Michael I Jordan

List of Publications by Year
in descending order

182 papers	**32,724** citations	26630 **56** h-index	14208 **128** g-index
196 all docs	**196** docs citations	**196** times ranked	**27110** citing authors

#	Article	IF	Citations
1	Evaluating Sensitivity to the Stick-Breaking Prior in Bayesian Nonparametrics (with Discussion). Bayesian Analysis, 2023, 18, .	3.0	3
2	Understanding the acceleration phenomenon via high-resolution differential equations. Mathematical Programming, 2022, 195, 79-148.	2.4	34
3	A control-theoretic perspective on optimal high-order optimization. Mathematical Programming, 2022, 195, 929-975.	2.4	9
4	A Python library for probabilistic analysis of single-cell omics data. Nature Biotechnology, 2022, 40, 163-166.	17.5	216
5	DestVI identifies continuums of cell types in spatial transcriptomics data. Nature Biotechnology, 2022, 40, 1360-1369.	17.5	75
6	SOUL: An Energy-Efficient Unsupervised Online Learning Seizure Detection Classifier. IEEE Journal of Solid-State Circuits, 2022, 57, 2532-2544.	5.4	9
7	Adaptivity of Stochastic Gradient Methods for Nonconvex Optimization. SIAM Journal on Mathematics of Data Science, 2022, 4, 634-648.	1.8	4
8	Geometric methods for sampling, optimization, inference, and adaptive agents. Handbook of Statistics, 2022, , 21-78.	0.6	4
9	Generalized Momentum-Based Methods: A Hamiltonian Perspective. SIAM Journal on Optimization, 2021, 31, 915-944.	2.0	11
10	On dissipative symplectic integration with applications to gradient-based optimization. Journal of Statistical Mechanics: Theory and Experiment, 2021, 2021, 043402.	2.3	11
11	Is there an analog of Nesterov acceleration for gradient-based MCMC?. Bernoulli, 2021, 27, .	1.3	13
12	A 1.5nJ/cls Unsupervised Online Learning Classifier for Seizure Detection. , 2021, , .		6
13	Probabilistic harmonization and annotation of single‐cell transcriptomics data with deep generative models. Molecular Systems Biology, 2021, 17, e9620.	7.2	211
14	Is Temporal Difference Learning Optimal? An Instance-Dependent Analysis. SIAM Journal on Mathematics of Data Science, 2021, 3, 1013-1040.	1.8	7
15	A Bayesian nonparametric approach to super-resolution single-molecule localization. Annals of Applied Statistics, 2021, 15, .	1.1	2
16	HopSkipJumpAttack: A Query-Efficient Decision-Based Attack. , 2020, , .		229
17	Unsupervised Online Learning for Long-Term High Sensitivity Seizure Detection. , 2020, 2020, 528-531.		1
18	On the Adaptivity of Stochastic Gradient-Based Optimization. SIAM Journal on Optimization, 2020, 30, 1473-1500.	2.0	6

#	Article	IF	Citations
19	Function-Specific Mixing Times and Concentration Away from Equilibrium. Bayesian Analysis, 2020, 15, .	3.0	2
20	Universal Domain Adaptation. , 2020, , 195-211.		4
21	A unified treatment of multiple testing with prior knowledge using the p-filter. Annals of Statistics, 2019, 47, .	2.6	30
22	A sequential algorithm for false discovery rate control on directed acyclic graphs. Biometrika, 2019, 106, 69-86.	2.4	14
23	Decoding from Pooled Data: Sharp Information-Theoretic Bounds. SIAM Journal on Mathematics of Data Science, 2019, 1, 161-188.	1.8	4
24	First-order methods almost always avoid strict saddle points. Mathematical Programming, 2019, 176, 311-337.	2.4	61
25	Universal Domain Adaptation. , 2019, , .		216
26	Sampling can be faster than optimization. Proceedings of the National Academy of Sciences of the United States of America, 2019, 116, 20881-20885.	7.1	50
27	Transferable Representation Learning with Deep Adaptation Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41, 3071-3085.	13.9	345
28	Decoding From Pooled Data: Phase Transitions of Message Passing. IEEE Transactions on Information Theory, 2019, 65, 572-585.	2.4	8
29	Communication-Efficient Distributed Statistical Inference. Journal of the American Statistical Association, 2019, 114, 668-681.	3.1	189
30	Latent Marked Poisson Process with Applications to Object Segmentation. Bayesian Analysis, 2018, 13, .	3.0	2
31	Minimax Optimal Procedures for Locally Private Estimation. Journal of the American Statistical Association, 2018, 113, 182-201.	3.1	167
32	Partial Transfer Learning with Selective Adversarial Networks. , 2018, , .		249
33	Deep generative modeling for single-cell transcriptomics. Nature Methods, 2018, 15, 1053-1058.	19.0	1,227
34	Posteriors, conjugacy, and exponential families for completely random measures. Bernoulli, 2018, 24, .	1.3	14
35	On kernel methods for covariates that are rankings. Electronic Journal of Statistics, 2018, 12, .	0.7	8
36	Saturating Splines and Feature Selection. Journal of Machine Learning Research, 2018, 18, .	62.4	0

