

# Tapio Salakoski

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/11911378/publications.pdf>

Version: 2024-02-01

67  
papers

3,070  
citations

361413

20  
h-index

182427

51  
g-index

70  
all docs

70  
docs citations

70  
times ranked

3679  
citing authors

#	ARTICLE	IF	CITATIONS
1	A large-scale evaluation of computational protein function prediction. <i>Nature Methods</i> , 2013, 10, 221-227.	19.0	789
2	An expanded evaluation of protein function prediction methods shows an improvement in accuracy. <i>Genome Biology</i> , 2016, 17, 184.	8.8	308
3	BioInfer: a corpus for information extraction in the biomedical domain. <i>BMC Bioinformatics</i> , 2007, 8, 50.	2.6	288
4	The CAFA challenge reports improved protein function prediction and new functional annotations for hundreds of genes through experimental screens. <i>Genome Biology</i> , 2019, 20, 244.	8.8	261
5	All-paths graph kernel for protein-protein interaction extraction with evaluation of cross-corpus learning. <i>BMC Bioinformatics</i> , 2008, 9, S2.	2.6	193
6	Regularized Machine Learning in the Genetic Prediction of Complex Traits. <i>PLoS Genetics</i> , 2014, 10, e1004754.	3.5	122
7	An experimental comparison of cross-validation techniques for estimating the area under the ROC curve. <i>Computational Statistics and Data Analysis</i> , 2011, 55, 1828-1844.	1.2	116
8	What about a simple language? Analyzing the difficulties in learning to program. <i>Computer Science Education</i> , 2006, 16, 211-227.	3.7	88
9	Large-Scale Event Extraction from Literature with Multi-Level Gene Normalization. <i>PLoS ONE</i> , 2013, 8, e55814.	2.5	83
10	Artificial intelligence in nursing: Priorities and opportunities from an international invitational thinkâ€tank of the Nursing and Artificial Intelligence Leadership Collaborative. <i>Journal of Advanced Nursing</i> , 2021, 77, 3707-3717.	3.3	67
11	Selection of a representative set of structures from brookhaven protein data bank. <i>Proteins: Structure, Function and Bioinformatics</i> , 1992, 14, 265-276.	2.6	55
12	University of Turku in the BioNLP'11 Shared Task. <i>BMC Bioinformatics</i> , 2012, 13, S4.	2.6	49
13	Comparison of automatic summarisation methods for clinical free text notes. <i>Artificial Intelligence in Medicine</i> , 2016, 67, 25-37.	6.5	42
14	EXTRACTING CONTEXTUALIZED COMPLEX BIOLOGICAL EVENTS WITH RICH GRAPH-BASED FEATURE SETS. <i>Computational Intelligence</i> , 2011, 27, 541-557.	3.2	38
15	Exploring Biomolecular Literature with EVEX: Connecting Genes through Events, Homology, and Indirect Associations. <i>Advances in Bioinformatics</i> , 2012, 2012, 1-12.	5.7	35
16	Lexical adaptation of link grammar to the biomedical sublanguage: a comparative evaluation of three approaches. <i>BMC Bioinformatics</i> , 2006, 7, S2.	2.6	34
17	Predicting patient acuity from electronic patient records. <i>Journal of Biomedical Informatics</i> , 2014, 51, 35-40.	4.3	31
18	Learning intransitive reciprocal relations with kernel methods. <i>European Journal of Operational Research</i> , 2010, 206, 676-685.	5.7	28

#	ARTICLE	IF	CITATIONS
19	Wrapper-based selection of genetic features in genome-wide association studies through fast matrix operations. <i>Algorithms for Molecular Biology</i> , 2012, 7, 11.	1.2	28
20	On study habits on an introductory course on programming. <i>Computer Science Education</i> , 2015, 25, 276-291.	3.7	26
21	General formulation and evaluation of agglomerative clustering methods with metric and non-metric distances. <i>Pattern Recognition</i> , 1993, 26, 1395-1406.	8.1	23
22	Multi-Perspective Study of Novice Learners Adopting the Visual Algorithm Simulation Exercise System TRAKLA2. <i>Informatics in Education</i> , 2005, 4, 49-68.	2.2	22
23	Evaluation of two dependency parsers on biomedical corpus targeted at protein-protein interactions. <i>International Journal of Medical Informatics</i> , 2006, 75, 430-442.	3.3	21
24	Multi-label learning under feature extraction budgets. <i>Pattern Recognition Letters</i> , 2014, 40, 56-65.	4.2	20
25	Combining hidden Markov models and latent semantic analysis for topic segmentation and labeling: Method and clinical application. <i>International Journal of Medical Informatics</i> , 2009, 78, e1-e6.	3.3	19
26	Programming Misconceptions in an Introductory Level Programming Course Exam. , 2016, , .		19
27	Training linear ranking SVMs in linearithmic time using red-black trees. <i>Pattern Recognition Letters</i> , 2011, 32, 1328-1336.	4.2	18
28	Accurate prediction of protein secondary structural class with fuzzy structural vectors. <i>Protein Engineering, Design and Selection</i> , 1995, 8, 505-512.	2.1	16
29	A Kernel-Based Framework for Learning Graded Relations From Data. <i>IEEE Transactions on Fuzzy Systems</i> , 2012, 20, 1090-1101.	9.8	16
30	Wide-scope biomedical named entity recognition and normalization with CRFs, fuzzy matching and character level modeling. <i>Database: the Journal of Biological Databases and Curation</i> , 2018, 2018, 1-10.	3.0	14
31	Analysis of link grammar on biomedical dependency corpus targeted at protein-protein interactions. , 2004, , .		14
32	Contextual weighting for Support Vector Machines in literature mining: an application to gene versus protein name disambiguation. <i>BMC Bioinformatics</i> , 2005, 6, 157.	2.6	13
33	Speeding Up Greedy Forward Selection for Regularized Least-Squares. , 2010, , .		13
34	Applying language technology to nursing documents: Pros and cons with a focus on ethics. <i>International Journal of Medical Informatics</i> , 2007, 76, S293-S301.	3.3	11
35	Neural Network and Random Forest Models in Protein Function Prediction. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2022, 19, 1772-1781.	3.0	11
36	Representative selection of proteins based on nuclear families. <i>Protein Engineering, Design and Selection</i> , 1995, 8, 501-503.	2.1	10

