Data Mining & Web Algorithms

Date / /
Page No.

Tutorial 1

Patil Amit G.
19104004
B11

Q.1. a) No
b) No
It of offers more to the traditional databases namely context and meaning to the data.
c) KDD is the root i.e. knowledge discovery processes. Data retrieval has lead to the term of data mining.

Q.2. a) No
b) ~~Yes~~ No
c) Yes
d) No
e) Yes
f) Yes
g) Yes
h) Yes
i) No

Q.3. Predicting best doctor for a perticular disease in a doctor-patient application.
By using data-mining techniques we can ~~use~~ leverage historical data and figure out who is the best doctor to cure perticular disease.
We need classification.

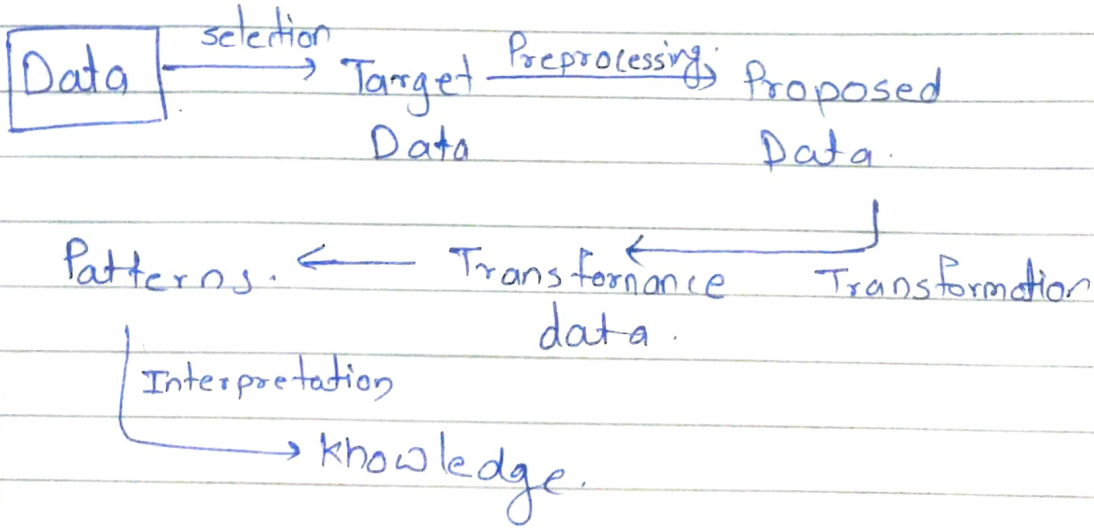
This can not be performed by data query processing or simple statistical analysis.

Q.4.

Database	Data Warehouse
① Organized collection of related data which stores data in a tabular format.	① A central location which stores consolidated data from multiple databases.
② Detailed data	② Summarized data.
③ Uses Online Transactional Processing (OLTP)	③ Uses online Analytical processing (OLAP)
④ Tables & joins are complex 'cause they are normalized	④ Tables & joins are simple 'cause they are decentralized.

Q.5. Data

Q.6.



1) Developing an understanding of.
   1) application domain.
   2) relavant prior knowledge.
   3) goals of the end-user.

2) Creating data set : focusing on subset of variables or data samples.

3) Data cleaning : Finding null values and abnormal values and removing them from dataset.
   1) Removel of outliers.

4) Data reduction & Projection. //