Alok Choudhary

List of Publications by Year
in descending order

87 papers	5,095 citations	430874 18 h-index	345221 36 g-index
88 all docs	88 docs citations	88 times ranked	4435 citing authors

#	Article	IF	Citations
1	A general-purpose machine learning framework for predicting properties of inorganic materials. Npj Computational Materials, 2016, 2, .	8.7	922
2	Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. APL Materials, 2016, 4, 053208.	5.1	712
3	Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. Construction and Building Materials, 2017, 157, 322-330.	7.2	666
4	Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. Physical Review B, 2017, 96, .	3.2	254
5	ElemNet: Deep Learning the Chemistry of Materials From Only Elemental Composition. Scientific Reports, 2018, 8, 17593.	3.3	242
6	Twitter Trending Topic Classification. , 2011, , .		211
7	Recent advances and applications of deep learning methods in materials science. Npj Computational Materials, 2022, 8, .	8.7	207
8	MineBench: A Benchmark Suite for Data Mining Workloads. , 2006, , .		175
9	Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning. Nature Communications, 2019, 10, 5316.	12.8	160
10	Improved parallel I/O via a two-phase run-time access strategy. Computer Architecture News, 1993, 21, 31-38.	2.5	159
11	Deep materials informatics: Applications of deep learning in materials science. MRS Communications, 2019, 9, 779-792.	1.8	137
12	An Extended Two-Phase Method for Accessing Sections of Out-of-Core Arrays. Scientific Programming, 1996, 5, 301-317.	0.7	89
13	Prediction of seebeck coefficient for compounds without restriction to fixed stoichiometry: A machine learning approach. Journal of Computational Chemistry, 2018, 39, 191-202.	3.3	65
14	Predictive analytics for crystalline materials: bulk modulus. RSC Advances, 2016, 6, 95246-95251.	3.6	62
15	An online tool for predicting fatigue strength of steel alloys based on ensemble data mining. International Journal of Fatigue, 2018, 113, 389-400.	5.7	60
16	Dynamically adapting file domain partitioning methods for collective I/O based on underlying parallel file system locking protocols. , 2008, , .		56
17	Cross-property deep transfer learning framework for enhanced predictive analytics on small materials data. Nature Communications, 2021, 12, 6595.	12.8	55
18	Extracting Grain Orientations from EBSD Patterns of Polycrystalline Materials Using Convolutional Neural Networks. Microscopy and Microanalysis, 2018, 24, 497-502.	0.4	46

#	Article	IF	Citations
19	Scaling parallel I/O performance through I/O delegate and caching system. , 2008, , .		45
20	NUMARCK: Machine Learning Algorithm for Resiliency and Checkpointing. , 2014, , .		43
21	Enabling active storage on parallel I/O software stacks. , 2010, , .		36
22	An Implementation and Evaluation of Client-Side File Caching for MPI-IO. , 2007, , .		33
23	Uncertain Range Queries for Necklaces. , 2010, , .		32
24	Property Prediction of Organic Donor Molecules for Photovoltaic Applications Using Extremely Randomized Trees. Molecular Informatics, 2019, 38, e1900038.	2.5	31
25	Enabling deeper learning on big data for materials informatics applications. Scientific Reports, 2021, 11, 4244.	3.3	29
26	A Comprehensive Approach to Image Spam Detection: From Server to Client Solution. IEEE Transactions on Information Forensics and Security, 2010, 5, 826-836.	6.9	28
27	Predicting Node Failure in High Performance Computing Systems from Failure and Usage Logs. , 2011, , .		28
28	Using Subfiling to Improve Programming Flexibility and Performance of Parallel Shared-file I/O. , 2009, , .		23
29	Parallel Deep Convolutional Neural Network Training by Exploiting the Overlapping of Computation and Communication. , 2017, , .		21
30	Detecting/preventing information leakage on the memory bus due to malicious hardware. , 2010, , .		20
31	Identifying HotSpots in Lung Cancer Data Using Association Rule Mining. , 2011, , .		20
32	A New Flexible MPI Collective I/O Implementation. , 2006, , .		19
33	A New Parallel Algorithm for Two-Pass Connected Component Labeling. , 2014, , .		17
34	Materials discovery: Understanding polycrystals from large-scale electron patterns. , 2016, , .		17
35	Machine Learning Models to Predict Performance of Computer System Design Alternatives. , 2008, , .		16
36	Parallel hierarchical clustering on shared memory platforms. , 2012, , .		16

