Masashi Sugiyama

List of Publications by Year in Descending Order

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	Spatiotemporal dynamics of odor representations in the human brain revealed by EEG decoding.. *Proceedings of the National Academy of Sciences of the United States of America*, **2022**, 119, e2114966119	11.5	2
134	Semisupervised Ordinal Regression Based on Empirical Risk Minimization. *Neural Computation*, **2021**, 33, 3361-3412	2.9	0
133	Classification From Pairwise Similarities/Dissimilarities and Unlabeled Data via Empirical Risk Minimization. *Neural Computation*, **2021**, 33, 1234-1268	2.9	3
132	Constraint learning for control tasks with limited duration barrier functions. *Automatica*, **2021**, 127, 109504	5.7	1
131	Artificial Neural Variability for Deep Learning: On Overfitting, Noise Memorization, and Catastrophic Forgetting. *Neural Computation*, **2021**, 33, 2163-2192	2.9	6
130	Direction Matters: On Influence-Preserving Graph Summarization and Max-Cut Principle for Directed Graphs. *Neural Computation*, **2021**, 33, 2128-2162	2.9	
129	Information-Theoretic Representation Learning for Positive-Unlabeled Classification. *Neural Computation*, **2021**, 33, 244-268	2.9	0
128	Binary classification with ambiguous training data. *Machine Learning*, **2020**, 109, 2369-2388	4	2
127	Stochastic Multichannel Ranking with Brain Dynamics Preferences. *Neural Computation*, **2020**, 32, 1499-1530	2.9	3
126	Classification from Triplet Comparison Data. *Neural Computation*, **2020**, 32, 659-681	2.9	5
125	Principled analytic classifier for positive-unlabeled learning via weighted integral probability metric. *Machine Learning*, **2020**, 109, 513-532	4	4
124	Active deep Q-learning with demonstration. *Machine Learning*, **2020**, 109, 1699-1725	4	6
123	Unsupervised key frame selection using information theory and colour histogram difference. *International Journal of Business Intelligence and Data Mining*, **2020**, 16, 324	0.3	
122	Polynomial-Time Algorithms for Multiple-Arm Identification with Full-Bandit Feedback. *Neural Computation*, **2020**, 32, 1733-1773	2.9	
121	Learning Efficient Tensor Representations with Ring-structured Networks **2019**,		19
120	Foreword: special issue for the journal track of the 10th Asian Conference on Machine Learning (ACML 2018). *Machine Learning*, **2019**, 108, 717-719	4	
119	Good arm identification via bandit feedback. *Machine Learning*, **2019**, 108, 721-745	4	2

118	On the Applicability of Registration Uncertainty. *Lecture Notes in Computer Science*, **2019**, 410-419	0.9	6
117	Clipped Matrix Completion: A Remedy for Ceiling Effects. *Proceedings of the AAAI Conference on Artificial Intelligence*, **2019**, 33, 5151-5158	5	1
116	Millionaire: a hint-guided approach for crowdsourcing. *Machine Learning*, **2019**, 108, 831-858	4	0
115	ECG-Based Concentration Recognition With Multi-Task Regression. *IEEE Transactions on Biomedical Engineering*, **2019**, 66, 101-110	5	5
114	Sufficient Dimension Reduction via Direct Estimation of the Gradients of Logarithmic Conditional Densities. *Neural Computation*, **2018**, 30, 477-504	2.9	1
113	Semi-supervised AUC optimization based on positive-unlabeled learning. *Machine Learning*, **2018**, 107, 767-794	4	10
112	Using the variogram for vector outlier screening: application to feature-based image registration. *International Journal of Computer Assisted Radiology and Surgery*, **2018**, 13, 1871-1880	3.9	8
111	Bias Reduction and Metric Learning for Nearest-Neighbor Estimation of Kullback-Leibler Divergence. *Neural Computation*, **2018**, 30, 1930-1960	2.9	9
110	A Feature-Driven Active Framework for Ultrasound-Based Brain Shift Compensation. *Lecture Notes in Computer Science*, **2018**, 30-38	0.9	8
109	Stochastic Divergence Minimization for Biterm Topic Models. *IEICE Transactions on Information and Systems*, **2018**, E101.D, 668-677	0.6	
108	Active Feature Acquisition with Supervised Matrix Completion **2018**,		7
107	Convex formulation of multiple instance learning from positive and unlabeled bags. *Neural Networks*, **2018**, 105, 132-141	9.1	5
106	Lung lesion detection in FDG-PET/CT with Gaussian process regression **2017**,		2
105	Support consistency of direct sparse-change learning in Markov networks. *Annals of Statistics*, **2017**, 45,	3.2	6
104	Direct Estimation of the Derivative of Quadratic Mutual Information with Application in Supervised Dimension Reduction. *Neural Computation*, **2017**, 29, 2076-2122	2.9	7
103	Tensor Networks for Dimensionality Reduction and Large-scale Optimization: Part 2 Applications and Future Perspectives. *Foundations and Trends in Machine Learning*, **2017**, 9, 249-429	32.9	74
102	Homotopy continuation approaches for robust SV classification and regression. *Machine Learning*, **2017**, 106, 1009-1038	4	5
101	Class-prior estimation for learning from positive and unlabeled data. *Machine Learning*, **2017**, 106, 463-492	4	26

