clinically important osteoporosis 5 influence the decision 6 to prescribe adrenocorticosteroids? Journal of Chronic Diseases, 37(5), 343–352. https://doi.org/10.1016/0021-9681(84)90100-0
109. Hair et al. (2021). An introduction to structural equation modeling. In Partial least squares structural equation modeling (PLS-SEM) using R: Classroom Companion: Business (pp. 1-25). Springer. https://doi.org/10.1007/978-3-030-80519-7\_1
110. Hair et al. (2024). Going beyond the untold facts in PLS–SEM and moving forward. European Journal of Marketing, 58(13), 81–106. https://doi.org/10.1108/EJM-08-2023-0645
111. Hajihosseinlou et al. (2024). A comprehensive evaluation of OPTICS, GMM and K-means clustering methodologies for geochemical anomaly detection connected with sample catchment basins. Geochemistry, 84(2), Article 126094. https://doi.org/10.1016/j.chemer.2024.126094
112. Han, H. (2017). A novel feature selection for RNA-seq analysis. Computational Biology and Chemistry, 71, 245–257. https://doi.org/10.1016/j.compbiolchem.2017.10.010
113. Han, J., & Kang, S. (2021). Active learning with missing values considering imputation uncertainty. Knowledge-Based Systems, 224, 107079. https://doi.org/10.1016/j.knosys.2021.107079
114. Hanbay, K. (2022). A new standard error based artificial bee colony algorithm and its applications in feature selection. Journal of King Saud University - Computer and Information Sciences, 34(7), 4554–4567. https://doi.org/10.1016/j.jksuci.2021.04.010
115. Hancer et al. (2018). Differential evolution for filter feature selection based on information theory and feature ranking. Knowledge-Based Systems, 140, 103–119. https://doi.org/10.1016/j.knosys.2017.10.028
116. Higham, P. A., & Higham, D. P. (2019). New improved gamma: Enhancing the accuracy of Goodman-Kruskal's gamma using ROC curves. Behavior Research Methods, 51(1), 108–125. https://doi.org/10.3758/s13428-018-1125-5
117. Hooshyar, D., & Yang, Y. (2024). Problems With SHAP and LIME in Interpretable AI for Education: A Comparative Study of Post-Hoc Explanations and Neural-Symbolic Rule Extraction. IEEE Access, 12, 137472-137490. https://doi.org/10.1109/ACCESS.2024.3463948
118. Horrace, W. C., & Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. Economics Letters, 90(3), 321-327. https://doi.org/10.1016/j.econlet.2005.08.024
119. Hou et al. (2022). Distance correlation application to gene co-expression network analysis. BMC Bioinformatics, 23(1), 81. https://doi.org/10.1186/s12859-022-04609-x
120. Hsieh et al. (2012). Caucasian male infants and boys with hypospadias exhibit reduced anogenital distance. Human Reproduction, 27(6), 1577–1580. https://doi.org/10.1093/humrep/des087
121. Huang et al. (2007). A hybrid genetic algorithm for feature selection wrapper based on mutual information. Pattern Recognition Letters, 28(13), 1 1825–1844. https://doi.org/10.1016/j.patrec.2007.05.011
122. Huang, X., & Marques-Silva, J. (2024). On the failings of Shapley values for explainability. International Journal of Approximate Reasoning, 171, 109112. https://doi.org/10.1016/j.ijar.2023.109112
123. Huang et al. (2025). Debiasing weighted multi-view k-means clustering based on causal regularization. Pattern Recognition, 160, 111195. https://doi.org/10.1016/j.patcog.2024.111195
124. Humberg et al. (2024). Estimating nonlinear effects of random slopes: A comparison of multilevel structural equation modeling with a two-step, a single-indicator, and a plausible values approach. Behavior Research Methods, 56(7), 7912–7938. https://doi.org/10.3758/s13428-024-02462-9
125. Huti et al. (2023). An investigation into race bias in random forest models based on breast DCE-MRI derived radiomics features. In Clinical Image Based Procedure Fairness AI Med Imaging Ethical Philos Issues Med Imaging (Vol. 14242, pp. 225-234). https://doi.org/10.1007/978-3-031-45249-9\_22
126. Irmer et al. (2024). Estimating power in complex nonlinear structural equation modeling including moderation effects: The powerNLSEM R-package. Behavior Research Methods, 56, 8897-8931. https://doi.org/10.3758/s13428-024-02476-3
127. Jacob, J., & Varadharajan, R. (2024). Robust Variance Inflation Factor: A Promising Approach for Collinearity Diagnostics in the Presence of Outliers. Sankhya B, 86, 845–871. https://doi.org/10.1007/s13571-024-00342-y
128. Jain, R., & Xu, W. (2021). HDSI: High dimensional selection with interactions algorithm on feature selection and testing. PLoS ONE, 16(2), e0246159. https://doi.org/10.1371/journal.pone.0246159
129. Janse et al. (2021). Conducting correlation analysis: Important limitations and pitfalls. Clinical Kidney Journal, 14(11), 2332-2337. https://doi.org/10.1093/ckj/sfab085
130. Jarantow et al. (2023). Introduction to the use of linear and nonlinear regression analysis in quantitative biological assays. Current Protocols, 3(6), e801. https://doi.org/10.1002/cpz1.801
131. Jeung et al. (2025). Graph neural networks and transfer entropy enhance forecasting of mesozooplankton community dynamics. Environmental Science and Ecotechnology, 23, 100514. https://doi.org/10.1016/j.ese.2024.100514
132. Jha et al. (2020). The Cloud UPDRS smartphone software in Parkinson’s study: cross-validation against blinded human raters. NPJ Parkinson’s Disease, 6(1), 36. https://doi.org/10.1038/s41531-020-00135-w
133. Jiang et al. (2024). T1 mapping-based radiomics in the identification of histological types of lung cancer: a reproducibility and feasibility study. BMC Medical Imaging, 24(1), 308. https://doi.org/10.1186/s12880-024-01487-y
134. Jin et al.(2021). Robust nonlinear structural equation modeling with interaction between exogenous and endogenous latent variables. Structural Equation Modeling: A Multidisciplinary Journal, 28(4), 547–556. https://doi.org/10.1080/10705511.2020.1857255
135. Jollans et al. (2019). Quantifying performance of machine learning methods for neuroimaging data. NeuroImage, 199, 351–365. https://doi.org/10.1016/j.neuroimage.2019.05.082
