

# Jian Huang

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/4894187/publications.pdf>

Version: 2024-02-01

145  
papers

9,330  
citations

76326

40  
h-index

46799

89  
g-index

147  
all docs

147  
docs citations

147  
times ranked

8878  
citing authors

#	ARTICLE	IF	CITATIONS
1	Regularized gene selection in cancer microarray meta-analysis. BMC Bioinformatics, 2009, 10, 1.	2.6	941
2	Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. Annals of Applied Statistics, 2011, 5, 232-253.	1.1	498
3	The sparsity and bias of the Lasso selection in high-dimensional linear regression. Annals of Statistics, 2008, 36, .	2.6	488
4	Homozygosity mapping with SNP arrays identifies <i>TRIM32</i> , an E3 ubiquitin ligase, as a Bardet-Biedl syndrome gene ( <i>BBS11</i> ). Proceedings of the National Academy of Sciences of the United States of America, 2006, 103, 6287-6292.	7.1	378
5	Variable selection in nonparametric additive models. Annals of Statistics, 2010, 38, 2282-2313.	2.6	357
6	Efficient estimation for the proportional hazards model with interval censoring. Annals of Statistics, 1996, 24, 540.	2.6	343
7	Asymptotic properties of bridge estimators in sparse high-dimensional regression models. Annals of Statistics, 2008, 36, .	2.6	337
8	MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. Nucleic Acids Research, 2012, 40, e61-e61.	14.5	330
9	A group bridge approach for variable selection. Biometrika, 2009, 96, 339-355.	2.4	257
10	A Selective Review of Group Selection in High-Dimensional Models. Statistical Science, 2012, 27, .	2.8	254
11	Penalized feature selection and classification in bioinformatics. Briefings in Bioinformatics, 2008, 9, 392-403.	6.5	203
12	Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors. Statistics and Computing, 2015, 25, 173-187.	1.5	191
13	Regulation of gene expression in the mammalian eye and its relevance to eye disease. Proceedings of the National Academy of Sciences of the United States of America, 2006, 103, 14429-14434.	7.1	190
14	Evidence supporting WNT2 as an autism susceptibility gene. American Journal of Medical Genetics Part A, 2001, 105, 406-413.	2.4	188
15	Maintenance of hematopoietic stem cells through regulation of Wnt and mTOR pathways. Nature Medicine, 2012, 18, 1778-1785.	30.7	188
16	Supervised group Lasso with applications to microarray data analysis. BMC Bioinformatics, 2007, 8, 60.	2.6	183
17	Penalized methods for bi-level variable selection. Statistics and Its Interface, 2009, 2, 369-380.	0.3	180
18	Immunohistochemical study of mucin carbohydrates and core proteins in human pancreatic tumors. Cancer, 1993, 71, 2191-2199.	4.1	170

