

Anne-Laure Boulesteix

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/3702281/publications.pdf>

Version: 2024-02-01

130
papers

12,005
citations

46918

47
h-index

29081

104
g-index

144
all docs

144
docs citations

144
times ranked

17391
citing authors

#	ARTICLE	IF	CITATIONS
1	Over-optimistic evaluation and reporting of novel cluster algorithms: an illustrative study. <i>Advances in Data Analysis and Classification</i> , 2023, 17, 211-238.	0.9	4
2	COMPANION: development of a patient-centred complexity and casemix classification for adult palliative care patients based on needs and resource use – a protocol for a cross-sectional multi-centre study. <i>BMC Palliative Care</i> , 2022, 21, 18.	0.8	3
3	Interaction forests: Identifying and exploiting interpretable quantitative and qualitative interaction effects. <i>Computational Statistics and Data Analysis</i> , 2022, 171, 107460.	0.7	7
4	Over-optimism in benchmark studies and the multiplicity of design and analysis options when interpreting their results. <i>Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery</i> , 2022, 12, .	4.6	9
5	Validation of cluster analysis results on validation data: A systematic framework. <i>Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery</i> , 2022, 12, .	4.6	29
6	Critical appraisal of artificial intelligence-based prediction models for cardiovascular disease. <i>European Heart Journal</i> , 2022, 43, 2921-2930.	1.0	50
7	Outcome of patients treated with extracorporeal life support in cardiogenic shock complicating acute myocardial infarction: 1-year result from the ECLS-Shock study. <i>Clinical Research in Cardiology</i> , 2021, 110, 1412-1420.	1.5	24
8	NetCoMi: network construction and comparison for microbiome data in R. <i>Briefings in Bioinformatics</i> , 2021, 22, .	3.2	222
9	Examining the robustness of observational associations to model, measurement and sampling uncertainty with the vibration of effects framework. <i>International Journal of Epidemiology</i> , 2021, 50, 266-278.	0.9	18
10	Improved Outcome Prediction Across Data Sources Through Robust Parameter Tuning. <i>Journal of Classification</i> , 2021, 38, 212-231.	1.2	5
11	Large-scale benchmark study of survival prediction methods using multi-omics data. <i>Briefings in Bioinformatics</i> , 2021, 22, .	3.2	53
12	The multiplicity of analysis strategies jeopardizes replicability: lessons learned across disciplines. <i>Royal Society Open Science</i> , 2021, 8, 201925.	1.1	29
13	On the optimistic performance evaluation of newly introduced bioinformatic methods. <i>Genome Biology</i> , 2021, 22, 152.	3.8	20
14	Sampling uncertainty versus method uncertainty: A general framework with applications to omics biomarker selection. <i>Biometrical Journal</i> , 2020, 62, 670-687.	0.6	12
15	Statistical learning approaches in the genetic epidemiology of complex diseases. <i>Human Genetics</i> , 2020, 139, 73-84.	1.8	14
16	Combining clinical and molecular data in regression prediction models: insights from a simulation study. <i>Briefings in Bioinformatics</i> , 2020, 21, 1904-1919.	3.2	11
17	A Replication Crisis in Methodological Research?. <i>Significance</i> , 2020, 17, 18-21.	0.3	21
18	Introduction to statistical simulations in health research. <i>BMJ Open</i> , 2020, 10, e039921.	0.8	24