136. Jung et al. (2020). Pseudo-Occlusion of the Internal Carotid Artery in Acute Ischemic Stroke: Clinical Outcome 2 after Mechanical Thrombectomy. Scientific Reports, 10(1), 2832. https://doi.org/10.1038/s41598-020-59609-9
137. Junges et al. (2024). Mitigating the Impact of Temperature Variations on Ultrasonic Guided Wave-Based Structural Health Monitoring 3 through Variational Autoencoders. Sensors, 24(5), 1494. https://doi.org/10.3390/s24051494
138. Kakuba, S., & Han, D. S. (2024). Addressing data scarcity in speech emotion recognition: A comprehensive review. ICT Express. https://doi.org/10.1016/j.icte.2024.11.003
139. Kamat et al. (2024). DeepTool: A deep learning framework for tool wear onset detection and remaining useful life prediction. MethodsX, 13, 102965. https://doi.org/10.1016/j.mex.2024.102965
140. Katal et al. (2024). AI in radiology: From promise to practice − A guide to effective integration. European Journal of Radiology, 181, 111798. https://doi.org/10.1016/j.ejrad.2024.111798
141. Kazienko, P., & Kajdanowicz, T. (2012). Label-dependent node classification in the network. Neurocomputing, 75(1), 199–209. https://doi.org/10.1016/j.neucom.2011.04.047
142. Kepplinger, D. (2023). Robust variable selection and estimation via adaptive elastic net S-estimators for linear regression. Computational Statistics & Data Analysis, 183, 107730. https://doi.org/10.1016/j.csda.2023.107730
143. Kerby et al. (2024). Learning local higher-order interactions with total correlation. In Proceedings of the 2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP) (pp. 1-6). IEEE. https://doi.org/10.1109/MLSP58920.2024.10734758
144. Keshavan et al. (2020). Neuroimaging in Schizophrenia. Neuroimaging Clinics of North America, 30(1), 73–83. https://doi.org/10.1016/j.nic.2019.09.007
145. Khan et al. (2024). Determining the optimal number of clusters by Enhanced Gap Statistic in K-mean algorithm. Egyptian Informatics Journal, 27, 100504. https://doi.org/10.1016/j.eij.2024.100504
146. Kipruto, E., & Sauerbrei, W. (2024). Post-Estimation Shrinkage in Full and Selected Linear Regression Models in Low-Dimensional Data Revisited. Biometrical Journal. Biometrische Zeitschrift, 66(7), e202300368. https://doi.org/10.1002/bimj.202300368
147. Kiviet, J. F., & Phillips, G. D. A. (1996). The bias of the ordinary least squares estimator in simultaneous equation models. Economics Letters, 53(2), 161-167. https://doi.org/10.1016/S0165-1765(96)00908-1
148. Krawczuk, J., & Łukaszuk, T. (2016). The feature selection bias problem in relation to high-dimensional gene data. Artificial Intelligence in Medicine, 66, 63-71. https://doi.org/10.1016/j.artmed.2015.11.001
149. Kretowska, M. (2018). Tree-based models for survival data with competing risks. Computer Methods and Programs in Biomedicine, 159, 185–198. https://doi.org/10.1016/j.cmpb.2018.03.017
150. Ku, W. L., & Min, H. (2024). Evaluating Machine Learning Stability in Predicting Depression and Anxiety Amidst Subjective Response Errors. Healthcare (Basel, Switzerland), 12(6), 625. https://doi.org/10.3390/healthcare12060625
151. Kumar, A. (2024). Comparative evaluation of linear and nonlinear regression models in predicting VLC channel response and BER performance. Journal of Optics. https://doi.org/10.1007/s12596-024-02361-4
152. Kumar et al. (2021). Shapley residuals: Quantifying the limits of the Shapley value for explanations. Advances in Neural Information Processing Systems, 34, 26598–26608.
153. Kunert-Graf et al. (2020). Partial Information Decomposition and the Information Delta: A Geometric Unification Disentangling Non-Pairwise Information. Entropy (Basel, Switzerland), 22(12), 1333. https://doi.org/10.3390/e22121333
154. Kursa, M. B. (2022). Kendall transformation brings a robust categorical representation of ordinal data. Scientific Reports, 12, 8341. https://doi.org/10.1038/s41598-022-12224-2
155. Lee et al. (2024). Unsupervised anomaly detection process using LLE and HDBSCAN by Style-GAN as a feature extractor. International Journal of Precision Engineering and Manufacturing, 25, 51–63. https://doi.org/10.1007/s12541-023-00908-2
156. Lenhof et al. (2024). Trust me if you can: A survey on reliability and interpretability of machine learning approaches for drug sensitivity prediction in cancer. Briefings in Bioinformatics, 25(5), bbae379. https://doi.org/10.1093/bib/bbae379
157. Lenz et al. (2016). Principal components analysis and the reported low intrinsic dimensionality of gene expression microarray data. Scientific Reports, 6, Article 25696. https://doi.org/10.1038/srep25696
158. Li et al. (2023). Bayesian forecast combination using time-varying features. International Journal of Forecasting, 39(3), 1287–1302. https://doi.org/10.1016/j.ijforecast.2022.06.002 7
159. Li et al. (2024). Functional connectivity via total correlation: Analytical results in visual areas. Neurocomputing, 571, 127143. https://doi.org/10.1016/j.neucom.2023.127143
160. Li et al. (2024). A new transfer entropy method for measuring directed connectivity from complex-valued fMRI data. Frontiers in Neuroscience, 18, 1423014. https://doi.org/10.3389/fnins.2024.1423014
161. Li et al. (2024). A scholars’ personality traits augmented multi-dimensional feature fusion scholarly journal recommendation model. Applied Soft Computing, 163, 111888. https://doi.org/10.1016/j.asoc.2024.111888 8
162. Li et al. (2021). Robust estimation and variable selection for the accelerated failure time model. Statistics in Medicine, 40(20), 4473-4491. https://doi.org/10.1002/sim.9042
163. Liang, J. E. (2024). Partial information decomposition for causal discovery with application to Internet of Things. IEEE Internet of Things Journal, 11(13), 24289-24299. https://doi.org/10.1109/JIOT.2024.3390449
164. Lin et al. (2022). Linear and nonlinear correlation estimators unveil undescribed taxa interactions in microbiome data. Nature Communications, 13, 4946. https://doi.org/10.1038/s41467-022-32243-x
