

Appliance of Data Mining In Knowledge Management

Harjinder Kaur (Lecturer)

School of Computer Applications, Lovely Professional University (Phagwara –[PB]. India)

Abstract

Mining of data is the process of digging out the constructive patterns from data warehouse which were further used for auxiliary analysis. It is one of the imperative steps of Knowledge Discovery Process. The preliminary part of this paper converse the impression of data mining next to various data mining tasks and the later part of this paper focuses on Knowledge Discovery Process .The final segment of this paper focuses on the applications of data mining in different areas of Knowledge management.

Keywords: Data Mining, Data warehouse, Knowledge Discovery Process, Knowledge Management.

I. Introduction

Data mining is one of the crucial steps of knowledge discovery process which is used to produce useful patterns from ample data which makes the analysis process easier. The following are the major tasks for data mining:

1) *Classification*: It's a two step process which is used for predicting class labels based on classifiers rules. In first step it build model for classification known as classifier based on training data and in second step it checks for the accuracy of the model based on some test data which is autonomous of training data. It is also known as supervised learning because class labels are already known.

2) *Regression*: With the help of some statistical method regression is used for predicting numerical data.

3) *Clustering*: It is also known as unsupervised learning because class labels are not known in advance. It is the process of partitioning the data into subsets which are considered as clusters. Objects in one cluster are similar in nature but they are dissimilar to other clusters.

4) *Summarization*: As a result of data mining we can get the summarized data which is in the form of interesting patterns which makes the analysis easier.

II. Phases of Data Mining Process [4]

The following are the various phases of Data Mining Process:

Problem Definition: The first phase is problem definition where the scope of the problem is defined along with the project objective which is then translated into problem definition.

- 1) *Data Exploration*: In this phase experts uses various exploration techniques such as find the mean and standard deviation of values in order to get accurate results. In this phase Meta data is also defined.
- 2) *Data Preparation*: In data preparation phase we collect data from heterogeneous sources where each and every source has its own format which results in data inconsistencies. So in this phase we have to clean the data and convert it into some acceptable format so that quality analysis should be made from that data.
- 3) *Modeling*: In this phase we have to build model based on some training sets and test data sets. Training sets are used to build the model and afterwards test data is used to test whether the model is working properly or not. We have to repeat the complete process until we get the optimal model.
- 4) *Evaluation*: After building a model we have to evaluate the model if the model works according to the specification then the model is acceptable otherwise we have to go back to modeling phase and have to change certain parameters till we get the optimal solution.
- 5) *Deployment*: In this phase we can deploy the model into operating environment after deploying we have to continuously update the model for its effective working.

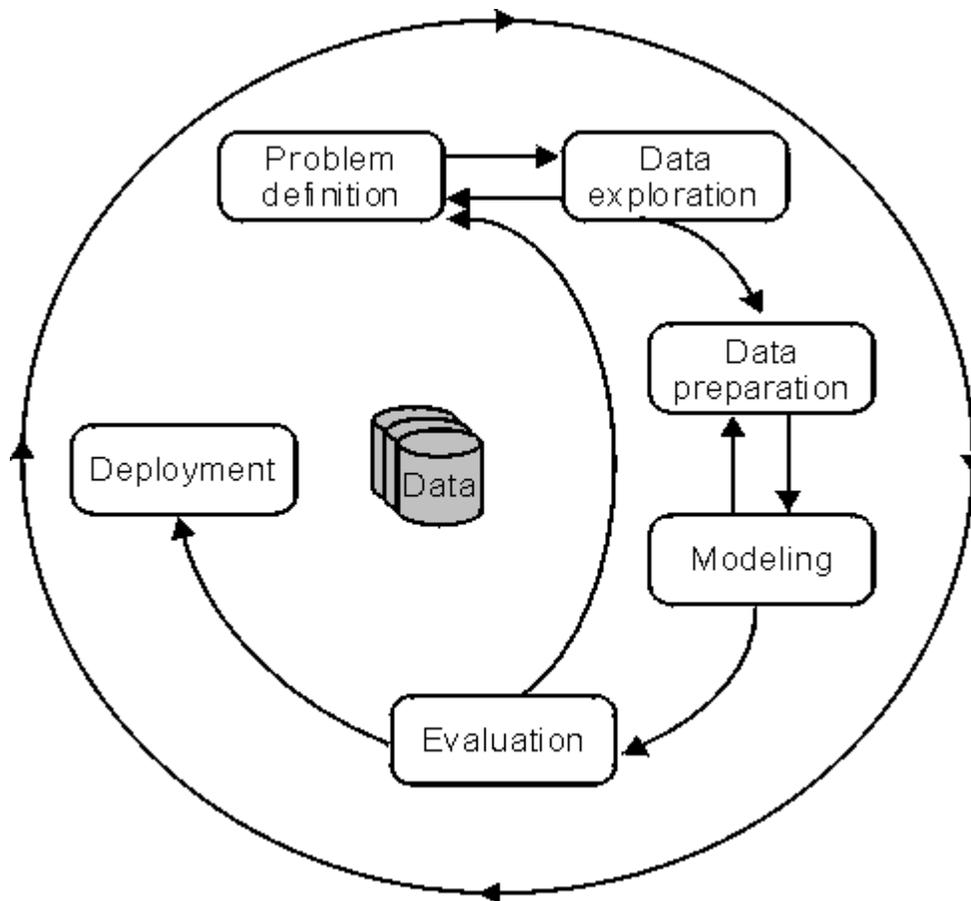


Figure1 Data Mining Process [4]

