Julian Luengo

List of Publications by Year
in descending order

88 papers	**6,662** citations	186209 **28** h-index	114418 **63** g-index
92 all docs	**92** docs citations	**92** times ranked	**5555** citing authors

#	Article	IF	Citations
1	Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. Information Sciences, 2010, 180, 2044-2064.	4.0	1,627
2	A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability. Soft Computing, 2009, 13, 959-977.	2.1	563
3	Data Preprocessing in Data Mining. Intelligent Systems Reference Library, 2015, , .	1.0	541
4	SMOTE–IPF: Addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering. Information Sciences, 2015, 291, 184-203.	4.0	413
5	A Survey of Discretization Techniques: Taxonomy and Empirical Analysis in Supervised Learning. IEEE Transactions on Knowledge and Data Engineering, 2013, 25, 734-750.	4.0	389
6	Big data preprocessing: methods and prospects. Big Data Analytics, 2016, 1, .	2.2	319
7	COVIDGR Dataset and COVID-SDNet Methodology for Predicting COVID-19 Based on Chest X-Ray Images. IEEE Journal of Biomedical and Health Informatics, 2020, 24, 3595-3605.	3.9	252
8	Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. Knowledge-Based Systems, 2016, 98, 1-29.	4.0	204
9	KEEL 3.0: An Open Source Software for Multi-Stage Analysis in Data Mining. International Journal of Computational Intelligence Systems, 2017, 10, 1238.	1.6	201
10	On the choice of the best imputation methods for missing values considering three groups of classification methods. Knowledge and Information Systems, 2012, 32, 77-108.	2.1	185
11	Addressing data complexity for imbalanced data sets: analysis of SMOTE-based oversampling and evolutionary undersampling. Soft Computing, 2011, 15, 1909-1936.	2.1	144
12	Genetics-Based Machine Learning for Rule Induction: State of the Art, Taxonomy, and Comparative Study. IEEE Transactions on Evolutionary Computation, 2010, 14, 913-941.	7.5	137
13	A study on the use of statistical tests for experimentation with neural networks: Analysis of parametric test conditions and non-parametric tests. Expert Systems With Applications, 2009, 36, 7798-7808.	4.4	127
14	Analyzing the presence of noise in multi-class problems: alleviating its influence with the One-vs-One decomposition. Knowledge and Information Systems, 2014, 38, 179-206.	2.1	105
15	Enabling Smart Data: Noise filtering in Big Data classification. Information Sciences, 2019, 479, 135-152.	4.0	103
16	Predicting noise filtering efficacy with data complexity measures for nearest neighbor classification. Pattern Recognition, 2013, 46, 355-364.	5.1	92
17	Transforming big data into smart data: An insight on the use of the k–nearest neighbors algorithm to obtain quality data. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2019, 9, e1289.	4.6	88
18	Towards highly accurate coral texture images classification using deep convolutional neural networks and data augmentation. Expert Systems With Applications, 2019, 118, 315-328.	4.4	85