165. Lin et al. (2019). A Super-Learner Model for Tumor Motion Prediction and Management in Radiation Therapy: Development and Feasibility Evaluation. Scientific Reports, 9(1), 14868. https://doi.org/10.1038/s41598-019-51338-y 9
166. Linardatos et al. (2020). Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy, 23(1), 18. https://doi.org/10.3390/e23010018
167. Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. Queue, 16(3), 31–57. https://doi.org/10.1145/3236386.3241340
168. Liu et al. (2021). Prediction of venous thromboembolism with machine learning techniques in young-middle-aged inpatients. Scientific Reports, 11(1), 12868. https://doi.org/10.1038/s41598-021-92287-9
169. Liu, J. (2025). Examination of nonlinear longitudinal processes with latent variables, latent processes, latent changes, and latent classes in the structural equation modeling framework: The R package nlpsem. Behavior Research Methods, 57(3), 87. https://doi.org/10.3758/s13428-025-02596-4
170. Liu et al. (2025). ESERNet: Learning spectrogram structure relationship for effective speech emotion recognition with swin transformer in classroom discourse analysis. Neurocomputing, 612, 128711. https://doi.org/10.1016/j.neucom.2024.128711
171. Liu et al. (2020). Can the development of a patient’s condition be predicted through intelligent inquiry under the e-health business mode? Sequential feature map-based disease risk prediction upon features selected from cognitive diagnosis big data. International Journal of Information Management, 50, 463–486. https://doi.org/10.1016/j.ijinfomgt.2019.05.006
172. Liu et al. (2021). Dilated Adversarial U-Net Network for automatic gross tumor volume segmentation of nasopharyngeal carcinoma. Applied Soft Computing, 111, 107722. https://doi.org/10.1016/j.asoc.2021.107722
173. Liu et al. (2021). An interactive filter-wrapper multi-objective evolutionary algorithm for feature selection. Swarm and Evolutionary Computation, 65, 100925. https://doi.org/10.1016/j.swevo.2021.100925
174. Loecher, M. (2024). Debiasing SHAP scores in random forests. AStA Advances in Statistical Analysis, 108, 427-440. https://doi.org/10.1007/s10182-023-00479-7
175. Lohrmann, C., & Luukka, P. (2022). Nonspecificity, strife and total uncertainty in supervised feature selection. Engineering Applications of Artificial Intelligence, 109, 104628. https://doi.org/10.1016/j.engappai.2021.104628
176. Lones, M. A. (2024). Avoiding common machine learning pitfalls. Patterns, 5(10), 101046. https://doi.org/10.1016/j.patter.2024.101046
177. Lovatti et al. (2019). Use of Random forest in the identification of important variables. Microchemical Journal, 145, 1129–1134. https://doi.org/10.1016/j.microc.2018.12.028
178. Lu et al. (2024). Label distribution feature selection based on hierarchical structure and neighborhood granularity. Information Fusion, 112, 102588. https://doi.org/10.1016/j.inffus.2024.102588
179. Ma et al. (2022). A feature fusion sequence learning approach for quantitative analysis of tremor symptoms based on digital handwriting. Expert Systems with Applications, 203, 117400. https://doi.org/10.1016/j.eswa.2022.117400
180. Ma et al. (2024). Linear and nonlinear causality in financial markets. Chaos, 34(11), 113125. https://doi.org/10.1063/5.0184267
181. Maggipinto et al. (2017). DTI measurements for Alzheimer’s classification. Physics in Medicine and Biology, 62(6), 2361–2375. https://doi.org/10.1088/1361-6560/aa5dbe
182. Mandler, H., & Weigand, B. (2024). A review and benchmark of feature importance methods for neural networks. ACM Computing Surveys, 56(12), 318. https://doi.org/10.1145/3679012
183. Manolov, R. (2023). Does the choice of a linear trend-assessment technique matter in the context of single-case data? Behavior Research Methods, 55(8), 4200–4221. https://doi.org/10.3758/s13428-022-02013-0
184. Markowetz, F., & Spang, R. (2005). Molecular diagnosis. Classification, model selection and performance evaluation. Methods of Information in Medicine, 44(3), 438–443. https://doi.org/10.1055/s-0038-1633990
185. Mateo et al. (2021). Automatic mass spectra recognition for Ultra High Vacuum systems using multilabel classification. Expert Systems with Applications, 178, 114959. https://doi.org/10.1016/j.eswa.2021.114959
186. Matusik et al. (2024). Parametric and nonparametric population pharmacokinetic analysis of fluconazole in critically ill patients and dosing simulations for Candida infections. Antimicrobial Agents and Chemotherapy, 68(11), e0099124. https://doi.org/10.1128/aac.00991-24
187. Mazumder, P., & Singh, P. (2022). Protected attribute guided representation learning for bias mitigation in limited data. Knowledge-Based Systems, 244, 108449. https://doi.org/10.1016/j.knosys.2022.108449
188. McLatchie, Y., & Vehtari, A. (2024). Efficient estimation and correction of selection-induced bias with order statistics. Statistical Computation, 34, 132. https://doi.org/10.1007/s11222-024-10442-4
189. Medjek et al. (2021). Fault-tolerant AI-driven Intrusion Detection System for the Internet of Things. International Journal of Critical Infrastructure Protection, 34, 100436. https://doi.org/10.1016/j.ijcip.2021.100436
190. Mehta et al. (2019). A high-bias, low-variance introduction to machine learning for physicists. Physics Reports, 810, 1-124. https://doi.org/10.1016/j.physrep.2019.03.001
191. Merino-Soto et al. (2022). Parametric and nonparametric analysis of the internal structure of the psychosocial work processes questionnaire (PROPSIT) as applied to workers. International Journal of Environmental Research and Public Health, 19(13), 7970. https://doi.org/10.3390/ijerph19137970
192. Metsämuuronen, J. (2021). Directional nature of Goodman–Kruskal gamma and some consequences: Identity of Goodman–Kruskal gamma and Somers delta, and their connection to Jonckheere–Terpstra test statistic. Behaviormetrika, 48, 283–307. https://doi.org/10.1007/s41237-021-00138-8