#	ARTICLE	IF	CITATIONS
19	Single-center versus multi-center data sets for molecular prognostic modeling: a simulation study. <i>Radiation Oncology</i> , 2020, 15, 109.	1.2	3
20	On the asymptotic behaviour of the variance estimator of a $\langle \text{mml:math xmlns:mml="http://www.w3.org/1998/Math/MathML" display="inline" id="d1e208" altimg="si7.svg" \rangle \langle \text{mml:mi} \rangle U \langle \text{mml:mi} \rangle \langle \text{mml:math} \rangle$ -statistic. <i>Journal of Statistical Planning and Inference</i> , 2020, 209, 101-111.	0.4	3
21	Use of Resampling Procedures to Investigate Issues of Model Building and Its Stability. , 2020, , 1-24.		1
22	Modelling Individual Response to Treatment and Its Uncertainty:A Review of Statistical Methods and Challenges for Future Research. <i>Boston Studies in the Philosophy and History of Science</i> , 2020, , 319-344.	0.4	0
23	Making complex prediction rules applicable for readers: Current practice in random forest literature and recommendations. <i>Biometrical Journal</i> , 2019, 61, 1314-1328.	0.6	13
24	A plea for taking all available clinical information into account when assessing the predictive value of omics data. <i>BMC Medical Research Methodology</i> , 2019, 19, 162.	1.4	10
25	Essential guidelines for computational method benchmarking. <i>Genome Biology</i> , 2019, 20, 125.	3.8	114
26	Extracorporeal Life Support in Cardiogenic Shock Complicating Acute Myocardial Infarction. <i>Journal of the American College of Cardiology</i> , 2019, 73, 2355-2357.	1.2	79
27	Immune function testing in sepsis patients receiving sodium selenite. <i>Journal of Critical Care</i> , 2019, 52, 208-212.	1.0	5
28	Hyperparameters and tuning strategies for random forest. <i>Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery</i> , 2019, 9, e1301.	4.6	651
29	Independent validation of a new reirradiation risk score (RRRS) for glioma patients predicting post-recurrence survival: A multicenter DTK/ROG analysis. <i>Radiotherapy and Oncology</i> , 2018, 127, 121-127.	0.3	37
30	On the choice and influence of the number of boosting steps for high-dimensional linear Cox-models. <i>Computational Statistics</i> , 2018, 33, 1195-1215.	0.8	20
31	Guest Editorialâ€”Special Collection Topic: Statistical Systems Theory in Cancer Modeling, Diagnosis, and Therapy. <i>Cancer Informatics</i> , 2018, 17, 117693511876094.	0.9	0
32	A computationally fast variable importance test for random forests for high-dimensional data. <i>Advances in Data Analysis and Classification</i> , 2018, 12, 885-915.	0.9	115
33	On the necessity and design of studies comparing statistical methods. <i>Biometrical Journal</i> , 2018, 60, 216-218.	0.6	66
34	Critical review of reporting of the data analysis step in metabolomics. <i>Metabolomics</i> , 2018, 14, 7.	1.4	52
35	Priority-Lasso: a simple hierarchical approach to the prediction of clinical outcome using multi-omics data. <i>BMC Bioinformatics</i> , 2018, 19, 322.	1.2	33
36	OC-0588: Validation of the reRT risk score (RRRS) in glioma patients: A multicenter DTK/ROG analysis. <i>Radiotherapy and Oncology</i> , 2018, 127, S307-S308.	0.3	0

