

Masashi Sugiyama

List of Publications by Citations

Source: <https://exaly.com/author-pdf/793759/masashi-sugiyama-publications-by-citations.pdf>

Version: 2024-04-28

This document has been generated based on the publications and citations recorded by exaly.com. For the latest version of this publication list, visit the link given above.

The third column is the impact factor (IF) of the journal, and the fourth column is the number of citations of the article.

135
papers

2,691
citations

27
h-index

47
g-index

141
ext. papers

3,382
ext. citations

3.6
avg, IF

5.48
L-index

#	Paper	IF	Citations
135	Change-point detection in time-series data by relative density-ratio estimation. <i>Neural Networks</i> , 2013 , 43, 72-83	9.1	217
134	Semi-supervised local Fisher discriminant analysis for dimensionality reduction. <i>Machine Learning</i> , 2010 , 78, 35-61	4	186
133	Direct importance estimation for covariate shift adaptation. <i>Annals of the Institute of Statistical Mathematics</i> , 2008 , 60, 699-746	1	173
132	High-dimensional feature selection by feature-wise kernelized Lasso. <i>Neural Computation</i> , 2014 , 26, 185-207	2.9	136
131	Density Ratio Estimation in Machine Learning 2012 ,		112
130	Statistical outlier detection using direct density ratio estimation. <i>Knowledge and Information Systems</i> , 2011 , 26, 309-336	2.4	98
129	Sequential change-point detection based on direct density-ratio estimation. <i>Statistical Analysis and Data Mining</i> , 2012 , 5, 114-127	1.4	93
128	Application of covariate shift adaptation techniques in brain-computer interfaces. <i>IEEE Transactions on Biomedical Engineering</i> , 2010 , 57, 1318-24	5	91
127	Tensor Networks for Dimensionality Reduction and Large-scale Optimization: Part 2 Applications and Future Perspectives. <i>Foundations and Trends in Machine Learning</i> , 2017 , 9, 249-429	32.9	74
126	Information-theoretic semi-supervised metric learning via entropy regularization. <i>Neural Computation</i> , 2014 , 26, 1717-62	2.9	56
125	The Degrees of Freedom of Partial Least Squares Regression. <i>Journal of the American Statistical Association</i> , 2011 , 106, 697-705	2.8	54
124	Density-ratio matching under the Bregman divergence: a unified framework of density-ratio estimation. <i>Annals of the Institute of Statistical Mathematics</i> , 2012 , 64, 1009-1044	1	48
123	Importance-weighted least-squares probabilistic classifier for covariate shift adaptation with application to human activity recognition. <i>Neurocomputing</i> , 2012 , 80, 93-101	5.4	47
122	Dual-Augmented Lagrangian Method for Efficient Sparse Reconstruction. <i>IEEE Signal Processing Letters</i> , 2009 , 16, 1067-1070	3.2	45
121	Statistical analysis of kernel-based least-squares density-ratio estimation. <i>Machine Learning</i> , 2012 , 86, 335-367	4	43
120	Relative density-ratio estimation for robust distribution comparison. <i>Neural Computation</i> , 2013 , 25, 1324-70	4.7	41
119	Subspace information criterion for model selection. <i>Neural Computation</i> , 2001 , 13, 1863-89	2.9	41

