# Library Management System Using Association Rule Mining

[1]V.Anitha Moses ,

Prof., Department of MCA,

Panimalar Engineering College, Chennai.

[2]Caroline Sunita Jeevan,

PG Scholar, Department of MCA,

Panimalar Engineering College, Chennai.

[3] .Dr. S. Kala,

Department of MCA,

Panimalar Engineering College, Chennai.

***Abstract***

Data mining is the latest tool available to explore the hidden information from the large amount of database, this can be done by applying the data mining in the library database. Data mining is usually employed on very large database. Library is a source of all knowledge and learning. Libraries are also generating large volume of data, but data mining techniques have to be used for dynamically analyzing the library database and to make strategic decisions for managing the library in an efficient manner. By applying association rule mining techniques, strategic decisions can be taken for library management.

**Keywords:** Data Mining - KDD - Clustering - Association Rule.

## I. INTRODUCTION

Data Mining, also popularly known as Knowledge Discovery in Databases (KDD), refers to the nontrivial extraction of implicit, previously unknown and potentially useful information from data in databases. While data mining and (or KDD) are frequently treated as synonyms, data mining is actually part of the knowledge discovery process. Data Mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, Data Mining is the process of finding correlations or patterns among dozens of fields in large relational databases. The following figure (Figure 1) shows data mining as a step in an iterative knowledge discovery process.

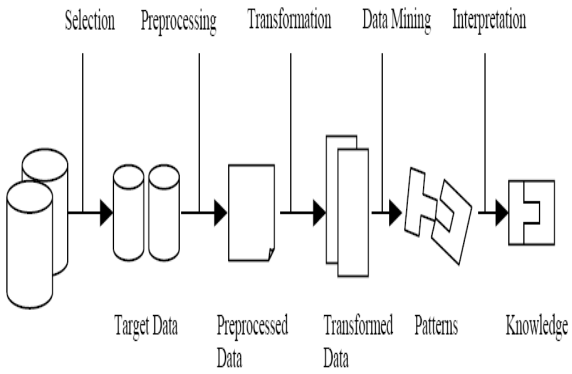



**Figure 1: Data Mining is the core of Knowledge Discovery process**

The data mining is actually a step in a larger KDD process. The KDD process employs data mining methods or algorithms to extract or identify knowledge according to some criteria or measure of interestingness, but it also includes steps that prepare

the data, such as preprocessing, sub-sampling, and transformations of the database [4].
The first step in the KDD process is to select data to be analyzed from the set of all available data. In many cases, the data is stored in transaction databases. These databases are quite large and extremely dynamic. Therefore a subset of the data must be selected from those databases, since it is unnecessary in the early stages to attempt to analyze all data. The KDD is an iterative process. Once the discovered knowledge is presented to the user, the evaluation measures can be enhanced, the mining can be further refined, new data can be selected or further transformed, or new data sources can be integrated, in order to get different, more appropriate results.

## II. EXISTING WORK

### A. *Statistical technique*

Statistical techniques were being used long before the term data mining was coined. However, statistical techniques are driven by the data and are used to discover patterns and build predictive models. Today people have to deal with up to terabytes of data and have to make sense of it and glean the important patterns from it. In statistics, prediction is usually synonymous with regression of some form. There are a variety of different types of regression in statistics but the basic idea is that a model is created that maps values from predictors in such a way that the lowest error occurs in making a prediction.
The simplest form of regression is Simple Linear Regression that just contains one predictor and a prediction. The relationship between the two can be mapped on a two dimensional space and the records plotted for the prediction values along the Y axis and the predictor values along the X axis. The simple linear regression model then could be viewed as the line that minimized the error rate between the actual prediction value and the point on the line [2]. Adding more predictors to the linear equation can produce more complicated lines that take more information into account and hence make a better prediction, and it is called multiple linear regressions.

**The Disadvantages of Statistical Technique**

Certainly statistics can do more than answer questions about the data but for most people today these are the questions that statistics cannot help answer. Consider that a large part of data the statistics is concerned with summarizing data, and more often than not, the problem that the summarization has to do with counting. Statistical Techniques cannot be useful without certain assumptions about data.

