Grigorios Tsoumakas

List of Publications by Year
in descending order

90 papers	6,254 citations	186265 28 h-index	98798 67 g-index
93 all docs	93 docs citations	93 times ranked	5688 citing authors

#	Article	IF	Citations
1	Multi-Label Classification. International Journal of Data Warehousing and Mining, 2007, 3, 1-13.	0.6	1,757
2	Mining Multi-label Data. , 2009, , 667-685.		652
3	Random k-Labelsets for Multilabel Classification. IEEE Transactions on Knowledge and Data Engineering, 2011, 23, 1079-1089.	5.7	637
4	Random k-Labelsets: An Ensemble Method for Multilabel Classification. Lecture Notes in Computer Science, 2007, , 406-417.	1.3	471
5	Multi-target regression via input space expansion: treating targets as inputs. Machine Learning, 2016, 104, 55-98.	5.4	232
6	On the Stratification of Multi-label Data. Lecture Notes in Computer Science, 2011, , 145-158.	1.3	197
7	Tracking recurring contexts using ensemble classifiers: an application to email filtering. Knowledge and Information Systems, 2010, 22, 371-391.	3.2	170
8	A Comprehensive Study Over VLAD and Product Quantization in Large-Scale Image Retrieval. IEEE Transactions on Multimedia, 2014, 16, 1713-1728.	7.2	103
9	An ensemble uncertainty aware measure forÂdirectedÂhillÂclimbing ensemble pruning. Machine Learning, 2010, 81, 257-282.	5.4	99
10	Multi-label classification of music by emotion. Eurasip Journal on Audio, Speech, and Music Processing, 2011, 2011, .	2.1	95
11	An Ensemble Pruning Primer. Studies in Computational Intelligence, 2009, , 1-13.	0.9	87
12	A review of keyphrase extraction. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2020, 10, e1339.	6.8	86
13	Selective fusion of heterogeneous classifiers. Intelligent Data Analysis, 2005, 9, 511-525.	0.9	77
14	Pruning an ensemble of classifiers via reinforcement learning. Neurocomputing, 2009, 72, 1900-1909.	5.9	77
15	A systematic review of multi-label feature selection and a new method based on label construction. Neurocomputing, 2016, 180, 3-15.	5.9	73
16	Clustering based multi-label classification for image annotation and retrieval. , 2009, , .		69
17	Dealing with class imbalance in classifier chains via random undersampling. Knowledge-Based Systems, 2020, 192, 105292.	7.1	66
18	A survey of machine learning techniques for food sales prediction. Artificial Intelligence Review, 2019, 52, 441-447.	15.7	60

#	Article	IF	Citations
19	An adaptive personalized news dissemination system. Journal of Intelligent Information Systems, 2009, 32, 191-212.	3.9	59
20	Local word vectors guiding keyphrase extraction. Information Processing and Management, 2018, 54, 888-902.	8.6	57
21	Multi-target Regression via Random Linear Target Combinations. Lecture Notes in Computer Science, 2014, , 225-240.	1.3	54
22	An empirical study on sea water quality prediction. Knowledge-Based Systems, 2008, 21, 471-478.	7.1	53
23	Greedy regression ensemble selection: Theory and an application to water quality prediction. Information Sciences, 2008, 178, 3867-3879.	6.9	51
24	The 9th annual MLSP competition: New methods for acoustic classification of multiple simultaneous bird species in a noisy environment. , 2013, , .		50
25	Clustering classifiers for knowledge discovery from physically distributed databases. Data and Knowledge Engineering, 2004, 49, 223-242.	3.4	49
26	Regression via Classification applied on software defect estimation. Expert Systems With Applications, 2008, 34, 2091-2101.	7.6	45
27	A Divide-and-Conquer Approach to the Summarization of Long Documents. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 3029-3040.	5.8	43
28	An Empirical Study of Multi-label Learning Methods for Video Annotation. , 2009, , .		37
29	Predicting Drug-Target Interactions With Multi-Label Classification and Label Partitioning. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2021, 18, 1596-1607.	3.0	34
30	Introduction to the special issue on learning from multi-label data. Machine Learning, 2012, 88, 1-4.	5.4	32
31	On the Combination of Textual and Semantic Descriptions for Automated Semantic Web Service Classification. IFIP Advances in Information and Communication Technology, 2009, , 95-104.	0.7	28
32	Dynamic ensemble pruning based on multi-label classification. Neurocomputing, 2015, 150, 501-512.	5.9	28
33	TRES: Identification of Discriminatory and Informative SNPs from Population Genomic Data: Figure 1.. Journal of Heredity, 2015, 106, 672-676.	2.4	26
34	Effective Voting of Heterogeneous Classifiers. Lecture Notes in Computer Science, 2004, , 465-476.	1.3	26
35	Obtaining Bipartitions from Score Vectors for Multi-Label Classification. , 2010, , .		24
36	Multi-label sampling based on local label imbalance. Pattern Recognition, 2022, 122, 108294.	8.1	23

