Department Of Statistics- MCA ,PU – Jai bardhan kanth

Data Mining
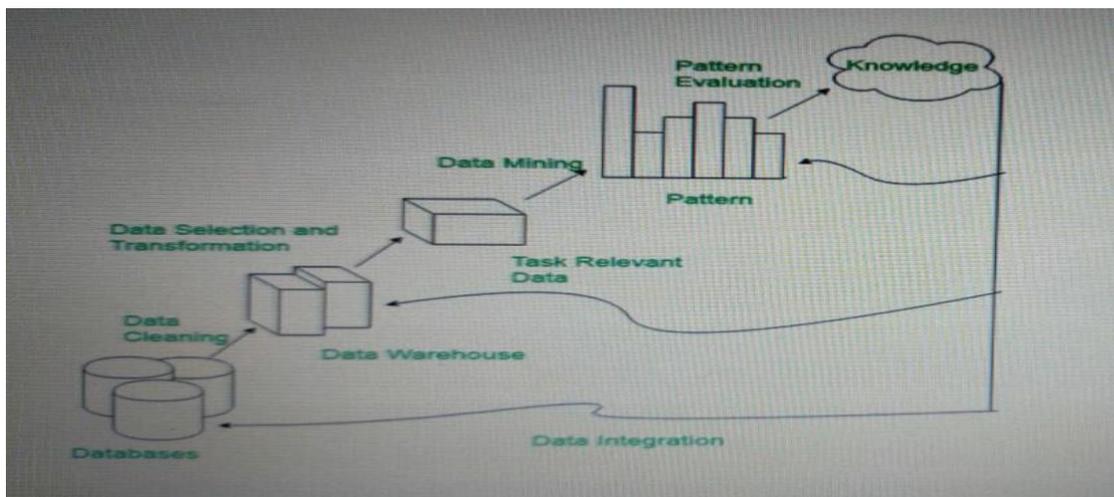
CS↔52

Date – 10th september 2020

⚙ **KDD** (**Knowledge Discovery Database)**

→ **KDD Process in data mining:-**

The volume of data increases everyday and we know that the data can handle business transaction, scientific computation etc .So we need a system that will be capable of extracting from the available data in the form of report , view and summary to make a better decision. The term KDD comes existence as a way to find knowledge in data. We needs high level applications for particular methods which able to extract knowledge from the describe knowledge from the given data. The main goal of KDD process is to extract knowledge from data in the context of large database. The process KDD using the techniques of data mining such as pre processing extracting and finally representing the data from the database.

▶ **These are the following steps in KDD to describe knowledge from the stored database:-**

   i.)   **Data Cleaning**
  ii.) **Data Integration**
 iii.) **Data Selection**
 iv.) **Data Transformation**
  v.) **Data Mining**
 vi.) **Pattern Evaluation**
vii.) **Knowledge Representation**

i.) Data cleaning:-

The noisy data removes from large volume data.This process is known as data cleaning. In this step extract only the relevant data from the entire storage .

ii.) Data integration:-

It may be possible that the data exist from heterogeneous data sources . So this step apply data migration tools, Data synchronization tools and ELT (Extract Load Transformation) tools etc which organize the data inside data warehouse.

iii.) Data selection:-

This step of KDD apply SQL or PL/ SQL query upon the data warehouse available after data integration which decides the data retrieval for knowledge. In this stage various types of tools such as neural network, Decision tree, Base algorithm, clustering and regression applied.

iv.) Data transformation:-

In this step appropriate data mapped by data mapping and code generation for the data description .

v.) Data mining:-

In this step data extracted potentially in the form of patterns and the purpose of this step is classification or characterization .

vi.) Pattern evaluation:-

In this step the defined pattern majors in the form of knowledge and the patterns appear in the form of summary for the visualization, that user can understand.In this stage a user can view the organised data in the form of image , audio, video or tabular form .

vii.) Knowledge representation:-

This step of KDD is also known as utilization step in which the data available in the form of report, summary from the large data stored inside the huge storage area.