118	Pool-based active learning in approximate linear regression. <i>Machine Learning</i> , 2009 , 75, 249-274	4	40
117	Sufficient dimension reduction via squared-loss mutual information estimation. <i>Neural Computation</i> , 2013 , 25, 725-58	2.9	39
116	Dimensionality reduction for density ratio estimation in high-dimensional spaces. <i>Neural Networks</i> , 2010 , 23, 44-59	9.1	36
115	Class Prior Estimation from Positive and Unlabeled Data. <i>IEICE Transactions on Information and Systems</i> , 2014 , E97.D, 1358-1362	0.6	35
114	Semi-supervised learning of class balance under class-prior change by distribution matching. <i>Neural Networks</i> , 2014 , 50, 110-9	9.1	33
113	A least-squares approach to anomaly detection in static and sequential data. <i>Pattern Recognition Letters</i> , 2014 , 40, 36-40	4.7	30
112	Variational Bayesian sparse additive matrix factorization. <i>Machine Learning</i> , 2013 , 92, 319-347	4	30
111	Tree-Based Ensemble Multi-Task Learning Method for Classification and Regression. <i>IEICE Transactions on Information and Systems</i> , 2014 , E97.D, 1677-1681	0.6	29
110	Density-difference estimation. <i>Neural Computation</i> , 2013 , 25, 2734-75	2.9	28
109	Semi-supervised speaker identification under covariate shift. <i>Signal Processing</i> , 2010 , 90, 2353-2361	4.4	27
108	Class-prior estimation for learning from positive and unlabeled data. <i>Machine Learning</i> , 2017 , 106, 463-492	4.2	26
107	Machine Learning with Squared-Loss Mutual Information. <i>Entropy</i> , 2013 , 15, 80-112	2.8	24
106	Direct density-ratio estimation with dimensionality reduction via least-squares hetero-distributional subspace search. <i>Neural Networks</i> , 2011 , 24, 183-98	9.1	24
105	Direct Divergence Approximation between Probability Distributions and Its Applications in Machine Learning. <i>Journal of Computing Science and Engineering</i> , 2013 , 7, 99-111	1.8	24
104	Single-shot surface profiling by local model fitting. <i>Applied Optics</i> , 2006 , 45, 7999-8005	1.7	23
103	Least-squares independent component analysis. <i>Neural Computation</i> , 2011 , 23, 284-301	2.9	22
102	Health checkup and telemedical intervention program for preventive medicine in developing countries: verification study. <i>Journal of Medical Internet Research</i> , 2015 , 17, e2	7.6	22
101	Information-maximization clustering based on squared-loss mutual information. <i>Neural Computation</i> , 2014 , 26, 84-131	2.9	21

100	\$\mathbb{F}\$-Divergence Estimation and Two-Sample Homogeneity Test Under Semiparametric Density-Ratio Models. <i>IEEE Transactions on Information Theory</i> , 2012 , 58, 708-720	2.8	20
99	Learning Efficient Tensor Representations with Ring-structured Networks 2019 ,		19
98	Theoretical and Experimental Analyses of Tensor-Based Regression and Classification. <i>Neural Computation</i> , 2016 , 28, 686-715	2.9	19
97	Analysis and improvement of policy gradient estimation. <i>Neural Networks</i> , 2012 , 26, 118-29	9.1	19
96	Geodesic Gaussian kernels for value function approximation. <i>Autonomous Robots</i> , 2008 , 25, 287-304	3	18
95	Geometry-aware principal component analysis for symmetric positive definite matrices. <i>Machine Learning</i> , 2017 , 106, 493-522	4	16
94	Averaging covariance matrices for EEG signal classification based on the CSP: An empirical study 2015 ,		16
93	Multi-parametric solution-path algorithm for instance-weighted support vector machines. <i>Machine Learning</i> , 2012 , 88, 297-330	4	16
92	Efficient sample reuse in policy gradients with parameter-based exploration. <i>Neural Computation</i> , 2013 , 25, 1512-47	2.9	16
91	A new algorithm of non-Gaussian component analysis with radial kernel functions. <i>Annals of the Institute of Statistical Mathematics</i> , 2007 , 59, 57-75	1	15
90	A batch ensemble approach to active learning with model selection. <i>Neural Networks</i> , 2008 , 21, 1278-86	9.1	15
89	Trading variance reduction with unbiasedness: the regularized subspace information criterion for robust model selection in kernel regression. <i>Neural Computation</i> , 2004 , 16, 1077-104	2.9	15
88	Canonical dependency analysis based on squared-loss mutual information. <i>Neural Networks</i> , 2012 , 34, 46-55	9.1	14
87	Artist Agent: A Reinforcement Learning Approach to Automatic Stroke Generation in Oriental Ink Painting. <i>IEICE Transactions on Information and Systems</i> , 2013 , E96.D, 1134-1144	0.6	14
86	Least-squares two-sample test. <i>Neural Networks</i> , 2011 , 24, 735-51	9.1	13
85	Reward-weighted regression with sample reuse for direct policy search in reinforcement learning. <i>Neural Computation</i> , 2011 , 23, 2798-832	2.9	13
84	Incremental active learning for optimal generalization. <i>Neural Computation</i> , 2000 , 12, 2909-40	2.9	13
83	Statistical outlier detection for diagnosis of cyber attacks in power state estimation 2016 ,		12