### B. *Clustering technique*

Clustering is concerned with grouping together objects that are similar to each other and dissimilar to the objects belonging to other clusters [3]. Clustering techniques is used by the end user to tag the customers in their database. Once this is done the business user can get a quick high level view of what is happening within the cluster. Clustering can be used for discovery or prediction. There are two main types of clustering techniques, those that create a hierarchy of clusters and those that do not.
Those techniques are: *Hierarchical Clustering Techniques and Partitional Clustering Techniques.*
**The Disadvantages of Clustering Technique**

The interpretation of how interesting a clustering is will inevitably be application-dependent and subjective to some degree. Clustering techniques suffer from the fact that once a merge or a split is committed, it cannot be undone or refined. Sometimes clustering is performed not so much to keep records together as to make it easier to see when one record sticks out from the rest.

## III. PROPOSED WORK

### A. *Association Rule*

An association rule tells us about the association between two or more items. For example, If we are given a set of items where items can be referred as books and a large collection of transactions (i.e., issue/return) which are subsets (baskets) of these items/books. The task is to find relationship between the presence of various items within these baskets. In order for the rules to be useful there are two pieces of information that must be supplied as well as the actual rule: *Support* is how often does the rule apply? and *Confidence* is How often is the rule is correct.

In fact association rule mining is a two-step process: Find all frequent itemsets / booksets - by definition, each of these itemsets will occur at least as frequently as a predetermined minimum support count, and then

generate strong association rules from the frequent itemsets – by definition, these rules must satisfy minimum support and minimum confidence.

**Advantages of Association Rule Technique**

Association rule algorithms can be formulated to look for sequential patterns. The methods of data acquisition and integration, and integrity checks are the most relevant to association rules.

***B. Modules***

The following modules in this project are

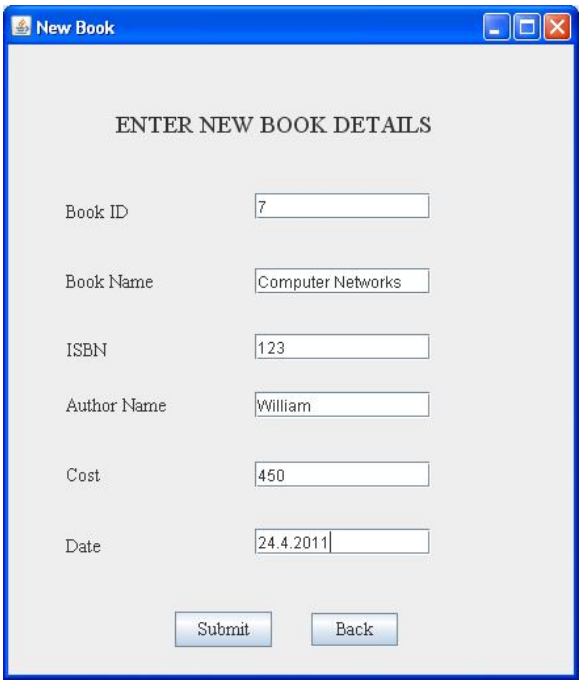
- Administrator
- Librarian
- User

**1. Administrator**

Administrator has the full authority to modify the library database like creating an account for a new user, deleting the account, adding a new book and deleting the book. Administrator monitors the library database updations.


- **Enter New Book Details**

The book id is a automatically generated whenever a new book is entered in the library. The various entries should be entered such as the book name, isbn, author name, cost, and date.