193. Mieth et al. (2021). DeepCOMBI: explainable artificial intelligence for the analysis and discovery in genome-wide association studies. NAR Genomics and Bioinformatics, 3(3), lqab065. https://doi.org/10.1093/nargab/lqab065
194. Mishra, A. K., & Das, B. (2024). A Hoeffding D statistic approach for detecting electricity theft. In Proceedings of the 2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET) (pp. 1–6). Hyderabad, India. https://doi.org/10.1109/SEFET61574.2024.10718232
195. Mohseni, N., & Elhaik, E. (2024). Biases of Principal Component Analysis (PCA) in Physical Anthropology Studies Require a Reevaluation of Evolutionary Insights. eLife, 13, RP94685. https://doi.org/10.7554/eLife.94685.2
196. Molnar et al. (2022). General pitfalls of model-agnostic interpretation methods for machine learning models. In A. Holzinger, R. Goebel, R. Fong, T. Moon, K. R. Müller, & W. Samek (Eds.), xxAI - Beyond Explainable AI (Vol. 13200, p. 4). Springer. https://doi.org/10.1007/978-3-031-04083-2\_4
197. Moon et al. (2019). Visualizing structure and transitions in high-dimensional biological data. Nature Biotechnology, 37(12), 1482–1492. https://doi.org/10.1038/s41587-019-0336-3Nahm, F. S. (2016). Nonparametric statistical tests for continuous data: The basic concept and the practical use. Korean Journal of Anesthesiology, 69(1), 8-14. https://doi.org/10.4097/kjae.2016.69.1.8
198. Nalenz et al. (2024). Learning de-biased regression trees and forests from complex samples. Machine Learning, 113, 3379–3398. https://doi.org/10.1007/s10994-023-06439-1
199. Nazer et al. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS Digital Health, 2(6), e0000278. https://doi.org/10.1371/journal.pdig.0000278
200. Nematzadeh et al. (2019). Frequency based feature selection method using whale algorithm. Genomics, 111(6), 1946–1955. https://doi.org/10.1016/j.ygeno.2019.01.006
201. Newson, R. (2006). Confidence Intervals for Rank Statistics: Somers’ D and Extensions. The Stata Journal, 6(3), 309–334. https://doi.org/10.1177/1536867X0600600302
202. Nguyen et al. (2020). A survey on swarm intelligence approaches to feature selection in data mining. Swarm and Evolutionary Computation, 54, 100663. https://doi.org/10.1016/j.swevo.2020.100663
203. Nguyen et al. (2018). Impact of missing data strategies in studies of parental employment and health: Missing items, missing waves, and missing mothers. Social Science & Medicine, 209, 160-168. https://doi.org/10.1016/j.socscimed.2018.03.009
204. Nguyen et al. (2015). Unbiased feature selection in learning random forests for high-dimensional data. The Scientific World Journal, 2015, Article 471371. https://doi.org/10.1155/2015/471371
205. Nogueira et al. (2022). FTIR spectroscopy as a point of care diagnostic tool for diabetes and periodontitis: A saliva analysis approach. Photodiagnosis and Photodynamic Therapy, 40, 103036. https://doi.org/10.1016/j.pdpdt.2022.103036
206. Nyamundanda et al. (2010). Probabilistic principal component analysis for metabolomic data. BMC Bioinformatics, 11, Article 571. https://doi.org/10.1186/1471-2105-11-571
207. O’Brien, T. E., & Silcox, J. W. (2024). Nonlinear regression modelling: A primer with applications and caveats. Bulletin of Mathematical Biology, 86, 40. https://doi.org/10.1007/s11538-024-01274-4
208. O’Driscoll, D., & Ramirez, D. E. (2015). Response surface designs using the generalized variance inflation factors. Cogent Mathematics, 2(1). https://doi.org/10.1080/23311835.2015.1053728
209. Okoye, K., & Hosseini, S. (2024). Correlation Tests in R: Pearson Cor, Kendall’s Tau, and Spearman’s Rho. In R Programming (pp. 205-220). Springer, Singapore. https://doi.org/10.1007/978-981-97-3385-9\_12
210. Owoeye et al. (2023). Linear and nonlinear regression modeling of the chemical, physical and quality variations in Cardaba banana (Musa acuminata x balbisiana – ABB) during ripening. Food Measure, 17, 12–23. https://doi.org/10.1007/s11694-022-01570-4
211. Pantazatos et al. (2012). Cortical functional connectivity decodes subconscious, task-irrelevant threat-related emotion processing. NeuroImage, 61(4), 1355-1363. https://doi.org/10.1016/j.neuroimage.2012.03.051
212. Parmeter, C. F., & Zhao, S. (2024). An alternative corrected ordinary least squares estimator for the stochastic frontier model. In S. C. Kumbhakar, R. C. Sickles, & H. J. Wang (Eds.), Advances in Applied Econometrics (Vol. 55, pp. 355-375). Springer, Cham. https://doi.org/10.1007/978-3-031-48385-1\_15
213. Paul et al. (2024). Machine learning approach to predict blood-secretory proteins and potential biomarkers for liver cancer using omics data. Journal of Proteomics, 309, 105298. https://doi.org/10.1016/j.jprot.2024.105298
214. Peralta, B., & Soto, A. (2014). Embedded local feature selection within mixture of experts. Information Sciences, 269, 176-187. https://doi.org/10.1016/j.ins.2014.01.008
215. Peralta et al. (2021). Data imputation and compression for Parkinson's disease clinical questionnaires. Artificial Intelligence in Medicine, 114, 102051. https://doi.org/10.1016/j.artmed.2021.102051
216. Perreault, S. (2024). Simultaneous computation of Kendall’s tau and its jackknife variance. Statistics & Probability Letters, 213, 110181. https://doi.org/10.1016/j.spl.2024.110181
217. Pfeifer et al. (2025). Tree smoothing: Post-hoc regularization of tree ensembles for interpretable machine learning. Information Sciences, 690, 121564. https://doi.org/10.1016/j.ins.2024.121564
218. Pinheiro-Guedes et al. (2024). Logistic regression: Limitations in the estimation of measures of association with binary health outcomes. Acta Médica Portuguesa, 37(10), 697-705. https://doi.org/10.20344/amp.21435
219. Politi et al. (2021). Nonparametric statistical tests: Friend or foe? Jornal Brasileiro de Pneumologia, 47(4), e20210292. https://doi.org/10.36416/1806-3756/e20210292
