Name: Atul Kumar Roll No.: RD2110B79
Reg. No.: 12102801 Course Code: CAP-446
Course Instructor: Dr. Geeta Sharma

Q1:- What are the challenges Associated with Data mining, explain them in brief.

Ans:- Data mining has attracted a great deal of attention in the information industry and in society as a whole. Data Mining and information disclosure are developing a critical innovation for researchers. But these are some challenges also:-

i) Complex Data: Real data is truly heterogeneous, and it very well may be media data, natural language text, time series, audio or video, images etc. It is truly hard to deal with these various types of data.

ii) Performance: The presentation of the data mining framework basically relies upon the productivity of techniques and algorithms utilized.

If the techniques and algorithms planned are not sufficient then the influence of the presentation of the data mining measure unfavorably.

iii) Security and Social Challenges

Private information about people and touchy information is gathered for the client's profiles so it requires impressive security. Information and the secret idea of information turning into a significant issue.

iv) Noisy and Incomplete Data

Data mining is the way toward obtaining information from huge volumes of data. This present reality information is noisy, incomplete and heterogeneous. This is because of human mistakes blunders or errors.

v) Distributed Data: True data is normally put away on various stages in distributed processing conditions. It is hard to carry all the data to a unified data archive because of technical and organizational

vi) Scalability and Efficiency of Algorithms: The data mining algorithm should be scalable and efficient to extricate information from tremendous measures of data in dataset.

vii) Mining Methodology Challenges These difficulties ase identified with data mining methods and their limits. Mining methods that cause the issue ase the control and handling of noise in data.

viii) Data Visualization: Data visualization is a vital cycle in data mining. It is very hard to address the information in a precise and straightforeward manner to the end-user. The output information and input data being very effective, successful and complex data perception methods should be applied to make it useful.

Q2: Explain the bottom-up approach of data warehouse and give its limitations also.

Ans:- In the bottom-up approach, the data marts are created first to provide reporting capability. A data mart addresses a single business area such as sales, finance etc. These data marts are then integrated to built a complete data warehouse. The integration of data marts is implemented using data warehouse bus architecture. In the bus architecture, a dimension is shared between facts in two or more data marts. This approach is given by Kinball as data marts are created first and provide a thin view of analyses and data warehouse is created after complete data marts have been created.

These are some limitations of bottom-up approach:-

i) This model is not strong as top-down approach as dimensional view of data marts is not consistent as it is in above approach.

ii) The position of the data ware-house and the data marts are reserved in the bottom-up approach.

Q3: Where you can implement the data warehouse? Explain its uses in some areas in detailed manner.

Ans: Data Warehousing can be implement anywhere where we have a huge amount of data and we want to see statistical result, visualization that help in decision making. The social networking websites like Facebook, Twitter, Linkedin etc are based on analyzing large data sets. Being a large amount of data, Data warehouse is needed. There are also many sectors where we can implement the data warehouse like in banking sector, government sector, private sector and many more companies.

There are many sectors which uses data warehousing like:-

a) Banking: Most of the banks these days use warehouses to see the spending patterns of account/cardholder. They also use this to provide them special offers, deals etc. Data warehousing also help banking sector in the following ways:

i) Identify the potential risk of default and manage and control collections.

ii) Performance analysis of each product, services, interchange and exchange rates

iii) Provide feedback to bankers regarding customer relationships and profitability.

iv) Data Warehousing also provide high scalability and the right banking continuity.

B) In E-commerce companies: E-commerce also uses data wasehousing that makes the decision making easy. It helps you to built real, practical product and content recommedations for your customer segments. It also help to improve theis bottom line. Data wasehouse allow business leaders to quickly access theis organization's historical activities and evalleate initiatives that have been successful or unseccessful in the past. This allows executives to see where they can adjust theis strategy to decrease costs, maximize efficiency and increase sales. It also improve the speed and efficiency of accessing different data sets and makes it easier for corporate decision-makers to derive insights that will guide the business strategies.