#	Article	IF	Citations
37	Distributed Data Mining. , 2009, , 709-715.		19
38	ETHOS: a multi-label hate speech detection dataset. Complex & Intelligent Systems, 2022, 8, 4663-4678.	6.5	19
39	Ensemble Pruning Using Reinforcement Learning. Lecture Notes in Computer Science, 2006, , 301-310.	1.3	18
40	Transferring task models in Reinforcement Learning agents. Neurocomputing, 2013, 107, 23-32.	5.9	17
41	Word embeddings and external resources for answer processing in biomedical factoid question answering. Journal of Biomedical Informatics, 2019, 92, 103118.	4.3	17
42	Improving Distantly-Supervised Relation Extraction Through BERT-Based Label and Instance Embeddings. IEEE Access, 2021, 9, 62574-62582.	4.2	17
43	A Triple-Random Ensemble Classification Method for Mining Multi-label Data. , 2010, , .		14
44	Large-scale online semantic indexing of biomedical articles via an ensemble of multi-label classification models. Journal of Biomedical Semantics, 2017, 8, 43.	1.6	14
45	Instance-Based Ensemble Pruning via Multi-Label Classification. , 2010, , .		13
46	Hierarchical partitioning of the output space in multi-label data. Data and Knowledge Engineering, 2018, 116, 42-60.	3.4	13
47	Discovering and Exploiting Deterministic Label Relationships in Multi-Label Learning. , 2015, , .		13
48	Label Construction for Multi-label Feature Selection. , 2014, , .		12
49	A prediction model of passenger demand using AVL and APC data from a bus fleet. , 2015, , .		12
50	Drug-target interaction prediction via an ensemble of weighted nearest neighbors with interaction recovery. Applied Intelligence, 2022, 52, 3705-3727.	5.3	11
51	Using multi-target feature evaluation to discover factors that affect business process behavior. Computers in Industry, 2018, 99, 253-261.	9.9	10
52	Multi-label active learning: key issues and a novel query strategy. Evolving Systems, 2019, 10, 63-78.	3.9	10
53	Beyond MeSH: Fine-grained semantic indexing of biomedical literature based on weak supervision. Information Processing and Management, 2020, 57, 102282.	8.6	10
54	Learning-to-Rank and Relevance Feedback for Literature Appraisal in Empirical Medicine. Lecture Notes in Computer Science, 2018, , 52-63.	1.3	9

