

Sayan Mukherjee

List of Publications by Year in Descending Order

Source: <https://exaly.com/author-pdf/5609377/sayan-mukherjee-publications-by-year.pdf>

Version: 2024-04-23

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.

90
papers

30,194
citations

38
h-index

101
g-index

101
ext. papers

40,106
ext. citations

7.4
avg, IF

6.43
L-index

#	Paper	IF	Citations
90	A topological data analytic approach for discovering biophysical signatures in protein dynamics.. <i>PLoS Computational Biology</i> , 2022 , 18, e1010045	5	1
89	Impact of anxiety and depression on progression to glaucoma among glaucoma suspects. <i>British Journal of Ophthalmology</i> , 2021 , 105, 1244-1249	5.5	6
88	A statistical pipeline for identifying physical features that differentiate classes of 3D shapes. <i>Annals of Applied Statistics</i> , 2021 , 15,	2.1	2
87	Measuring and mitigating PCR bias in microbiota datasets. <i>PLoS Computational Biology</i> , 2021 , 17, e1009113	5.13	7
86	The Geometry of Synchronization Problems and Learning Group Actions. <i>Discrete and Computational Geometry</i> , 2021 , 65, 150-211	0.6	2
85	Likelihood ratio statistics for gene set enrichment in Alzheimer's disease pathways. <i>Alzheimers and Dementia</i> , 2021 , 17, 561-573	1.2	1
84	Naught all zeros in sequence count data are the same. <i>Computational and Structural Biotechnology Journal</i> , 2020 , 18, 2789-2798	6.8	31
83	Predicting Clinical Outcomes in Glioblastoma: An Application of Topological and Functional Data Analysis. <i>Journal of the American Statistical Association</i> , 2020 , 115, 1139-1150	2.8	25
82	Label propagation defines signaling networks associated with recurrently mutated cancer genes. <i>Scientific Reports</i> , 2019 , 9, 9401	4.9	0
81	Phylofactorization: a graph partitioning algorithm to identify phylogenetic scales of ecological data. <i>Ecological Monographs</i> , 2019 , 89, e01353	9	33
80	Estimating Rates of Progression and Predicting Future Visual Fields in Glaucoma Using a Deep Variational Autoencoder. <i>Scientific Reports</i> , 2019 , 9, 18113	4.9	16
79	A unifying framework for interpreting and predicting mutualistic systems. <i>Nature Communications</i> , 2019 , 10, 242	17.4	15
78	Fast Moment Estimation for Generalized Latent Dirichlet Models. <i>Journal of the American Statistical Association</i> , 2018 , 113, 1528-1540	2.8	0
77	Development and Assessment of Fully Automated and Globally Transitive Geometric Morphometric Methods, With Application to a Biological Comparative Dataset With High Interspecific Variation. <i>Anatomical Record</i> , 2018 , 301, 636-658	2.1	20
76	Bayesian Approximate Kernel Regression with Variable Selection. <i>Journal of the American Statistical Association</i> , 2018 , 113, 1710-1721	2.8	16
75	Dynamic linear models guide design and analysis of microbiota studies within artificial human guts. <i>Microbiome</i> , 2018 , 6, 202	16.6	30
74	Geometric Representations of Random Hypergraphs. <i>Journal of the American Statistical Association</i> , 2017 , 112, 363-383	2.8	3

73	Differential expression analysis for RNAseq using Poisson mixed models. <i>Nucleic Acids Research</i> , 2017 , 45, e106	20.1	44
72	Topological consistency via kernel estimation. <i>Bernoulli</i> , 2017 , 23,	1.6	19
71	A phylogenetic transform enhances analysis of compositional microbiota data. <i>ELife</i> , 2017 , 6,	8.9	144
70	Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. <i>PLoS Genetics</i> , 2017 , 13, e1006869	6	53
69	Melanoma Therapeutic Strategies that Select against Resistance by Exploiting MYC-Driven Evolutionary Convergence. <i>Cell Reports</i> , 2017 , 21, 2796-2812	10.6	46
68	HOMINID: a framework for identifying associations between host genetic variation and microbiome composition. <i>GigaScience</i> , 2017 , 6, 1-7	7.6	5
67	Phylogenetic factorization of compositional data yields lineage-level associations in microbiome datasets. <i>PeerJ</i> , 2017 , 5, e2969	3.1	76
66	Response to Shen et al. <i>American Journal of Human Genetics</i> , 2016 , 99, 1220-1221	11	
65	Efficient Genome-Wide Sequencing and Low-Coverage Pedigree Analysis from Noninvasively Collected Samples. <i>Genetics</i> , 2016 , 203, 699-714	4	58
64	Random walks on simplicial complexes and harmonics. <i>Random Structures and Algorithms</i> , 2016 , 49, 379-405	4.05	30
63	Fast Principal-Component Analysis Reveals Convergent Evolution of ADH1B in Europe and East Asia. <i>American Journal of Human Genetics</i> , 2016 , 98, 456-472	11	201
62	The Information Geometry of Mirror Descent. <i>IEEE Transactions on Information Theory</i> , 2015 , 61, 1451-1458	14.58	17
61	A new fully automated approach for aligning and comparing shapes. <i>Anatomical Record</i> , 2015 , 298, 249-261	261	64
60	The topology of probability distributions on manifolds. <i>Probability Theory and Related Fields</i> , 2015 , 161, 651-686	1.4	29
59	Statistical inference for dynamical systems: A review. <i>Statistics Surveys</i> , 2015 , 9,	1.7	11
58	Citizen Science as a New Tool in Dog Cognition Research. <i>PLoS ONE</i> , 2015 , 10, e0135176	3.7	36
57	GSAASeqSP: a toolset for gene set association analysis of RNA-Seq data. <i>Scientific Reports</i> , 2014 , 4, 63474.9	4.9	48
56	A Cheeger-type inequality on simplicial complexes. <i>Advances in Applied Mathematics</i> , 2014 , 56, 56-77	0.8	22