#	ARTICLE	IF	CITATIONS
37	Mediation analysis reveals common mechanisms of RUNX1 point mutations and RUNX1/RUNX1T1 fusions influencing survival of patients with acute myeloid leukemia. <i>Scientific Reports</i> , 2018, 8, 11293.	1.6	9
38	Random forest versus logistic regression: a large-scale benchmark experiment. <i>BMC Bioinformatics</i> , 2018, 19, 270.	1.2	446
39	Effects of the Cardio First Angel™ on chest compression performance. <i>Technology and Health Care</i> , 2018, 26, 69-80.	0.5	6
40	A computationally fast variable importance test for random forests for high-dimensional data. <i>Advances in Data Analysis and Classification</i> , 2018, 12, 885.	0.9	14
41	Improving cross-study prediction through add-on batch effect adjustment or add-on normalization. <i>Bioinformatics</i> , 2017, 33, 397-404.	1.8	18
42	Detection of influential points as a byproduct of resampling-based variable selection procedures. <i>Computational Statistics and Data Analysis</i> , 2017, 116, 19-31.	0.7	4
43	IPF-LASSO: Integrative L_1 -Penalized Regression with Penalty Factors for Prediction Based on Multi-Omics Data. <i>Computational and Mathematical Methods in Medicine</i> , 2017, 2017, 1-14.	0.7	62
44	Towards evidence-based computational statistics: lessons from clinical research on the role and design of real-data benchmark studies. <i>BMC Medical Research Methodology</i> , 2017, 17, 138.	1.4	43
45	On Fishing for Significance and Statistician's Degree of Freedom in the Era of Big Molecular Data. , 2017, , 155-170.		4
46	Benchmarking for Clustering Methods Based on Real Data: A Statistical View. <i>Studies in Classification, Data Analysis, and Knowledge Organization</i> , 2017, , 73-82.	0.1	3
47	The Residual-Based Predictiveness Curve: A Visual Tool to Assess the Performance of Prediction Models. <i>Biometrics</i> , 2016, 72, 392-401.	0.8	4
48	Categorical variables with many categories are preferentially selected in bootstrap-based model selection procedures for multivariable regression models. <i>Biometrical Journal</i> , 2016, 58, 652-673.	0.6	10
49	Subsampling Versus Bootstrapping in Resampling-Based Model Selection for Multivariable Regression. <i>Biometrics</i> , 2016, 72, 272-280.	0.8	70
50	Interleukin-22 is elevated in lavage from patients with lung cancer and other pulmonary diseases. <i>BMC Cancer</i> , 2016, 16, 409.	1.1	19
51	Assessment of predictive performance in incomplete data by combining internal validation and multiple imputation. <i>BMC Medical Research Methodology</i> , 2016, 16, 144.	1.4	49
52	Pitfalls of hypothesis tests and model selection on bootstrap samples: Causes and consequences in biometrical applications. <i>Biometrical Journal</i> , 2016, 58, 447-473.	0.6	17
53	Gram-negative and -positive bacteria differentiation in blood culture samples by headspace volatile compound analysis. <i>Journal of Biological Research</i> , 2016, 23, 3.	2.2	13
54	Combining location-and-scale batch effect adjustment with data cleaning by latent factor adjustment. <i>BMC Bioinformatics</i> , 2016, 17, 27.	1.2	37

#	ARTICLE	IF	CITATIONS
55	Improved prediction of complex diseases by common genetic markers: state of the art and further perspectives. <i>Human Genetics</i> , 2016, 135, 259-272.	1.8	37
56	Random forest for ordinal responses: Prediction and variable selection. <i>Computational Statistics and Data Analysis</i> , 2016, 96, 57-73.	0.7	125
57	Which Resampling-Based Error Estimator for Benchmark Studies? A Power Analysis with Application to PLS-LDA. <i>Springer Proceedings in Mathematics and Statistics</i> , 2016, , 45-57.	0.1	1
58	Exhaled breath volatile organic and inorganic compound composition in end-stage renal disease. <i>Clinical Nephrology</i> , 2016, 86, 132-140.	0.4	9
59	Publication Bias in Methodological Computational Research. <i>Cancer Informatics</i> , 2015, 14s5, CIN.S30747.	0.9	15
60	Live music reduces stress levels in very low birthweight infants. <i>Acta Paediatrica, International Journal of Paediatrics</i> , 2015, 104, 360-367.	0.7	36
61	A measure of the impact of CV incompleteness on prediction error estimation with application to PCA and normalization. <i>BMC Medical Research Methodology</i> , 2015, 15, 95.	1.4	15
62	Gender mismatch in allograft aortic valve surgery. <i>Interactive Cardiovascular and Thoracic Surgery</i> , 2015, 21, 329-335.	0.5	2
63	Ten Simple Rules for Reducing Overoptimistic Reporting in Methodological Computational Research. <i>PLoS Computational Biology</i> , 2015, 11, e1004191.	1.5	43
64	Letter to the Editor: On the term 'interaction' and related phrases in the literature on Random Forests. <i>Briefings in Bioinformatics</i> , 2015, 16, 338-345.	3.2	48
65	Letter to the Editor: On Reviews and Papers on New Methods. <i>Briefings in Bioinformatics</i> , 2015, 16, 365-366.	3.2	0
66	A Statistical Framework for Hypothesis Testing in Real Data Comparison Studies. <i>American Statistician</i> , 2015, 69, 201-212.	0.9	29
67	On stability issues in deriving multivariable regression models. <i>Biometrical Journal</i> , 2015, 57, 531-555.	0.6	43
68	Investigating the prediction ability of survival models based on both clinical and omics data: two case studies. <i>Statistics in Medicine</i> , 2014, 33, 5310-5329.	0.8	36
69	Added predictive value of omics data: specific issues related to validation illustrated by two case studies. <i>BMC Medical Research Methodology</i> , 2014, 14, 117.	1.4	6
70	Cross-study validation for the assessment of prediction algorithms. <i>Bioinformatics</i> , 2014, 30, i105-i112.	1.8	75
71	Machine learning versus statistical modeling. <i>Biometrical Journal</i> , 2014, 56, 588-593.	0.6	65
72	An AUC-based permutation variable importance measure for random forests. <i>BMC Bioinformatics</i> , 2013, 14, 119.	1.2	179

