

# Jinyan Li

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/8059075/publications.pdf>

Version: 2024-02-01

228  
papers

7,718  
citations

101496

36  
h-index

79644

73  
g-index

239  
all docs

239  
docs citations

239  
times ranked

7142  
citing authors

#	ARTICLE	IF	CITATIONS
1	A binary PSO-based ensemble under-sampling model for rebalancing imbalanced training data. Journal of Supercomputing, 2022, 78, 7428-7463.	2.4	17
2	FUNMarker: Fusion Network-Based Method to Identify Prognostic and Heterogeneous Breast Cancer Biomarkers. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2021, 18, 2483-2491.	1.9	10
3	Genetic source completeness of HIV-1 circulating recombinant forms (CRFs) predicted by multi-label learning. Bioinformatics, 2021, 37, 750-758.	1.8	3
4	Targeted Therapy of <i>TERT</i> -Rearranged Neuroblastoma with BET Bromodomain Inhibitor and Proteasome Inhibitor Combination Therapy. Clinical Cancer Research, 2021, 27, 1438-1451.	3.2	20
5	Dynamic swarm class rebalancing for the process mining of rare events. Journal of Supercomputing, 2021, 77, 7549-7583.	2.4	0
6	Transformation of FASTA files into feature vectors for unsupervised compression of short reads databases. Journal of Bioinformatics and Computational Biology, 2021, 19, 2050048.	0.3	4
7	Multigene editing: current approaches and beyond. Briefings in Bioinformatics, 2021, 22, .	3.2	3
8	Deep learning detection of anomalous patterns from bus trajectories for traffic insight analysis. Knowledge-Based Systems, 2021, 217, 106833.	4.0	13
9	Infectious disease mRNA vaccines and a review on epitope prediction for vaccine design. Briefings in Functional Genomics, 2021, 20, 289-303.	1.3	16
10	Single-cell multi-omics sequencing: application trends, COVID-19, data analysis issues and prospects. Briefings in Bioinformatics, 2021, 22, .	3.2	14
11	Instance-based error correction for short reads of disease-associated genes. BMC Bioinformatics, 2021, 22, 142.	1.2	1
12	Aberration-corrected ultrafine analysis of miRNA reads at single-base resolution: a <i>k</i> -mer lattice approach. Nucleic Acids Research, 2021, 49, e106-e106.	6.5	4
13	Hamming-shifting graph of genomic short reads: Efficient construction and its application for compression. PLoS Computational Biology, 2021, 17, e1009229.	1.5	5
14	Substituting clinical features using synthetic medical phrases: Medical text data augmentation techniques. Artificial Intelligence in Medicine, 2021, 120, 102167.	3.8	1
15	OUP accepted manuscript. Briefings in Functional Genomics, 2021, , .	1.3	2
16	Bi-Level Error Correction for PacBio Long Reads. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2020, 17, 899-905.	1.9	0
17	Prediction of Taxi Destinations Using a Novel Data Embedding Method and Ensemble Learning. IEEE Transactions on Intelligent Transportation Systems, 2020, 21, 68-78.	4.7	50
18	Accurate prediction of DNA N4-methylcytosine sites via boost-learning various types of sequence features. BMC Genomics, 2020, 21, 627.	1.2	20