#	Article	IF	Citations
37	Mining Massive Amounts of Genomic Data: A Semiparametric Topic Modeling Approach. Journal of the American Statistical Association, 2017, 112, 921-932.	3.1	3
38	Distributed optimization with arbitrary local solvers. Optimization Methods and Software, 2017, 32, 813-848.	2.4	111
39	Perturbed Iterate Analysis for Asynchronous Stochastic Optimization. SIAM Journal on Optimization, 2017, 27, 2202-2229.	2.0	50
40	Real-Time Machine Learning. , 2017, , .		31
41	A Marked Poisson Process Driven Latent Shape Model for 3D Segmentation of Reflectance Confocal Microscopy Image Stacks of Human Skin. IEEE Transactions on Image Processing, 2017, 26, 172-184.	9.8	11
42	Optimal prediction for sparse linear models? Lower bounds for coordinate-separable M-estimators. Electronic Journal of Statistics, 2017, 11, .	0.7	29
43	Decoding from pooled data: Phase transitions of message passing. , 2017, , .		1
44	QuTE: Decentralized multiple testing on sensor networks with false discovery rate control. , 2017, , .		5
45	On the computational complexity of high-dimensional Bayesian variable selection. Annals of Statistics, 2016, 44, .	2.6	68
46	A variational perspective on accelerated methods in optimization. Proceedings of the National Academy of Sciences of the United States of America, 2016, 113, E7351-E7358.	7.1	193
47	Optimism-driven exploration for nonlinear systems. , 2015, , .		15
48	Automating model search for large scale machine learning. , 2015, , .		91
49	Machine Learning and Databases. , 2015, , .		9
50	Combinatorial Clustering and the Beta Negative Binomial Process. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37, 290-306.	13.9	23
51	Nested Hierarchical Dirichlet Processes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37, 256-270.	13.9	111
52	Machine learning: Trends, perspectives, and prospects. Science, 2015, 349, 255-260.	12.6	4,904
53	Optimal Rates for Zero-Order Convex Optimization: The Power of Two Function Evaluations. IEEE Transactions on Information Theory, 2015, 61, 2788-2806.	2.4	168
54	Computational Thinking, Inferential Thinking and "Big Data". , 2015, , .		2

#	ARTICLE	IF	CITATIONS
55	Matrix concentration inequalities via the method of exchangeable pairs. Annals of Probability, 2014, 42, .	1.8	71
56	Privacy Aware Learning. Journal of the ACM, 2014, 61, 1-57.	2.2	96
57	Scaling up crowd-sourcing to very large datasets. Proceedings of the VLDB Endowment, 2014, 8, 125-136.	3.8	109
58	Knowing when you're wrong. , 2014, , .		106
59	A Scalable Bootstrap for Massive Data. Journal of the Royal Statistical Society Series B: Statistical Methodology, 2014, 76, 795-816.	2.2	237
60	SM a SH: a benchmarking toolkit for human genome variant calling. Bioinformatics, 2014, 30, 2787-2795.	4.1	40
61	Iterative Discovery of Multiple AlternativeClustering Views. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36, 1340-1353.	13.9	47
62	Matrix-Variate Dirichlet Process Priors with Applications. Bayesian Analysis, 2014, 9, .	3.0	4
63	Joint modeling of multiple time series via the beta process with application to motion capture segmentation. Annals of Applied Statistics, 2014, 8, .	1.1	62
64	Nonparametric link prediction in large scale dynamic networks. Electronic Journal of Statistics, 2014, 8, .	0.7	18
65	Changepoint Analysis for Efficient Variant Calling. Lecture Notes in Computer Science, 2014, , 20-34.	1.3	0
66	A general bootstrap performance diagnostic. , 2013, , .		12
67	Local privacy and statistical minimax rates. , 2013, , .		11
68	Learning Dependency-Based Compositional Semantics. Computational Linguistics, 2013, 39, 389-446.	3.3	268
69	Cluster Forests. Computational Statistics and Data Analysis, 2013, 66, 178-192.	1.2	37
70	Bayesian semiparametric Wiener system identification. Automatica, 2013, 49, 2053-2063.	5.0	55
71	MLI: An API for Distributed Machine Learning. , 2013, , .		105
72	Computational and statistical tradeoffs via convex relaxation. Proceedings of the National Academy of Sciences of the United States of America, 2013, 110, E1181-90.	7.1	143