#	Article	IF	Citations
55	PASER: a curricula synthesis system based on automated problem solving. International Journal of Teaching and Case Studies, 2007, 1, 159.	0.1	8
56	Web Robot Detection: A Semantic Approach. , 2018, , .		8
57	WISE 2014 Challenge: Multi-label Classification of Print Media Articles to Topics. Lecture Notes in Computer Science, 2014, , 541-548.	1.3	8
58	Short-Term Renewable Energy Forecasting in Greece Using Prophet Decomposition and Tree-Based Ensembles. Communications in Computer and Information Science, 2021, , 227-238.	0.5	7
59	Synthetic Oversampling of Multi-label Data Based on Local Label Distribution. Lecture Notes in Computer Science, 2020, , 180-193.	1.3	7
60	LioNets: Local Interpretation of Neural Networks Through Penultimate Layer Decoding. Communications in Computer and Information Science, 2020, , 265-276.	0.5	7
61	An interoperable and scalable Web-based system for classifier sharing and fusion. Expert Systems With Applications, 2007, 33, 716-724.	7.6	6
62	Zero-Shot Classification of Biomedical Articles with Emerging MeSH Descriptors. , 2020, , .		6
63	Content-aware web robot detection. Applied Intelligence, 2020, 50, 4017-4028.	5.3	5
64	Active Learning Algorithms for Multi-label Data. IFIP Advances in Information and Communication Technology, 2016, , 267-279.	0.7	4
65	What is all this new MeSH about?. International Journal on Digital Libraries, 2021, 22, 319-337.	1.5	4
66	Extracting Semantic Relationships in Greek Literary Texts. Sustainability, 2021, 13, 9391.	3.2	4
67	Subset Labeled LDA: A Topic Model for Extreme Multi-label Classification. Lecture Notes in Computer Science, 2018, , 152-162.	1.3	4
68	E-mail Mining. , 2007, , 220-243.		4
69	Hatebusters: A Web Application for Actively Reporting YouTube Hate Speech. , 2018, , .		4
70	Multi-label Modality Classification for Figures in Biomedical Literature. , 2017, , .		3
71	A Knowledge-Based Web Information System for the Fusion of Distributed Classifers. , 2004, , 268-304.		3
72	LioNets: a neural-specific local interpretation technique exploiting penultimate layer information. Applied Intelligence, 2023, 53, 2538-2563.	5.3	3

#	Article	IF	Citations
73	Conclusive local interpretation rules for random forests. Data Mining and Knowledge Discovery, 0, , .	3.7	3
74	Classifying Biomedical Figures by Modality via Multi-Label Learning. IEEE Journal of Biomedical and Health Informatics, 2019, 23, 2230-2237.	6.3	2
75	A Multi-instance Multi-label Weakly Supervised Approach for Dealing with Emerging MeSH Descriptors. Lecture Notes in Computer Science, 2021, , 397-407.	1.3	2
76	Transferring Models in Hybrid Reinforcement Learning Agents. International Federation for Information Processing, 2011, , 162-171.	0.4	2
77	Machine Learning for Adaptive Planning. , 2005, , 90-120.		2
78	Ensemble Feature Selection using Rank Aggregation Methods for Population Genomic Data. , 2016, , .		1
79	Integrating multiple immunogenetic data sources for feature extraction and mining somatic hypermutation patterns: the case of â€œtowards analysisâ€•in chronic lymphocytic leukaemia. BMC Bioinformatics, 2016, 17, 173.	2.6	1
80	An Empirical Comparison of Methods for Multi-label Data Stream Classification. Advances in Intelligent Systems and Computing, 2017, , 151-159.	0.6	1
81	Instance-Based Zero-Shot learning for semi-Automatic MeSH indexing. Pattern Recognition Letters, 2021, 151, 62-68.	4.2	1
82	Using the k-Nearest Problems for Adaptive Multicriteria Planning. Lecture Notes in Computer Science, 2004, , 132-141.	1.3	1
83	Multi-label Learning Approaches for Music Instrument Recognition. Lecture Notes in Computer Science, 2011, , 734-743.	1.3	1
84	Yes/No Question Answering in BioASQ 2019. Communications in Computer and Information Science, 2020, , 661-669.	0.5	1
85	Segmento. , 2016, , .		0
86	Machine Learning Methods for Customer's Payment Acceptance Prediction in an Electricity Distribution Company. , 2017, , .		0
87	Semantic Indexing of 19th-Century Greek Literature Using 21st-Century Linguistic Resources. Sustainability, 2021, 13, 8878.	3.2	0
88	Land Evaluation - An Artificial Intelligence Approach. , 2001, , 158-166.		0
89	Feature Evaluation Metrics for Population Genomic Data. Lecture Notes in Computer Science, 2014, , 436-441.	1.3	0
90	Artificial fine-tuning tasks for yes/no question answering. Natural Language Engineering, 2024, 30, 73-95.	2.5	0