III. Knowledge Discovery Process

Knowledge discovery process is iterative sequence of following steps: [2]

- 1) *Data Cleaning*: it is the process of removing noisy and inconsistent data and smoothing out the outliers.
- 2) *Data Integration*: It is the process of gathering data from different heterogeneous data sources.
- 3) *Data Selection*: Process of selecting relevant data from database which is helpful in data analysis.
- 4) *Data Transformation*: In this process data is transformed into smaller values either by using aggregate functions or by normalizing it which is helpful in data mining.
- 5) *Data Mining*: In this process intelligent methods are applied to get interesting patterns which are helpful in analysis.
- 6) *Pattern Evaluation*: This process involves evaluation of interesting patterns which are helpful for analysis.
- 7) *Knowledge presentation*: Process in which knowledge representation techniques are used to represent knowledge to the end use.

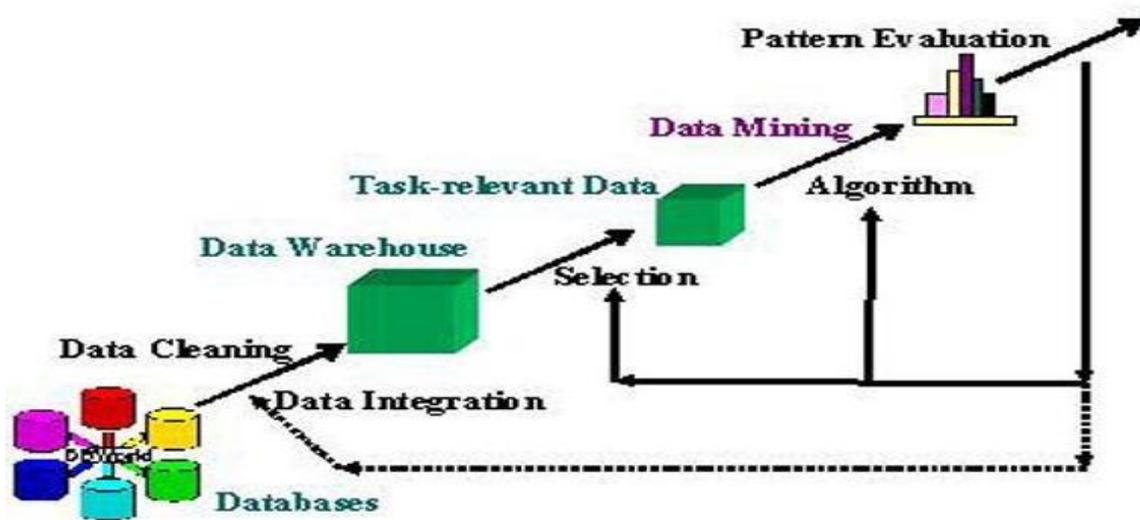


Figure2 shows the knowledge discovery process [2]

- IV. Data mining applications used in Knowledge management [3]

As knowledge is everywhere so we can find data mining techniques are applicable to almost every field ranging from banks, medical, Web search engines and there are many more.

Banks: Data mining by using classification techniques helps the loan officer to identify which customers are risky and safe according to the analysis of the bank loan officer. We can build a classifier based on some training data which has some classification rules and based on those rules we can identify a class label whether a particular customer is safe or risky. Eg If the classifier is having the following rules:

If customer > 100000 then safe

If customer income < 50000 then risky

Now suppose we have to predict whether customer x with income 70000 is risky or is safe. So based on classification rules we can easily predict the class label for customer x that x is risky.

Medical: Suppose a person is suffering from fever and a medical practitioner wants to diagnose whether the person is suffering from malaria, typhoid or dengue. We have to build a model for this based on training data. The model has some classification rules and based on those rules medical practitioner easily identify the person is suffering from which type of fever.

Web Search engines: Web search Engine is a server which finds voluminous data from the web. Various data mining techniques are used in different process of search engine like searching, crawling etc. for searching data from web

by applying data mining techniques is known as web mining.

V. Conclusion:

This paper represents a step toward a common framework that will ultimately provide a unifying vision of the common objectives and data mining techniques used in Knowledge Discovery Process. The prime goal was to elucidate the relation between data mining and knowledge discovery process. This paper focuses on the various stages of KDD process along with its processing. The applications of data mining is enclosed almost every industry. The data mining techniques like prediction, classification etc. helps in the discovery of patterns which helps the industry to identify the future trends.

VI. References:

- 1) http://www.intechopen.com/books/data_mining_and_knowledge_discovery_in_real_life_applications/a_data_mining_amp_knowledge_discovery_process_model”
- 2) Yanxia Zhang and Yongheng Zhao, “Data mining in astronomy”
- 3) <http://www.zentut.com/data-mining/data-mining-processes>.
- 4) Paulraj Ponnian, .Data Warehousing Fundamentals. John Wiley.
- 5) Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*. Springer Science & Business Media.

