Eibe Frank

List of Publications by Year
in descending order

85 papers	28,763 citations	117571 34 h-index	102432 66 g-index
90 all docs	90 docs citations	90 times ranked	24649 citing authors

#	Article	IF	Citations
1	The WEKA data mining software. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 2009, 11, 10-18.	3.2	14,483
2	Classifier chains for multi-label classification. Machine Learning, 2011, 85, 333-359.	3.4	1,483
3	Logistic Model Trees. Machine Learning, 2005, 59, 161-205.	3.4	981
4	Data mining in bioinformatics using Weka. Bioinformatics, 2004, 20, 2479-2481.	1.8	793
5	Gene selection from microarray data for cancer classificationâ€"a machine learning approach. Computational Biology and Chemistry, 2005, 29, 37-46.	1.1	336
6	Classifier Chains for Multi-label Classification. Lecture Notes in Computer Science, 2009, , 254-269.	1.0	321
7	Using Model Trees for Classification. Machine Learning, 1998, 32, 63-76.	3.4	316
8	A Simple Approach to Ordinal Classification. Lecture Notes in Computer Science, 2001, , 145-156.	1.0	291
9	Sentiment Knowledge Discovery in Twitter Streaming Data. Lecture Notes in Computer Science, 2010, , 1-15.	1.0	288
10	A review of multi-instance learning assumptions. Knowledge Engineering Review, 2010, 25, 1-25.	2.1	270
11	Evaluating the Replicability of Significance Tests for Comparing Learning Algorithms. Lecture Notes in Computer Science, 2004, , 3-12.	1.0	214
12	Multinomial Naive Bayes for Text Categorization Revisited. Lecture Notes in Computer Science, 2004, , 488-499.	1.0	205
13	Weka-A Machine Learning Workbench for Data Mining. , 2009, , 1269-1277.		189
14	Human-competitive tagging using automatic keyphrase extraction. , 2009, , .		179
15	Technical Note: Naive Bayes for Regression. Machine Learning, 2000, 41, 5-25.	3.4	164
16	Accelerating the XGBoost algorithm using GPU computing. PeerJ Computer Science, 0, 3, e127.	2.7	162
17	Improving browsing in digital libraries with keyphrase indexes. Decision Support Systems, 1999, 27, 81-104.	3.5	157
18	Large-scale attribute selection using wrappers. , 2009, , .		155

#	Article	IF	Citations
19	Logistic Regression and Boosting for Labeled Bags of Instances. Lecture Notes in Computer Science, 2004, , 272-281.	1.0	139
20	Interactive machine learning: letting users build classifiers. International Journal of Human Computer Studies, 2001, 55, 281-292.	3.7	137
21	Introducing Machine Learning Concepts with WEKA. Methods in Molecular Biology, 2016, 1418, 353-378.	0.4	131
22	Multiclass Alternating Decision Trees. Lecture Notes in Computer Science, 2002, , 161-172.	1.0	109
23	Naive Bayes for Text Classification with Unbalanced Classes. Lecture Notes in Computer Science, 2006, , 503-510.	1.0	105
24	Weka. , 2005, , 1305-1314.		101
25	A Two-Level Learning Method for Generalized Multi-instance Problems. Lecture Notes in Computer Science, 2003, , 468-479.	1.0	93
26	Clustering Documents with Active Learning Using Wikipedia. , 2008, , .		89
27	One-Class Classification by Combining Density and Class Probability Estimation. Lecture Notes in Computer Science, 2008, , 505-519.	1.0	88
28	WekaDeeplearning4j: A deep learning package for Weka based on Deeplearning4j. Knowledge-Based Systems, 2019, 178, 48-50.	4.0	82
29	Deep learning in diabetic foot ulcers detection: A comprehensive evaluation. Computers in Biology and Medicine, 2021, 135, 104596.	3.9	75
30	Logistic Model Trees. Lecture Notes in Computer Science, 2003, , 241-252.	1.0	72
31	Fast Perceptron Decision Tree Learning from Evolving Data Streams. Lecture Notes in Computer Science, 2010, , 299-310.	1.0	61
32	DNA methylation-associated colonic mucosal immune and defense responses in treatment-naÃ¯ve pediatric ulcerative colitis. Epigenetics, 2014, 9, 1131-1137.	1.3	59
33	Clustering Documents Using a Wikipedia-Based Concept Representation. Lecture Notes in Computer Science, 2009, , 628-636.	1.0	58
34	Ensembles of nested dichotomies for multi-class problems. , 2004, , .		54
35	Learning a conceptâ€based document similarity measure. Journal of the Association for Information Science and Technology, 2012, 63, 1593-1608.	2.6	53
36	Building a Twitter opinion lexicon from automatically-annotated tweets. Knowledge-Based Systems, 2016, 108, 65-78.	4.0	51

