

Data warehouse and miningUnit 4 - Data mining techniques (of Post)OLAP (Online Analytical Processing)

OLAP is a category of software that allows users to analyze information from multiple database systems at the same time.

It is a technology that enables analysts to extract and view business data from different points of view. OLAP stands for Online Analytical Processing.

Analysts frequently need to group, aggregate and join data. These operations in relational databases are resource intensive. With OLAP data can be pre-calculated and pre-aggregated, making analysis faster.

OLAP cube is a data structure optimized for very quick data analysis. OLAP cube consists of numeric facts called measures which are categorized by dimension. OLAP cube is called the hypercube.

A datawarehouse would extract information from multiple data sources and formats like text files, excel sheets, multi-media files etc.

The extracted data is cleaned and transformed. Data is loaded into an OLAP server where information is pre-calculated in advance for further analysis.