

Efficient Ordering Policy for Imperfect Quality Items Using Association Rule Mining

Mandeep Mittal

Department of Computer Science Engineering, Amity School of Engineering and Technology, India

Sarla Pareek

Apaji Institute of Mathematics and Applied Computer Technology, Banasthali University, India

Reshu Agarwal

Apaji Institute of Mathematics and Applied Computer Technology, Banasthali University, India

INTRODUCTION

Data mining is the process of finding interesting patterns from huge databases. It is the main part of the knowledge discovery in database (KDD) process. The KDD process may consist of following steps: data cleaning, data transformation, pattern searching, finding presentation, finding interpretation and finding evaluation. Emergence of data mining and knowledge discovery results in fast development of information and database technologies.

Further, (Mannila & Raiha, 1987) have done work in the database community on inferring functional dependencies, and efficient inference algorithms. However, association rule mining received a great attention (Agrawal et al., 1993). It is used to find fascinating rules from large collections of data which express an association or relation between items or sets of items. The most important algorithm for generating association rules is apriori-algorithm. This algorithm is designed to operate on databases containing transactions. It works to find the items for frequent item-set based on minimum support and generate association rules based on threshold confidence. Figure 1, explains the flowchart of apriori-algorithm. Now, we will explain the meaning of the terms minimum support and confidence. Given the item-set I containing items $\{i_1, i_2, i_3, \dots, i_n\}$, support for item i_1 is defined as the frequency of its occurrence in total number of transactions. It is given by formula:

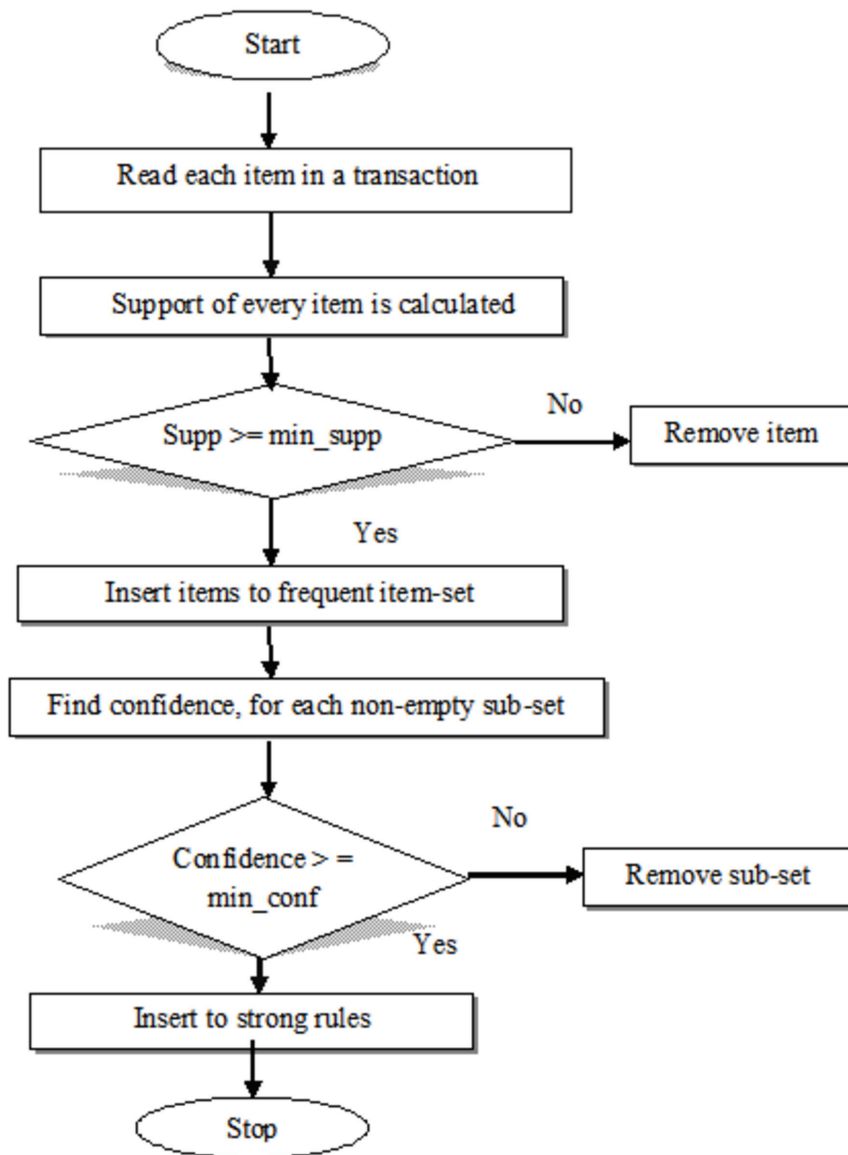
$$\text{Support}(i_1) = \frac{\text{Frequency of } i_1}{\text{Total number of Transactions}}$$