#	ARTICLE	IF	CITATIONS
19	Sequencing dropout-and-batch effect normalization for single-cell mRNA profiles: a survey and comparative analysis. <i>Briefings in Bioinformatics</i> , 2020, 22, .	3.2	4
20	Allowing mutations in maximal matches boosts genome compression performance. <i>Bioinformatics</i> , 2020, 36, 4675-4681.	1.8	13
21	Guest Editorial for the 29th International Conference on Genome Informatics (GIW 2018). <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2020, 17, 726-727.	1.9	0
22	A Dictionary-based Oversampling Approach to Clinical Document Classification on Small and Imbalanced Dataset. , 2020, , .		1
23	An Ontology-based Two-Stage Approach to Medical Text Classification with Feature Selection by Particle Swarm Optimisation. , 2019, , .		11
24	The long noncoding RNA lncNB1 promotes tumorigenesis by interacting with ribosomal protein RPL35. <i>Nature Communications</i> , 2019, 10, 5026.	5.8	67
25	Sequence-based prediction of protein-protein interaction sites by simplified long short-term memory network. <i>Neurocomputing</i> , 2019, 357, 86-100.	3.5	71
26	Fast detection of maximal exact matches via fixed sampling of query $k$ -mers and Bloom filtering of index $k$ -mers. <i>Bioinformatics</i> , 2019, 35, 4560-4567.	1.8	21
27	Dual feature selection and rebalancing strategy using metaheuristic optimization algorithms in X-ray image datasets. <i>Multimedia Tools and Applications</i> , 2019, 78, 20913-20933.	2.6	17
28	Inverse similarity and reliable negative samples for drug side-effect prediction. <i>BMC Bioinformatics</i> , 2019, 19, 554.	1.2	26
29	Sketch distance-based clustering of chromosomes for large genome database compression. <i>BMC Genomics</i> , 2019, 20, 978.	1.2	3
30	DDI-PULearn: a positive-unlabeled learning method for large-scale prediction of drug-drug interactions. <i>BMC Bioinformatics</i> , 2019, 20, 661.	1.2	30
31	Old drug repositioning and new drug discovery through similarity learning from drug-target joint feature spaces. <i>BMC Bioinformatics</i> , 2019, 20, 605.	1.2	10
32	Construction of competing endogenous RNA networks from paired RNA-seq data sets by pointwise mutual information. <i>BMC Genomics</i> , 2019, 20, 943.	1.2	5
33	Identification of lung cancer gene markers through kernel maximum mean discrepancy and information entropy. <i>BMC Medical Genomics</i> , 2019, 12, 183.	0.7	3
34	Index suffix "prefix overlaps by $(w, k)$ -minimizer to generate long contigs for reads compression. <i>Bioinformatics</i> , 2019, 35, 2066-2074.	1.8	27
35	Detection of Anomalous Traffic Patterns and Insight Analysis from Bus Trajectory Data. <i>Lecture Notes in Computer Science</i> , 2019, , 307-321.	1.0	1
36	Stratifying Risk of Coronary Artery Disease Using Discriminative Knowledge-Guided Medical Concept Pairings from Clinical Notes. <i>Lecture Notes in Computer Science</i> , 2019, , 457-473.	1.0	2

#	ARTICLE	IF	CITATIONS
37	CRISPR/Cas9 cleavage efficiency regression through boosting algorithms and Markov sequence profiling. <i>Bioinformatics</i> , 2018, 34, 3069-3077.	1.8	39
38	Similarity Majority Under-Sampling Technique for Easing Imbalanced Classification Problem. <i>Communications in Computer and Information Science</i> , 2018, , 3-23.	0.4	2
39	Discovering pan-correlation patterns from time course data sets by efficient mining algorithms. <i>Computing (Vienna/New York)</i> , 2018, 100, 421-437.	3.2	4
40	Novel overlapping subgraph clustering for the detection of antigen epitopes. <i>Bioinformatics</i> , 2018, 34, 2061-2068.	1.8	6
41	AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. <i>Scientific Reports</i> , 2018, 8, 1697.	1.6	181
42	Delineation of the frequency and boundary of chromosomal copy number variations in paediatric neuroblastoma. <i>Cell Cycle</i> , 2018, 17, 749-758.	1.3	13
43	Adaptive multi-objective swarm fusion for imbalanced data classification. <i>Information Fusion</i> , 2018, 39, 1-24.	11.7	72
44	A suite of swarm dynamic multi-objective algorithms for rebalancing extremely imbalanced datasets. <i>Applied Soft Computing Journal</i> , 2018, 69, 784-805.	4.1	6
45	Introduction to Selected Papers from GIW2018. <i>Journal of Bioinformatics and Computational Biology</i> , 2018, 16, 1802005.	0.3	0
46	Identification of pre-microRNAs by characterizing their sequence order evolution information and secondary structure graphs. <i>BMC Bioinformatics</i> , 2018, 19, 521.	1.2	8
47	Predicting adverse drug reactions of combined medication from heterogeneous pharmacologic databases. <i>BMC Bioinformatics</i> , 2018, 19, 517.	1.2	18
48	Hot spot prediction in protein-protein interactions by an ensemble system. <i>BMC Systems Biology</i> , 2018, 12, 132.	3.0	24
49	dbMPIKT: a database of kinetic and thermodynamic mutant protein interactions. <i>BMC Bioinformatics</i> , 2018, 19, 455.	1.2	26
50	Predicting Drug Targets from Heterogeneous Spaces using Anchor Graph Hashing and Ensemble Learning. , 2018, , .		6
51	An isomiR expression panel based novel breast cancer classification approach using improved mutual information. <i>BMC Medical Genomics</i> , 2018, 11, 118.	0.7	16
52	Connectivity Based Method for Clustering Microbial Communities from Metagenomics Data of Water and Soil Samples. , 2018, , .		0
53	Mining statistically-solid k-mers for accurate NGS error correction. <i>BMC Genomics</i> , 2018, 19, 912.	1.2	15
54	Recognition of CRISPR/Cas9 off-target sites through ensemble learning of uneven mismatch distributions. <i>Bioinformatics</i> , 2018, 34, i757-i765.	1.8	38