#	Article	IF	Citations
37	Predicting Library of Congress classifications from Library of Congress subject headings. Journal of the Association for Information Science and Technology, 2004, 55, 214-227.	2.6	48
38	Accurate photometric redshift probability density estimation â€“ method comparison and application. Monthly Notices of the Royal Astronomical Society, 2015, 452, 3710-3725.	1.6	45
39	Determining Word-Emotion Associations from Tweets by Multi-label Classification. , 2016, , .		37
40	Ensembles of Balanced Nested Dichotomies for Multi-class Problems. Lecture Notes in Computer Science, 2005, , 84-95.	1.0	36
41	Accuracy of machine learning models versus â€œhand craftedâ€•expert systems â€“ A credit scoring case study. Expert Systems With Applications, 2009, 36, 5264-5271.	4.4	35
42	A Study of Hierarchical and Flat Classification of Proteins. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2010, 7, 563-571.	1.9	32
43	Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological characteristics of dairy cows to detect lameness. Journal of Dairy Science, 2013, 96, 7043-7053.	1.4	29
44	An Empirical Comparison of Exact Nearest Neighbour Algorithms. Lecture Notes in Computer Science, 2007, , 140-151.	1.0	28
45	Alternating model trees. , 2015, , .		26
46	Using Classification to Evaluate the Output of Confidence-Based Association Rule Mining. Lecture Notes in Computer Science, 2004, , 538-549.	1.0	24
47	Adaptive XGBoost for Evolving Data Streams. , 2020, , .		24
48	Fragment generation and support vector machines for inducing SARs. SAR and QSAR in Environmental Research, 2002, 13, 509-523.	1.0	23
49	Ensembles of Restricted Hoeffding Trees. ACM Transactions on Intelligent Systems and Technology, 2012, 3, 1-20.	2.9	23
50	Artificial neural network is highly predictive of outcome in paediatric acute liver failure. Pediatric Transplantation, 2013, 17, 535-542.	0.5	23
51	Conditional Density Estimation with Class Probability Estimators. Lecture Notes in Computer Science, 2009, , 65-81.	1.0	23
52	Unsupervised Discretization Using Tree-Based Density Estimation. Lecture Notes in Computer Science, 2005, , 240-251.	1.0	19
53	Discriminating Against New Classes: One-class versus Multi-class Classification. Lecture Notes in Computer Science, 2008, , 325-336.	1.0	17
54	Improving on Bagging with Input Smearing. Lecture Notes in Computer Science, 2006, , 97-106.	1.0	15