#	ARTICLE	IF	CITATIONS
19	Regularized ROC method for disease classification and biomarker selection with microarray data. <i>Bioinformatics</i> , 2005, 21, 4356-4362.	4.1	152
20	SCAD-penalized regression in high-dimensional partially linear models. <i>Annals of Statistics</i> , 2009, 37, .	2.6	145
21	Sieve Estimation for the Proportional-Odds Failure-Time Regression Model with Interval Censoring. <i>Journal of the American Statistical Association</i> , 1997, 92, 960-967.	3.1	134
22	Regularized Estimation in the Accelerated Failure Time Model with High-Dimensional Covariates. <i>Biometrics</i> , 2006, 62, 813-820.	1.4	134
23	Interval Censored Survival Data: A Review of Recent Progress. <i>Lecture Notes in Statistics</i> , 1997, , 123-169.	0.2	128
24	A Concave Pairwise Fusion Approach to Subgroup Analysis. <i>Journal of the American Statistical Association</i> , 2017, 112, 410-423.	3.1	126
25	Efficient estimation of the partly linear additive Cox model. <i>Annals of Statistics</i> , 1999, 27, 1536.	2.6	122
26	Combining multidimensional genomic measurements for predicting cancer prognosis: observations from TCGA. <i>Briefings in Bioinformatics</i> , 2015, 16, 291-303.	6.5	122
27	A Spline-Based Semiparametric Maximum Likelihood Estimation Method for the Cox Model with Interval-Censored Data. <i>Scandinavian Journal of Statistics</i> , 2010, 37, 338-354.	1.4	109
28	Pivotal role for glycogen synthase kinase $\beta$ 3 in hematopoietic stem cell homeostasis in mice. <i>Journal of Clinical Investigation</i> , 2009, 119, 3519-29.	8.2	109
29	Variable selection and estimation in high-dimensional varying-coefficient models. <i>Statistica Sinica</i> , 2011, 21, 1515-1540.	0.3	98
30	Sieve Estimation for the Proportional-Odds Failure-Time Regression Model With Interval Censoring. <i>Journal of the American Statistical Association</i> , 1997, 92, 960.	3.1	93
31	Semismooth Newton Coordinate Descent Algorithm for Elastic-Net Penalized Huber Loss Regression and Quantile Regression. <i>Journal of Computational and Graphical Statistics</i> , 2017, 26, 547-557.	1.7	86
32	Consistent group selection in high-dimensional linear regression. <i>Bernoulli</i> , 2010, 16, 1369-1384.	1.3	84
33	Variable selection in the accelerated failure time model via the bridge method. <i>Lifetime Data Analysis</i> , 2010, 16, 176-195.	0.9	84
34	Oracle inequalities for the lasso in the Cox model. <i>Annals of Statistics</i> , 2013, 41, 1142-1165.	2.6	75
35	The sparse Laplacian shrinkage estimator for high-dimensional regression. <i>Annals of Statistics</i> , 2011, 39, 2021-2046.	2.6	72
36	Power to Detect Linkage Based on Multiple Sets of Data in the Presence of Locus Heterogeneity: Comparative Evaluation of Model-Based Linkage Methods for Affected Sib Pair Data. <i>Human Heredity</i> , 2001, 51, 199-208.	0.8	65

#	ARTICLE	IF	CITATIONS
37	Semiparametric Estimation Methods for Panel Count Data Using Monotone $B$ -Splines. Journal of the American Statistical Association, 2009, 104, 1060-1070.	3.1	64
38	Combining Multiple Markers for Classification Using ROC. Biometrics, 2007, 63, 751-757.	1.4	63
39	Gene Expression Analysis of Photoreceptor Cell Loss in Bbs4-Knockout Mice Reveals an Early Stress Gene Response and Photoreceptor Cell Damage. , 2007, 48, 3329.		57
40	Identification of gene-environment interactions in cancer studies using penalization. Genomics, 2013, 102, 189-194.	2.9	47
41	Comparison of $\tilde{\text{Model-Free}}$ and $\tilde{\text{Model-Based}}$ Linkage Statistics in the Presence of Locus Heterogeneity: Single Data Set and Multiple Data Set Applications. Human Heredity, 2001, 51, 217-225.	0.8	45
42	Integrative analysis and variable selection with multiple high-dimensional data sets. Biostatistics, 2011, 12, 763-775.	1.5	39
43	Model Selection for Cox Models with Time-Varying Coefficients. Biometrics, 2012, 68, 419-428.	1.4	38
44	A semiparametric approach for the nonparametric transformation survival model with multiple covariates. Biostatistics, 2007, 8, 197-211.	1.5	37
45	Semiparametric Regression Pursuit. Statistica Sinica, 2012, 22, 1403-1426.	0.3	37
46	Statistical Evaluation of Age-at-Onset Anticipation: A New Test and Evaluation of Its Behavior in Realistic Applications. American Journal of Human Genetics, 1998, 62, 1212-1227.	6.2	35
47	Additive risk survival model with microarray data. BMC Bioinformatics, 2007, 8, 192.	2.6	34
48	Integrative analysis of $\tilde{\text{omics}}$ data using penalty functions. Wiley Interdisciplinary Reviews: Computational Statistics, 2015, 7, 99-108.	3.9	34
49	Statin Use and the Risk of Kidney Disease With Long-Term Follow-Up (8.4-Year Study). American Journal of Cardiology, 2016, 117, 647-655.	1.6	34
50	Two-Locus Heterogeneity Cannot Be Distinguished from Two-Locus Epistasis on the Basis of Affected-Sib-Pair Data. American Journal of Human Genetics, 2003, 73, 223-232.	6.2	32
51	A new statistical test for age-of-onset anticipation: Application to bipolar disorder. Genetic Epidemiology, 1997, 14, 1091-1096.	1.3	31
52	Integrative analysis of multiple cancer prognosis studies with gene expression measurements. Statistics in Medicine, 2011, 30, 3361-3371.	1.6	31
53	Integrative Analysis of Cancer Diagnosis Studies with Composite Penalization. Scandinavian Journal of Statistics, 2014, 41, 87-103.	1.4	31
54	Oracle inequalities for sparse additive quantile regression in reproducing kernel Hilbert space. Annals of Statistics, 2018, 46, .	2.6	31