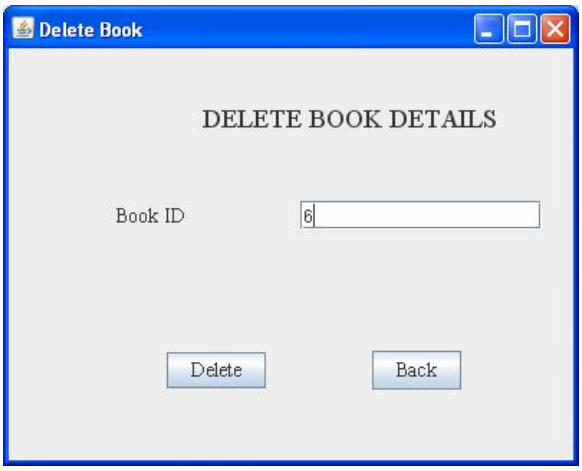

- **New Account Creation**

The user is classified into two types

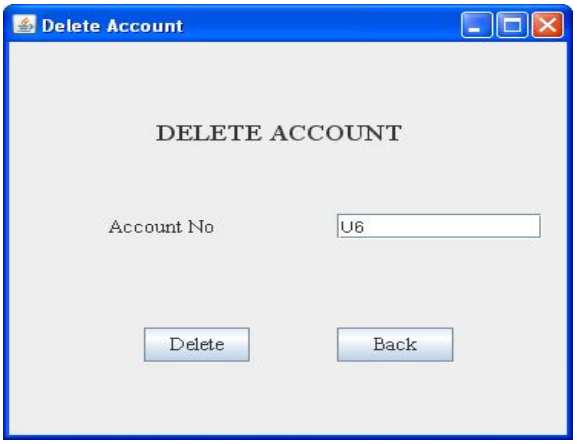
1. Other user.