82	Learning under nonstationarity: covariate shift and class-balance change. <i>Wiley Interdisciplinary Reviews: Computational Statistics</i> , 2013 , 5, 465-477	1.4	12
81	A unified method for optimizing linear image restoration filters. <i>Signal Processing</i> , 2002 , 82, 1773-1787	4.4	12
80	Computational complexity of kernel-based density-ratio estimation: a condition number analysis. <i>Machine Learning</i> , 2013 , 90, 431-460	4	11
79	Importance-weighted covariance estimation for robust common spatial pattern. <i>Pattern Recognition Letters</i> , 2015 , 68, 139-145	4.7	10
78	Semi-supervised AUC optimization based on positive-unlabeled learning. <i>Machine Learning</i> , 2018 , 107, 767-794	4	10
77	Target-less camera-LiDAR extrinsic calibration using a bagged dependence estimator 2016 ,		9
76	Modal Regression via Direct Log-Density Derivative Estimation. <i>Lecture Notes in Computer Science</i> , 2016 , 108-116	0.9	9
75	Bias Reduction and Metric Learning for Nearest-Neighbor Estimation of Kullback-Leibler Divergence. <i>Neural Computation</i> , 2018 , 30, 1930-1960	2.9	9
74	Computationally Efficient Estimation of Squared-Loss Mutual Information with Multiplicative Kernel Models. <i>IEICE Transactions on Information and Systems</i> , 2014 , E97.D, 968-971	0.6	9
73	Direct learning of sparse changes in Markov networks by density ratio estimation. <i>Neural Computation</i> , 2014 , 26, 1169-97	2.9	9
72	Using the variogram for vector outlier screening: application to feature-based image registration. <i>International Journal of Computer Assisted Radiology and Surgery</i> , 2018 , 13, 1871-1880	3.9	8
71	Bandit-Based Task Assignment for Heterogeneous Crowdsourcing. <i>Neural Computation</i> , 2015 , 27, 2447-759		8
70	Feature Selection via l1-Penalized Squared-Loss Mutual Information. <i>IEICE Transactions on Information and Systems</i> , 2013 , E96.D, 1513-1524	0.6	8
69	Theory and algorithm for learning with dissimilarity functions. <i>Neural Computation</i> , 2009 , 21, 1459-84	2.9	8
68	A Feature-Driven Active Framework for Ultrasound-Based Brain Shift Compensation. <i>Lecture Notes in Computer Science</i> , 2018 , 30-38	0.9	8
67	Direct Estimation of the Derivative of Quadratic Mutual Information with Application in Supervised Dimension Reduction. <i>Neural Computation</i> , 2017 , 29, 2076-2122	2.9	7
66	Conditional density estimation with dimensionality reduction via squared-loss conditional entropy minimization. <i>Neural Computation</i> , 2015 , 27, 228-54	2.9	7
65	Model-based reinforcement learning with dimension reduction. <i>Neural Networks</i> , 2016 , 84, 1-16	9.1	7

64	Semi-supervised information-maximization clustering. <i>Neural Networks</i> , 2014 , 57, 103-11	9.1	7
63	Model-based policy gradients with parameter-based exploration by least-squares conditional density estimation. <i>Neural Networks</i> , 2014 , 57, 128-40	9.1	7
62	Least-squares independence regression for non-linear causal inference under non-Gaussian noise. <i>Machine Learning</i> , 2014 , 96, 249-267	4	7
61	On Kernel Parameter Selection in Hilbert-Schmidt Independence Criterion. <i>IEICE Transactions on Information and Systems</i> , 2012 , E95.D, 2564-2567	0.6	7
60	Active Feature Acquisition with Supervised Matrix Completion 2018 ,		7
59	Support consistency of direct sparse-change learning in Markov networks. <i>Annals of Statistics</i> , 2017 , 45,	3.2	6
58	On the Applicability of Registration Uncertainty. <i>Lecture Notes in Computer Science</i> , 2019 , 410-419	0.9	6
57	Active deep Q-learning with demonstration. <i>Machine Learning</i> , 2020 , 109, 1699-1725	4	6
56	Artificial Neural Variability for Deep Learning: On Overfitting, Noise Memorization, and Catastrophic Forgetting. <i>Neural Computation</i> , 2021 , 33, 2163-2192	2.9	6
55	Homotopy continuation approaches for robust SV classification and regression. <i>Machine Learning</i> , 2017 , 106, 1009-1038	4	5
54	Cross-Domain Matching with Squared-Loss Mutual Information. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2015 , 37, 1764-76	13.3	5
53	Classification from Triplet Comparison Data. <i>Neural Computation</i> , 2020 , 32, 659-681	2.9	5
52	Direct Density Derivative Estimation. <i>Neural Computation</i> , 2016 , 28, 1101-40	2.9	5
51	Computationally Efficient Class-Prior Estimation under Class Balance Change Using Energy Distance. <i>IEICE Transactions on Information and Systems</i> , 2016 , E99.D, 176-186	0.6	5
50	ECG-Based Concentration Recognition With Multi-Task Regression. <i>IEEE Transactions on Biomedical Engineering</i> , 2019 , 66, 101-110	5	5
49	Convex formulation of multiple instance learning from positive and unlabeled bags. <i>Neural Networks</i> , 2018 , 105, 132-141	9.1	5
48	Direct Density Ratio Estimation with Convolutional Neural Networks with Application in Outlier Detection. <i>IEICE Transactions on Information and Systems</i> , 2015 , E98.D, 1073-1079	0.6	4
47	Direct Approximation of Quadratic Mutual Information and Its Application to Dependence-Maximization Clustering. <i>IEICE Transactions on Information and Systems</i> , 2013 , E96.D, 2282-2285	9.6	4

