Overview of Data Mining Applications

Ms. Gurpreet Kaur

Assistant Professor, University Institute of Computing (UIC), Chandigarh University, Gharaun, Mohali, India

Abstract: In this Paper we have discussed about data mining technologies. As we aware that there are large organization, each organization generate large amount of data. The data warehouse is helpful in making the effective decisions. The huge amount of data is available in the form of terabyte. To manage and make decision of such type of huge amount of data we need a technique called data mining.

Keywords: Data mining Task, Data mining Life cycle, Types of Data Mining System, Application of Data Mining System.

I. INTRODUCTION

Today we are using different technologies to adequate in the society. Every day we are using the large amount of data and these data are in different fields like document, Records, video etc. All these data has different formats. When the customer will require the data should be retrieved from database and make better decision. This technique is called KDD (Knowledge Discovery Database).

II. THE DATA MINING TASKS

Summarization:
Summarization is the gernalization of data set of task relevant data is summarized and abstracted, resulting smaller set which gives a gernal overview of data.

Classification
Classification is derivation of function or model which determines the class of an object based on its attributes.

Prediction
This model permits the value of one variable to be predicted from the known value of other variable.

Discovering Patterns and Rules
This task is used to find the hidden pattern. In a cluster a number of patterns of different size and cluster are available. This is done by data mining algorithm like (K-Means).These are called clustering algorithm.

III. TYPES OF DATA MINING SYSTEM

Classification according to kind of database mined: Data base system can be classified according to different criteria such as datamodels, type of data etc.

Classification of data mining systems according to the type of data source mined: In an organization a large amount of data's are available where we need to classify these data but these are available in a similar fashion. We need to classify these data according to its type (maybe audio/video, text format etc)

Classification according to data model drawn on: This classification categorizes data mining system based on data model involved such as relational database, object oriented database, data warehouse etc.

Classification according to application adapted: We can classify the data mining system according to application adapted. These applications are as follows finance, Telecommunication, stock market, email.

IV. DATA MINING LIFE CYCLE

The life cycle of a data mining project consists of six phases.

The main Phases are:

Bussiness Understanding:
First it is required to understand business objectives clearly and find out what are the business needs.

Data Understanding:
Data understanding phase starts with initial data collection which we collect from available data sources. Some important activities must be performed including data load and data integration in order to make the data collection successfully.

Data Preparation:
The outcome of the data preparation phase is final data set. Once available data source are identified, they need to be selected, cleaned, constructed and formatted in desired form.

Modeling:
In this phase, various modelling techniques are selected and applied.

Evaluation:
In the evaluation phase, the model result must be evaluated. In this phase new business requirement must be raised due to new pattern.

Deployment:

The knowledge or information which gain through data mining process, needs to be presented in such a way that stakeholder can use it when they want it.

V. DATA MINING APPLICATION

Data Mining is used in Market Analysis and Management:

1. Data mining help to determine what kind of people buy what kind of Product.
2. Data mining helps in identifying the best Product for different customer
3. Data mining perform association between product and sale.
4. Data Mining helps to find customers who share the same characteristics.
5. Data mining provide us various summary reports.

Data Mining Application in Insurance:

Data mining enables to forecast which customer will potentially purchase new polices. Data mining helps to detect fraudlent behavior.

Data Mining Application in Fraud Detection:

It is used to detect fraud. In fraud telephone calls, it helps to find the destination of call, duration of call, time of day or week etc.

Data Mining Applications in Healthcare:

Data mining applications in health can have usefulness. However, the success of healthcare data mining requires the availability of clean healthcare data. In this respect, it is critical that the healthcare industry look into how data can be better captured, stored, prepared and mined. Possible directions include the standardization of clinical vocabulary and the sharing of data across organizations to enhance the benefits of healthcare data mining applications.

Data Mining Application in Education:

Data Mining can be used by an institution to take accurate decisions and also predict the result of the student. With the result the institutions can focus what to teach and how to teach.

VI. CONCLUSION

In this paper we discuss various data mining applications. This review would be helpful to researchers to focus on the various issues of data mining. In future course, we will discuss the various classification algorithms and significance of evolutionary computing (genetic programming). Most of the previous studies on data mining applications in various fields use the variety of data types range from text to images and stores in variety of databases and data structures. The different methods of data mining are used to extract the patterns and thus the knowledge from this variety databases. Selection of data

and methods for data mining is an important task in this process and needs the knowledge of the domain.. The domain experts shall be guided by the system to effectively apply their knowledge for the use of data mining systems to generate required knowledge. The domain experts are required to determine the variety of data that should be collected in the specific problem domain, selection of specific data for data mining, cleaning and transformation of data, extracting patterns for knowledge generation and finally interpretation of the patterns and knowledge generation. Most of the domain specific data mining applications show accuracy above 90%. The generic data mining applications are having the limitations. From the study of various data mining applications it is observed that, no application called generic application is 100 % generic. The intelligent interfaces and intelligent agents up to some extent make the application generic but have limitations.

REFERENCES

[1] Introduction to Data Mining and Knowledge Discovery, Third Edition ISBN: 1-892095-02-5, Two Crows Corporation, 10500 Falls Road, Potomac, MD 20854 (U.S.A.), 1999.
[2] Larose, D. T., "Discovering Knowledge in Data: An Introduction to Data Mining", ISBN 0-471-66657-2, ohn Wiley & Sons, Inc, 2005.
[3] Dunham, M. H., Sridhar S., "Data Mining: Introductory and Advanced Topics", Pearson Education, New Delhi, ISBN: 81-7758-785-4, 1st Edition, 2006
[4] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. and Wirth, R... "CRISP-DM 1.0 : Step-by-step data mining guide, NCR Systems Engineering Copenhagen (USA and Denmark), DaimlerChrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringenen Bank Group B.V (The Netherlands), 2000".