Trend Analysis of E-Commerce Data using Hadoop Ecosystem

Volume 147
Number 6

Year of Publication: 2016

Authors:
Rama Satish K. V., N. P. Kavya

10.5120/ijca2016911109

Abstract

Trend Analysis is the custom of collecting information and attempting to spot a trend, or pattern, in the information. Trend analysis is often used to estimate future events, it could be used to approximate uncertain events in the past. Technical analysts and Technicians also uses market indicators of many types. Processing or analyzing the Trend in huge amount of data or extracting meaningful information is a challenging task. As the enterprises faced issues of gathering large chunks of data and analyzing the Trend. They found that the data cannot be processed using any of the existing centralized architectures. One of the best open source tools used in the market to harness the distributed architecture in order to solve the data processing and analyzing problems is Apache Hadoop and Hive for querying best results. This paper addresses an experimental work on Trend analysis problem of big data and its optimal solution using Hadoop ecosystem, using parallel processing framework to process large data sets using Map Reduce programming and Apache Hive is a data warehouse infrastructure which is built on top of Hadoop for providing data summarization, querying and analysis.
References

1. Impetus white paper, March, 2011, “Planning Hadoop/NoSQL Projects for 2011” by Technologies
2. ChenHauWang Dept. of Comput. Sci., Nat. Chiao-Tung Univ., Hsinchu, Taiwan Ching TsorngTsai ; Chia Chen Fan ; Shyan Ming Yuan A Hadoop Based Weblog Analysis System.
6. Dan Meng ; Huili Gong ; Xiaojuan Li ; Demin Zhou Geoinformatics (GEOINFORMATICS), 2012 20th International Conference on DOI:10.1109/Geoinformatics.2012.6270336 Publication Year: 2012 , Page(s): 1 – 6.
7. Jianguo Cui ; Pengyuan Zhao ; Shiliang Dong ; Liqiu Liu ; Rui Lv ; Zhonghai Li Electrical and Control Engineering (ICECE), 2011 International Conference on
8. DOI: 10.1109/ICECENG.2011.6057830Publication Year: 2011 , Page(s): 3339 – 3342.

Index Terms

Computer Science Information Sciences

Keywords
Trend Analysis, E-Commerce data, Apache Hadoop, Hive.