#	ARTICLE	IF	CITATIONS
55	Clustering threshold gradient descent regularization: with applications to microarray studies. <i>Bioinformatics</i> , 2007, 23, 466-472.	4.1	30
56	CoMM-S2: a collaborative mixed model using summary statistics in transcriptome-wide association studies. <i>Bioinformatics</i> , 2020, 36, 2009-2016.	4.1	30
57	A Two-Way Semilinear Model for Normalization and Analysis of cDNA Microarray Data. <i>Journal of the American Statistical Association</i> , 2005, 100, 814-829.	3.1	29
58	A bayesian approach to replication of linkage findings. <i>Genetic Epidemiology</i> , 1999, 17, S749-54.	1.3	28
59	Deciphering the associations between gene expression and copy number alteration using a sparse double Laplacian shrinkage approach. <i>Bioinformatics</i> , 2015, 31, 3977-3983.	4.1	27
60	Linkage Detection Adaptive to Linkage Disequilibrium: The Disequilibrium Maximum-Likelihood Binomial Test for Affected-Sibship Data. <i>American Journal of Human Genetics</i> , 1999, 65, 1741-1759.	6.2	26
61	Exploration of Heterogeneous Treatment Effects via Concave Fusion. <i>International Journal of Biostatistics</i> , 2020, 16, .	0.7	26
62	The Posterior Probability of Linkage Allowing for Linkage Disequilibrium and a New Estimate of Disequilibrium between a Trait and a Marker. <i>Human Heredity</i> , 2005, 59, 210-219.	0.8	25
63	Incorporating group correlations in genome-wide association studies using smoothed group Lasso. <i>Biostatistics</i> , 2013, 14, 205-219.	1.5	24
64	A Penalized Robust Method for Identifying Gene-Environment Interactions. <i>Genetic Epidemiology</i> , 2014, 38, 220-230.	1.3	24
65	Estimation and Selection via Absolute Penalized Convex Minimization And Its Multistage Adaptive Applications. <i>Journal of Machine Learning Research</i> , 2012, 13, 1839-1864.	62.4	24
66	Monoclonal antibody directed against colon cancer mucin has high specificity for malignancy. <i>International Journal of Cancer</i> , 1993, 54, 467-474.	5.1	22
67	Identifying gene-gene interactions using penalized tensor regression. <i>Statistics in Medicine</i> , 2018, 37, 598-610.	1.6	22
68	Accounting for linkage disequilibrium in genome-wide association studies: a penalized regression method. <i>Statistics and Its Interface</i> , 2013, 6, 99-115.	0.3	22
69	Identification of cancer genomic markers via integrative sparse boosting. <i>Biostatistics</i> , 2012, 13, 509-522.	1.5	21
70	Promoting Similarity of Sparsity Structures in Integrative Analysis With Penalization. <i>Journal of the American Statistical Association</i> , 2017, 112, 342-350.	3.1	20
71	A robust two-way semi-linear model for normalization of cDNA microarray data. <i>BMC Bioinformatics</i> , 2005, 6, 14.	2.6	18
72	Regularized binormal ROC method in disease classification using microarray data. <i>BMC Bioinformatics</i> , 2006, 7, 253.	2.6	18