#	ARTICLE	IF	CITATIONS
55	Version Space Completeness for Novel Hypothesis Induction in Biomedical Applications. , 2018, , .		0
56	Imbalance learning for the prediction of N6-Methylation sites in mRNAs. BMC Genomics, 2018, 19, 574.	1.2	48
57	Prediction of 8-state protein secondary structures by a novel deep learning architecture. BMC Bioinformatics, 2018, 19, 293.	1.2	92
58	Benchmarking Swarm Rebalancing Algorithm for Relieving Imbalanced Machine Learning Problems. International Series on Computer Entertainment and Media Technology, 2018, , 1-40.	0.7	0
59	A bioinformatics potpourri. BMC Genomics, 2018, 19, 920.	1.2	1
60	Emerging Pattern Based Classification. , 2018, , 1293-1294.		1
61	Distinction Between Ships and Icebergs in SAR Images Using Ensemble Loss Trained Convolutional Neural Networks. Lecture Notes in Computer Science, 2018, , 216-223.	1.0	0
62	Emerging Patterns. , 2018, , 1294-1297.		0
63	Applications of Emerging Patterns for Microarray Gene Expression Data Analysis. , 2018, , 131-132.		0
64	Exploring Consensus RNA Substructural Patterns Using Subgraph Mining. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2017, 14, 1134-1146.	1.9	10
65	MapReduce for accurate error correction of next-generation sequencing data. Bioinformatics, 2017, 33, 3844-3851.	1.8	15
66	High-speed and high-ratio referential genome compression. Bioinformatics, 2017, 33, 3364-3372.	1.8	31
67	predCar-site: Carbonylation sites prediction in proteins using support vector machine with resolving data imbalanced issue. Analytical Biochemistry, 2017, 525, 107-113.	1.1	30
68	Septic shock prediction for ICU patients via coupled HMM walking on sequential contrast patterns. Journal of Biomedical Informatics, 2017, 66, 19-31.	2.5	76
69	Using propensity scores to predict the kinases of unannotated phosphopeptides. Knowledge-Based Systems, 2017, 135, 60-76.	4.0	7
70	Protein binding hot spots prediction from sequence only by a new ensemble learning method. Amino Acids, 2017, 49, 1773-1785.	1.2	35
71	Rare Event Prediction Using Similarity Majority Under-Sampling Technique. Communications in Computer and Information Science, 2017, , 23-39.	0.4	7
72	Structure embedding for knowledge base completion and analytics. , 2017, , .		2

#	ARTICLE	IF	CITATIONS
73	Elitist Binary Wolf Search Algorithm for Heuristic Feature Selection in High-Dimensional Bioinformatics Datasets. <i>Scientific Reports</i> , 2017, 7, 4354.	1.6	23
74	Staged heterogeneity learning to identify conformational B-cell epitopes from antigen sequences. <i>BMC Genomics</i> , 2017, 18, 113.	1.2	5
75	Cross disease analysis of co-functional microRNA pairs on a reconstructed network of disease-gene-microRNA tripartite. <i>BMC Bioinformatics</i> , 2017, 18, 193.	1.2	12
76	50/50 Expressional Odds of Retention Signifies the Distinction between Retained Introns and Constitutively Spliced Introns in <i>Arabidopsis thaliana</i> . <i>Frontiers in Plant Science</i> , 2017, 8, 1728.	1.7	2
77	An Optimized Drug Similarity Framework for Side-effect Prediction. , 2017, , .		3
78	Adaptive Swarm Balancing Algorithms for rare-event prediction in imbalanced healthcare data. <i>PLoS ONE</i> , 2017, 12, e0180830.	1.1	34
79	Chromosome preference of disease genes and vectorization for the prediction of non-coding disease genes. <i>Oncotarget</i> , 2017, 8, 78901-78916.	0.8	2
80	SOHPRED: a new bioinformatics tool for the characterization and prediction of human S-sulfonylation sites. <i>Molecular BioSystems</i> , 2016, 12, 2849-2858.	2.9	37
81	Predicting short-term ICU outcomes using a sequential contrast motif based classification framework. , 2016, 2016, 5612-5615.		4
82	Adaptive swarm cluster-based dynamic multi-objective synthetic minority oversampling technique algorithm for tackling binary imbalanced datasets in biomedical data classification. <i>BioData Mining</i> , 2016, 9, 37.	2.2	34
83	Using constrained information entropy to detect rare adverse drug reactions from medical forums. , 2016, 2016, 2460-2463.		5
84	Identification of WD40 repeats by secondary structure-aided profileâ€“profile alignment. <i>Journal of Theoretical Biology</i> , 2016, 398, 122-129.	0.8	7
85	Accurate detection for a wide range of mutation and editing sites of microRNAs from small RNA high-throughput sequencing profiles. <i>Nucleic Acids Research</i> , 2016, 44, e123-e123.	6.5	43
86	Depth-First Search Encoding of RNA Substructures. <i>Lecture Notes in Computer Science</i> , 2016, , 328-334.	1.0	0
87	Using contrast patterns between true complexes and random subgraphs in PPI networks to predict unknown protein complexes. <i>Scientific Reports</i> , 2016, 6, 21223.	1.6	30
88	Grouping miRNAs of similar functions via weighted information content of gene ontology. <i>BMC Bioinformatics</i> , 2016, 17, 507.	1.2	13
89	A Sequence-Based Dynamic Ensemble Learning System for Protein Ligand-Binding Site Prediction. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2016, 13, 901-912.	1.9	32
90	Improving the classification performance of biological imbalanced datasets by swarm optimization algorithms. <i>Journal of Supercomputing</i> , 2016, 72, 3708-3728.	2.4	36

