Chapter 20

Application of Data Mining Techniques on Library Circulation Data for Library Material Acquisition and Budget Allocation

Md. Hossain
North South University, Bangladesh

Rashedur M. Rahman
North South University, Bangladesh

ABSTRACT

This chapter offers a model for automated library material utilization that is based on knowledge discovery using association rules. Processing the circulation data of the library to extract the statistics and association utilization of the materials for departments is a great achievement that makes the analysis easier for calculating material utilization. Moreover, processing the circulation data of the library, two important dimensions, namely concentration and connection (Kao, Chang, & Lin, 2003), could be explored among departments and library members. This can make the analysis easier by calculating weights in those two important dimensions to make the decision about budget allocation. This chapter analyses the circulation data of North South University Library and suggests that efficient management and budget allocation can be achieved by using the above-mentioned metrics.

1. INTRODUCTION

Data Mining is a fabulous technique to acquire knowledge from any archive of scattered data. It can easily help us taking any decision based on the collected knowledge. Allocation of budget for a library is important to manage and maintain it properly. Besides, good allocation ensures the maximum and optimum usability of the library materials for the sake of the members of the library. In this occasion...
Application of Data Mining Techniques on Library Circulation Data

data mining can play a vital role to allocate a perfect or nearly perfect budget for a library based on the circulation data. There are a number of research studies reported in the literature that used data mining techniques in the last few years. A large amount of time usually the librarians spend to acquire resources for numbers of library users. However, prior to this panic task, library authority must determine how to list up all the materials that will meet the needs of most of the users. Many studies involved in the literature have observed the increased use of the data mining techniques like association rule in the past few years. In (Hamaker, 1995) it was indicated that usage information showed something different from the collectors’ recommendations. More importantly, information such as ‘what were the materials most utilized by the students of a particular department’ would be highly useful for acquiring materials (Kao, Chang, & Lin, 2003). In (Budd & Adams, 1989), it was emphasized that circulation statistics could be one of the most significantly references for library managerial decisions. This paper introduces some easier techniques of statistical analysis and goal programming to extract the knowledge which can give us the power to take the decision about the allocation of budget.

In this research we analyze North South University (NSU) Library database for possible allocation of academic library acquisition budget via circulation database mining. We derived weights of acquisition budget allocation by composing the descriptive knowledge via utilization concentration and the suitability via utilization connection (Kao, Chang, & Lin, 2003) for departments concerned. The value of concentration and connection is calculated from the number of records, the distribution of material categories used, and the relation between categories and subjects. We obtained the degree of concentration for a department by measuring information entropy using ID3 algorithm introduced by Quinlan (1986). Finally, we obtain budget allocation table for each department using corresponding weights. We also derived material utilization by composing statistics utilization and association utilization from the knowledge collected from the circulation database. We obtained the degree of support for categories of a department by measuring the association in transactions using association rule mining algorithm and the confidence from the records found in circulation database. We obtained budget allocation table for each department corresponding to material utilization.

The rest of the chapter is organized as follows: Section 2 highlights some related papers that describe similar kind of researches. Section 3 shows the diagram of the database model, Section 4 describes the model in detail. Section 5 presents our research findings and finally Section 6 draws the conclusion.

2. RELATED WORKS

Many research works have been involved into budget allocation for library materials acquisition. Most of them are in common style and with a common objective, some have slightly different in objectives. We report findings from some of the researches related to our work below.

Kao, Chang, & Lin (2003) worked with the circulation database of Kun Shan University of Technology where they used ID3 (Quinlan, 1986) algorithm of Data Mining. Wu, Lee, & Kao (2004) used circulation statistics mechanism and an association rule was applied to discover knowledge. There are several models approached by several researchers like ABAMDM (Kao, Chang, & Lin, 2003), KDBM_{LMA} (Wu, Lee, & Kao, 2004) to help derive the utilization of library material categories. Although many approaches and research reports have been extensively used to help library material acquisitions, the knowledge contained in circulation databases has rarely been used to investigate in-depth how the acquired materials are being used. Wu & Lee (2005) presented a decision support model for library material acquisition and budget allocation using the knowledge derived from circulation databases.