220. Potharlanka, J. L., & Bhat M, N. (2024). Feature importance feedback with Deep Q process in ensemble-based metaheuristic feature selection algorithms. Scientific Reports, 14(1), 2923. https://doi.org/10.1038/s41598-024-53141-w
221. Prasad, S., & Bruce, L. M. (2008). Limitations of Principal Components Analysis for Hyperspectral Target Recognition. IEEE Geoscience and Remote Sensing Letters, 5(4), 625-629. https://doi.org/10.1109/LGRS.2008.2001282
222. Qi et al. (2024). A deep neural network prediction method for diabetes based on Kendall's correlation coefficient and attention mechanism. PLoS One, 19(7), e0306090. https://doi.org/10.1371/journal.pone.0306090
223. Qian et al. (2022). Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree. Expert Systems with Applications, 190, 116202. https://doi.org/10.1016/j.eswa.2021.116202
224. Qu et al. (2023). Is the radiomics-clinical combined model helpful in distinguishing between pancreatic cancer and mass-forming pancreatitis? European Journal of Radiology, 164, 110857. https://doi.org/10.1016/j.ejrad.2023.110857
225. Racette et al. (2010). Combining functional and structural tests improves the diagnostic accuracy of relevance vector machine classifiers. Journal of Glaucoma, 19(3), 167-175.1 https://doi.org/10.1097/IJG.0b013e3181a98b85
226. Raudys, S., & Pikelis, V. (1982). Collective selection of the best version of a pattern recognition system. Pattern Recognition Letters, 1(1), 7-13. https://doi.org/10.1016/0167-8655(82)90044-7
227. Rifada et al. (2022). Estimation of nonparametric ordinal logistic regression model using generalized additive models (GAM) method based on local scoring algorithm. AIP Conference Proceedings, 2668(1), 070013. https://doi.org/10.1063/5.0111771
228. Ringle et al. (2023). A perspective on using partial least squares structural equation modelling in data articles. Data in Brief, 48, 109074. https://doi.org/10.1016/j.dib.2023.109074
229. Ros et al. (2023). PDBI: A partitioning Davies-Bouldin index for clustering evaluation. Neurocomputing, 528, 178-199. https://doi.org/10.1016/j.neucom.2023.01.043
230. Roussos et al. (2022). Identifying and characterising sources of variability in digital outcome measures in Parkinson's disease. NPJ Digital Medicine, 5(1), 93. Published 2022 Jul 15. https://doi.org/10.1038/s41746-022-00643-4
231. Saccenti et al. (2020). Corruption of the Pearson correlation coefficient by measurement error and its estimation, bias, and correction under 1 different error models. Scientific Reports, 10, 438. https://doi.org/10.1038/s41598-019-57247-4
232. Sahran et al. (2018). Absolute cosine-based SVM-RFE feature selection method for prostate histopathological grading. Artificial Intelligence in Medicine, 87, 78-90. https://doi.org/10.1016/j.artmed.2018.04.002
233. Sahu et al. (2020). Linear and nonlinear mechanical responses can be quite different in models for biological tissues. Soft Matter, 16(7), 1850-1856. https://doi.org/10.1039/c9sm01068h
234. Salles et al. (2021). A bias-variance analysis of state-of-the-art random forest text classifiers. Advances in Data Analysis and Classification, 15, 379-405. https://doi.org/10.1007/s11634-020-00409-4
235. Salmerón-Gómez et al. (2025). A redefined variance inflation factor: Overcoming the limitations of the variance inflation factor. Computational Economics, 65, 337–363. https://doi.org/10.1007/s10614-024-10575-8
236. Samuel et al. (1988). Liver Transplantation for Protoporphyria: Evidence for the Predominant Role of the Erythropoietic Tissue in Protoporphyrin Overproduction. Gastroenterology, 95(3), 816-819. https://doi.org/10.1016/S0016-5085(88)80033-7
237. Sanjalawe, Y., & Althobaiti, T. (2023). DDoS Attack Detection in Cloud Computing Based on Ensemble Feature Selection and Deep Learning. Computers, Materials and Continua, 75(2), 3571-3588. https://doi.org/10.32604/cmc.2023.037386
238. Sarstedt et al. (2019). Structural model robustness checks in PLS-SEM. Tourism Economics, 26(4), 531–554. https://doi.org/10.1177/1354816618823921 (Original work published 2020)
239. Schober, P., & Vetter, T. R. (2020). Nonparametric statistical methods in medical research. Anesthesia & Analgesia, 131(6), 1862-1863. https://doi.org/10.1213/ANE.0000000000005101
240. Schwarzer et al. (2021). Predicting genotoxicity of viral vectors for stem cell gene therapy using gene expression-based machine learning. Molecular Therapy, 29(12), 3383-3397. https://doi.org/10.1016/j.ymthe.2021.06.017
241. et al. (2023). High-frequency fecal indicator bacteria (FIB) observations to assess water quality drivers at an enclosed beach. PLoS One, 18(6), e0286029. Published 2023 Jun 2. https://doi.org/10.1371/journal.pone.0286029
242. Sefidian, A. M., & Daneshpour, N. (2019). Missing value imputation using a novel grey based fuzzy c-means, mutual information based feature selection, and regression model. Expert Systems with Applications, 115, 68-94. https://doi.org/10.1016/j.eswa.2018.07.057
243. Sefidian, A. M., & Daneshpour, N. (2020). Estimating missing data using novel correlation maximization based methods. Applied Soft Computing, 91, 106249. https://doi.org/10.1016/j.asoc.2020.106249
244. Semwal, R., & Varadwaj, P. K. (2020). HumDLoc: Human Protein Subcellular Localization Prediction Using Deep Neural Network. Current Genomics, 21(7), 546-557. https://doi.org/10.2174/1389202921999200528160534
245. Shahapure, K. R., & Nicholas, C. (2020). Cluster quality analysis using Silhouette Score. In Proceedings of the 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA) (pp. 747-748). Sydney, NSW, Australia. https://doi.org/10.1109/DSAA49011.2020.00096
246. Shao et al. (2023). Underground haulage network design using HDBSCAN and RRT algorithms built on Dubins path. Mining, Metallurgy & Exploration, 40, 773–786. https://doi.org/10.1007/s42461-023-00777-3
247. Shen et al. (2024). Statistical significance of clustering with multidimensional scaling. Journal of Computational and Graphical Statistics, 33(1), 219-230. https://doi.org/10.1080/10618600.2023.2219708