#	ARTICLE	IF	CITATIONS
73	Correcting the Optimal Resampling-Based Error Rate by Estimating the Error Rate of Wrapper Algorithms. <i>Biometrics</i> , 2013, 69, 693-702.	0.8	18
74	Molecular Differences between Chronic and Aggressive Periodontitis. <i>Journal of Dental Research</i> , 2013, 92, 1081-1088.	2.5	77
75	Homografts in aortic position: does blood group incompatibility have an impact on patient outcomes? <i>Interactive Cardiovascular and Thoracic Surgery</i> , 2013, 16, 619-624.	0.5	4
76	On representative and illustrative comparisons with real data in bioinformatics: response to the letter to the editor by Smith <i>et al.</i> <i>Bioinformatics</i> , 2013, 29, 2664-2666.	1.8	22
77	What is the "normal" fetal heart rate?. <i>PeerJ</i> , 2013, 1, e82.	0.9	112
78	Application of Microarray Analysis on Computer Cluster and Cloud Platforms. <i>Methods of Information in Medicine</i> , 2013, 52, 65-71.	0.7	3
79	A Plea for Neutral Comparison Studies in Computational Sciences. <i>PLoS ONE</i> , 2013, 8, e61562.	1.1	86
80	Complexity Selection with Cross-validation for Lasso and Sparse Partial Least Squares Using High-Dimensional Data. <i>Studies in Classification, Data Analysis, and Knowledge Organization</i> , 2013, , 261-268.	0.1	7
81	Iterative Reconstruction of High-Dimensional Gaussian Graphical Models Based on a New Method to Estimate Partial Correlations under Constraints. <i>PLoS ONE</i> , 2013, 8, e60536.	1.1	1
82	Random forest Gini importance favours SNPs with large minor allele frequency: impact, sources and recommendations. <i>Briefings in Bioinformatics</i> , 2012, 13, 292-304.	3.2	92
83	Breath isoprene concentrations in persons undergoing general anesthesia and in healthy volunteers. <i>Journal of Breath Research</i> , 2012, 6, 046004.	1.5	13
84	Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. <i>Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery</i> , 2012, 2, 493-507.	4.6	516
85	An eight-gene expression signature for the prediction of survival and time to treatment in chronic lymphocytic leukemia. <i>Leukemia</i> , 2011, 25, 1639-1645.	3.3	59
86	AUC-RF: A New Strategy for Genomic Profiling with Random Forest. <i>Human Heredity</i> , 2011, 72, 121-132.	0.4	122
87	Inhibition of radiation induced migration of human head and neck squamous cell carcinoma cells by blocking of EGF receptor pathways. <i>BMC Cancer</i> , 2011, 11, 388.	1.1	50
88	Use of pretransformation to cope with extreme values in important candidate features. <i>Biometrical Journal</i> , 2011, 53, 673-688.	0.6	3
89	Stability Investigations of Multivariable Regression Models Derived from Low- and High-Dimensional Data. <i>Journal of Biopharmaceutical Statistics</i> , 2011, 21, 1206-1231.	0.4	104
90	Editorial. <i>Briefings in Bioinformatics</i> , 2011, 12, 187-188.	3.2	3