#	Article	IF	Citations
73	Molecular function prediction for a family exhibiting evolutionary tendencies toward substrate specificity swapping: Recurrence of tyrosine aminotransferase activity in the Iα subfamily. Proteins: Structure, Function and Bioinformatics, 2013, 81, 1593-1609.	2.6	6
74	Local Privacy and Statistical Minimax Rates. , 2013, , .		310
75	Distributed Low-Rank Subspace Segmentation. , 2013, , .		30
76	Evolutionary inference via the Poisson Indel Process. Proceedings of the National Academy of Sciences of the United States of America, 2013, 110, 1160-1166.	7.1	34
77	Cluster and Feature Modeling from Combinatorial Stochastic Processes. Statistical Science, 2013, 28, .	2.8	25
78	On statistics, computation and scalability. Bernoulli, 2013, 19, .	1.3	65
79	The asymptotics of ranking algorithms. Annals of Statistics, 2013, 41, .	2.6	9
80	Feature Allocations, Probability Functions, and Paintboxes. Bayesian Analysis, 2013, 8, .	3.0	30
81	Phylogenetic Inference via Sequential Monte Carlo. Systematic Biology, 2012, 61, 579-593.	5.6	64
82	Beta Processes, Stick-Breaking and Power Laws. Bayesian Analysis, 2012, 7, .	3.0	40
83	Active spectral clustering via iterative uncertainty reduction. , 2012, , .		29
84	Qualcomm Context-Awareness Symposium Sets Research Agenda for Context-Aware Smartphones. IEEE Pervasive Computing, 2012, 11, 76-79.	1.3	6
85	A semiparametric Bayesian approach to Wiener system identification*. IFAC Postprint Volumes IPPV / International Federation of Automatic Control, 2012, 45, 1137-1142.	0.4	8
86	Ergodic Mirror Descent. SIAM Journal on Optimization, 2012, 22, 1549-1578.	2.0	37
87	Visually Relating Gene Expression and in vivo DNA Binding Data. , 2011, , .		0
88	Nonparametric Combinatorial Sequence Models. Journal of Computational Biology, 2011, 18, 1649-1660.	1.6	1
89	A sticky HDP-HMM with application to speaker diarization. Annals of Applied Statistics, 2011, 5, .	1.1	225
90	Learning Low-Dimensional Signal Models. IEEE Signal Processing Magazine, 2011, 28, 39-51.	5.6	22

#	Article	IF	Citations
91	Bayesian Nonparametric Inference of Switching Dynamic Linear Models. IEEE Transactions on Signal Processing, 2011, 59, 1569-1585.	5.3	162
92	Support union recovery in high-dimensional multivariate regression. Annals of Statistics, 2011, 39, .	2.6	217
93	Bayesian inference for queueing networks and modeling of internet services. Annals of Applied Statistics, 2011, 5, .	1.1	32
94	Genome-scale phylogenetic function annotation of large and diverse protein families. Genome Research, 2011, 21, 1969-1980.	5.5	54
95	Managing data transfers in computer clusters with orchestra. , 2011, , .		301
96	Ergodic mirror descent. , 2011, , .		4
97	Supervised hierarchical Pitman-Yor process for natural scene segmentation. , 2011, , .		10
98	Nonparametric Combinatorial Sequence Models. Lecture Notes in Computer Science, 2011, , 516-530.	1.3	1
99	Hierarchical Bayesian nonparametric models with applications. , 2010, , 158-207.		149
100	Leo Breiman. Annals of Applied Statistics, 2010, 4, .	1.1	1
101	Joint covariate selection and joint subspace selection for multiple classification problems. Statistics and Computing, 2010, 20, 231-252.	1.5	325
102	Estimating Divergence Functionals and the Likelihood Ratio by Convex Risk Minimization. IEEE Transactions on Information Theory, 2010, 56, 5847-5861.	2.4	264
103	Feature space resampling for protein conformational search. Proteins: Structure, Function and Bioinformatics, 2010, 78, 1583-1593.	2.6	15
104	The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies. Journal of the ACM, 2010, 57, 1-30.	2.2	390
105	Sufficient dimension reduction for visual sequence classification. , 2010, , .		15
106	Active site prediction using evolutionary and structural information. Bioinformatics, 2010, 26, 617-624.	4.1	67
107	Neighbor-Dependent Ramachandran Probability Distributions of Amino Acids Developed from a Hierarchical Dirichlet Process Model. PLoS Computational Biology, 2010, 6, e1000763.	3.2	145
108	Characterizing, modeling, and generating workload spikes for stateful services. , 2010, , .		145