55	Frühet Means for Distributions of Persistence Diagrams. <i>Discrete and Computational Geometry</i> , 2014 , 52, 44-70	0.6	74
54	Statistical analysis of crystallization database links protein physico-chemical features with crystallization mechanisms. <i>PLoS ONE</i> , 2014 , 9, e101123	3.7	15
53	A digital network approach to infer sex behavior in emerging HIV epidemics. <i>PLoS ONE</i> , 2014 , 9, e101416	3.7	4
52	Core and region-enriched networks of behaviorally regulated genes and the singing genome. <i>Science</i> , 2014 , 346, 1256780	33.3	81
51	Persistent homology transform for modeling shapes and surfaces. <i>Information and Inference</i> , 2014 , 3, 310-344	2.4	54
50	Novel distal eQTL analysis demonstrates effect of population genetic architecture on detecting and interpreting associations. <i>Genetics</i> , 2014 , 198, 879-93	4	10
49	Partial Factor Modeling: Predictor-Dependent Shrinkage for Linear Regression. <i>Journal of the American Statistical Association</i> , 2013 , 108, 999-1008	2.8	8
48	A comparative study of covariance selection models for the inference of gene regulatory networks. <i>Journal of Biomedical Informatics</i> , 2013 , 46, 894-904	10.2	12
47	Assessing the radiation response of lung cancer with different gene mutations using genetically engineered mice. <i>Frontiers in Oncology</i> , 2013 , 3, 72	5.3	26
46	Kernel Sliced Inverse Regression: Regularization and Consistency. <i>Abstract and Applied Analysis</i> , 2013 , 2013, 1-11	0.7	5
45	DNase-seq predicts regions of rotational nucleosome stability across diverse human cell types. <i>Genome Research</i> , 2013 , 23, 1118-29	9.7	21
44	Dissecting high-dimensional phenotypes with bayesian sparse factor analysis of genetic covariance matrices. <i>Genetics</i> , 2013 , 194, 753-67	4	41
43	Cyclin-dependent kinases are regulators and effectors of oscillations driven by a transcription factor network. <i>Molecular Cell</i> , 2012 , 45, 669-79	17.6	46
42	Genetics of gene expression responses to temperature stress in a sea urchin gene network. <i>Molecular Ecology</i> , 2012 , 21, 4547-62	5.7	60
41	Local Homology Transfer and Stratification Learning 2012 ,		18
40	Integrating genetic and gene expression evidence into genome-wide association analysis of gene sets. <i>Genome Research</i> , 2012 , 22, 386-97	9.7	77
39	Genetic effects on mating success and partner choice in a social mammal. <i>American Naturalist</i> , 2012 , 180, 113-29	3.7	18
38	A predictive framework for integrating disparate genomic data types using sample-specific gene set enrichment analysis and multi-task learning. <i>PLoS ONE</i> , 2012 , 7, e44635	3.7	4