2. Staff user.


- **Delete Book Details**

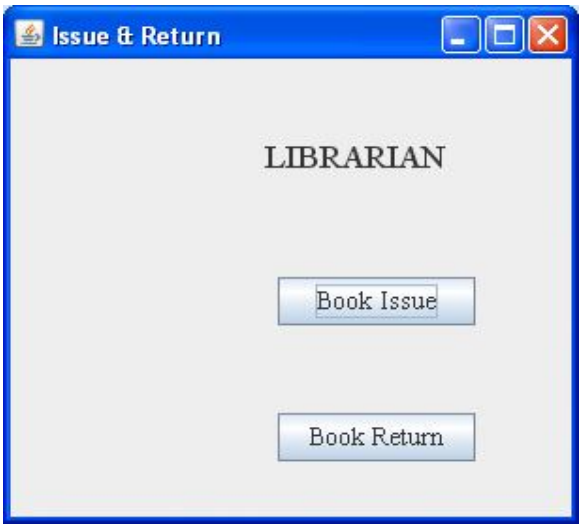



- **Delete Account**

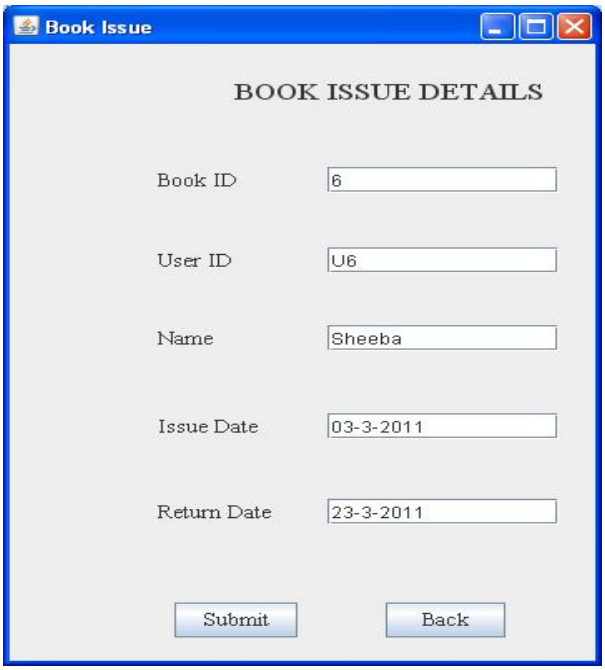



**2. Librarian module**

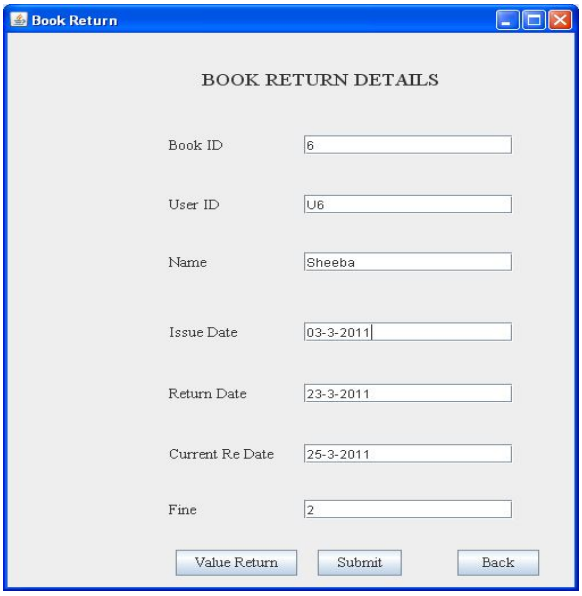
Librarian handles the library action like to issue the book and return back the book. Librarian collects the daily report and submits it to the administrator.


- **Book Issue Details**




- **Book Return Details**




### 3. User module

In the user module every authorized user has a unique login username and password. All these should be provided by the administrator. Through the login user, user can receive books from the library and return books to the library. All updation takes place in the library database.


**4. Staff module**

All the staff have a unique login username and password. Staff have some additional feature compare to other user likewise they can take two more books and time extension compare to other user.


## IV. RESULT

Association rules can be implemented in a two-step process: find all frequent item sets: by definition, each of these item sets will represent the transaction, and then generate strong association rules from the frequent item sets: by definition, these rules must satisfy minimum support and minimum confidence.

**TABLE 1**

**ASSOCIATION RULE MINING**

| Rule | Support | Confidence |
|---|---|---|
| In a semester if a book "Mobile Computing" was issued to a member then a book ".Net" also issued to the same member. | 5% | 75% |
| In a semester if Issue-Count of a book is between 20 to 30 | 2% | 90% |

## V. CONCLUSION

This paper describes the processes of selected techniques using association rule mining. It has been realized that all association mining techniques accomplish its goals perfectly, but this technique has its own characteristics and specifications that demonstrate its accuracy, proficiency and preference. It claims that new research solutions are needed for the problem of categorical data mining techniques, and presenting our ideas for future work. There is no effective technique in data mining. It is therefore recommended that these techniques should be used in cooperation with each other. For efficiently and effectively doing the library administration and extending library services, the need of library automation and digital library occur. But simply automating the library or developing digital library is not the only solution unless and until it is not able to explore the hidden information from the large amount of database. This can be done by applying the data mining in the library database.

## VI. REFERENCES

[1]. M. S. Chen, J. Han, and P. S. Yu. Data mining: An overview from a database perspective. IEEE Trans. Knowledge and Data Engineering, 8:866-883, 1996.
[2]. U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, 1996.
[3]. W. J. Frawley, G. Piatetsky-Shapiro and C. J. Matheus, Knowledge Discovery in Databases: An

Overview. In G. Piatetsky-Shapiro et al. (eds.), Knowledge Discovery in Databases. AAAI/MIT Press, 1991.
[4]. J. Han and M. Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2000.
[5]. T. Imielinski and H. Mannila. A database perspective on knowledge discovery. Communications of ACM, 39:58-64, 1996.
[6]. G. Piatetsky-Shapiro, U. M. Fayyad, and P. Smyth. From data mining to knowledge discovery: An overview. In U.M. Fayyad, et al. (eds.), Advances in Knowledge Discovery and Data Mining, 1-35. AAAI/MIT Press, 1996.
[7]. G. Piatetsky-Shapiro and W. J. Frawley. Knowledge Discovery in Databases. AAAI/MIT Press, 1991.
[8]. Agrawal, R. and Srikant, R. "Mining Sequential Patterns", Proceeding 1995: International Conference on Data Engineering, Taipei, Taiwan, March 1995, pp. 3-14.
[9]. Berson, A., Smith, S. and Thearling, K. "An Overview of Data Mining Techniques", White Paper from Internet, 2005.
[10]. Chen, M.S, Han, J. and. Yu, P.S. "Data Mining: An Overview from a Database Perspective", IEEE Transaction on Knowledge and Data Engineering, Vol.8, 1996, pp. 866-883.
[11]. U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy.
Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, 1996.