#	ARTICLE	IF	CITATIONS
91	Added predictive value of high-throughput molecular data to clinical data and its validation. Briefings in Bioinformatics, 2011, 12, 215-229.	3.2	46
92	Differential expression of the chemokines GRO-2, GRO-3, and interleukin-8 in colon cancer and their impact on metastatic disease and survival. International Journal of Colorectal Disease, 2010, 25, 573-581.	1.0	98
93	Testing the additional predictive value of high-dimensional molecular data. BMC Bioinformatics, 2010, 11, 78.	1.2	31
94	Mid-term outcomes of patients with PCI prior to CABG in comparison to patients with primary CABG. Vascular Health and Risk Management, 2010, 6, 495.	1.0	18
95	Over-optimism in bioinformatics: an illustration. Bioinformatics, 2010, 26, 1990-1998.	1.8	90
96	Over-optimism in bioinformatics research. Bioinformatics, 2010, 26, 437-439.	1.8	110
97	Centrosome abnormalities in head and neck squamous cell carcinoma (HNSCC). Acta Oto-Laryngologica, 2009, 129, 205-213.	0.3	12
98	Stability and aggregation of ranked gene lists. Briefings in Bioinformatics, 2009, 10, 556-568.	3.2	138
99	Regularized estimation of large-scale gene association networks using graphical Gaussian models. BMC Bioinformatics, 2009, 10, 384.	1.2	169
100	Optimal classifier selection and negative bias in error rate estimation: an empirical study on high-dimensional prediction. BMC Medical Research Methodology, 2009, 9, 85.	1.4	56
101	Survival prediction using gene expression data: A review and comparison. Computational Statistics and Data Analysis, 2009, 53, 1590-1603.	0.7	98
102	Comments on: Augmenting the bootstrap to analyze high dimensional genomic data. Test, 2008, 17, 31-35.	0.7	3
103	Time and rate of sinus formation in pilonidal sinus disease. International Journal of Colorectal Disease, 2008, 23, 359-364.	1.0	83
104	Methylene Blue halves the long-term recurrence rate in acute pilonidal sinus disease. International Journal of Colorectal Disease, 2008, 23, 181-187.	1.0	59
105	Penalized Partial Least Squares with applications to B-spline transformations and functional data. Chemometrics and Intelligent Laboratory Systems, 2008, 94, 60-69.	1.8	55
106	Conditional variable importance for random forests. BMC Bioinformatics, 2008, 9, 307.	1.2	2,129
107	CMA " a comprehensive Bioconductor package for supervised classification with high dimensional data. BMC Bioinformatics, 2008, 9, 439.	1.2	84
108	Influence of diatoms on copepod reproduction. II. Uncorrelated effects of diatom-derived $\hat{1}\pm, \hat{1}^2, \hat{1}^3, \hat{1}^4$ -unsaturated aldehydes and polyunsaturated fatty acids on <i>Calanus helgolandicus</i> in the field. Progress in Oceanography, 2008, 77, 30-44.	1.5	48