#	ARTICLE	IF	CITATIONS
91	Hypotension Risk Prediction via Sequential Contrast Patterns of ICU Blood Pressure. IEEE Journal of Biomedical and Health Informatics, 2016, 20, 1416-1426.	3.9	31
92	Hierarchical classification in text mining for sentiment analysis of online news. Soft Computing, 2016, 20, 3411-3420.	2.1	44
93	Coordinating Discernibility and Independence Scores of Variables in a 2D Space for Efficient and Accurate Feature Selection. Lecture Notes in Computer Science, 2016, , 116-127.	1.0	4
94	Efficient Mining of Pan-Correlation Patterns from Time Course Data. Lecture Notes in Computer Science, 2016, , 234-249.	1.0	1
95	Adaptive Multi-objective Swarm Crossover Optimization for Imbalanced Data Classification. Lecture Notes in Computer Science, 2016, , 374-390.	1.0	8
96	Solving the Under-Fitting Problem for Decision Tree Algorithms by Incremental Swarm Optimization in Rare-Event Healthcare Classification. Journal of Medical Imaging and Health Informatics, 2016, 6, 1102-1110.	0.2	19
97	The long noncoding RNA MALAT1 promotes tumor-driven angiogenesis by up-regulating pro-angiogenic gene expression. Oncotarget, 2016, 7, 8663-8675.	0.8	97
98	Applications of Emerging Patterns for Microarray Gene Expression Data Analysis. , 2016, , 1-2.		0
99	Emerging Patterns. , 2016, , 1-3.		0
100	Deriving Public Sector Workforce Insights: A Case Study Using Australian Public Sector Employment Profiles. Lecture Notes in Computer Science, 2016, , 764-774.	1.0	0
101	Optimizing SMOTE by Metaheuristics with Neural Network and Decision Tree. , 2015, , .		28
102	Inference of gene interaction networks using conserved subsequential patterns from multiple time course gene expression datasets. BMC Genomics, 2015, 16, S4.	1.2	1
103	Connecting rules from paired miRNA and mRNA expression data sets of HCV patients to detect both inverse and positive regulatory relationships. BMC Genomics, 2015, 16, S11.	1.2	21
104	Co-Occurring Atomic Contacts for the Characterization of Protein Binding Hot Spots. PLoS ONE, 2015, 10, e0144486.	1.1	8
105	Identification of lung cancer miRNA-miRNA co-regulation networks through a progressive data refining approach. Journal of Theoretical Biology, 2015, 380, 271-279.	0.8	18
106	Positive-unlabeled learning for the prediction of conformational B-cell epitopes. BMC Bioinformatics, 2015, 16, S12.	1.2	22
107	Laplacian normalization and random walk on heterogeneous networks for disease-gene prioritization. Computational Biology and Chemistry, 2015, 57, 21-28.	1.1	35
108	Burial Level Change Defines a High Energetic Relevance for Protein Binding Interfaces. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2015, 12, 410-421.	1.9	1