46	SemiCCA: Efficient Semi-supervised Learning of Canonical Correlations. <i>IPSJ Online Transactions</i> , 2013 , 6, 37-44		4
45	Theoretical and Experimental Evaluation of the Subspace Information Criterion. <i>Machine Learning</i> , 2002 , 48, 25-50	4	4
44	Predictive Approaches for Low-Cost Preventive Medicine Program in Developing Countries 2015 ,		4
43	Principled analytic classifier for positive-unlabeled learning via weighted integral probability metric. <i>Machine Learning</i> , 2020 , 109, 513-532	4	4
42	Stochastic Multichannel Ranking with Brain Dynamics Preferences. <i>Neural Computation</i> , 2020 , 32, 1499-1530	3	
41	Regularized Multitask Learning for Multidimensional Log-Density Gradient Estimation. <i>Neural Computation</i> , 2016 , 28, 1388-410	2.9	3
40	Trial and Error: Using Previous Experiences as Simulation Models in Humanoid Motor Learning. <i>IEEE Robotics and Automation Magazine</i> , 2016 , 23, 96-105	3.4	3
39	Constrained Least-Squares Density-Difference Estimation. <i>IEICE Transactions on Information and Systems</i> , 2014 , E97.D, 1822-1829	0.6	3
38	Statistical Analysis of Distance Estimators with Density Differences and Density Ratios. <i>Entropy</i> , 2014 , 16, 921-942	2.8	3
37	Boosting and margin theory. <i>Frontiers of Electrical and Electronic Engineering</i> , 2012 , 7, 127-133		3
36	Winning the Kaggle Algorithmic Trading Challenge with the Composition of Many Models and Feature Engineering. <i>IEICE Transactions on Information and Systems</i> , 2013 , E96.D, 742-745	0.6	3
35	On Generalization Performance and Non-Convex Optimization of Extended Support Vector Machine. <i>New Generation Computing</i> , 2009 , 27, 259-279	0.9	3
34	A Multipurpose Linear Component Analysis Method Based on Modulated Hebb-Oja Learning Rule. <i>IEEE Signal Processing Letters</i> , 2008 , 15, 677-680	3.2	3
33	Classification From Pairwise Similarities/Dissimilarities and Unlabeled Data via Empirical Risk Minimization. <i>Neural Computation</i> , 2021 , 33, 1234-1268	2.9	3
32	Lung lesion detection in FDG-PET/CT with Gaussian process regression 2017 ,		2
31	Good arm identification via bandit feedback. <i>Machine Learning</i> , 2019 , 108, 721-745	4	2
30	A fault detection technique for the steel manufacturing process based on a normal pattern library. <i>IFAC-PapersOnLine</i> , 2015 , 48, 871-876	0.7	2
29	Registration of infrared transmission images using squared-loss mutual information. <i>Precision Engineering</i> , 2015 , 39, 187-193	2.9	2