#	ARTICLE	IF	CITATIONS
73	Identification of genes associated with multiple cancers via integrative analysis. BMC Genomics, 2009, 10, 535.	2.8	18
74	Identification of non-Hodgkin's lymphoma prognosis signatures using the CTGDR method. Bioinformatics, 2010, 26, 15-21.	4.1	18
75	Incorporating higher-order representative features improves prediction in network-based cancer prognosis analysis. BMC Medical Genomics, 2011, 4, 5.	1.5	18
76	Incorporating Network Structure in Integrative Analysis of Cancer Prognosis Data. Genetic Epidemiology, 2013, 37, 173-183.	1.3	18
77	Integrative Analysis of High-throughput Cancer Studies With Contrasted Penalization. Genetic Epidemiology, 2014, 38, 144-151.	1.3	17
78	A Unified Primal Dual Active Set Algorithm for Nonconvex Sparse Recovery. Statistical Science, 2021, 36, .	2.8	17
79	The Mnet method for variable selection. Statistica Sinica, 2016, , .	0.3	17
80	Genetic Linkage Analysis of a Dichotomous Trait Incorporating a Tightly Linked Quantitative Trait in Affected Sib Pairs. American Journal of Human Genetics, 2003, 72, 949-960.	6.2	16
81	Gene network-based cancer prognosis analysis with sparse boosting. Genetical Research, 2012, 94, 205-221.	0.9	15
82	The cross-validated AUC for MCP-logistic regression with high-dimensional data. Statistical Methods in Medical Research, 2013, 22, 505-518.	1.5	15
83	Integrative analysis of prognosis data on multiple cancer subtypes. Biometrics, 2014, 70, 480-488.	1.4	15
84	A tissue-specific collaborative mixed model for jointly analyzing multiple tissues in transcriptome-wide association studies. Nucleic Acids Research, 2020, 48, e109-e109.	14.5	15
85	Genetic anticipation and breast cancer: a prospective follow-up study. Breast Cancer Research and Treatment, 1999, 55, 21-28.	2.5	14
86	The Null Distribution of the Heterogeneity Lod Score Does Depend on the Assumed Genetic Model for the Trait. Human Heredity, 2001, 52, 217-222.	0.8	14
87	Integrative analysis of multiple cancer genomic datasets under the heterogeneity model. Statistics in Medicine, 2013, 32, 3509-3521.	1.6	13
88	Reply to Cordell and Farrall. American Journal of Human Genetics, 2003, 73, 1471-1473.	6.2	12
89	Sparse group penalized integrative analysis of multiple cancer prognosis datasets. Genetical Research, 2013, 95, 68-77.	0.9	12
90	Similarity of markers identified from cancer gene expression studies: observations from GEO. Briefings in Bioinformatics, 2014, 15, 671-684.	6.5	12

