

{tag}

Conference X-PLORE 2014
by IJCA Journal

{/tag}

IJCA Proceedings on National Level Technical
© 2014

XPLORE 2014

Year of Publication: 2014

Authors:

Uzma Shafaque

Parag D. Thakare

{bibtex}xplore1417.bib{/bibtex}

Abstract

Environment of Big Data produces a large amount of data, in which it need to be analyzed and patterns have to be extracted, to gain knowledge. In this era of big data, with boom of data both structured data and unstructured data, in different field such as Engineering, Genomics, Biology, Meteorology, Environmental research and many more, it has become difficult to manage, process and analyze patterns using architectures and databases that are traditional. So, we should understand a proper architecture to gain knowledge about the Big Data. In this paper a review of various algorithms necessary for handling such large data set is given. These algorithms give us various methods implemented to handle Big Data.

ences

- F. Michel, "How Many Photos Are Uploaded to Flickr Every Day and Month?"<http://www.flickr.com/photos/franckmichel/6855169886/>, 2012.
- "IBM What Is Big Data: Bring Big Data to the Enterprise," <http://www-01.ibm.com/software/data/bigdata/>, IBM, 2012.
- Feifei Li, Suman Nath "Scalable data summarization on big data", Distributed and Parallel Databases An International Journal, 15 February 2014.
- United Nations Global Pulse, 2012, Big Data for Development: Challenges & Opportunities, May 2012
- Office of Science and Technology Policy | Executive Office of the President, 2012, Fact Sheet: Big Data across the Federal Government, March 29, 2012, [www.WhiteHouse.gov/OSTP](http://www.whitehouse.gov/OSTP)
- Office of Science and Technology Policy | Executive Office of the President, 2012, Obama Administration Unveils "Big Data" Initiative: Announces \$200 Million in New R&D Investments, March 29, 2012, [www.WhiteHouse.gov/OSTP](http://www.whitehouse.gov/OSTP)
- McKinsey Global Institute, 2011, Big Data: the Next Frontier for Innovation, Competition, and Productivity, May 2011
- Rajaraman A, Ullman J D "Mining of Massive Datasets, Cambridge University Press, 2011
- Edmon Begoli, James Horey, "Design Principles for Effective Knowledge Discovery from Big Data", Joint Working Conference on Software Architecture & 6th European Conference on Software Architecture, 2012
- Ivanka Valova, Monique Noirhomme, "Processing Of Large Data Sets: Evolution, Opportunities And Challenges", Proceedings of PCaPAC08
- Neha Saxena, Niket Bhargava, Urmila Mahor, Nitin Dixit, "An Efficient Technique on Cluster Based Master Slave Architecture Design", Fourth International Conference on Computational Intelligence and Communication Networks, 2012
- R. Caruana and A. Niculescu-Mizil, An Empirical Comparison of Supervised Learning Algorithms," in Proceedings of the 23rd international conference on Machine learning, ICML '06, (New York, NY, USA), pp. 161-168, ACM, 2006.
- R. Caruana, N. Karampatziakis, and A. Yessenalina, An Empirical Evaluation of Supervised Learning in High Dimensions," in Proceedings of the 25th international
- Mr. D. V. Patil, Prof. Dr. R. S. Bichkar, "A Hybrid Evolutionary Approach To Construct Optimal Decision Trees with Large Data Sets", IEEE, 2006
- Guillermo Sinchez-Diaz, Jose Ruiz-Shulcloper, "A Clustering Method for Very Large Mixed Data Sets", IEEE, 2001
- Mehmet Koyuturk, Ananth Grama, and Naren Ramakrishnan, "Compression, Clustering, and Pattern Discovery in very High-Dimensional Discrete-Attribute Data Sets", IEEE Transactions On Knowledge And Data Engineering, April 2005, Vol. 17, No. 4
- Emily Namey, Greg Guest, Lucy Thairu, Laura Johnson, "Data Reduction Techniques for Large Qualitative Data Sets", 2007
- Moshe Looks, Andrew Levine, G. Adam Covington, Ronald P. Loui, John W. Lockwood, Young H. Cho, "Streaming Hierarchical Clustering for Concept Mining",

IEEE, 2007

- Yen-ling Lu, chin-shyurngfahn, "Hierarchical Artificial Neural Networks For Recognizing High Similar Large Data Sets. "; Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, August 2007
- Archana Singh, MeghaChaudhary, Dr (Prof.) Ajay Rana, GauravDubey, "Online Mining of data to Generate Association Rule Mining in Large Databases"; International Conference on Recent Trends in Information Systems, 2011
- David N. Reshef et al. ,"Detecting Novel Associations in Large Data Sets"; Science AAAS, 2011, Science 334
- Shuliang Wang, WenyanGan, Deyi Li, Deren Li "Data Field For Hierarchical Clustering"; International Journal of Data Warehousing and Mining, Dec. 2011
- Tatiana V. Karpinets, ByungH. Park, Edward C. Uberbacher, "Analyzing large biological datasets with association network"; Nucleic Acids Research, 2012
- M. Vijayalakshmi, M. Renuka Devi, "A Survey Of Different Issues Of Different Clustering Algorithms Used In Large Data Sets"; International Journal Of Advanced Research In Computer Science And Software Engineering, March 2012
- Xindong Wu, XingquanZhu,Gong-Qing Wu, and Wei Ding, "Data Mining with Big Data," IEEE Transactions On Knowledge And Data Engineering, Vol. 26, No. 1, January 2014.

Computer Science

Index Terms

Data Mining

Keywords

Big Data Data Mining Map-reduce Crowdsourcing Algorithm.