#	ARTICLE	IF	CITATIONS
109	Bayesian Nonparametric Methods for Learning Markov Switching Processes. IEEE Signal Processing Magazine, 2010, , .	5.6	43
110	Major Advances and Emerging Developments of Graphical Models [From the Guest Editors]. IEEE Signal Processing Magazine, 2010, 27, 17-138.	5.6	12
111	Convex and Semi-Nonnegative Matrix Factorizations. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32, 45-55.	13.9	951
112	Automatic exploration of datacenter performance regimes. , 2009, , .		29
113	Detecting large-scale system problems by mining console logs. , 2009, , .		737
114	Learning from measurements in exponential families. , 2009, , .		36
115	Joint estimation of gene conversion rates and mean conversion tract lengths from population SNP data. Bioinformatics, 2009, 25, i231-i239.	4.1	20
116	Genomic privacy and limits of individual detection in a pool. Nature Genetics, 2009, 41, 965-967.	21.4	153
117	Predicting Multiple Metrics for Queries: Better Decisions Enabled by Machine Learning. Proceedings - International Conference on Data Engineering, 2009, , .	0.0	189
118	Online System Problem Detection by Mining Patterns of Console Logs. , 2009, , .		134
119	Fast approximate spectral clustering. , 2009, , .		295
120	Phylogenetic molecular function annotation. Journal of Physics: Conference Series, 2009, 180, 012024.	0.4	12
121	Kernel dimension reduction in regression. Annals of Statistics, 2009, 37, .	2.6	162
122	On surrogate loss functions and f-divergences. Annals of Statistics, 2009, 37, .	2.6	50
123	Nonparametric Bayesian Identification of Jump Systems with Sparse Dependencies. IFAC Postprint Volumes IPPV / International Federation of Automatic Control, 2009, 42, 1591-1596.	0.4	2
124	A Flexible and Efficient Algorithm for Regularized Fisher Discriminant Analysis. Lecture Notes in Computer Science, 2009, , 632-647.	1.3	12
125	Learning semantic correspondences with less supervision. , 2009, , .		152
126	Association Mapping and Significance Estimation via the Coalescent. American Journal of Human Genetics, 2008, 83, 675-683.	6.2	11

#	ARTICLE	IF	CITATIONS
127	On Optimal Quantization Rules for Some Problems in Sequential Decentralized Detection. IEEE Transactions on Information Theory, 2008, 54, 3285-3295.	2.4	7
128	A critical assessment of Mus musculus gene function prediction using integrated genomic evidence. Genome Biology, 2008, 9, S2.	9.6	214
129	Consistent probabilistic outputs for protein function prediction. Genome Biology, 2008, 9, S6.	9.6	68
130	Nonnegative Matrix Factorization for Combinatorial Optimization: Spectral Clustering, Graph Matching, and Clique Finding. , 2008, , .		73
131	Union support recovery in high-dimensional multivariate regression. , 2008, , .		17
132	On the inference of ancestries in admixed populations. Genome Research, 2008, 18, 668-675.	5.5	44
133	ICMLA 2008 Invited Speakers. , 2008, , .		0
134	A Dual Receptor Crosstalk Model of G-Protein-Coupled Signal Transduction. PLoS Computational Biology, 2008, 4, e1000185.	3.2	38
135	An asymptotic analysis of generative, discriminative, and pseudolikelihood estimators. , 2008, , .		43
136	An HDP-HMM for systems with state persistence. , 2008, , .		137
137	Multiway Spectral Clustering: A Margin-Based Perspective. Statistical Science, 2008, 23, .	2.8	29
138	On the Inference of Ancestries in Admixed Populations. , 2008, , 424-433.		1
139	Regression on manifolds using kernel dimension reduction. , 2007, , .		59
140	A permutation-augmented sampler for DP mixture models. , 2007, , .		3
141	Bayesian Haplotype Inference via the Dirichlet Process. Journal of Computational Biology, 2007, 14, 267-284.	1.6	32
142	A Direct Formulation for Sparse PCA Using Semidefinite Programming. SIAM Review, 2007, 49, 434-448.	9.5	524
143	Nonparametric estimation of the likelihood ratio and divergence functionals. , 2007, , .		15
144	A Randomization Test for Controlling Population Stratification in Whole-Genome Association Studies. American Journal of Human Genetics, 2007, 81, 895-905.	6.2	48