#	ARTICLE	IF	CITATIONS
91	Robust Decoding from 1-Bit Compressive Sampling with Ordinary and Regularized Least Squares. SIAM Journal of Scientific Computing, 2018, 40, A2062-A2086.	2.8	12
92	High-dimensional integrative analysis with homogeneity and sparsity recovery. Journal of Multivariate Analysis, 2019, 174, 104529.	1.0	12
93	Combined Multipoint Analysis of Multiple Asthma Data Sets Based on the Posterior Probability of Linkage. Genetic Epidemiology, 2001, 21, S73-8.	1.3	11
94	Concave 1-norm group selection. Biostatistics, 2015, 16, 252-267.	1.5	11
95	Analyzing Association Mapping in Pedigree-Based GWAS Using a Penalized Multitrait Mixed Model. Genetic Epidemiology, 2016, 40, 382-393.	1.3	11
96	Identification of Predictive Pathways for Non-Hodgkin Lymphoma Prognosis. Cancer Informatics, 2010, 9, CIN.S6315.	1.9	10
97	Asymptotic properties of Lasso in high-dimensional partially linear models. Science China Mathematics, 2016, 59, 769-788.	1.7	10
98	A Semismooth Newton Algorithm for High-Dimensional Nonconvex Sparse Learning. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31, 2993-3006.	11.3	10
99	Subgroup analysis in the heterogeneous Cox model. Statistics in Medicine, 2021, 40, 739-757.	1.6	10
100	Majorization minimization by coordinate descent for concave penalized generalized linear models. Statistics and Computing, 2014, 24, 871-883.	1.5	9
101	Sieve Estimation of Cox Models with Latent Structures. Biometrics, 2016, 72, 1086-1097.	1.4	9
102	Accommodating missingness in environmental measurements in gene-environment interaction analysis. Genetic Epidemiology, 2017, 41, 523-554.	1.3	9
103	Integrative prescreening in analysis of multiple cancer genomic studies. BMC Bioinformatics, 2012, 13, 168.	2.6	8
104	A Forward and Backward Stagewise algorithm for nonconvex loss functions with adaptive Lasso. Computational Statistics and Data Analysis, 2018, 124, 235-251.	1.2	8
105	Histopathological imaging-based cancer heterogeneity analysis via penalized fusion with model averaging. Biometrics, 2021, 77, 1397-1408.	1.4	8
106	Identification of cancer-associated gene clusters and genes via clustering penalization. Statistics and Its Interface, 2009, 2, 1-11.	0.3	8
107	Inference for Low-Dimensional Covariates in a High-Dimensional Accelerated Failure Time Model. Statistica Sinica, 2019, 29, 877-894.	0.3	8
108	Semiparametric prognosis models in genomic studies. Briefings in Bioinformatics, 2010, 11, 385-393.	6.5	7

#	ARTICLE	IF	CITATIONS
109	A group adaptive elastic-net approach for variable selection in high-dimensional linear regression. Science China Mathematics, 2018, 61, 173-188.	1.7	7
110	LPG: A four-group probabilistic approach to leveraging pleiotropy in genome-wide association studies. BMC Genomics, 2018, 19, 503.	2.8	7
111	Genome-wide linkage analysis of blood pressure under locus heterogeneity. BMC Genetics, 2003, 4, S78.	2.7	6
112	Regularized regression method for genome-wide association studies. BMC Proceedings, 2011, 5, S67.	1.6	6
113	Integrative Analysis of Cancer Prognosis Data With Multiple Subtypes Using Regularized Gradient Descent. Genetic Epidemiology, 2012, 36, 829-838.	1.3	6
114	Combining clinical and genomic covariates via Cov-TGDR. Cancer Informatics, 2007, 3, 371-8.	1.9	6
115	Robust Semiparametric Microarray Normalization and Significance Analysis. Biometrics, 2006, 62, 555-561.	1.4	5
116	GSDAR: a fast Newton algorithm for $\ell_0$ regularized generalized linear models with statistical guarantee. Computational Statistics, 2022, 37, 507-533.	1.5	5
117	Analysis of Genome-Wide Association Studies with Multiple Outcomes Using Penalization. PLoS ONE, 2012, 7, e51198.	2.5	5
118	Combining Clinical and Genomic Covariates via Cov-TGDR. Cancer Informatics, 2007, 3, 117693510700300.	1.9	4
119	Identification of breast cancer prognosis markers via integrative analysis. Computational Statistics and Data Analysis, 2012, 56, 2718-2728.	1.2	4
120	Concave group methods for variable selection and estimation in high-dimensional varying coefficient models. Science China Mathematics, 2014, 57, 2073-2090.	1.7	4
121	Bayesian inference for high-dimensional linear regression under mnet priors. Canadian Journal of Statistics, 2016, 44, 180-197.	0.9	4
122	Estimating high-dimensional additive Cox model with time-dependent covariate processes. Scandinavian Journal of Statistics, 2018, 45, 900-922.	1.4	4
123	Fitting sparse linear models under the sufficient and necessary condition for model identification. Statistics and Probability Letters, 2021, 168, 108925.	0.7	4
124	$\hat{\alpha}_0$ -Regularized high-dimensional accelerated failure time model. Computational Statistics and Data Analysis, 2022, 170, 107430.	1.2	4
125	A Deep Generative Approach to Conditional Sampling. Journal of the American Statistical Association, 2023, 118, 1-12.	3.1	4
126	Linearized maximum rank correlation estimation. Biometrika, 2023, 110, 187-203.	2.4	4