The relationship between items is expressed in terms of confidence. Confidence is defined as conditional probability as confidence ($i_1 \rightarrow i_2$) refers to frequency of purchasing i_2 when i_1 is purchased. It is defined by the formula:

$$\text{Confidence}(i_1 \rightarrow i_2) = \frac{\text{Support of } i_1 \cup i_2}{\text{Support of } i_1}$$

Further, many variants of mining association rules are studied to explore more mining capabilities, such as incremental updating (Lee et al., 2001), mining of generalized and multi-level rules (Srikant & Agrawal, 1995), mining of multi-dimensional rules (Ng & Han, 1994) and temporal association rule discovery (Ale & Rossi, 2000). Temporal association rule mining leads to discovery of association rules that may hold during some time intervals but not during others as explained by (Li et al., 2003). The general temporal association rule is termed to be frequent within its maximum common exhibition period (MCEP) if and only if its support is not smaller than the minimum support (min_supp), and its confidence is not smaller

Figure 1. Flow chart of Apriori-algorithm



than the minimum confidence (min_conf) as defined by (Lee et al., 2003). Generally the exhibition period of an item-set is expressed in terms of maximal common exhibition period. $\text{MCEP}(P)$ denotes the MCEP value of P . The MCEP value of an item-set P is the shortest MCEP among the items in item-set P . For example, value of $\text{MCEP}(P) = (1, 3)$ in Table 1.

Consequently, the general temporal association rule mining algorithm can be explained in three steps:

1. Find every frequent maximal temporal item-set (TIs) with their support values.
2. Find the support values of every corresponding temporal sub-item-set (SIs) of frequent TIs.

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/efficient-ordering-policy-for-imperfect-quality-items-using-association-rule-mining/112392

Related Content

Hybrid Clustering using Elitist Teaching Learning-Based Optimization: An Improved Hybrid Approach of TLBO

D.P. Kanungo, Janmenjoy Nayak, Bighnaraj Naik and H.S. Behera (2016). *International Journal of Rough Sets and Data Analysis* (pp. 1-19).

www.irma-international.org/article/hybrid-clustering-using-elitist-teaching-learning-based-optimization/144703

Should the Cloud Computing Definition Include a Big Data Perspective?

Rafik Ouanouki, Abraham Gomez Morales and Alain April (2015). *Encyclopedia of Information Science and Technology, Third Edition* (pp. 1088-1095).

www.irma-international.org/chapter/should-the-cloud-computing-definition-include-a-big-data-perspective/112504

Designing Personalised Learning Resources for Disabled Students Using an Ontology-Driven Community of Agents

Julius T. Nganji and Mike Brayshaw (2013). *Information Systems Research and Exploring Social Artifacts: Approaches and Methodologies* (pp. 81-102).

www.irma-international.org/chapter/designing-personalised-learning-resources-disabled/70711

Discovering Patterns using Process Mining

Ishak Meddah and Belkadi Khaled (2016). *International Journal of Rough Sets and Data Analysis* (pp. 21-31).

www.irma-international.org/article/discovering-patterns-using-process-mining/163101

Formal Verification Methods

Osman Hasan and Sofiene Tahar (2015). *Encyclopedia of Information Science and Technology, Third Edition* (pp. 7162-7170).

www.irma-international.org/chapter/formal-verification-methods/112414