#	ARTICLE	IF	CITATIONS
109	Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning. <i>Computers and Electronics in Agriculture</i> , 2015, 119, 123-132.	3.7	120
110	Computational Identification of Protein Pupylation Sites by Using Profile-Based Composition of k-Spaced Amino Acid Pairs. <i>PLoS ONE</i> , 2015, 10, e0129635.	1.1	65
111	Advances of Applying Metaheuristics to Data Mining Techniques. <i>Advances in Data Mining and Database Management Book Series</i> , 2015, , 75-103.	0.4	4
112	Using sequential patterns as features for classification models to make accurate predictions on ICU events. <i>Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference</i> , 2015, 2015, 8157-60.	0.5	0
113	Influenza A HA's conserved epitopes and broadly neutralizing antibodies: A prediction method. <i>Journal of Bioinformatics and Computational Biology</i> , 2014, 12, 1450023.	0.3	5
114	Coupling Graphs, Efficient Algorithms and B-Cell Epitope Prediction. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2014, 11, 7-16.	1.9	6
115	Hierarchical Classification in Text Mining for Sentiment Analysis. , 2014, , .		6
116	Bird Flu Outbreak Prediction via Satellite Tracking. <i>IEEE Intelligent Systems</i> , 2014, 29, 10-17.	4.0	2
117	Rule discovery and distance separation to detect reliable miRNA biomarkers for the diagnosis of lung squamous cell carcinoma. <i>BMC Genomics</i> , 2014, 15, S16.	1.2	10
118	Tertiary structure-based prediction of conformational B-cell epitopes through B factors. <i>Bioinformatics</i> , 2014, 30, i264-i273.	1.8	33
119	PPI-RO: a two-stage method for protein-protein interaction extraction based on interaction relation ontology. <i>International Journal of Data Mining and Bioinformatics</i> , 2014, 10, 98.	0.1	0
120	Polyline-sourced Geodesic Voronoi Diagrams on Triangle Meshes. <i>Computer Graphics Forum</i> , 2014, 33, 161-170.	1.8	14
121	k-Partite cliques of protein interactions: A novel subgraph topology for functional coherence analysis on PPI networks. <i>Journal of Theoretical Biology</i> , 2014, 340, 146-154.	0.8	9
122	Feature Selection in Life Science Classification: Metaheuristic Swarm Search. <i>IT Professional</i> , 2014, 16, 24-29.	1.4	48
123	Integrating water exclusion theory into $\hat{\Gamma}^2$ contacts to predict binding free energy changes and binding hot spots. <i>BMC Bioinformatics</i> , 2014, 15, 57.	1.2	8
124	Use B-factor related features for accurate classification between protein binding interfaces and crystal packing contacts. <i>BMC Bioinformatics</i> , 2014, 15, S3.	1.2	56
125	Modeling Asymmetry and Tail Dependence among Multiple Variables by Using Partial Regular Vine. , 2014, , .		5
126	Binding Affinity Prediction for Protein-Ligand Complexes Based on $\langle i \rangle \hat{\Gamma}^2 \langle i \rangle$ Contacts and B Factor. <i>Journal of Chemical Information and Modeling</i> , 2013, 53, 3076-3085.	2.5	53



#	ARTICLE	IF	CITATIONS
127	Structural analysis on mutation residues and interfacial water molecules for human TIM disease understanding. <i>BMC Bioinformatics</i> , 2013, 14, S11.	1.2	14
128	Accurate prediction of hot spot residues through physicochemical characteristics of amino acid sequences. <i>Proteins: Structure, Function and Bioinformatics</i> , 2013, 81, 1351-1362.	1.5	43
129	Sentiment Analysis of Online News Using MALLET. , 2013, , .		14
130	Protein Binding Interfaces and Their Binding Hot Spot Prediction: A Survey. <i>Translational Bioinformatics</i> , 2013, , 79-106.	0.0	1
131	Optimal Allocation of High Dimensional Assets through Canonical Vines. <i>Lecture Notes in Computer Science</i> , 2013, , 366-377.	1.0	2
132	Model the complex dependence structures of financial variables by using canonical vine. , 2012, , .		5
133	Conservation of water molecules in protein binding interfaces. <i>International Journal of Bioinformatics Research and Applications</i> , 2012, 8, 228.	0.1	3
134	Detection of Outlier Residues for Improving Interface Prediction in Protein Heterocomplexes. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2012, 9, 1155-1165.	1.9	22
135	Progressive dry-core-wet-rim hydration trend in a nested-ring topology of protein binding interfaces. <i>BMC Bioinformatics</i> , 2012, 13, 51.	1.2	7
136	Z-score biological significance of binding hot spots of protein interfaces by using crystal packing as the reference state. <i>Biochimica Et Biophysica Acta - Proteins and Proteomics</i> , 2012, 1824, 1457-1467.	1.1	2
137	B-cell epitope prediction through a graph model. <i>BMC Bioinformatics</i> , 2012, 13, S20.	1.2	29
138	Disease gene identification by random walk on multigraphs merging heterogeneous genomic and phenotype data. <i>BMC Genomics</i> , 2012, 13, S27.	1.2	83
139	Structural and Functional Analysis of Multi-Interface Domains. <i>PLoS ONE</i> , 2012, 7, e50821.	1.1	5
140	Mining Iterative Generators and Representative Rules for Software Specification Discovery. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2011, 23, 282-296.	4.0	10
141	Exploring the wild birds's migration data for the disease spread study of H5N1: a clustering and association approach. <i>Knowledge and Information Systems</i> , 2011, 27, 227-251.	2.1	17
142	DBAC: A simple prediction method for protein binding hot spots based on burial levels and deeply buried atomic contacts. <i>BMC Systems Biology</i> , 2011, 5, S5.	3.0	12
143	Distinct interfacial biclique patterns between ssDNA-binding proteins and those with dsDNAs. <i>Proteins: Structure, Function and Bioinformatics</i> , 2011, 79, 598-610.	1.5	4
144	A case study on financial ratios via cross-graph quasi-bicliques. <i>Information Sciences</i> , 2011, 181, 201-216.	4.0	15