#	ARTICLE	IF	CITATIONS
127	The Score Statistic of the LD-Lod Analysis: Detecting Linkage Adaptive to Linkage Disequilibrium. <i>Human Heredity</i> , 2001, 52, 83-98.	0.8	3
128	Penalized multivariate linear mixed model for longitudinal genome-wide association studies. <i>BMC Proceedings</i> , 2014, 8, S73.	1.6	3
129	On the sign consistency of the Lasso for the high-dimensional Cox model. <i>Journal of Multivariate Analysis</i> , 2018, 167, 79-96.	1.0	3
130	Robust post-selection inference of high-dimensional mean regression with heavy-tailed asymmetric or heteroskedastic errors. <i>Journal of Econometrics</i> , 2022, 230, 416-431.	6.5	3
131	Hierarchical cancer heterogeneity analysis based on histopathological imaging features. <i>Biometrics</i> , 2022, 78, 1579-1591.	1.4	3
132	Model pursuit and variable selection in the additive accelerated failure time model. <i>Statistical Papers</i> , 2020, , 1.	1.2	2
133	Reply to Horvath et al.. <i>American Journal of Human Genetics</i> , 2000, 67, 534-535.	6.2	1
134	Genome Scan for Linkage to Asthma Using a Linkage Disequilibrium $\times$ Lod Score Test. <i>Genetic Epidemiology</i> , 2001, 21, S303-7.	1.3	1
135	A lower bound based smoothed quasi-Newton algorithm for group bridge penalized regression. <i>Communications in Statistics Part B: Simulation and Computation</i> , 2017, 46, 4694-4707.	1.2	1
136	Generalized Newton $\hat{c}$ Raphson algorithm for high dimensional LASSO regression. <i>Statistics and Its Interface</i> , 2021, 14, 339-350.	0.3	1
137	Projection $\hat{c}$ based and cross $\hat{c}$ validated estimation in high $\hat{c}$ dimensional Cox model. <i>Scandinavian Journal of Statistics</i> , 0, , .	1.4	1
138	A new statistical test for age $\hat{c}$ of $\hat{c}$ onset anticipation: Application to bipolar disorder. <i>Genetic Epidemiology</i> , 1997, 14, 1091-1096.	1.3	1
139	Penalized Integrative Analysis of High-Dimensional Omics Data. , 0, , 174-204.		0
140	A rank-based approach to estimating monotone individualized two treatment regimes. <i>Computational Statistics and Data Analysis</i> , 2020, 151, 107015.	1.2	0
141	A SEMIPARAMETRIC METHOD FOR MAPPING QUANTITATIVE TRAIT LOCI. , 2003, , 262-271.		0
142	IL-21 Plus CpG ODN Induces Granzyme B-Dependent Induction of Apoptosis in CD5-Positive B Cells Including B-CLL Cells.. <i>Blood</i> , 2006, 108, 2823-2823.	1.4	0
143	PSNA: A pathwise semismooth Newton algorithm for sparse recovery with optimal local convergence and oracle properties. <i>Signal Processing</i> , 2022, 194, 108432.	3.7	0
144	Bayesian hierarchical finite mixture of regression for histopathological imaging $\hat{c}$ based cancer data analysis. <i>Statistics in Medicine</i> , 2022, , .	1.6	0

#	ARTICLE	IF	CITATIONS
145	Robust Signal Recovery for High-Dimensional Linear Log-Contrast Models with Compositional Covariates. <i>Journal of Business and Economic Statistics</i> , 2023, 41, 957-967.	2.9	0