

Zengyou He

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/2168065/publications.pdf>

Version: 2024-02-01

77
papers

3,238
citations

257450

24
h-index

149698

56
g-index

78
all docs

78
docs citations

78
times ranked

3355
citing authors

#	ARTICLE	IF	CITATIONS
1	Discovering cluster-based local outliers. <i>Pattern Recognition Letters</i> , 2003, 24, 1641-1650.	4.2	689
2	Stable feature selection for biomarker discovery. <i>Computational Biology and Chemistry</i> , 2010, 34, 215-225.	2.3	251
3	Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis. <i>BMC Bioinformatics</i> , 2009, 10, 4.	2.6	220
4	Improved and Promising Identification of Human MicroRNAs by Incorporating a High-Quality Negative Set. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2014, 11, 192-201.	3.0	207
5	SNPHarvester: a filtering-based approach for detecting epistatic interactions in genome-wide association studies. <i>Bioinformatics</i> , 2009, 25, 504-511.	4.1	182
6	On the Impact of Dissimilarity Measure in k-Modes Clustering Algorithm. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2007, 29, 503-507.	13.9	172
7	FP-outlier: Frequent pattern based outlier detection. <i>Computer Science and Information Systems</i> , 2005, 2, 103-118.	1.0	130
8	Squeezer: An efficient algorithm for clustering categorical data. <i>Journal of Computer Science and Technology</i> , 2002, 17, 611-624.	1.5	120
9	Protein inference: a review. <i>Briefings in Bioinformatics</i> , 2012, 13, 586-614.	6.5	96
10	Technical, bioinformatical and statistical aspects of liquid chromatography-mass spectrometry (LC-MS) and capillary electrophoresis-mass spectrometry (CE-MS) based clinical proteomics: A critical assessment. <i>Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences</i> , 2009, 877, 1250-1258.	2.3	80
11	An Optimization Model for Outlier Detection in Categorical Data. <i>Lecture Notes in Computer Science</i> , 2005, , 400-409.	1.3	70
12	Mining action rules from scratch. <i>Expert Systems With Applications</i> , 2005, 29, 691-699.	7.6	64
13	A cluster ensemble method for clustering categorical data. <i>Information Fusion</i> , 2005, 6, 143-151.	19.1	63
14	Mining class outliers: concepts, algorithms and applications in CRM. <i>Expert Systems With Applications</i> , 2004, 27, 681-697.	7.6	62
15	k-ANMI: A mutual information based clustering algorithm for categorical data. <i>Information Fusion</i> , 2008, 9, 223-233.	19.1	58
16	Scalable algorithms for clustering large datasets with mixed type attributes. <i>International Journal of Intelligent Systems</i> , 2005, 20, 1077-1089.	5.7	55
17	A Fast Greedy Algorithm for Outlier Mining. <i>Lecture Notes in Computer Science</i> , 2006, , 567-576.	1.3	47
18	G-ANMI: A mutual information based genetic clustering algorithm for categorical data. <i>Knowledge-Based Systems</i> , 2010, 23, 144-149.	7.1	44

#	ARTICLE	IF	CITATIONS
19	A Frequent Pattern Discovery Method for Outlier Detection. Lecture Notes in Computer Science, 2004, , 726-732.	1.3	35
20	Motif-All: discovering all phosphorylation motifs. BMC Bioinformatics, 2011, 12, S22.	2.6	32
21	Discriminative pattern mining and its applications in bioinformatics. Briefings in Bioinformatics, 2015, 16, 884-900.	6.5	29
22	Outlier Detection Integrating Semantic Knowledge. Lecture Notes in Computer Science, 2002, , 126-131.	1.3	27
23	Attribute value weighting in k-modes clustering. Expert Systems With Applications, 2011, 38, 15365-15369.	7.6	27
24	A linear programming model for protein inference problem in shotgun proteomics. Bioinformatics, 2012, 28, 2956-2962.	4.1	22
25	Protein complexes identification based on go attributed network embedding. BMC Bioinformatics, 2018, 19, 535.	2.6	22
26	Significance-based discriminative sequential pattern mining. Expert Systems With Applications, 2019, 122, 54-64.	7.6	22
27	Improving Mass Spectrometry Peak Detection Using Multiple Peak Alignment Results. Journal of Proteome Research, 2008, 7, 123-129.	3.7	21
28	Network inference from AP-MS data: computational challenges and solutions. Briefings in Bioinformatics, 2015, 16, 658-674.	6.5	21
29	Conditional discriminative pattern mining: Concepts and algorithms. Information Sciences, 2017, 375, 1-15.	6.9	21
30	TCSOM: Clustering Transactions Using Self-Organizing Map. Neural Processing Letters, 2005, 22, 249-262.	3.2	19
31	Mining conditional discriminative sequential patterns. Information Sciences, 2019, 478, 524-539.	6.9	19
32	ProteinLasso: A Lasso regression approach to protein inference problem in shotgun proteomics. Computational Biology and Chemistry, 2013, 43, 46-54.	2.3	18
33	Computing exact permutation p-values for association rules. Information Sciences, 2016, 346-347, 146-162.	6.9	15
34	A Unified Subspace Outlier Ensemble Framework for Outlier Detection. Lecture Notes in Computer Science, 2005, , 632-637.	1.3	15
35	A Partial Set Covering Model for Protein Mixture Identification Using Mass Spectrometry Data. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2011, 8, 368-380.	3.0	13
36	Reconstruction of the Protein-Protein Interaction Network for Protein Complexes Identification by Walking on the Protein Pair Fingerprints Similarity Network. Frontiers in Genetics, 2018, 9, 272.	2.3	13