#	Article	IF	Citations
37	Delegation-Based I/O Mechanism for High Performance Computing Systems. IEEE Transactions on Parallel and Distributed Systems, 2012, 23, 271-279.	5.6	16
38	IOPin: Runtime Profiling of Parallel I/O in HPC Systems. , 2012, , .		15
39	Evaluating voltage islands in CMPs under process variations. , 2007, , .		14
40	A Formation Energy Predictor for Crystalline Materials Using Ensemble Data Mining. , 2016, , .		14
41	Improved Scaling of Molecular Network Calculations: The Emergence of Molecular Domains. Journal of Physical Chemistry Letters, 2017, 8, 415-421.	4.6	14
42	An Architectural Characterization Study of Data Mining and Bioinformatics Workloads. , 2006, , .		13
43	Combining I/O operations for multiple array variables in parallel netCDF. , 2009, , .		13
44	HIGH UTILITY ITEMSETS MINING. International Journal of Information Technology and Decision Making, 2010, 09, 905-934.	3.9	13
45	Efficient pairwise statistical significance estimation for local sequence alignment using GPU. , 2011, , .		13
46	Evaluating the effects of cache redundancy on profit. , 2008, , .		11
47	High Performance Data Mining Using R on Heterogeneous Platforms. , 2011, , .		11
48	Transfer Learning Using Ensemble Neural Networks for Organic Solar Cell Screening. , 2019, , .		11
49	Parallel pairwise statistical significance estimation of local sequence alignment using Message Passing Interface library. Concurrency Computation Practice and Experience, 2011, 23, 2269-2279.	2.2	10
50	Dynamic Directories: A mechanism for reducing on-chip interconnect power in multicores. , 2012, , .		10
51	Reducing I/O variability using dynamic I/O path characterization in petascale storage systems. Journal of Supercomputing, 2017, 73, 2069-2097.	3.6	10
52	A flexible I/O arbitration framework for netCDFâ€based big data processing workflows on highâ€end supercomputers. Concurrency Computation Practice and Experience, 2017, 29, e4161.	2.2	10
53	Compiler-Directed Energy Optimization for Parallel Disk Based Systems. IEEE Transactions on Parallel and Distributed Systems, 2007, 18, 1241-1257.	5.6	9
54	Deep learning based domain knowledge integration for small datasets: Illustrative applications in materials informatics. , 2019, , .		9

#	ARTICLE	IF	CITATIONS
55	Microarchitectures for Managing Chip Revenues under Process Variations. IEEE Computer Architecture Letters, 2008, 7, 5-8.	1.5	8
56	Improving MPI Collective I/O for High Volume Non-Contiguous Requests With Intra-Node Aggregation. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 2682-2695.	5.6	8
57	A high-performance application data environment for large-scale scientific computations. IEEE Transactions on Parallel and Distributed Systems, 2003, 14, 1262-1274.	5.6	7
58	Operating System Controlled Processor-Memory Bus Encryption. , 2008, , .		7
59	Lung transplant outcome prediction using UNOS data. , 2013, , .		7
60	Automated Tracing of I/O Stack. Lecture Notes in Computer Science, 2010, , 72-81.	1.3	7
61	Improving MPI Independent Write Performance Using A Two-Stage Write-Behind Buffering Method. , 2007, , .		6
62	Community Dynamics and Analysis of Decadal Trends in Climate Data. , 2011, , .		6
63	Supporting computational data model representation with high-performance I/O in parallel netCDF. , 2011, , .		6
64	Poster: A lung cancer mortality risk calculator based on SEER data. , 2011, , .		6
65	Parallel DTFE Surface Density Field Reconstruction. , 2016, , .		6
66	Improving the Average Response Time in Collective I/O. Lecture Notes in Computer Science, 2011, , 71-80.	1.3	6
67	Enzyme Function Classification Using Protein Sequence Features and Random Forest. , 2009, , .		5
68	Improving All-to-Many Personalized Communication in Two-Phase I/O. , 2020, , .		5
69	An Efficient FPGA Implementation of Principle Component Analysis based Network Intrusion Detection System. , 2008, , .		4
70	Excavating social circles via user interests. Social Network Analysis and Mining, 2014, 4, 1.	2.8	4
71	Martensite Start Temperature Predictor for Steels Using Ensemble Data Mining. , 2019, , .		4
72	pFANGS: Parallel high speed sequence mapping for Next Generation 454-roche Sequencing reads. , 2010, , .		3

#	Article	IF	Citations
73	Parallel Implementation of Lossy Data Compression for Temporal Data Sets. , 2016, , .		3
74	SILVERBACK+: scalable association mining via fast list intersection for columnar social data. Knowledge and Information Systems, 2017, 50, 969-997.	3.2	3
75	GPU-accelerated Monte Carlo simulations of dense stellar systems. , 2012, , .		2
76	Analyzing Informal Caregiving Expression in Social Media. , 2017, , .		2
77	Data-Driven Insights from Predictive Analytics on Heterogeneous Experimental Data of Industrial Magnetic Materials. , 2019, , .		2
78	Large-scale file systems with the flexibility of databases. ACM Computing Surveys, 1996, 28, 207.	23.0	2
79	A distributed multi-storage I/O system for data intensive scientific computing. Parallel Computing, 2003, 29, 1623-1643.	2.1	1
80	PinterNet: A thematic label curation tool for large image datasets. , 2016, , .		1
81	Towards Identifying Informal Caregivers of Alzheimerâ€™s and Dementia Patients in Social Media. , 2017, , .		1
82	Heterogeneous Feature Fusion Based Machine Learning on Shallow-Wide and Heterogeneous-Sparse Industrial Datasets. Lecture Notes in Computer Science, 2021, , 566-577.	1.3	1
83	SIGRNN: Synthetic Minority Instances Generation in Imbalanced Datasets using a Recurrent Neural Network. , 2021, , .		1
84	Enhancing Phase Mapping for High-throughput X-ray Diffraction Experiments using Fuzzy Clustering. , 2021, , .		1
85	Sensing, Triggers and Mobile (Meta)Data. , 2010, , .		0
86	Achieving Target MTTF by Duplicating Reliability-Critical Components in High Performance Computing Systems. , 2011, , .		0
87	IOPro: a parallel I/O profiling and visualization framework for high-performance storage systems. Journal of Supercomputing, 2015, 71, 840-870.	3.6	0