100	Geometry-aware principal component analysis for symmetric positive definite matrices. *Machine Learning*, **2017**, 106, 493-522	4	16
99	Model-based reinforcement learning with dimension reduction. *Neural Networks*, **2016**, 84, 1-16	9.1	7
98	Target-less camera-LiDAR extrinsic calibration using a bagged dependence estimator **2016**,		9
97	Statistical outlier detection for diagnosis of cyber attacks in power state estimation **2016**,		12
96	Modal Regression via Direct Log-Density Derivative Estimation. *Lecture Notes in Computer Science*, **2016**, 108-116	0.9	9
95	Direct Density Derivative Estimation. *Neural Computation*, **2016**, 28, 1101-40	2.9	5
94	Regularized Multitask Learning for Multidimensional Log-Density Gradient Estimation. *Neural Computation*, **2016**, 28, 1388-410	2.9	3
93	Theoretical and Experimental Analyses of Tensor-Based Regression and Classification. *Neural Computation*, **2016**, 28, 686-715	2.9	19
92	Trial and Error: Using Previous Experiences as Simulation Models in Humanoid Motor Learning. *IEEE Robotics and Automation Magazine*, **2016**, 23, 96-105	3.4	3
91	An Online Policy Gradient Algorithm for Markov Decision Processes with Continuous States and Actions. *Neural Computation*, **2016**, 28, 563-93	2.9	0
90	Computationally Efficient Class-Prior Estimation under Class Balance Change Using Energy Distance. *IEICE Transactions on Information and Systems*, **2016**, E99.D, 176-186	0.6	5
89	Dependence maximization localization: a novel approach to 2D street-map-based robot localization. *Advanced Robotics*, **2016**, 30, 1431-1445	1.7	1
88	Importance-weighted covariance estimation for robust common spatial pattern. *Pattern Recognition Letters*, **2015**, 68, 139-145	4.7	10
87	Online Direct Density-Ratio Estimation Applied to Inlier-Based Outlier Detection. *Neural Computation*, **2015**, 27, 1899-914	2.9	
86	A fault detection technique for the steel manufacturing process based on a normal pattern library. *IFAC-PapersOnLine*, **2015**, 48, 871-876	0.7	2
85	Cross-Domain Matching with Squared-Loss Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2015**, 37, 1764-76	13.3	5
84	Conditional density estimation with dimensionality reduction via squared-loss conditional entropy minimization. *Neural Computation*, **2015**, 27, 228-54	2.9	7
83	Registration of infrared transmission images using squared-loss mutual information. *Precision Engineering*, **2015**, 39, 187-193	2.9	2