#	ARTICLE	IF	CITATIONS
37	U-Compare bio-event meta-service: compatible BioNLP event extraction services. BMC Bioinformatics, 2011, 12, 481.	2.6	10
38	Conditional Ranking on Relational Data. Lecture Notes in Computer Science, 2010, , 499-514.	1.3	10
39	Matrix representations, linear transformations, and kernels for disambiguation in natural language. Machine Learning, 2009, 74, 133-158.	5.4	9
40	Define and Visualize Your First Programming Language. , 2008, , .		8
41	Towards automated processing of clinical Finnish: Sublanguage analysis and a rule-based parser. International Journal of Medical Informatics, 2009, 78, e7-e12.	3.3	8
42	Comparing the collaborative and independent viewing of program visualizations. , 2011, , .		8
43	Supporting the use of standardized nursing terminologies with automatic subject heading prediction: a comparison of sentence-level text classification methods. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 81-88.	4.4	8
44	Supporting Communication and Decision Making in Finnish Intensive Care with Language Technology. Journal of Healthcare Engineering, 2010, 1, 595-613.	1.9	7
45	Application of the EVEX resource to event extraction and network construction: Shared Task entry and result analysis. BMC Bioinformatics, 2015, 16, S3.	2.6	7
46	Assisting nurses in care documentation: from automated sentence classification to coherent document structures with subject headings. Journal of Biomedical Semantics, 2020, 11, 10.	1.6	7
47	How does collaboration affect algorithm learning? A case study using TRAKLA2 algorithm visualization tool. , 2010, , .		6
48	On Learning and Cross-Validation with Decomposed Nyström Approximation of Kernel Matrix. Neural Processing Letters, 2011, 33, 17-30.	3.2	4
49	Locality-Convolution Kernel and Its Application to Dependency Parse Ranking. Lecture Notes in Computer Science, 2006, , 610-618.	1.3	4
50	Assessment of metal ion concentration in water with structured feature selection. Chemosphere, 2017, 185, 1063-1071.	8.2	3
51	Efficient Hold-Out for Subset of Regressors. Lecture Notes in Computer Science, 2009, , 350-359.	1.3	3
52	An Improved Training Algorithm for the Linear Ranking Support Vector Machine. Lecture Notes in Computer Science, 2011, , 134-141.	1.3	3
53	Feature selection for regularized least-squares: New computational short-cuts and fast algorithmic implementations. , 2010, , .		2
54	Representative noise-free complete-link classification with application to protein structures. Pattern Recognition, 1997, 30, 467-482.	8.1	1

#	ARTICLE	IF	CITATIONS
55	Locality kernels for sequential data and their applications to parse ranking. Applied Intelligence, 2009, 31, 81-88.	5.3	1
56	Applying Permutation Tests for Assessing the Statistical Significance of Wrapper Based Feature Selection. , 2010, , .		1
57	Fast and parallelized greedy forward selection of genetic variants in Genome-wide association studies. , 2011, , .		1
58	Secondary use of electronic health records: Availability aspects in two Nordic countries. Health Information Management Journal, 2019, 48, 144-151.	1.2	1
59	Incorporating External Information in Bayesian Classifiers Via Linear Feature Transformations. Lecture Notes in Computer Science, 2006, , 399-410.	1.3	1
60	Text Classification Model Explainability for Keyword Extraction â€“ Towards Keyword-Based Summarization of Nursing Care Episodes. Studies in Health Technology and Informatics, 2022, , .	0.3	1
61	Evaluation of Protein Hydropathy Scales. , 2007, , .		0
62	Greedy Regularized Least-Squares for Multi-task Learning. , 2011, , .		0
63	Learning Valued Relations from Data. Advances in Intelligent and Soft Computing, 2011, , 257-268.	0.2	0
64	Avoiding Hazards â€“ What Can Health Care Learn from Aviation?. Communications in Computer and Information Science, 2012, , 119-127.	0.5	0
65	Properties of Object-Level Cross-Validation Schemes for Symmetric Pair-Input Data. Lecture Notes in Computer Science, 2014, , 384-393.	1.3	0
66	Learning Low Cost Multi-target Models by Enforcing Sparsity. Lecture Notes in Computer Science, 2015, , 252-261.	1.3	0
67	Clustering Nursing Sentences â€“ Comparing Three Sentence Embedding Methods. Studies in Health Technology and Informatics, 2022, , .	0.3	0