

# A survey of data partitioning and sampling methods to

Big Data Mining and Analytics

3, 85-101

DOI: [10.26599/bdma.2019.9020015](https://doi.org/10.26599/bdma.2019.9020015)

Citation Report

#	ARTICLE	IF	CITATIONS
1	Distributed Data Strategies to Support Large-Scale Data Analysis Across Geo-Distributed Data Centers. IEEE Access, 2020, 8, 178526-178538.	2.6	28
2	Improvement of an Online Education Model with the Integration of Machine Learning and Data Analysis in an LMS. Applied Sciences (Switzerland), 2020, 10, 5371.	1.3	75
3	Social Group Recommendation With TrAdaBoost. IEEE Transactions on Computational Social Systems, 2020, 7, 1278-1287.	3.2	11
4	Data Analysis as a Tool for the Application of Adaptive Learning in a University Environment. Applied Sciences (Switzerland), 2020, 10, 7016.	1.3	14
5	Covid-19 Plasma Monitoring Based on Clustering a Large Set of Recovered Patient Data. , 2020, , .		1
6	Big data analytics in telecommunications: Governance, architecture and use cases. Journal of King Saud University - Computer and Information Sciences, 2022, 34, 2758-2770.	2.7	14
7	Sampling for Big Data Profiling: A Survey. IEEE Access, 2020, 8, 72713-72726.	2.6	17
8	Realizing the Heterogeneity: A Self-Organized Federated Learning Framework for IoT. IEEE Internet of Things Journal, 2021, 8, 3088-3098.	5.5	106
9	RDMA-Enabled Concurrency Control Protocols for Transactions in the Cloud Era. IEEE Transactions on Cloud Computing, 2023, 11, 798-810.	3.1	7
10	Machine Translation Utilizing the Frequent-Item Set Concept. Sensors, 2021, 21, 1493.	2.1	3
11	An enhanced visual approach for accessing the clustering tendency of big data. Distributed and Parallel Databases, 0, , 1.	1.0	0
12	Machine learning based intelligent posture design of driver. Journal of Physics: Conference Series, 2021, 1802, 032131.	0.3	1
13	Pre-filtering based summarization for data partitioning in distributed stream processing. Concurrency Computation Practice and Experience, 2021, 33, e6338.	1.4	3
14	Academic Activities Recommendation System for Sustainable Education in the Age of COVID-19. Informatics, 2021, 8, 29.	2.4	8
15	Explaining the Factors Affecting the Relative Value of Services and Their Role in the Performance-Based Payment System in Teaching-Therapeutic Centers in the Six national-Wide region of the Country. Preventive Care in Nursing and Midwifery Journal, 2021, 11, 22-29.	0.1	0
16	Reinforcement Learning for Security-Aware Workflow Application Scheduling in Mobile Edge Computing. Security and Communication Networks, 2021, 2021, 1-13.	1.0	3
17	Analysis of the State of Learning in University Students with the Use of a Hadoop Framework. Future Internet, 2021, 13, 140.	2.4	4
18	Application of Statistical Sampling Survey in the Context of Big Data Era : "Based on China's 1% population sample survey in 2015. , 2021, , .		0