37	Discovering genetic variants in Crohn's disease by exploring genomic regions enriched of weak association signals. <i>Digestive and Liver Disease</i> , 2011 , 43, 623-31	3.3	4
36	Probability measures on the space of persistence diagrams. <i>Inverse Problems</i> , 2011 , 27, 124007	2.3	83
35	Estimating variable structure and dependence in multitask learning via gradients. <i>Machine Learning</i> , 2011 , 83, 265-287	4	5
34	RS-SNP: a random-set method for genome-wide association studies. <i>BMC Genomics</i> , 2011 , 12, 166	4.5	1
33	Localized Sliced Inverse Regression. <i>Journal of Computational and Graphical Statistics</i> , 2010 , 19, 843-860	1.4	8
32	Making Mountains out of Molehills: Moving from Single Gene to Pathway Based Models of Colon Cancer Progression 2010 , 73-87		
31	Evidence-ranked motif identification. <i>Genome Biology</i> , 2010 , 11, R19	18.3	73
30	Learning gradients on manifolds. <i>Bernoulli</i> , 2010 , 16,	1.6	21
29	On the reproducibility of results of pathway analysis in genome-wide expression studies of colorectal cancers. <i>Journal of Biomedical Informatics</i> , 2010 , 43, 397-406	10.2	10
28	Characterizing the developmental pathways TTF-1, NKX2-8, and PAX9 in lung cancer. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2009 , 106, 5312-7	11.5	38
27	Genomic features that predict allelic imbalance in humans suggest patterns of constraint on gene expression variation. <i>Molecular Biology and Evolution</i> , 2009 , 26, 2047-59	8.3	7
26	Comparative study of gene set enrichment methods. <i>BMC Bioinformatics</i> , 2009 , 10, 275	3.6	78
25	Cross species genomic analysis identifies a mouse model as undifferentiated pleomorphic sarcoma/malignant fibrous histiocytoma. <i>PLoS ONE</i> , 2009 , 4, e8075	3.7	54
24	Modeling cancer progression via pathway dependencies. <i>PLoS Computational Biology</i> , 2008 , 4, e28	5	54
23	Age-specific differences in oncogenic pathway deregulation seen in human breast tumors. <i>PLoS ONE</i> , 2008 , 3, e1373	3.7	65
22	An integrated approach to the prediction of chemotherapeutic response in patients with breast cancer. <i>PLoS ONE</i> , 2008 , 3, e1908	3.7	71
21	Statistical Assessment of MSigDB Gene Sets in Colon Cancer. <i>Lecture Notes in Computer Science</i> , 2008 , 206-213	0.9	
20	Gene expression programs of human smooth muscle cells: tissue-specific differentiation and prognostic significance in breast cancers. <i>PLoS Genetics</i> , 2007 , 3, 1770-84	6	48

19	The Use of Unlabeled Data in Predictive Modeling. <i>Statistical Science</i> , 2007 , 22, 189	2.4	21
18	Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. <i>Lancet Oncology</i> , 2007 , 8, 1071-1078	21.7	170
17	Analysis of sample set enrichment scores: assaying the enrichment of sets of genes for individual samples in genome-wide expression profiles. <i>Bioinformatics</i> , 2006 , 22, e108-16	7.2	54
16	Evidence of influence of genomic DNA sequence on human X chromosome inactivation. <i>PLoS Computational Biology</i> , 2006 , 2, e113	5	72
15	Gene expression changes and molecular pathways mediating activity-dependent plasticity in visual cortex. <i>Nature Neuroscience</i> , 2006 , 9, 660-8	25.5	174
14	A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. <i>New England Journal of Medicine</i> , 2006 , 355, 570-80	59.2	512
13	Learning theory: stability is sufficient for generalization and necessary and sufficient for consistency of empirical risk minimization. <i>Advances in Computational Mathematics</i> , 2006 , 25, 161-193	1.6	73
12	Permutation Tests for Classification. <i>Lecture Notes in Computer Science</i> , 2005 , 501-515	0.9	45
11	Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2005 , 102, 15545-50	11.5	24578
10	An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. <i>Nature Genetics</i> , 2005 , 37, 48-55	36.3	361
9	Risk bounds for mixture density estimation. <i>ESAIM - Probability and Statistics</i> , 2005 , 9, 220-229	0.4	29
8	General conditions for predictivity in learning theory. <i>Nature</i> , 2004 , 428, 419-22	50.4	148
7	Estimating dataset size requirements for classifying DNA microarray data. <i>Journal of Computational Biology</i> , 2003 , 10, 119-42	1.7	191
6	Optimal gene expression analysis by microarrays. <i>Cancer Cell</i> , 2002 , 2, 353-61	24.3	128
5	Choosing Multiple Parameters for Support Vector Machines. <i>Machine Learning</i> , 2002 , 46, 131-159	4	1320
4	Learning Subspaces of Different Dimensions. <i>Journal of Computational and Graphical Statistics</i> , 1-35	1.4	
3	HOMINID: A framework for identifying associations between host genetic variation and microbiome composition		2
2	Dynamic linear models guide design and analysis of microbiota studies within artificial human guts		2

1 Statistical Considerations in the Design and Analysis of Longitudinal Microbiome Studies

5