Unit-5

Application of Data housing and Data mining

Data warehouse

↳ It is a repository of an organizations electronically stored data and are designed to facilate reporting and analysis.

↳ It also emphasises on the means to retrieve and analyze data, to abstract, transform and load data and to manage data dictionary.

↳ Mainly data warehousing includes business intelligence tools to abstract, transform and load data into repository and tools to manage and retrieve meta data (data about data).

↳ A data warehouse is a collection of computer based information that is critical to successful execution of enterprise initiatives.

Advantages of Data warehouse

1. More cost effective decision making
2. Better enterprise intelligence
3. Enhance customer service
4. Business re-engineering

Data mining

Data mining is the process of abstracting hidden predictive information from a large database. As more data are gathered, with the amount of data doubling every year, data mining is becoming an increasingly an important tool (technique) to transform the data into information.

It is commonly used in wide range of profiling practices such as marketing, fraud detection and scientific discovery. Data mining can be applied to data set of any size. Data mining tools predict future trends and behaviours allowing business to make proactive knowledge driven decision.

Data mining somethings also known as data or knowledge discovery. is the process of analyzing data from different perspectives and summarizing it into useful information. Technically, data mining is a process of finding co-relation or patterns among dozen of fields in large relational database.

Data Mining consists of five major elements :

Data mining consists of five major elements they are listed below :-

(i) Abstract, transform and load transaction data onto the data warehouse system.

(ii) Store and manage the data in a multi-dimensional data base system.

(iii) Provide data access to business analyst and information technology professional.

(iv) Analyze the data by application software.

(v) Present the data in a useful format such as table or graph.

Advantages of data mining

1. Automated prediction of trends and behaviour.

2. Automated discovery of previously unknown pattern.

3. Database can be larger in both depth and breadth.

Data warehousing and Data mining in E-government

Introduction

Data warehousing and data mining are the important means of preparing the government to face the challenges of the new millennium (era). This technologies have extensible potential application in the government. Mostly in various central government sectors such as agricultural, rural development, health and energy and also in state government activities. This technology should be implemented throughly in the development sector in order to carryout the task and activities smootly.

National Data warehouse

A large number of national data warehouse can be identified from the existing data sources within the central government ministries. National Data warehouse contains the data regarding the particular nation under different headings and subject areas. Such data warehouse helps the government of particular nation in order to make different plans and policies and implement them successfully. and helps the government in future.

Census Data

The data collected, gathered, compiled and normalized after census is known as census data. The census data are also stored in data warehouse so that it can be used in future. As the census compilation is performed once in ten years, the data is quasi static and therefore no refreshing of the warehouse needs to be done on the periodic basis. Only the new data needs to be either appended to the data warehouse or alternatively a new data warehouse can be built. There exist many other subject areas (eg. migration table) within the census purview (concern) which may be amenable and appropriate for dataware development; OLAP (Online (change garna sakna) Analytical Processing) and data mining application on which work can be taken off in the future.