28	Binary classification with ambiguous training data. <i>Machine Learning</i> , 2020 , 109, 2369-2388	4	2
27	Unsupervised Dimension Reduction via Least-Squares Quadratic Mutual Information. <i>IEICE Transactions on Information and Systems</i> , 2014 , E97.D, 2806-2809	0.6	2
26	Direct conditional probability density estimation with sparse feature selection. <i>Machine Learning</i> , 2015 , 100, 161-182	4	2
25	Improved algorithm for multiwavelength single-shot interferometric surface profiling: speeding up the multiwavelength-integrated local model fitting method by local information sharing. <i>Applied Optics</i> , 2013 , 52, 4042-53	1.7	2
24	Multiwavelength-integrated local model fitting method for interferometric surface profiling. <i>Applied Optics</i> , 2012 , 51, 6700-7	1.7	2
23	Spatiotemporal dynamics of odor representations in the human brain revealed by EEG decoding.. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2022 , 119, e2114966119 ^{11.5}		2
22	Sufficient Dimension Reduction via Direct Estimation of the Gradients of Logarithmic Conditional Densities. <i>Neural Computation</i> , 2018 , 30, 477-504	2.9	1
21	Improving importance estimation in pool-based batch active learning for approximate linear regression. <i>Neural Networks</i> , 2012 , 36, 73-82	9.1	1
20	Noise adaptive optimization of matrix initialization for frequency-domain independent component analysis 2013 , 23, 1-8		1
19	Clustering Unclustered Data: Unsupervised Binary Labeling of Two Datasets Having Different Class Balances 2013 ,		1
18	Salient Object Detection Based on Direct Density-ratio Estimation. <i>IPSJ Online Transactions</i> , 2013 , 6, 96-103		1
17	Designing Various Multivariate Analysis at Will via Generalized Pairwise Expression. <i>IPSJ Online Transactions</i> , 2013 , 6, 45-54		1
16	Early Stopping Heuristics in Pool-Based Incremental Active Learning for Least-Squares Probabilistic Classifier. <i>IEICE Transactions on Information and Systems</i> , 2012 , E95.D, 2065-2073	0.6	1
15	Constraint learning for control tasks with limited duration barrier functions. <i>Automatica</i> , 2021 , 127, 109504		1
14	Dependence maximization localization: a novel approach to 2D street-map-based robot localization. <i>Advanced Robotics</i> , 2016 , 30, 1431-1445	1.7	1
13	Clipped Matrix Completion: A Remedy for Ceiling Effects. <i>Proceedings of the AAAI Conference on Artificial Intelligence</i> , 2019 , 33, 5151-5158	5	1
12	An Online Policy Gradient Algorithm for Markov Decision Processes with Continuous States and Actions. <i>Neural Computation</i> , 2016 , 28, 563-93	2.9	0
11	Semisupervised Ordinal Regression Based on Empirical Risk Minimization. <i>Neural Computation</i> , 2021 , 33, 3361-3412	2.9	0

10	Millionaire: a hint-guided approach for crowdsourcing. <i>Machine Learning</i> , 2019 , 108, 831-858	4	o
9	Information-Theoretic Representation Learning for Positive-Unlabeled Classification. <i>Neural Computation</i> , 2021 , 33, 244-268	2.9	o
8	Foreword: special issue for the journal track of the 10th Asian Conference on Machine Learning (ACML 2018). <i>Machine Learning</i> , 2019 , 108, 717-719	4	
7	Online Direct Density-Ratio Estimation Applied to Inlier-Based Outlier Detection. <i>Neural Computation</i> , 2015 , 27, 1899-914	2.9	
6	Multi-Task Approach to Reinforcement Learning for Factored-State Markov Decision Problems. <i>IEICE Transactions on Information and Systems</i> , 2012 , E95.D, 2426-2437	0.6	
5	Computationally Efficient Multi-Label Classification by Least-Squares Probabilistic Classifiers. <i>IEICE Transactions on Information and Systems</i> , 2013 , E96.D, 1871-1874	0.6	
4	Unsupervised key frame selection using information theory and colour histogram difference. <i>International Journal of Business Intelligence and Data Mining</i> , 2020 , 16, 324	0.3	
3	Polynomial-Time Algorithms for Multiple-Arm Identification with Full-Bandit Feedback. <i>Neural Computation</i> , 2020 , 32, 1733-1773	2.9	
2	Direction Matters: On Influence-Preserving Graph Summarization and Max-Cut Principle for Directed Graphs. <i>Neural Computation</i> , 2021 , 33, 2128-2162	2.9	
1	Stochastic Divergence Minimization for Biterm Topic Models. <i>IEICE Transactions on Information and Systems</i> , 2018 , E101.D, 668-677	0.6	