Opportunity Cost Estimation Using Clustering and Association Rule Mining

Reshu Agarwal, Amity Institute of Information Technology, Amity University, Noida, India

ABSTRACT
Information mining strategies are most appropriate for the classification, useful patterns extraction and predications which are imperative for business support and decision making. However, an efficient method for evaluating the penalty cost has not been proposed. In this article, considering the cross-selling effect, a quantitative approach to estimate the opportunity cost based on association rules in each cluster is proposed. This article helps in better decision making for improving sales, services and quality, which is useful mechanism for business support, investment, and surveillance. A numerical illustration is utilized to clarify the new approach. Further, to understand the effect of above approach in the real scenario, experiments are conducted on a real-world dataset.

KEYWORDS
Association Rule Mining, Clustering, Data Mining, Inventory Control, Loss-Profit, Opportunity Cost

INTRODUCTION
Data Mining is the process of extracting information from large data sets through the use of algorithms and techniques drawn from the field of statistics, machine learning and database management systems (Feelders, Daniels, & Holsheimer, 2000; Raju & Schumacker, 2016). Traditional data analysis methods often involve manual work and interpretation of data that is slow, expensive, and highly subjective. Data Mining, popularly called as knowledge discovery in large data, enables firms and organizations to make calculated decisions by assembling, accumulating, analyzing and accessing corporate data (Edvardsson, 2017). Various methods of data mining such as item-set mining, mining sequential pattern, association rule, classification, prediction, clustering, regression etc. are applied for the above areas (Han & Kamber, 2006). Further, the problem of mining association rules from transactional database was introduced by Agrawal, Imielinski, and Swami (1993). The concept aims to find frequent patterns, interesting correlations, and associations among sets of items in the transaction databases or other data repositories. Association rule are the statements that find the relationship between data in any database. Association rule has two parts “Antecedent” and “Consequent.” For example: {bread} => {eggs}. Here bread is the antecedent and egg is the consequent. Antecedent is the item that is found in the database, and consequent is the item that is found in combination with the first. Gautam & Pardasani (2010) presented an efficient version of Apriori algorithm for mining multilevel association rules in large databases for finding maximum frequent item-set at lower level of abstraction. There are two important basic measures for association rules, support and confidence. Support of an
association rule is defined as the percentage of records that contain items \( A \cup B \) to the total number of records in the database. Confidence of an association rule is defined as the percentage of the number of transactions that contain items \( A \cup B \) to the total number of records that contain item \( A \). The proposed algorithm can derive the multiple-level association rules under different supports in simple and effective way. To explain association rule mining, consider in a grocery store, bread is purchased with butter 30% of the time and that milk is purchased with it 40% of the time. Based on these associations, special displays of milk and bread are placed near the butter which is on sale. The management decided not to put these items on sale. These actions are aimed at increasing overall sales volume by taking advantage of frequency with which these items are purchased together. There are two association rules found in this example. The first one states that when butter is purchased, bread is purchased 30% of the time. The second one states that 40% of the time when butter is purchased so is milk. The discovered association rules can be used by management to increase the effectiveness (and reduce the cost) associated with advertising, marketing, inventory, and stock location on the floor.

Moreover, clustering is a technique of finding groups of objects such that the objects in one group will be similar to one another and different from the objects in another group. In business intelligence clustering can be used to organize a large number of customers into groups where customers within a group share similar characteristics. This facilitates the development of business strategies for enhanced customer relationship management. Broder, Glassman, Manasse, and Zweig (1997) define clusters as maximal connected components of some pair-wise similarity of transactions, thus suffers from the breakdown of the transitivity of pair-wise similarity. Guha, Rastogi, & Shim (2000) proposed the common neighbors of two transactions as a measure of pair-wise similarity. Wang, Xu, and Liu (1999) method does not use any notion of pair-wise similarity. They cluster transactions that contain similar items. The difference is that clustering emphasizes the dissimilarity of clusters. The rationale behind clustering transactions prior to mining association rules is that the latter is performed on partitions that are essentially distinct from each other. Both association rule mining and clustering techniques helps in effective inventory management.

The problem of inventory control is one of the most important in organizational management. As a rule, there is no standard solution – the conditions at each company or firm are unique and include many different features and limitations. Proper and timely determination of the optimal inventory control strategy allows freeing a significant amount of assets, frozen in the form of stocks, which ultimately increases the efficiency of resource use. Even though there are literally millions of different types of products manufactured in our society, there are only two fundamental decisions that one has to make when controlling inventory:

1. How large should an inventory replenishment order be?
2. When should an inventory replenishment order be placed?

The objectives of inventory management often reduce the problem if it is more profitable to do quickly but more expensive or slower but cheaper. Such a strategy will be optimal inventory control, which minimizes the sum of milestones costs associated with the production, storage and inventory shortage per unit of time or for a specific (including infinite) amount of time.

Traditionally, ABC analysis has been based on the criterion of dollar volume and on the principle that there is a relatively small number of items-category A -that account for the bulk of the dollar volume. At the other extreme, a large number of items-category C -account for a small share of the dollar volume. Category B items are between categories A and C, both in number and dollar volume. By this criterion, A items are those of both high-value and high-demand and C items are low-value and low-demand. However, over the last 30 years, there has been an accumulation of research questioning this focus on a single criterion—the dollar volume. It has been pointed out that other criteria can be important; among these are lead time, item criticality, durability, scarcity, reparability, stock-ability,
10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/article/opportunity-cost-estimation-using-clustering-and-association-rule-mining/237152?camid=4v1


www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

An Eco-System Architectural Model for Delivering Educational Services to Children With Learning Problems in Basic Mathematics
Miguel Angel Ortiz Esparza, Jaime Muñoz Arteaga, José Eder Guzman Mendoza, Juana Canul-Reich and Julien Broisin (2019). International Journal of Information Technologies and Systems Approach (pp. 61-81).

www.igi-global.com/article/an-eco-system-architectural-model-for-delivering-educational-services-to-children-with-learning-problems-in-basic-mathematics/230305?camid=4v1a

A Generic Approach for the Semantic Annotation of Conceptual Models Using a Service-Oriented Architecture
Hans-Georg Fill, Daniela Schremser and Dimitris Karagiannis (2013). International Journal of Knowledge Management (pp. 76-88).

www.igi-global.com/article/generic-approach-semantic-annotation-conceptual/77328?camid=4v1a
Narrative
www.igi-global.com/chapter/narrative/17014?camid=4v1a

Using Patterns to Capture Tacit Knowledge and Enhance Knowledge Transfer in Virtual Teams
www.igi-global.com/chapter/using-patterns-capture-tacit-knowledge/54257?camid=4v1a