82	Bandit-Based Task Assignment for Heterogeneous Crowdsourcing. *Neural Computation*, **2015**, 27, 2447-75	2.9	8
81	Direct Density Ratio Estimation with Convolutional Neural Networks with Application in Outlier Detection. *IEICE Transactions on Information and Systems*, **2015**, E98.D, 1073-1079	0.6	4
80	Averaging covariance matrices for EEG signal classification based on the CSP: An empirical study **2015**,		16
79	Direct conditional probability density estimation with sparse feature selection. *Machine Learning*, **2015**, 100, 161-182	4	2
78	Predictive Approaches for Low-Cost Preventive Medicine Program in Developing Countries **2015**,		4
77	Health checkup and telemedical intervention program for preventive medicine in developing countries: verification study. *Journal of Medical Internet Research*, **2015**, 17, e2	7.6	22
76	High-dimensional feature selection by feature-wise kernelized Lasso. *Neural Computation*, **2014**, 26, 185-207	2.9	136
75	Information-maximization clustering based on squared-loss mutual information. *Neural Computation*, **2014**, 26, 84-131	2.9	21
74	Semi-supervised learning of class balance under class-prior change by distribution matching. *Neural Networks*, **2014**, 50, 110-9	9.1	33
73	A least-squares approach to anomaly detection in static and sequential data. *Pattern Recognition Letters*, **2014**, 40, 36-40	4.7	30
72	Semi-supervised information-maximization clustering. *Neural Networks*, **2014**, 57, 103-11	9.1	7
71	Class Prior Estimation from Positive and Unlabeled Data. *IEICE Transactions on Information and Systems*, **2014**, E97.D, 1358-1362	0.6	35
70	Constrained Least-Squares Density-Difference Estimation. *IEICE Transactions on Information and Systems*, **2014**, E97.D, 1822-1829	0.6	3
69	Unsupervised Dimension Reduction via Least-Squares Quadratic Mutual Information. *IEICE Transactions on Information and Systems*, **2014**, E97.D, 2806-2809	0.6	2
68	Computationally Efficient Estimation of Squared-Loss Mutual Information with Multiplicative Kernel Models. *IEICE Transactions on Information and Systems*, **2014**, E97.D, 968-971	0.6	9
67	Tree-Based Ensemble Multi-Task Learning Method for Classification and Regression. *IEICE Transactions on Information and Systems*, **2014**, E97.D, 1677-1681	0.6	29
66	Model-based policy gradients with parameter-based exploration by least-squares conditional density estimation. *Neural Networks*, **2014**, 57, 128-40	9.1	7
65	Statistical Analysis of Distance Estimators with Density Differences and Density Ratios. *Entropy*, **2014**, 16, 921-942	2.8	3

64	Direct learning of sparse changes in Markov networks by density ratio estimation. *Neural Computation*, **2014**, 26, 1169-97	2.9	9
63	Information-theoretic semi-supervised metric learning via entropy regularization. *Neural Computation*, **2014**, 26, 1717-62	2.9	56
62	Least-squares independence regression for non-linear causal inference under non-Gaussian noise. *Machine Learning*, **2014**, 96, 249-267	4	7
61	Variational Bayesian sparse additive matrix factorization. *Machine Learning*, **2013**, 92, 319-347	4	30
60	Computational complexity of kernel-based density-ratio estimation: a condition number analysis. *Machine Learning*, **2013**, 90, 431-460	4	11
59	Learning under nonstationarity: covariate shift and class-balance change. *Wiley Interdisciplinary Reviews: Computational Statistics*, **2013**, 5, 465-477	1.4	12
58	Change-point detection in time-series data by relative density-ratio estimation. *Neural Networks*, **2013**, 43, 72-83	9.1	217
57	Sufficient dimension reduction via squared-loss mutual information estimation. *Neural Computation* , **2013**, 25, 725-58	2.9	39
56	Efficient sample reuse in policy gradients with parameter-based exploration. *Neural Computation*, **2013**, 25, 1512-47	2.9	16
55	Relative density-ratio estimation for robust distribution comparison. *Neural Computation*, **2013**, 25, 1324-70	2.9	41
54	Noise adaptive optimization of matrix initialization for frequency-domain independent component analysis **2013**, 23, 1-8		1
53	Clustering Unclustered Data: Unsupervised Binary Labeling of Two Datasets Having Different Class Balances **2013**,		1
52	Improved algorithm for multiwavelength single-shot interferometric surface profiling: speeding up the multiwavelength-integrated local model fitting method by local information sharing. *Applied Optics*, **2013**, 52, 4042-53	1.7	2
51	Machine Learning with Squared-Loss Mutual Information. *Entropy*, **2013**, 15, 80-112	2.8	24
50	Density-difference estimation. *Neural Computation*, **2013**, 25, 2734-75	2.9	28
49	Direct Approximation of Quadratic Mutual Information and Its Application to Dependence-Maximization Clustering. *IEICE Transactions on Information and Systems*, **2013**, E96.D, 2282-2285	0.6	4
48	Winning the Kaggle Algorithmic Trading Challenge with the Composition of Many Models and Feature Engineering. *IEICE Transactions on Information and Systems*, **2013**, E96.D, 742-745	0.6	3
47	Computationally Efficient Multi-Label Classification by Least-Squares Probabilistic Classifiers. *IEICE Transactions on Information and Systems*, **2013**, E96.D, 1871-1874	0.6	