#	ARTICLE	IF	CITATIONS
19	A study on the use of imputation methods for experimentation with Radial Basis Function Network classifiers handling missing attribute values: The good synergy between RBFNs and EventCovering method. Neural Networks, 2010, 23, 406-418.	3.3	81
20	On the characterization of noise filters for self-training semi-supervised in nearest neighbor classification. Neurocomputing, 2014, 132, 30-41.	3.5	81
21	Tackling the problem of classification with noisy data using Multiple Classifier Systems: Analysis of the performance and robustness. Information Sciences, 2013, 247, 1-20.	4.0	77
22	INFFC: An iterative class noise filter based on the fusion of classifiers with noise sensitivity control. Information Fusion, 2016, 27, 19-32.	11.7	73
23	Evaluating the classifier behavior with noisy data considering performance and robustness: The Equalized Loss of Accuracy measure. Neurocomputing, 2016, 176, 26-35.	3.5	60
24	An automatic extraction method of the domains of competence for learning classifiers using data complexity measures. Knowledge and Information Systems, 2015, 42, 147-180.	2.1	47
25	Domains of competence of fuzzy rule based classification systems with data complexity measures: A case of study using a fuzzy hybrid genetic based machine learning method. Fuzzy Sets and Systems, 2010, 161, 3-19.	1.6	45
26	Big Data Preprocessing. , 2020, , .		42
27	CNC-NOS: Class noise cleaning by ensemble filtering and noise scoring. Knowledge-Based Systems, 2018, 140, 27-49.	4.0	33
28	Missing data imputation for fuzzy rule-based classification systems. Soft Computing, 2012, 16, 863-881.	2.1	31
29	Fast and Scalable Approaches to Accelerate the Fuzzy k-Nearest Neighbors Classifier for Big Data. IEEE Transactions on Fuzzy Systems, 2020, 28, 874-886.	6.5	31
30	Statistical computation of feature weighting schemes through data estimation for nearest neighbor classifiers. Pattern Recognition, 2014, 47, 3941-3948.	5.1	28
31	Using the One-vs-One decomposition to improve the performance of class noise filters via an aggregation strategy in multi-class classification problems. Knowledge-Based Systems, 2015, 90, 153-164.	4.0	26
32	Coral species identification with texture or structure images using a two-level classifier based on Convolutional Neural Networks. Knowledge-Based Systems, 2019, 184, 104891.	4.0	26
33	Dealing with Missing Values. Intelligent Systems Reference Library, 2015, , 59-105.	1.0	26
34	A Study on the Noise Label Influence inÂBoosting Algorithms: AdaBoost, GBM and XGBoost. Lecture Notes in Computer Science, 2017, , 268-280.	1.0	26
35	A tutorial on the segmentation of metallographic images: Taxonomy, new MetalDAM dataset, deep learning-based ensemble model, experimental analysis and challenges. Information Fusion, 2022, 78, 232-253.	11.7	24
36	Shared domains of competence of approximate learning models using measures of separability of classes. Information Sciences, 2012, 185, 43-65.	4.0	22

#	Article	IF	Citations
37	Evolutionary selection of hyperrectangles in nested generalized exemplar learning. Applied Soft Computing Journal, 2011, 11, 3032-3045.	4.1	19
38	Emerging topics and challenges of learning from noisy data in nonstandard classification: a survey beyond binary class noise. Knowledge and Information Systems, 2019, 60, 63-97.	2.1	18
39	Managing Borderline and Noisy Examples in Imbalanced Classification by Combining SMOTE with Ensemble Filtering. Lecture Notes in Computer Science, 2014, , 61-68.	1.0	17
40	Preprocessing methodology for time series: An industrial world application case study. Information Sciences, 2020, 514, 385-401.	4.0	16
41	Feature Selection. Intelligent Systems Reference Library, 2015, , 163-193.	1.0	14
42	Exact fuzzy k-nearest neighbor classification for big datasets. , 2017, , .		13
43	Smartdata: Data preprocessing to achieve smart data in R. Neurocomputing, 2019, 360, 1-13.	3.5	13
44	From Big to Smart Data: Iterative ensemble filter for noise filtering in Big Data classification. International Journal of Intelligent Systems, 2019, 34, 3260-3274.	3.3	12
45	Fuzzy Rule Based Classification Systems versus crisp robust learners trained in presence of class noise's effects: A case of study. , 2011, , .		11
46	From Big Data to Smart Data with the K-Nearest Neighbours Algorithm. , 2016, , .		11
47	The NoiseFiltersR Package: Label Noise Preprocessing in R. R Journal, 2017, 9, 219.	0.7	11
48	Dealing with Noisy Data. Intelligent Systems Reference Library, 2015, , 107-145.	1.0	10
49	Instance Selection. Intelligent Systems Reference Library, 2015, , 195-243.	1.0	10
50	Data Preparation Basic Models. Intelligent Systems Reference Library, 2015, , 39-57.	1.0	10
51	DILS: Constrained clustering through dual iterative local search. Computers and Operations Research, 2020, 121, 104979.	2.4	9
52	Label noise filtering techniques to improve monotonic classification. Neurocomputing, 2019, 353, 83-95.	3.5	8
53	A First Study on Decomposition Strategies with Data with Class Noise Using Decision Trees. Lecture Notes in Computer Science, 2012, , 25-35.	1.0	8
54	A First Approach to Nearest Hyperrectangle Selection by Evolutionary Algorithms. , 2009, , .		7

