Logical Itemset Mining Implementation on Hadoop

International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

Volume 135
Number 9

Year of Publication: 2016

Authors:
Karan Jawalkar, Avinash Patil, Shreemay Panhalkar, Raj Pande

Abstract

Frequent Itemset Mining (FISM) finds the large and frequently occurring items from the datasets using Apriori algorithm. The FISM framework does not address two major properties that are Mixture-of property (more than one customer intent) and Projection-of property. To overcome the problems of irrelevant and non-actionable data and also to address the properties mentioned above, Logical Itemset Mining (LISM) framework is introduced. LISM finds logical itemsets from the data which helps in eliminating non-actionable data but at the same time keeps data which is logically connected. LISM not only finds logically connected items but also items which are rarely occurring but logically connected are also discovered. LISM also addresses the Mixture of property and Projection of property which are not very well addressed in FISM.

References
1. R. Agrawal, T. Imielinski, and A. N. Swami, Mining association rules between sets of items in large databases, SIGMOD, pp. 207-216, 1993

Index Terms

Computer Science

Information Sciences

Keywords

FISM, LISM, M/R Job, FLASK, LUCENE