#	ARTICLE	IF	CITATIONS
145	Structural analysis of the hot spots in the binding between H1N1 HA and the 2D1 antibody: do mutations of H1N1 from 1918 to 2009 affect much on this binding?. <i>Bioinformatics</i> , 2011, 27, 2529-2536.	1.8	8
146	Antibody-Specified B-Cell Epitope Prediction in Line with the Principle of Context-Awareness. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2011, 8, 1483-1494.	1.9	36
147	DomSVR: domain boundary prediction with support vector regression from sequence information alone. <i>Amino Acids</i> , 2010, 39, 713-726.	1.2	23
148	Protein binding hot spots and the residue-residue pairing preference: a water exclusion perspective. <i>BMC Bioinformatics</i> , 2010, 11, 244.	1.2	27
149	Sequence-based identification of interface residues by an integrative profile combining hydrophobic and evolutionary information. <i>BMC Bioinformatics</i> , 2010, 11, 402.	1.2	40
150	Prediction of protein long-range contacts using an ensemble of genetic algorithm classifiers with sequence profile centers. <i>BMC Structural Biology</i> , 2010, 10, S2.	2.3	16
151	Mining for the antibody-antigen interacting associations that predict the B cell epitopes. <i>BMC Structural Biology</i> , 2010, 10, S6.	2.3	51
152	Propensity vectors of low $\alpha$ -ASA residue pairs in the distinction of protein interactions. <i>Proteins: Structure, Function and Bioinformatics</i> , 2010, 78, 589-602.	1.5	22
153	Geometrically centered region: A $\alpha$ -ASA model of protein binding hot spots not excluding water molecules. <i>Proteins: Structure, Function and Bioinformatics</i> , 2010, 78, 3304-3316.	1.5	19
154	PATTERN SPACE MAINTENANCE FOR DATA UPDATES AND INTERACTIVE MINING*. <i>Computational Intelligence</i> , 2010, 26, 282-317.	2.1	4
155	Maximization of negative correlations in time-course gene expression data for enhancing understanding of molecular pathways. <i>Nucleic Acids Research</i> , 2010, 38, e1-e1.	6.5	31
156	Water Bioinformatics: An Association between Estrogen Degradation and 16S rRNA Motifs. <i>International Conference on Bioinformatics and Biomedical Engineering: [proceedings] International Conference on Bioinformatics and Biomedical Engineering</i> , 2010, , .	0.0	0
157	Negative correlations in collaboration. , 2010, , .		5
158	Modeling Protein Interacting Groups by Quasi-Bicliques: Complexity, Algorithm, and Application. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2010, 7, 354-364.	1.9	28
159	Insights into Bacterial Genome Composition through Variable Target GC Content Profiling. <i>Journal of Computational Biology</i> , 2010, 17, 79-96.	0.8	3
160	Birds Bring Flues? Mining Frequent and High Weighted Cliques from Birds Migration Networks. <i>Lecture Notes in Computer Science</i> , 2010, , 359-369.	1.0	4
161	High Functional Coherence in k-Partite Protein Cliques of Protein Interaction Networks. , 2009, , .		4
162	'Double water exclusion': a hypothesis refining the O-ring theory for the hot spots at protein interfaces. <i>Bioinformatics</i> , 2009, 25, 743-750.	1.8	56