#	ARTICLE	IF	CITATIONS
37	Decoy-free protein-level false discovery rate estimation. <i>Bioinformatics</i> , 2014, 30, 675-681.	4.1	12
38	Data construction for phosphorylation site prediction. <i>Briefings in Bioinformatics</i> , 2014, 15, 839-855.	6.5	12
39	Optimization-Based Peptide Mass Fingerprinting for Protein Mixture Identification. <i>Journal of Computational Biology</i> , 2010, 17, 221-235.	1.6	10
40	Mining Class Outliers: Concepts, Algorithms and Applications. <i>Lecture Notes in Computer Science</i> , 2004, , 589-599.	1.3	10
41	Score regularization for peptide identification. <i>BMC Bioinformatics</i> , 2011, 12, S2.	2.6	9
42	Mining Conditional Phosphorylation Motifs. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2014, 11, 915-927.	3.0	9
43	A graph-traversal approach to identify influential nodes in a network. <i>Patterns</i> , 2021, 2, 100321.	5.9	9
44	Mining top-k strongly correlated item pairs without minimum correlation threshold. <i>International Journal of Knowledge-Based and Intelligent Engineering Systems</i> , 2006, 10, 105-112.	1.0	8
45	Peak bagging for peptide mass fingerprinting. <i>Bioinformatics</i> , 2008, 24, 1293-1299.	4.1	8
46	BagReg: Protein inference through machine learning. <i>Computational Biology and Chemistry</i> , 2015, 57, 12-20.	2.3	8
47	Protein inference: A protein quantification perspective. <i>Computational Biology and Chemistry</i> , 2016, 63, 21-29.	2.3	8
48	A two-step framework for inferring direct protein-protein interaction network from AP-MS data. <i>BMC Systems Biology</i> , 2017, 11, 82.	3.0	8
49	Significance-Based Essential Protein Discovery. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2022, 19, 633-642.	3.0	8
50	An FP-Tree Based Approach for Mining All Strongly Correlated Item Pairs. <i>Lecture Notes in Computer Science</i> , 2005, , 735-740.	1.3	7
51	Computing exact P-values for community detection. <i>Data Mining and Knowledge Discovery</i> , 2020, 34, 833-869.	3.7	7
52	Decision Tree for Sequences. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2021, , 1-1.	5.7	7
53	Improving peptide identification with single-stage mass spectrum peaks. <i>Bioinformatics</i> , 2009, 25, 2969-2974.	4.1	6
54	The choice of null distributions for detecting gene-gene interactions in genome-wide association studies. <i>BMC Bioinformatics</i> , 2011, 12, S26.	2.6	6

#	ARTICLE	IF	CITATIONS
55	Bioinformatic Analysis of Data Generated from MALDI Mass Spectrometry for Biomarker Discovery. Topics in Current Chemistry, 2012, 331, 193-209.	4.0	6
56	Protein Complexes Identification with Family-Wise Error Rate Control. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2020, 17, 2062-2073.	3.0	6
57	Permutation methods for testing the significance of phosphorylation motifs. Statistics and Its Interface, 2012, 5, 61-73.	0.3	5
58	Clustering categorical data streams. Journal of Computational Methods in Sciences and Engineering, 2011, 11, 185-192.	0.2	4
59	Detecting Statistically Significant Communities. IEEE Transactions on Knowledge and Data Engineering, 2020, , 1-1.	5.7	4
60	Essential Protein Recognition via Community Significance. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2021, 18, 2788-2794.	3.0	4
61	A Combinatorial Perspective of the Protein Inference Problem. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2013, 10, 1542-1547.	3.0	3
62	On the statistical significance of protein complex. Quantitative Biology, 2018, 6, 313-320.	0.5	3
63	Protein Inference. Advances in Experimental Medicine and Biology, 2016, 919, 237-242.	1.6	2
64	Reinforce: An Ensemble Approach for Inferring PPI Network from AP-MS Data. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2019, 16, 365-376.	3.0	2
65	Improving categorical data clustering algorithm by weighting uncommon attribute value matches. Computer Science and Information Systems, 2006, 3, 23-32.	1.0	2
66	Dynamic Self-configuration of User QoS for Next Generation Network. , 2009, , .		1
67	Comments on 'MMFP: A Maximal Motif Finder for Phosphoproteomics Datasets'. Bioinformatics, 2012, 28, 2211-2212.	4.1	1
68	Peptide Reranking with Protein-Peptide Correspondence and Precursor Peak Intensity Information. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2012, 9, 1212-1219.	3.0	1
69	Reference-Based Sequence Classification. IEEE Access, 2020, 8, 218199-218214.	4.2	1
70	Instance-Based Classification Through Hypothesis Testing. IEEE Access, 2021, 9, 17485-17494.	4.2	1
71	On the statistical significance of communities from weighted graphs. Scientific Reports, 2021, 11, 20304.	3.3	1
72	Optimization-Based Peptide Mass Fingerprinting for Protein Mixture Identification. Nature Precedings, 2008, , .	0.1	0

#	ARTICLE	IF	CITATIONS
73	Phosphorylation motif discovery. , 2015, , 19-27.		0
74	Protein inference in shotgun proteomics. , 2015, , 39-49.		0
75	Biomarker discovery. , 2015, , 69-78.		0
76	RPML: A Learning-Based Approach for Reranking Protein-Spectrum Matches. Lecture Notes in Computer Science, 2018, , 559-564.	1.3	0
77	On the Statistical Significance of a Community Structure. IEEE Transactions on Knowledge and Data Engineering, 2021, , 1-1.	5.7	0