46	Feature Selection via l1-Penalized Squared-Loss Mutual Information. *IEICE Transactions on Information and Systems*, **2013**, E96.D, 1513-1524	0.6	8
45	Artist Agent: A Reinforcement Learning Approach to Automatic Stroke Generation in Oriental Ink Painting. *IEICE Transactions on Information and Systems*, **2013**, E96.D, 1134-1144	0.6	14
44	Salient Object Detection Based on Direct Density-ratio Estimation. *IPSJ Online Transactions*, **2013**, 6, 96-103		1
43	SemiCCA: Efficient Semi-supervised Learning of Canonical Correlations. *IPSJ Online Transactions*, **2013**, 6, 37-44		4
42	Designing Various Multivariate Analysis at Will via Generalized Pairwise Expression. *IPSJ Online Transactions*, **2013**, 6, 45-54		1
41	Direct Divergence Approximation between Probability Distributions and Its Applications in Machine Learning. *Journal of Computing Science and Engineering*, **2013**, 7, 99-111	1.8	24
40	Importance-weighted least-squares probabilistic classifier for covariate shift adaptation with application to human activity recognition. *Neurocomputing*, **2012**, 80, 93-101	5.4	47
39	Analysis and improvement of policy gradient estimation. *Neural Networks*, **2012**, 26, 118-29	9.1	19
38	f -Divergence Estimation and Two-Sample Homogeneity Test Under Semiparametric Density-Ratio Models. *IEEE Transactions on Information Theory*, **2012**, 58, 708-720	2.8	20
37	Statistical analysis of kernel-based least-squares density-ratio estimation. *Machine Learning*, **2012**, 86, 335-367	4	43
36	Improving importance estimation in pool-based batch active learning for approximate linear regression. *Neural Networks*, **2012**, 36, 73-82	9.1	1
35	Canonical dependency analysis based on squared-loss mutual information. *Neural Networks*, **2012**, 34, 46-55	9.1	14
34	Density Ratio Estimation in Machine Learning **2012**,		112
33	Multi-Task Approach to Reinforcement Learning for Factored-State Markov Decision Problems. *IEICE Transactions on Information and Systems*, **2012**, E95.D, 2426-2437	0.6	
32	Sequential change-point detection based on direct density-ratio estimation. *Statistical Analysis and Data Mining*, **2012**, 5, 114-127	1.4	93
31	Multi-parametric solution-path algorithm for instance-weighted support vector machines. *Machine Learning*, **2012**, 88, 297-330	4	16
30	Boosting and margin theory. *Frontiers of Electrical and Electronic Engineering*, **2012**, 7, 127-133		3
29	Density-ratio matching under the Bregman divergence: a unified framework of density-ratio estimation. *Annals of the Institute of Statistical Mathematics*, **2012**, 64, 1009-1044	1	48