#	ARTICLE	IF	CITATIONS
163	PADS: a simple yet effective pattern-aware dynamic search method for fast maximal frequent pattern mining. Knowledge and Information Systems, 2009, 20, 375-391.	2.1	19
164	Mining maximal quasi-bicliques: Novel algorithm and applications in the stock market and protein networks. Statistical Analysis and Data Mining, 2009, 2, 255-273.	1.4	29
165	Efficient mining of distance-based subspace clusters. Statistical Analysis and Data Mining, 2009, 2, 427-444.	1.4	14
166	Sequence-based B-cell epitope prediction by using associations in antibody-antigen structural complexes. , 2009, , .		1
167	Prediction of protein long-range contacts using GaMC approach with sequence profile centers. , 2009, , .		0
168	Emerging Patterns. , 2009, , 985-988.		1
169	Emerging Pattern Based Classification. , 2009, , 985-985.		1
170	Discovery of Migration Habitats and Routes of Wild Bird Species by Clustering and Association Analysis. Lecture Notes in Computer Science, 2009, , 288-301.	1.0	13
171	Maintenance of Frequent Patterns. , 2009, , 273-293.		1
172	Discovery of Protein Interaction Sites. , 2009, , 683-688.		0
173	Applications of Emerging Patterns for Microarray Gene Expression Data Analysis. , 2009, , 107-107.		0
174	Mining Conditional Contrast Patterns. , 2009, , 294-310.		4
175	Clustering orthologous proteins across phylogenetically distant species. Proteins: Structure, Function and Bioinformatics, 2008, 71, 1113-1122.	1.5	7
176	A new concise representation of frequent itemsets using generators and a positive border. Knowledge and Information Systems, 2008, 17, 35-56.	2.1	29
177	Mining and Ranking Generators of Sequential Patterns. , 2008, , .		33
178	Maximal Quasi-Bicliques with Balanced Noise Tolerance: Concepts and Co-clustering Applications. , 2008, , .		31
179	Quasi-bicliques: Complexity and Binding Pairs. Lecture Notes in Computer Science, 2008, , 255-264.	1.0	16
180	Assessing and Predicting Protein Interactions Using Both Local and Global Network Topological Metrics. , 2008, , .		31

#	ARTICLE	IF	CITATIONS
181	A pHMM-ANN based discriminative approach to promoter identification in prokaryote genomic contexts. <i>Nucleic Acids Research</i> , 2007, 35, e12-e12.	6.5	18
182	Mining statistically important equivalence classes and delta-discriminative emerging patterns. , 2007, , .		83
183	Distance Based Subspace Clustering with Flexible Dimension Partitioning. , 2007, , .		29
184	Maximal Biclique Subgraphs and Closed Pattern Pairs of the Adjacency Matrix: A One-to-One Correspondence and Mining Algorithms. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2007, 19, 1625-1637.	4.0	101
185	Evolution and Maintenance of Frequent Pattern Space When Transactions Are Removed. , 2007, , 489-497.		4
186	Hybrid Outcome Prediction Model for Severe Traumatic Brain Injury. <i>Journal of Neurotrauma</i> , 2007, 24, 136-146.	1.7	52
187	Strong Compound-Risk Factors: Efficient Discovery Through Emerging Patterns and Contrast Sets. <i>IEEE Transactions on Information Technology in Biomedicine</i> , 2007, 11, 544-552.	3.6	24
188	Multidimensional Decision Support Indicator (mDSI) for Time Series Stock Trend Prediction. , 2007, , 841-848.		2
189	INTERACTING AMINO ACID PREFERENCES OF 3D PATTERN PAIRS AT THE BINDING SITES OF TRANSIENT AND OBLIGATE PROTEIN COMPLEXES. , 2007, , .		4
190	Discovering Patterns of DNA Methylation: Rule Mining with Rough Sets and Decision Trees, and Comethylation Analysis. <i>Lecture Notes in Computer Science</i> , 2007, , 389-397.	1.0	1
191	MINING SUPER-SECONDARY STRUCTURE MOTIFS FROM 3D PROTEIN STRUCTURES: A SEQUENCE ORDER INDEPENDENT APPROACH. , 2007, , .		1
192	Mining Maximal Quasi-Bicliques to Co-Cluster Stocks and Financial Ratios for Value Investment. <i>IEEE International Conference on Data Mining</i> , 2006, , .	0.0	39
193	Positive Borders or Negative Borders: How to Make Lossless Generator Based Representations Concise. , 2006, , .		8
194	Discovering motif pairs at interaction sites from protein sequences on a proteome-wide scale. <i>Bioinformatics</i> , 2006, 22, 989-996.	1.8	54
195	Mining Temporal Indirect Associations. <i>Lecture Notes in Computer Science</i> , 2006, , 425-434.	1.0	14
196	Bayesian Approaches to Ranking Sequential Patterns Interestingness. <i>Lecture Notes in Computer Science</i> , 2006, , 241-250.	1.0	0
197	Using fixed point theorems to model the binding in protein-protein interactions. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2005, 17, 1079-1087.	4.0	3
198	Mining border descriptions of emerging patterns from dataset pairs. <i>Knowledge and Information Systems</i> , 2005, 8, 178-202.	2.1	106