#	Article	IF	Citations
145	Learning Multiscale Representations of Natural Scenes Using Dirichlet Processes. , 2007, , .		44
146	Image Denoising with Nonparametric Hidden Markov Trees. , 2007, , .		10
147	Graphical Models, Exponential Families, and Variational Inference. Foundations and Trends in Machine Learning, 2007, 1, 1-305.	69.0	1,196
148	Solving Consensus and Semi-supervised Clustering Problems Using Nonnegative Matrix Factorization. , 2007, , .		136
149	Hierarchical Dirichlet Processes. Journal of the American Statistical Association, 2006, 101, 1566-1581.	3.1	2,215
150	Variational inference for Dirichlet process mixtures. Bayesian Analysis, 2006, 1, 121.	3.0	939
151	Nonparametric empirical Bayes for the Dirichlet process mixture model. Statistics and Computing, 2006, 16, 5-14.	1.5	78
152	Bayesian multi-population haplotype inference via a hierarchical dirichlet process mixture. , 2006, , .		18
153	Statistical debugging. , 2006, , .		103
154	A graphical model for predicting protein molecular function. , 2006, , .		7
155	Convexity, Classification, and Risk Bounds. Journal of the American Statistical Association, 2006, 101, 138-156.	3.1	529
156	Log-determinant relaxation for approximate inference in discrete Markov random fields. IEEE Transactions on Signal Processing, 2006, 54, 2099-2109.	5.3	40
157	On optimal quantization rules for sequential decision problems. , 2006, , .		8
158	Word alignment via quadratic assignment. , 2006, , .		29
159	A latent variable model for chemogenomic profiling. Bioinformatics, 2005, 21, 3286-3293.	4.1	53
160	Genome-Wide Requirements for Resistance to Functionally Distinct DNA-Damaging Agents. PLoS Genetics, 2005, 1, e24.	3.5	144
161	Protein Molecular Function Prediction by Bayesian Phylogenomics. PLoS Computational Biology, 2005, 1, e45.	3.2	162
162	A kernel-based learning approach to ad hoc sensor network localization. ACM Transactions on Sensor Networks, 2005, 1, 134-152.	3.6	145

#	ARTICLE	IF	CITATIONS
163	Sulfur and Nitrogen Limitation in Escherichia coli K-12: Specific Homeostatic Responses. Journal of Bacteriology, 2005, 187, 1074-1090.	2.2	96
164	Subtree power analysis and species selection for comparative genomics. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102, 7900-7905.	7.1	19
165	Lessons from Escherichia coli genes similarly regulated in response to nitrogen and sulfur limitation. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102, 3453-3458.	7.1	39
166	Nonparametric decentralized detection using kernel methods. IEEE Transactions on Signal Processing, 2005, 53, 4053-4066.	5.3	53
167	Extensions of the Informative Vector Machine. Lecture Notes in Computer Science, 2005, , 56-87.	1.3	19
168	Scalable statistical bug isolation. ACM SIGPLAN Notices, 2005, 40, 15-26.	0.2	234
169	A statistical framework for genomic data fusion. Bioinformatics, 2004, 20, 2626-2635.	4.1	568
170	Graphical Models. Statistical Science, 2004, 19, 140.	2.8	375
171	Multiple-sequence functional annotation and the generalized hidden Markov phylogeny. Bioinformatics, 2004, 20, 1850-1860.	4.1	44
172	Chemogenomic profiling: Identifying the functional interactions of small molecules in yeast. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101, 793-798.	7.1	460
173	Multiple kernel learning, conic duality, and the SMO algorithm. , 2004, , .		872
174	An Introduction to MCMC for Machine Learning. Machine Learning, 2003, 50, 5-43.	5.4	1,641
175	Modeling annotated data. , 2003, , .		597
176	Bug isolation via remote program sampling. ACM SIGPLAN Notices, 2003, 38, 141-154.	0.2	138
177	Bayesian parameter estimation via variational methods. Statistics and Computing, 2000, 10, 25-37.	1.5	383
178	Mixed Memory Markov Models: Decomposing Complex Stochastic Processes as Mixtures of Simpler Ones. Machine Learning, 1999, 37, 75-87.	5.4	102
179	An Introduction to Variational Methods for Graphical Models. Machine Learning, 1999, 37, 183-233.	5.4	1,889
180	Probabilistic Independence Networks for Hidden Markov Probability Models. Neural Computation, 1997, 9, 227-269.	2.2	168

#	ARTICLE	IF	CITATIONS
181	Factorial Hidden Markov Models. Machine Learning, 1997, 29, 245-273.	5.4	674
182	Private Prediction Sets. , 0, , .		1