#	Article	IF	Citations
55	Using KEEL software as a educational tool: A case of study teaching data mining. , 2011, , .		7
56	Synthetic Sample Generation for Label Distribution Learning. Information Sciences, 2021, 544, 197-213.	4.0	7
57	Addressing Data-Complexity for Imbalanced Data-Sets: A Preliminary Study on the Use of Preprocessing for C4.5. , 2009, , .		6
58	An analysis on the use of pre-processing methods in evolutionary fuzzy systems for subgroup discovery. Expert Systems With Applications, 2012, 39, 11404-11412.	4.4	5
59	Enhancing instance-level constrained clustering through differential evolution. Applied Soft Computing Journal, 2021, 108, 107435.	4.1	5
60	ME-MEOA/D$_{CC}$: Multiobjective constrained clustering through decomposition-based memetic elitism. Swarm and Evolutionary Computation, 2021, 66, 100939.	4.5	5
61	Multiple instance classification: Bag noise filtering for negative instance noise cleaning. Information Sciences, 2021, 579, 388-400.	4.0	5
62	A first study on the noise impact in classes for Fuzzy Rule Based Classification Systems. , 2010, , .		4
63	Discretization. Intelligent Systems Reference Library, 2015, , 245-283.	1.0	4
64	A preliminary study on Hybrid Spill-Tree Fuzzy k-Nearest Neighbors for big data classification. , 2018, , .		4
65	A First Study on the Use of Boosting for Class Noise Reparation. Lecture Notes in Computer Science, 2016, , 549-559.	1.0	3
66	A Study on the Use of Statistical Tests for Experimentation with Neural Networks. , 2007, , 72-79.		3
67	Implementation and Integration of Algorithms into the KEEL Data-Mining Software Tool. Lecture Notes in Computer Science, 2009, , 562-569.	1.0	3
68	The impact of heterogeneous distance functions on missing data imputation and classification performance. Engineering Applications of Artificial Intelligence, 2022, 111, 104791.	4.3	3
69	A preliminary study on missing data imputation in evolutionary fuzzy systems of subgroup discovery. , 2012, , .		2
70	Smart Data. , 2020, , 45-51.		2
71	Data Reduction for Big Data. , 2020, , 81-99.		2
72	Data Reduction. Intelligent Systems Reference Library, 2015, , 147-162.	1.0	2

#	Article	IF	Citations
73	Big Data Preprocessing as the Bridge between Big Data and Smart Data: BigDaPSpark and BigDaPFlink Libraries. , 2019, , .		2
74	Data Sets and Proper Statistical Analysis of Data Mining Techniques. Intelligent Systems Reference Library, 2015, , 19-38.	1.0	2
75	A Data Mining Software Package Including Data Preparation and Reduction: KEEL. Intelligent Systems Reference Library, 2015, , 285-313.	1.0	2
76	A First Approach on Big Data Missing Values Imputation. , 2019, , .		2
77	Big Data Discretization. , 2020, , 121-146.		2
78	An extraction method for the characterization of the Fuzzy Rule Based Classification Systems' behavior using data complexity measures: A case of study with FH-GBML. , 2010, , .		1
79	The influence of noise on the evolutionary fuzzy systems for subgroup discovery. Soft Computing, 2016, 20, 4313-4330.	2.1	1
80	A First Study on the Use of Noise Filtering to Clean the Bags in Multi-Instance Classification. , 2018, , .		1
81	Improving constrained clustering via decomposition-based multiobjective optimization with memetic elitism. , 2020, , .		1
82	Domains of Competence of Artificial Neural Networks Using Measures of Separability of Classes. Lecture Notes in Computer Science, 2009, , 81-88.	1.0	0
83	An Experimental Case of Study on the Behavior of Multiple Classifier Systems with Class Noise Datasets. Lecture Notes in Computer Science, 2013, , 568-577.	1.0	0
84	Improving the Behavior of the Nearest Neighbor Classifier against Noisy Data with Feature Weighting Schemes. Lecture Notes in Computer Science, 2014, , 597-606.	1.0	0
85	Final Thoughts: From Big Data to Smart Data. , 2020, , 183-186.		0
86	Agglomerative Constrained Clustering Through Similarity and Distance Recalculation. Lecture Notes in Computer Science, 2020, , 424-436.	1.0	0
87	Big Data Software. , 2020, , 161-182.		0
88	3SHACC: Three stages hybrid agglomerative constrained clustering. Neurocomputing, 2022, 490, 441-461.	3.5	0