#	ARTICLE	IF	CITATIONS
199	Structural geography of the space of emerging patterns. <i>Intelligent Data Analysis</i> , 2005, 9, 567-588.	0.4	5
200	Use of extreme patient samples for outcome prediction from gene expression data. <i>Bioinformatics</i> , 2005, 21, 3377-3384.	1.8	28
201	DNAFSMiner: a web-based software toolbox to recognize two types of functional sites in DNA sequences. <i>Bioinformatics</i> , 2005, 21, 671-673.	1.8	45
202	Discovery of stable and significant binding motif pairs from PDB complexes and protein interaction datasets. <i>Bioinformatics</i> , 2005, 21, 314-324.	1.8	15
203	Efficient selection of discriminative genes from microarray gene expression data for cancer diagnosis. <i>IEEE Transactions on Circuits and Systems Part 1: Regular Papers</i> , 2005, 52, 1909-1918.	0.1	16
204	Relative risk and odds ratio. , 2005, , .		45
205	Guest Editors' Introduction: Data Mining in Bioinformatics. <i>IEEE Intelligent Systems</i> , 2005, 20, 16-18.	4.0	74
206	Selection of patient samples and genes for outcome prediction. , 2004, , 382-92.		2
207	Twelve C2H2 zinc-finger genes on human chromosome 19 can be each translated into the same type of protein after frameshifts. <i>Bioinformatics</i> , 2004, 20, 1-4.	1.8	38
208	Incremental Maintenance on the Border of the Space of Emerging Patterns. <i>Data Mining and Knowledge Discovery</i> , 2004, 9, 89-116.	2.4	16
209	DeEPs: A New Instance-Based Lazy Discovery and Classification System. <i>Machine Learning</i> , 2004, 54, 99-124.	3.4	93
210	A Tree-Based Approach to the Discovery of Diagnostic Biomarkers for Ovarian Cancer. <i>Lecture Notes in Computer Science</i> , 2004, , 682-691.	1.0	3
211	Discovery of significant rules for classifying cancer diagnosis data. <i>Bioinformatics</i> , 2003, 19, ii93-ii102.	1.8	99
212	Simple rules underlying gene expression profiles of more than six subtypes of acute lymphoblastic leukemia (ALL) patients. <i>Bioinformatics</i> , 2003, 19, 71-78.	1.8	106
213	Using Rules to Analyse Bio-medical Data: A Comparison between C4.5 and PCL. <i>Lecture Notes in Computer Science</i> , 2003, , 254-265.	1.0	28
214	Bioinformatics Adventures in Database Research. <i>Lecture Notes in Computer Science</i> , 2003, , 31-46.	1.0	9
215	Identifying good diagnostic gene groups from gene expression profiles using the concept of emerging patterns. <i>Bioinformatics</i> , 2002, 18, 725-734.	1.8	139
216	Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. <i>Cancer Cell</i> , 2002, 1, 133-143.	7.7	1,756

#	ARTICLE	IF	CITATIONS
217	Geography of Differences between Two Classes of Data. Lecture Notes in Computer Science, 2002, , 325-337.	1.0	10
218	Making Use of the Most Expressive Jumping Emerging Patterns for Classification. Knowledge and Information Systems, 2001, 3, 131-145.	2.1	83
219	Combining the Strength of Pattern Frequency and Distance for Classification. Lecture Notes in Computer Science, 2001, , 455-466.	1.0	15
220	Making Use of the Most Expressive Jumping Emerging Patterns for Classification. Lecture Notes in Computer Science, 2000, , 220-232.	1.0	40
221	Emerging Patterns and Classification. Lecture Notes in Computer Science, 2000, , 15-32.	1.0	5
222	Efficient mining of emerging patterns. , 1999, , .		714
223	Efficient Mining of High Confidence Association Rules without Support Thresholds. Lecture Notes in Computer Science, 1999, , 406-411.	1.0	34
224	CAEP: Classification by Aggregating Emerging Patterns. Lecture Notes in Computer Science, 1999, , 30-42.	1.0	247
225	Interestingness of discovered association rules in terms of neighborhood-based unexpectedness. Lecture Notes in Computer Science, 1998, , 72-86.	1.0	52
226	Solving the fragmentation problem of decision trees by discovering boundary emerging patterns. , 0, , .		0
227	Diagnostic Rules Induced by an Ensemble Method for Childhood Leukemia. , 0, , .		0
228	Simultaneous compression of multiple error-corrected short-read sets for faster data transmission and better <i>de novo</i> assemblies. Briefings in Functional Genomics, 0, , .	1.3	0