#	ARTICLE	IF	CITATIONS
19	A Comprehensive Survey of Privacy-preserving Federated Learning. ACM Computing Surveys, 2022, 54, 1-36.	16.1	191
20	Deterministic cooperative hybrid ring-mesh network coding for big data transmission over lossy channels in 5G networks. Eurasip Journal on Wireless Communications and Networking, 2021, 2021, .	1.5	11
21	Toward feature selection in big data preprocessing based on hybrid cloud-based model. Journal of Supercomputing, 0, , 1.	2.4	6
22	Security-Aware computation offloading for Mobile edge computing-Enabled smart city. Journal of Cloud Computing: Advances, Systems and Applications, 2021, 10, .	2.1	7
23	Visibility estimation via deep label distribution learning in cloud environment. Journal of Cloud Computing: Advances, Systems and Applications, 2021, 10, .	2.1	6
24	Intelligent Simulation of Children's Psychological Path Selection Based on Chaotic Neural Network Algorithm. Computational Intelligence and Neuroscience, 2021, 2021, 1-9.	1.1	1
25	Human Skeleton Detection and Extraction in Dance Video Based on PSO-Enabled LSTM Neural Network. Computational Intelligence and Neuroscience, 2021, 2021, 1-10.	1.1	5
26	Distributed and Privacy Preserving Graph Data Collection in Internet of Thing Systems. IEEE Internet of Things Journal, 2022, 9, 9301-9309.	5.5	8
27	An Intelligent Approach for Data Analysis and Decision Making in Big Data: A Case Study on E-commerce Industry. International Journal of Advanced Computer Science and Applications, 2021, 12, .	0.5	6
28	A Survey of Emotion Analysis in Text Based on Deep Learning. , 2020, , .		9
29	High-Dimensional Dataset Simplification by Laplace-Beltrami Operator. Lecture Notes in Computer Science, 2021, , 568-585.	1.0	0
30	Binomial Distribution based K-means for Graph Partitioning Approach in Partially Reconfigurable Computing system. , 2021, , .		0
31	A blockchain index structure based on subchain query. Journal of Cloud Computing: Advances, Systems and Applications, 2021, 10, .	2.1	10
32	A survey on deep learning for textual emotion analysis in social networks. Digital Communications and Networks, 2022, 8, 745-762.	2.7	61
33	Stratified linear systematic sampling based clustering approach for detection of financial risk group by mining of big data. International Journal of Systems Assurance Engineering and Management, 2022, 13, 1239-1253.	1.5	10
34	Query Optimization Algorithm of Replication Join Based on Sampling Partition. Journal of Physics: Conference Series, 2020, 1693, 012074.	0.3	0
35	Personalized Learning Design of Ideology and Politics of Distance Education Courses Based on Big Data. , 2020, , .		3
36	Coauthorship Network Mining for Scholar Communication and Collaboration Path Recommendation. Security and Communication Networks, 2021, 2021, 1-11.	1.0	0

#	ARTICLE	IF	CITATIONS
37	An edge computing oriented unified cryptographic key management service for financial context. <i>Wireless Networks</i> , 0, , .	2.0	3
38	Overcomplete pre-learned dictionary for incomplete data SAR imaging towards pervasive aerial and satellite vision. <i>Wireless Networks</i> , 0, , 1.	2.0	2
39	A horizontal partitioning-based method for frequent pattern mining in transport timetable. <i>Expert Systems</i> , 0, , e12881.	2.9	0
40	A Distributed Data Storage Strategy Based on LOPs. <i>Arabian Journal for Science and Engineering</i> , 0, , 1.	1.7	1
41	Analysis of Sports Injury Estimation Model Based on Mutation Fuzzy Neural Network. <i>Computational Intelligence and Neuroscience</i> , 2021, 2021, 1-10.	1.1	9
42	Principal component analysis based data collection for sustainable internet of things enabled Cyber-Physical Systems. <i>Microprocessors and Microsystems</i> , 2022, 88, 104032.	1.8	3
43	An evaluation model for Cloud-based Data mining Systems with Hadoop. , 2020, , .		1
44	Data Mid-platform Service System Design Method for Distributed New Energy. , 2021, , .		0
45	Optimization of Data Mining and Analysis System for Chinese Language Teaching Based on Convolutional Neural Network. <i>Computational Intelligence and Neuroscience</i> , 2021, 2021, 1-8.	1.1	1
46	Online computation offloading for deadline-aware tasks in edge computing. <i>Wireless Networks</i> , 0, , 1.	2.0	1
47	A Comprehensive Survey on Nanophotonic Neural Networks: Architectures, Training Methods, Optimization, and Activations Functions. <i>Sensors</i> , 2022, 22, 720.	2.1	3
48	A distributed learning based sentiment analysis methods with Web applications. <i>World Wide Web</i> , 2022, 25, 1905-1922.	2.7	7
49	Semantic Image Analysis on Social Networks and Data Processing. <i>Advances in Business Information Systems and Analytics Book Series</i> , 2022, , 189-214.	0.3	0
51	GAN for Load Estimation and Traffic-Aware Network Selection for 5G Terminals. <i>IEEE Internet of Things Journal</i> , 2022, 9, 16353-16362.	5.5	3
52	A Hybrid Method to Measure Distribution Consistency of Mixed-Attribute Datasets. <i>IEEE Transactions on Artificial Intelligence</i> , 2023, 4, 182-196.	3.4	2
53	Cloud Computing-aided Multi-type Data Fusion with Correlation for Education. <i>Wireless Networks</i> , 0, , 1.	2.0	1
54	Visual Communication Design and Wireless Data Transmission Technology for Blockchain Big Data Information Presentation. <i>Wireless Communications and Mobile Computing</i> , 2022, 2022, 1-14.	0.8	1
55	A blockchain and IoT-based lightweight framework for enabling information transparency in supply chain finance. <i>Digital Communications and Networks</i> , 2022, 8, 576-587.	2.7	36