248. Shen et al. (2022). Use of machine learning to identify functional connectivity changes in a clinical cohort of patients at risk for dementia. Frontiers in Aging Neuroscience, 14, 962319. Published 2022 Sep 1. https://doi.org/10.3389/fnagi.2022.962319
249. Shi et al. (2019). Variable selection and validation in multivariate modelling. Bioinformatics, 35(6), 972-980. https://doi.org/10.1093/bioinformatics/bty710
250. Shi et al. (2024). Mutual information as a measure of mixing efficiency in viscous fluids. Physical Review Research, 6(2), L022050. https://doi.org/10.1103/PhysRevResearch.6.L022050
251. Shin, H., & Markey, M. K. (2006). A machine learning perspective on the development of clinical decision support systems utilizing mass spectra of blood samples. Journal of Biomedical Informatics, 39(2), 227-248. https://doi.org/10.1016/j.jbi.2005.04.002
252. Shutaywi, M., & Kachouie, N. N. (2021). Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering. Entropy, 23(6), 759. https://doi.org/10.3390/e23060759
253. Smart et al. (2022). Clinical predictors of antipsychotic treatment resistance: Development and internal validation of a prognostic prediction model by the STRATA-G consortium. Schizophrenia Research, 250, 1-9. https://doi.org/10.1016/j.schres.2022.09.009
254. Smith et al. (2024). Lost in the Forest: Encoding categorical variables and the absent levels problem. Data Mining and Knowledge Discovery, 38, 1889-1908. https://doi.org/10.1007/s10618-024-01019-w
255. Song et al. (2015). Cancer classification in the genomic era: five contemporary problems. Human Genomics, 9, 27. Published 2015 Oct 19. https://doi.org/10.1186/s40246-015-0049-8
256. Song et al. (2021). Feature selection using bare-bones particle swarm optimization with mutual information. Pattern Recognition, 112, 107804. https://doi.org/10.1016/j.patcog.2020.107804
257. Stamate et al. (2021). PDKit: A data science toolkit for the digital assessment of Parkinson's Disease. PLoS Computational Biology, 17(3), e1008833. Published 2021 Mar 12. https://doi.org/10.1371/journal.pcbi.1008833
258. Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. Personality and Individual Differences, 42(5), 893-898. https://doi.org/10.1016/j.paid.2006.09.017
259. Steiner, P. M., & Kim, Y. (2016). The mechanics of omitted variable bias: Bias amplification and cancellation of offsetting biases. Journal of Causal Inference, 4(2), 20160009. https://doi.org/10.1515/jci-2016-0009
260. Stiglic et al. (2015). Comprehensible Predictive Modeling Using Regularized Logistic Regression and Comorbidity Based Features. PLoS One, 10(12), e0144439. Published 2015 Dec 8. https://doi.org/10.1371/journal.pone.0144439
261. Stiglic et al. (2010). Finding optimal classifiers for small feature sets in genomics and proteomics. Neurocomputing, 73(13-15), 2346-2352. https://doi.org/10.1016/j.neucom.2010.02.024
262. Stojanova et al. (2013). Dealing with spatial autocorrelation when learning predictive clustering trees. Ecological Informatics, 13, 22-39. https://doi.org/10.1016/j.ecoinf.2012.10.006
263. Strobl et al. (2007). Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics, 8, 25. https://doi.org/10.1186/1471-2105-8-25
264. Suárez-Marcote et al. (2024). Towards federated feature selection: Logarithmic division for resource-conscious methods. Neurocomputing, 596, 128099. https://doi.org/10.1016/j.neucom.2024.128099
265. Sun, J., & Li, H. (2008). Data mining method for listed companies’ financial distress prediction. Knowledge-Based Systems, 21(1), 1-5. https://doi.org/10.1016/j.knosys.2006.11.003
266. Tagiling et al. (2025). Gastric accommodation testing using hybrid nuclear imaging volumetry and combined high-resolution manometry-nutrient drink test: A pilot study in healthy individuals. Neurogastroenterology & Motility. Advance online publication. https://doi.org/10.1111/nmo.15006
267. Takefuji, Y. (2024). Mitigating biases in feature selection and importance assessments in predictive models using LASSO regression. Oral Oncology, 159, 107090. https://doi.org/10.1016/j.oraloncology.2024.107090
268. Tang et al. (2024). Prediction models for COVID-19 disease outcomes. Emerging Microbes & Infections, 13(1), 2361791. https://doi.org/10.1080/22221751.2024.2361791
269. Tang et al. (2024). Accurate and visualiable discrimination of Chenpi age using 2D-CNN and Grad-CAM++ based on infrared spectral images. Food Chemistry: X, 23, 101759. https://doi.org/10.1016/j.fochx.2024.101759
270. Tang et al. (2024). An intelligent airflow perception model for metal mines based on CNN-LSTM architecture. Process Safety and Environmental Protection, 187, 1234-1247. https://doi.org/10.1016/j.psep.2024.05.044
271. Tarabichi et al. (2015). Revisiting the transcriptional analysis of primary tumours and associated nodal metastases with enhanced biological and statistical controls: application to thyroid cancer. British Journal of Cancer, 112(10), 1665-1674. https://doi.org/10.1038/bjc.2014.665
272. Taylor et al. (2021). Predicting subclinical psychotic-like experiences on a continuum using machine learning. NeuroImage, 241, 118329. https://doi.org/10.1016/j.neuroimage.2021.118329
273. Tekchandani et al. (2020). Performance improvement of mediastinal lymph node severity detection using GAN and Inception network. Computer Methods and Programs in Biomedicine, 194, 105478. https://doi.org/10.1016/j.cmpb.2020.105478
274. Thakur, D., & Biswas, S. (2024). Permutation importance based modified guided regularized random forest in human activity recognition with smartphone. Engineering Applications of Artificial Intelligence, 129, 107681. https://doi.org/10.1016/j.engappai.2023.107681
275. Thanjavur et al. (2021). Deep Learning Recurrent Neural Network for Concussion Classification in Adolescents Using Raw Electroencephalography Signals: Toward a Minimal Number of Sensors. Frontiers in Human Neuroscience, 15, 734501. Published 2021 Nov 24. https://doi.org/10.3389/fnhum.2021.734501