#	ARTICLE	IF	CITATIONS
109	Microarray-Based Prediction of Tumor Response to Neoadjuvant Radiochemotherapy of Patients With Locally Advanced Rectal Cancer. <i>Clinical Gastroenterology and Hepatology</i> , 2008, 6, 53-61.	2.4	128
110	Microarray-based classification and clinical predictors: on combined classifiers and additional predictive value. <i>Bioinformatics</i> , 2008, 24, 1698-1706.	1.8	76
111	Evaluating Microarray-based Classifiers: An Overview. <i>Cancer Informatics</i> , 2008, 6, CIN.S408.	0.9	70
112	Effects of Aprotinin Dosage on Renal Function. <i>Anesthesiology</i> , 2008, 108, 189-198.	1.3	31
113	WilcoxCV: an R package for fast variable selection in cross-validation. <i>Bioinformatics</i> , 2007, 23, 1702-1704.	1.8	18
114	Methylation of Tumor-Related Genes in Neoadjuvant-Treated Gastric Cancer: Relation to Therapy Response and Clinicopathologic and Molecular Features. <i>Clinical Cancer Research</i> , 2007, 13, 5095-5102.	3.2	35
115	Aprotinin and Anaphylaxis: Analysis of 12,403 Exposures to Aprotinin in Cardiac Surgery. <i>Annals of Thoracic Surgery</i> , 2007, 84, 1144-1150.	0.7	65
116	Multiple Testing for SNP-SNP Interactions. <i>Statistical Applications in Genetics and Molecular Biology</i> , 2007, 6, Article37.	0.2	15
117	The Normal Fetal Heart Rate Study: Analysis Plan. <i>Nature Precedings</i> , 2007, , .	0.1	2
118	The Normal Fetal Heart Rate Study: Analysis Plan. <i>Nature Precedings</i> , 2007, , .	0.1	1
119	Unbiased split selection for classification trees based on the Gini Index. <i>Computational Statistics and Data Analysis</i> , 2007, 52, 483-501.	0.7	201
120	Maximally selected Chi-squared statistics and non-monotonic associations: An exact approach based on two cutpoints. <i>Computational Statistics and Data Analysis</i> , 2007, 51, 6295-6306.	0.7	12
121	Bias in random forest variable importance measures: Illustrations, sources and a solution. <i>BMC Bioinformatics</i> , 2007, 8, 25.	1.2	2,328
122	Identification of interaction patterns and classification with applications to microarray data. <i>Computational Statistics and Data Analysis</i> , 2006, 50, 783-802.	0.7	11
123	Maximally Selected Chi-square Statistics for Ordinal Variables. <i>Biometrical Journal</i> , 2006, 48, 451-462.	0.6	25
124	Maximally Selected Chi-Square Statistics and Binary Splits of Nominal Variables. <i>Biometrical Journal</i> , 2006, 48, 838-848.	0.6	20
125	Reader's Reaction to "Dimension Reduction for Classification with Gene Expression Microarray Data" by Dai et al (2006). <i>Statistical Applications in Genetics and Molecular Biology</i> , 2006, 5, Article16.	0.2	9
126	Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. <i>Briefings in Bioinformatics</i> , 2006, 8, 32-44.	3.2	611

#	ARTICLE	IF	CITATIONS
127	Predicting transcription factor activities from combined analysis of microarray and ChIP data: a partial least squares approach. <i>Theoretical Biology and Medical Modelling</i> , 2005, 2, 23.	2.1	98
128	La genómica funcional: herramientas para mejorar la sanidad y el bienestar del ganado. <i>OIE Revue Scientifique Et Technique</i> , 2005, 24, 355-377.	0.5	12
129	PLS Dimension Reduction for Classification with Microarray Data. <i>Statistical Applications in Genetics and Molecular Biology</i> , 2004, 3, 1-30.	0.2	160
130	A CART-based approach to discover emerging patterns in microarray data. <i>Bioinformatics</i> , 2003, 19, 2465-2472.	1.8	74