#	Article	IF	Citations
55	Propositionalisation of Multi-instance Data Using Random Forests. Lecture Notes in Computer Science, 2013, , 362-373.	1.0	14
56	Beyond Trees: Adopting MITI to Learn Rules and Ensemble Classifiers for Multi-Instance Data. Lecture Notes in Computer Science, 2011, , 41-50.	1.0	14
57	Racing Committees for Large Datasets. Lecture Notes in Computer Science, 2002, , 153-164.	1.0	12
58	A Toolbox for Learning from Relational Data with Propositional and Multi-instance Learners. Lecture Notes in Computer Science, 2004, , 1017-1023.	1.0	12
59	Revisiting Multiple-Instance Learning Via Embedded Instance Selection. Lecture Notes in Computer Science, 2008, , 300-310.	1.0	11
60	From Opinion Lexicons to Sentiment Classification of Tweets and Vice Versa: A Transfer Learning Approach. , 2016, , .		10
61	Building Ensembles of Adaptive Nested Dichotomies with Random-Pair Selection. Lecture Notes in Computer Science, 2016, , 179-194.	1.0	10
62	Speeding Up and Boosting Diverse Density Learning. Lecture Notes in Computer Science, 2010, , 102-116.	1.0	10
63	From Unlabelled Tweets to Twitter-specific Opinion Words. , 2015, , .		9
64	Visualizing Class Probability Estimators. Lecture Notes in Computer Science, 2003, , 168-179.	1.0	9
65	Online Estimation of Discrete Densities. , 2013, , .		8
66	Improving Naive Bayes for Regression with Optimized Artificial Surrogate Data. Applied Artificial Intelligence, 2020, 34, 484-514.	2.0	8
67	The Applicability of Ambient Sensors as Proximity Evidence for NFC Transactions. , 2017, , .		7
68	A comparison of methods for estimating prediction intervals in NIR spectroscopy: Size matters. Chemometrics and Intelligent Laboratory Systems, 2011, 109, 139-145.	1.8	6
69	On the Effectiveness of Ambient Sensing for Detecting NFC Relay Attacks. , 2017, , .		5
70	Determining Progression in Glaucoma Using Visual Fields. Lecture Notes in Computer Science, 2001, , 136-147.	1.0	5
71	Large-Scale Automatic Species Identification. Lecture Notes in Computer Science, 2017, , 301-312.	1.0	5
72	Methods for Eliciting Informative Prior Distributions: A Critical Review. Decision Analysis, 2022, 19, 189-204.	1.2	5

#	Article	IF	Citations
73	Transferring sentiment knowledge between words and tweets. Web Intelligence, 2018, 16, 203-220.	0.1	4
74	Ensembles of Nested Dichotomies with Multiple Subset Evaluation. Lecture Notes in Computer Science, 2019, , 81-93.	1.0	4
75	Learning Through Utility Optimization in Regression Tasks. , 2017, , .		3
76	Online estimation of discrete, continuous, and conditional joint densities using classifier chains. Data Mining and Knowledge Discovery, 2018, 32, 561-603.	2.4	3
77	Hidden Features: Experiments with Feature Transfer for Fine-Grained Multi-Class and One-Class Image Categorization. , 2018, , .		3
78	On Calibration of Nested Dichotomies. Lecture Notes in Computer Science, 2019, , 69-80.	1.0	3
79	Additive Regression Applied to a Large-Scale Collaborative Filtering Problem. Lecture Notes in Computer Science, 2008, , 435-446.	1.0	3
80	A data mining approach to evaluate suitability of dissolved oxygen sensor observations for lake metabolism analysis. Limnology and Oceanography: Methods, 2018, 16, 787-801.	1.0	2
81	Good Vibrations: Artificial Ambience-Based Relay Attack Detection. , 2018, , .		2
82	Experiments in cross-domain few-shot learning for image classification. Journal of the Royal Society of New Zealand, 2023, 53, 169-191.	1.0	2
83	Efficiently correcting machine learning: considering the role of example ordering in human-in-the-loop training of image classification models. , 2022, , .		1
84	Analysing chromatographic data using data mining to monitor petroleum content in water. Environmental Science and Engineering, 2009, , 278-290.	0.1	0
85	Bandwidth-Optimal Random Shuffling for GPUs. ACM Transactions on Parallel Computing, 2022, 9, 1-20.	1.2	0