#	ARTICLE	IF	CITATIONS
56	A review of optimization methods for computation offloading in edge computing networks. Digital Communications and Networks, 2023, 9, 450-461.	2.7	28
57	A Research on the Development and Application of Blockchain Technology in Industrial Internet of Things. , 2021, , .		1
58	RSP-Hist: Approximate Histograms for Big Data Exploration on Hadoop Clusters. , 2021, , .		2
59	What Is the Level of People's Acceptance for Electric Taxis and Buses? Exploring Citizens' Perceptions of Transportation Electrification to Pay Additional Fees. World Electric Vehicle Journal, 2022, 13, 3.	1.6	1
60	Trusted Service Evaluation for Mobile Edge Users: Challenges and Reviews. Complexity, 2021, 2021, 1-10.	0.9	0
61	Fast wireless sensor for anomaly detection based on data stream in an edge-computing-enabled smart greenhouse. Digital Communications and Networks, 2022, 8, 498-507.	2.7	16
62	Dynamic Distributed and Parallel Machine Learning algorithms for big data mining processing. Data Technologies and Applications, 2022, 56, 558-601.	0.9	1
63	Evaluation of Performance for Big Data Security Using Advanced Cryptography Policy. , 2021, , .		0
64	Deadline-constrained cost-energy aware workflow scheduling in cloud. Concurrency Computation Practice and Experience, 2022, 34, .	1.4	4
65	Recommendation System for Privacy-Preserving Education Technologies. Computational Intelligence and Neuroscience, 2022, 2022, 1-8.	1.1	2
66	Min-max kurtosis mean distance based k-means initial centroid initialization method for big genomic data clustering. Evolutionary Intelligence, 0, , .	2.3	3
67	Adaptive Educational Hypermedia System Based on Variational Bayesian Petri Net. Mathematical Problems in Engineering, 2022, 2022, 1-7.	0.6	1
68	An Advanced Deep Attention Collaborative Mechanism for Secure Educational Email Services. Computational Intelligence and Neuroscience, 2022, 2022, 1-9.	1.1	3
69	Exploring Consumers' Attitudes towards Food Products Derived by New Plant Breeding Techniques. Sustainability, 2022, 14, 5995.	1.6	10
70	A New Stratified Block Model to Process Large-Scale Data for a Small Cluster. Lecture Notes on Data Engineering and Communications Technologies, 2022, , 253-263.	0.5	1
71	A Lightweight Enterprise Messaging Service at the Edge. , 2021, , .		0
72	Design of fuzzy rule-based models with fuzzy relational factorization. Expert Systems With Applications, 2022, 206, 117904.	4.4	6
73	A General Framework for Class Label Specific Mutual Information Feature Selection Method. IEEE Transactions on Information Theory, 2022, 68, 7996-8014.	1.5	3