276. Thomson et al. (2015). Blood-based identification of non-responders to anti-TNF therapy in rheumatoid arthritis. BMC Medical Genomics, 8, 26. Published 2015 Jun 3. http://doi.org/10.1186/s12920-015-0100-6
277. Timmons et al. (2023). A Call to Action on Assessing and Mitigating Bias in Artificial Intelligence Applications for Mental Health. Perspectives on Psychological Science, 18(5), 1062-1096. https://doi.org/10.1177/17456916221134490
278. Tomalin et al. (2020). Early Quantification of Systemic Inflammatory Proteins Predicts Long-Term Treatment Response to Tofacitinib and Etanercept. Journal of Investigative Dermatology, 140(5), 1026-1034. https://doi.org/10.1016/j.jid.2019.09.023
279. Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: strengths, limitations, and misconceptions. Annual Review of Clinical Psychology, 1, 31–65. https://doi.org/10.1146/annurev.clinpsy.1.102803.144239
280. Torres Moral et al. (2022). Methods for Stratification and Validation Cohorts: A Scoping Review. Journal of Personalized Medicine, 12(5), 688. Published 2022 Apr 26. https://doi.org/10.3390/jpm12050688
281. Tracy et al. (2019). RESCUE: imputing dropout events in single-cell RNA-sequencing data. BMC Bioinformatics, 20(1), 388. Published 2019 Jul 12. https://doi.org/10.1186/s12859-019-2977-0
282. Tran et al. (2014). The role of monocytes in the development of Tuberculosis-associated Immune Reconstitution Inflammatory Syndrome. Immunobiology, 219(1), 37-44. https://doi.org/10.1016/j.imbio.2013.07.004
283. Tserkis et al. (2025). Quantifying total correlations in quantum systems through the Pearson correlation coefficient. Physics Letters A, 543, 130432. https://doi.org/10.1016/j.physleta.2025.130432
284. Ugirumurera et al. (2024). Addressing bias in bagging and boosting regression models. Scientific Reports, 14(1), 18452. https://doi.org/10.1038/s41598-024-68907-5
285. Ugrinowitsch et al. (2004). Limitations of ordinary least squares models in analyzing repeated measures data. Medicine and Science in Sports and Exercise, 36(12), 2144–2148. https://doi.org/10.1249/01.mss.0000147580.40591.75
286. Umeki et al. (2025). Evaluation of information flows in the RAS-MAPK system using transfer entropy measurements. eLife, 14, e104432. https://doi.org/10.7554/eLife.104432
287. Ünal, B. (2022). Causality analysis for COVID-19 among countries using effective transfer entropy. Entropy, 24(8), 1115. https://doi.org/10.3390/e24081115
288. Vable et al. (2019). Performance of Matching Methods as Compared With Unmatched Ordinary Least Squares Regression Under Constant Effects. American Journal of Epidemiology, 188(7), 1345–1354. https://doi.org/10.1093/aje/kwz093
289. van Koppen et al. (2018). Uncovering a predictive molecular signature for the onset of NASH-related fibrosis in a translational NASH mouse model. Cellular and Molecular Gastroenterology and Hepatolog2y, 5(1), 83-98.e10. https://doi.org/10.1016/j.jcmgh.2017.10.001
290. van Maanen et al. (2019). Fast and slow errors: Logistic regression to identify patterns in accuracy–response time relationships. Behavior Research Methods, 51, 2378–2389. https://doi.org/10.3758/s13428-018-1110-z
291. Varley et al. (2023). Partial entropy decomposition reveals higher-order information structures in human brain activity. Proceedings of the National Academy of Sciences of the United States of America, 120(30), e2300888120. https://doi.org/10.1073/pnas.2300888120
292. Vos et al. (2024). Making pathologists ready for the new AI era: changes in required competencies. Modern Pathology, 100657. https://doi.org/10.1016/j.modpat.2024.100657
293. Waernbaum, I., & Pazzagli, L. (2023). Model misspecification and bias for inverse probability weighting estimators of average causal effects. Biometrical Journal. Biometrische Zeitschrift, 65(2), e2100118. https://doi.org/10.1002/bimj.202100118
294. Wallace et al. (2023). Use and misuse of random forest variable importance metrics in medicine: demonstrations through incident stroke prediction. BMC Medical Research Methodology, 23(1), 144. https://doi.org/10.1186/s12874-023-01965-x
295. Wang et al. (2012). Interaction-based feature selection and classification for high-dimensional biological data. Bioinformatics, 28(21), 2834-2842. https://doi.org/10.1093/bioinformatics/bts531
296. Wang et al. (2023). Semi-supervised inference for nonparametric logistic regression. Statistics in Medicine, 42(15), 2573–2589. https://doi.org/10.1002/sim.9737
297. Wei et al. (2023). Using Machine Learning Methods to Study Colorectal Cancer Tumor Micro-Environment and Its Biomarkers. International Journal of Molecular Sciences, 24(13), 11133. Published 2023 Jul 6. https://doi.org/10.3390/ijms241311133
298. Weintraub et al. (2021). Using machine learning analyses of speech to classify levels of expressed emotion in parents of youth with mood disorders. Journal of Psychiatric Research, 136, 39-46. https://doi.org/10.1016/j.jpsychires.2021.01.019
299. Wesolowski et al. (2010). Limitations of ordinary least squares fitting of gamma variate functions to plasma-clearance curves. J Nucl Med, 51(Supplement 2), No.1674. https://doi.org/10.1007/s10928-010-9167-z
300. Wongoutong, C. (2024). The impact of neglecting feature scaling in k-means clustering. PLoS One, 19(12), e0310839. https://doi.org/10.1371/journal.pone.0310839
301. Wood et al. (2024). Model-agnostic variable importance for predictive uncertainty: An entropy-based approach. Data Mining and Knowledge Discovery, 38, 4184-4216. https://doi.org/10.1007/s10618-024-01070-7
302. Work et al. (1989). Limitations of a conventional logistic regression model based on left ventricular ejection fraction in predicting coronary events after myocardial infarction. American Journal of Cardiology, 64(12), 702-707. https://doi.org/10.1016/0002-9149(89)90751-0
303. Wüthrich, K., & Zhu, Y. (2023). Omitted variable bias of Lasso-based inference methods: A finite sample analysis. Review of Economics and Statistics, 105(4), 982–997. https://doi.org/10.1162/rest\_a\_01128