28	Multiwavelength-integrated local model fitting method for interferometric surface profiling. *Applied Optics*, **2012**, 51, 6700-7	1.7	2
27	On Kernel Parameter Selection in Hilbert-Schmidt Independence Criterion. *IEICE Transactions on Information and Systems*, **2012**, E95.D, 2564-2567	0.6	7
26	Early Stopping Heuristics in Pool-Based Incremental Active Learning for Least-Squares Probabilistic Classifier. *IEICE Transactions on Information and Systems*, **2012**, E95.D, 2065-2073	0.6	1
25	The Degrees of Freedom of Partial Least Squares Regression. *Journal of the American Statistical Association*, **2011**, 106, 697-705	2.8	54
24	Statistical outlier detection using direct density ratio estimation. *Knowledge and Information Systems*, **2011**, 26, 309-336	2.4	98
23	Direct density-ratio estimation with dimensionality reduction via least-squares hetero-distributional subspace search. *Neural Networks*, **2011**, 24, 183-98	9.1	24
22	Least-squares two-sample test. *Neural Networks*, **2011**, 24, 735-51	9.1	13
21	Reward-weighted regression with sample reuse for direct policy search in reinforcement learning. *Neural Computation*, **2011**, 23, 2798-832	2.9	13
20	Least-squares independent component analysis. *Neural Computation*, **2011**, 23, 284-301	2.9	22
19	Semi-supervised local Fisher discriminant analysis for dimensionality reduction. *Machine Learning*, **2010**, 78, 35-61	4	186
18	Application of covariate shift adaptation techniques in brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, **2010**, 57, 1318-24	5	91
17	Semi-supervised speaker identification under covariate shift. *Signal Processing*, **2010**, 90, 2353-2361	4.4	27
16	Dimensionality reduction for density ratio estimation in high-dimensional spaces. *Neural Networks*, **2010**, 23, 44-59	9.1	36
15	Theory and algorithm for learning with dissimilarity functions. *Neural Computation*, **2009**, 21, 1459-84	2.9	8
14	Pool-based active learning in approximate linear regression. *Machine Learning*, **2009**, 75, 249-274	4	40
13	On Generalization Performance and Non-Convex Optimization of Extended ν-Support Vector Machine. *New Generation Computing*, **2009**, 27, 259-279	0.9	3
12	Dual-Augmented Lagrangian Method for Efficient Sparse Reconstruction. *IEEE Signal Processing Letters*, **2009**, 16, 1067-1070	3.2	45
11	A Multipurpose Linear Component Analysis Method Based on Modulated Hebb-Oja Learning Rule. *IEEE Signal Processing Letters*, **2008**, 15, 677-680	3.2	3

10	Direct importance estimation for covariate shift adaptation. *Annals of the Institute of Statistical Mathematics*, **2008**, 60, 699-746	1	173
9	Geodesic Gaussian kernels for value function approximation. *Autonomous Robots*, **2008**, 25, 287-304	3	18
8	A batch ensemble approach to active learning with model selection. *Neural Networks*, **2008**, 21, 1278-86	9.1	15
7	A new algorithm of non-Gaussian component analysis with radial kernel functions. *Annals of the Institute of Statistical Mathematics*, **2007**, 59, 57-75	1	15
6	Single-shot surface profiling by local model fitting. *Applied Optics*, **2006**, 45, 7999-8005	1.7	23
5	Trading variance reduction with unbiasedness: the regularized subspace information criterion for robust model selection in kernel regression. *Neural Computation*, **2004**, 16, 1077-104	2.9	15
4	A unified method for optimizing linear image restoration filters. *Signal Processing*, **2002**, 82, 1773-1787	4.4	12
3	Theoretical and Experimental Evaluation of the Subspace Information Criterion. *Machine Learning*, **2002**, 48, 25-50	4	4
2	Subspace information criterion for model selection. *Neural Computation*, **2001**, 13, 1863-89	2.9	41
1	Incremental active learning for optimal generalization. *Neural Computation*, **2000**, 12, 2909-40	2.9	13