#	ARTICLE	IF	CITATIONS
74	Analysis of methods of electricity consumption forecasting for a coal industry enterprise. IOP Conference Series: Earth and Environmental Science, 2022, 1045, 012035.	0.2	1
75	Proposal of a Model for the Analysis of the State of the Use of ICT in Education Applied to Technological Institutes of Higher Education. Computers, 2022, 11, 112.	2.1	0
76	Spatiotemporal data partitioning for distributed random forest algorithm: Air quality prediction using imbalanced big spatiotemporal data on spark distributed framework. Environmental Technology and Innovation, 2022, 27, 102776.	3.0	8
77	Determining the Level of Use of the Industry 4.0 Solutions in the COVID-19 Pandemic Era: Results of Empirical Research. Sustainability, 2022, 14, 8844.	1.6	3
78	Novel Decentralized Security Architecture for the Centralized Storage System in Hadoop using Blockchain Technology. , 2022, , .		8
79	Forecasting Electricity Consumption of Electrical Machines of a Coal Industry Enterprise Using the Wavelet Transform. , 2022, , .		1
80	Investigation of Load Schedules of Electrical Machines of a Mining Enterprise Using Wavelet Analysis. , 2022, , .		1
81	EDSUC: A robust ensemble data summarization method for effective medical diagnosis. Digital Communications and Networks, 2022, , .	2.7	0
82	Analysis of electricity consumption of electrical machines of a coal industry enterprise using the wavelet transform. IOP Conference Series: Earth and Environmental Science, 2022, 1070, 012002.	0.2	0
83	Big Data Analytics: Deep Content-Based Prediction with Sampling Perspective. Computer Systems Science and Engineering, 2023, 45, 531-544.	1.9	0
84	Intelligent Identification and Order-Sensitive Correction Method of Outliers from Multi-Data Source Based on Historical Data Mining. Electronics (Switzerland), 2022, 11, 2819.	1.8	2
85	Performance Evaluation of Machine-Learning Models for Self-Healing in 5G Networks. International Journal of Interdisciplinary Telecommunications and Networking, 2022, 14, 1-14.	0.2	0
86	Weighted Reservoir Sampling On Evolving Streams: A Sampling Algorithmic Framework For Stream Event Identification. , 2022, , .		2
87	A Data-Driven Method for Residential Transformers Winding Connections and Capacities Recognition in the Distribution Management System. , 2022, , .		0
88	A Cluster-Based Partition Method of Remote Sensing Data for Efficient Distributed Image Processing. Remote Sensing, 2022, 14, 4964.	1.8	0
89	AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. Artificial Intelligence Review, 2023, 56, 4929-5021.	9.7	63
90	Maintenance Decision System Design Based on Equipment Support Data System. , 2022, , .		0
91	Predicting the Volume Phase Transition Temperature of Multi-Responsive Poly( <i>N</i> -isopropylacrylamide)-Based Microgels Using a Cluster-Based Partial Least Squares Modeling Approach. ACS Applied Polymer Materials, 2022, 4, 9160-9175.	2.0	2