304. Xie et al. (2020). Single-Cell Classification Using Mass Spectrometry through Interpretable Machine Learning. Analytical Chemistry, 92(13), 9338-9347. https://doi.org/10.1021/acs.analchem.0c01660
305. Xu et al. (2021). Combing machine learning and elemental profiling for geographical authentication of Chinese Geographical Indication (GI) rice. NPJ Science of Food, 5(1), 18. Published 2021 Jul 8. https://doi.org/10.1038/s41538-021-00100-8
306. Xu et al. (2024). Evaluation of mean shift, ComBat, and CycleGAN for harmonizing brain connectivity matrices across sites. In Proceedings of SPIE (Vol. 12926, Article 129261X). https://doi.org/10.1117/12.3005563
307. Xu et al. (2025). On summed nonparametric dependence measures in high dimensions, fixed or large samples. Computational Statistics & Data Analysis, 205, 108109. https://doi.org/10.1016/j.csda.2024.108109
308. Xu et al. (2019). Review of classical dimensionality reduction and sample selection methods for large-scale data processing. Neurocomputing, 328, 5-15. https://doi.org/10.1016/j.neucom.2018.02.100
309. Yang et al. (2024). A fast dual-module hybrid high-dimensional feature selection algorithm. Information Sciences, 681, 121185. https://doi.org/10.1016/j.ins.2024.121185
310. Yang et al. (2024). An improved binary particle swarm optimization algorithm for clinical cancer biomarker identification in microarray data. Computer Methods and Programs in Biomedicine, 244, 107987. https://doi.org/10.1016/j.cmpb.2023.107987
311. Yang, Y., & Webb, G. I. (2009). Discretization for naive-Bayes learning: Managing discretization bias and variance. Machine Learning, 74, 39–74. https://doi.org/10.1007/s10994-008-5083-5
312. Yao, Y., & Ochoa, A. (2023). Limitations of principal components in quantitative genetic association models for human studies. eLife, 12, Article e79238. https://doi.org/10.7554/eLife.79238
313. Yip, S. S., & Aerts, H. J. (2016). Applications and limitations of radiomics. Physics in Medicine and Biology, 61(13), R150-R166. https://doi.org/10.1088/0031-9155/61/13/R150
314. Yu et al. (2024). Nomogram for predicting in-hospital mortality in trauma patients undergoing resuscitative endovascular balloon occlusion of the aorta: a retrospective multicenter study. Scientific Reports, 14(1), 9164. https://doi.org/10.1038/s41598-024-59861-3
315. Yu et al. (2021). Deep exploration of random forest model boosts the interpretability of machine learning studies of complicated immune responses and lung burden of nanoparticles. Science Advances, 7(22), eabf4130. Published 2021 May 26. https://doi.org/10.1126/sciadv.abf4130
316. Yu, H., & Hutson, A. D. (2024). A robust Spearman correlation coefficient permutation test. Communications in Statistics: Theory and Methods, 53(6), 2141-2153. https://doi.org/10.1080/03610926.2022.2121144
317. Yu et al. (2023). A bidirectional dynamic grouping multi-objective evolutionary algorithm for feature selection on high-dimensional classification. Information Sciences, 648, 119619. https://doi.org/10.1016/j.ins.2023.119619
318. Zajac, G., & Ignatiev, A. (1979). High temperature optical and structural degradation of black chrome coatings. Solar Energy Materials, 2(2), 239-247. https://doi.org/10.1016/0165-1633(79)90021-2
319. Zarei et al. (2021). Bias correction of global ensemble precipitation forecasts by Random Forest method. Earth Science Informatics, 14, 677-689. https://doi.org/10.1007/s12145-021-00577-7
320. Zemariam et al. (2024). Employing supervised machine learning algorithms for classification and prediction of anemia among youth girls in Ethiopia. Scientific Reports, 14(1), 9080. Published 2024 Apr 20. https://doi.org/10.1038/s41598-024-60027-4
321. Zhang et al. (2024). Detecting financial contagion using a new nonparametric measure of asymmetric comovements. International Review of Economics and Finance, 89, 284-296. https://doi.org/10.1016/j.iref.2023.07.067
322. Zhang et al. (2014). Modeling pathologic response of esophageal cancer to chemoradiation therapy using spatial-temporal 18F-FDG PET features, clinical parameters, and demographics. International Journal of Radiation Oncology•Biology•Physics, 88(1), 195-203. https://doi.org/10.1016/j.ijrobp.2013.09.037
323. Zhang et al. (2021). A Radiomics Nomogram for Preoperative Prediction of Clinical Occult Lymph Node Metastasis in cT1-2N0M0 Solid Lung Adenocarcinoma. Cancer Management and Research, 13, 8157-8167. Published 2021 Oct 28. https://doi.org/10.2147/CMAR.S330824
324. Zhang et al. (2022). Modulation format identification using the Calinski–Harabasz index. Applied Optics, 61(3), 851-857. https://doi.org/10.1364/ao.448043
325. Zhang et al. (2023). Feature selection based on neighborhood rough sets and Gini index. PeerJ Computer Science, 9, e1711. Published 2023 Dec 12. https://doi.org/10.7717/peerj-cs.1711
326. Zhang et al. (2017). A return-cost-based binary firefly algorithm for feature selection. Information Sciences, 418-419, 561-574. https://doi.org/10.1016/j.ins.2017.08.047
327. Zhao et al. (2022). Multiple imputation method of missing credit risk assessment data based on generative adversarial networks. Applied Soft Computing, 126, 109273. https://doi.org/10.1016/j.asoc.2022.109273
328. Zhou et al. (2015). The performance of corporate financial distress prediction models with features selection guided by domain knowledge and data mining approaches. Knowledge-Based Systems, 85, 52-61. https://doi.org/10.1016/j.knosys.2015.04.017
329. Zhou et al. (2024). Rank-based indices for testing independence between two high-dimensional vectors. The Annals of Statistics, 52(1), 184-206. https://doi.org/10.1214/23-aos2339
330. Zuur et al. (2009). Limitations of linear regression applied on ecological data. In Mixed effects models and extensions in ecology with R. Statistics for biology and health (pp. 43-67). Springer, New York, NY. https://doi.org/10.1007/978-0-387-87458-6\_2