#	ARTICLE	IF	CITATIONS
92	Federated Learning for Supply Chain Demand Forecasting. <i>Mathematical Problems in Engineering</i> , 2022, 2022, 1-8.	0.6	3
93	A Systematic Multi-level Evaluation Framework and Approach for Electricity Market Operation Based on Big Data. , 2022, , .		0
94	Parallel processing proposal by clustering integration of low-cost microcomputers. <i>Procedia Computer Science</i> , 2022, 214, 100-107.	1.2	2
95	When bioprocess engineering meets machine learning: A survey from the perspective of automated bioprocess development. <i>Biochemical Engineering Journal</i> , 2023, 190, 108764.	1.8	9
98	A new ranking-based stability measure for feature selection algorithms. <i>Soft Computing</i> , 0, , .	2.1	0
99	Analysis of full reference quality metrics for image transmission over a MIMO OWC channel under varying turbulent conditions. <i>International Journal of Communication Systems</i> , 2023, 36, , .	1.6	3
100	Big Data Framework Classification for Public E-Governance Using Machine Learning Techniques. , 2022, 48, 112-122.		1
101	Data Analysis Model for the Evaluation of the Factors That Influence the Teaching of University Students. <i>Computers</i> , 2023, 12, 30.	2.1	1
102	Introduction to Distributed Nearest Hash: On Further Optimizing Cloud Based Distributed kNN Variant. <i>Procedia Computer Science</i> , 2023, 218, 1571-1580.	1.2	5
103	Survey of Distributed Computing Frameworks for Supporting Big Data Analysis. <i>Big Data Mining and Analytics</i> , 2023, 6, 154-169.	7.5	7
104	Big Data Storage Index Mechanism Based on Hierarchical Indexing and Concurrent Updating. , 2022, , .		0
105	Sampling and noise filtering methods for recommender systems: A literature review. <i>Engineering Applications of Artificial Intelligence</i> , 2023, 122, 106129.	4.3	1
106	Diversified and compatible web APIs recommendation based on game theory in IoT. <i>Digital Communications and Networks</i> , 2023, , .	2.7	2
107	A Comprehensive Review on Data Partitioning and Sampling Techniques for Processing Big Data. , 2022, , , .		0
108	Shipping code towards data in an inter-region serverless environment to leverage latency. <i>Journal of Supercomputing</i> , 0, , .	2.4	1
109	Approximate Clustering Ensemble Method for Big Data. <i>IEEE Transactions on Big Data</i> , 2023, 9, 1142-1155.	4.4	4
110	TeraHeap: Reducing Memory Pressure in Managed Big Data Frameworks. , 2023, , .		3
112	An ensemble method for estimating the number of clusters in a big data set using multiple random samples. <i>Journal of Big Data</i> , 2023, 10, , .	6.9	2

#	ARTICLE	IF	CITATIONS
113	Case Design and Analysis of Digital Media Courses Incorporating Aesthetic Education. , 2022, , .		0
114	A Systematic Literature Review on Security in Cloud Technology. , 2023, , .		0
118	Data Collection, Presentation and Analysis. , 2023, , 115-138.		0
119	Research on Privacy Protection Methods for Data Mining. Communications in Computer and Information Science, 2023, , 609-620.	0.4	0
122	Dimensions and Hadoop of Big Data. A Review. , 2023, , 323-333.		0
123	Analysis of Data using hybridized K-means clustering with PSO-JAYA algorithm. , 2023, , .		0
126	Machine Learning and Data Mining Algorithms for Geospatial Big Data. Springer Remote Sensing/photogrammetry, 2023, , 207-226.	0.4	0
127	Assessment of Data Quality Through Multi-granularity Data Profiling. Lecture Notes in Computer Science, 2023, , 195-209.	1.0	0
134	FPGA-Integrated Bag of Little Bootstraps Accelerator for Approximate Database Query Processing. Lecture Notes in Computer Science, 2023, , 115-130.	1.0	0
140	Motion Aware Data Sampling Using Sequential Frames for Deep Learning Models. , 2023, , .		0
143	Future Expectations About Big Data Analytics. Advances in Business Information Systems and Analytics Book Series, 2024, , 451-465.	0.3	0
148	An Enhanced Machine Learning-Based Analysis of Teaching and Learning Process for Higher Education System. Lecture Notes in Business Information Processing, 2024, , 321-332.	0.8	0
150	Multi-FDMF: An Agile Management Framework of Multi-domain Data in Decentralized Heterogeneous Environments. , 2023, , .		0