DMBI.

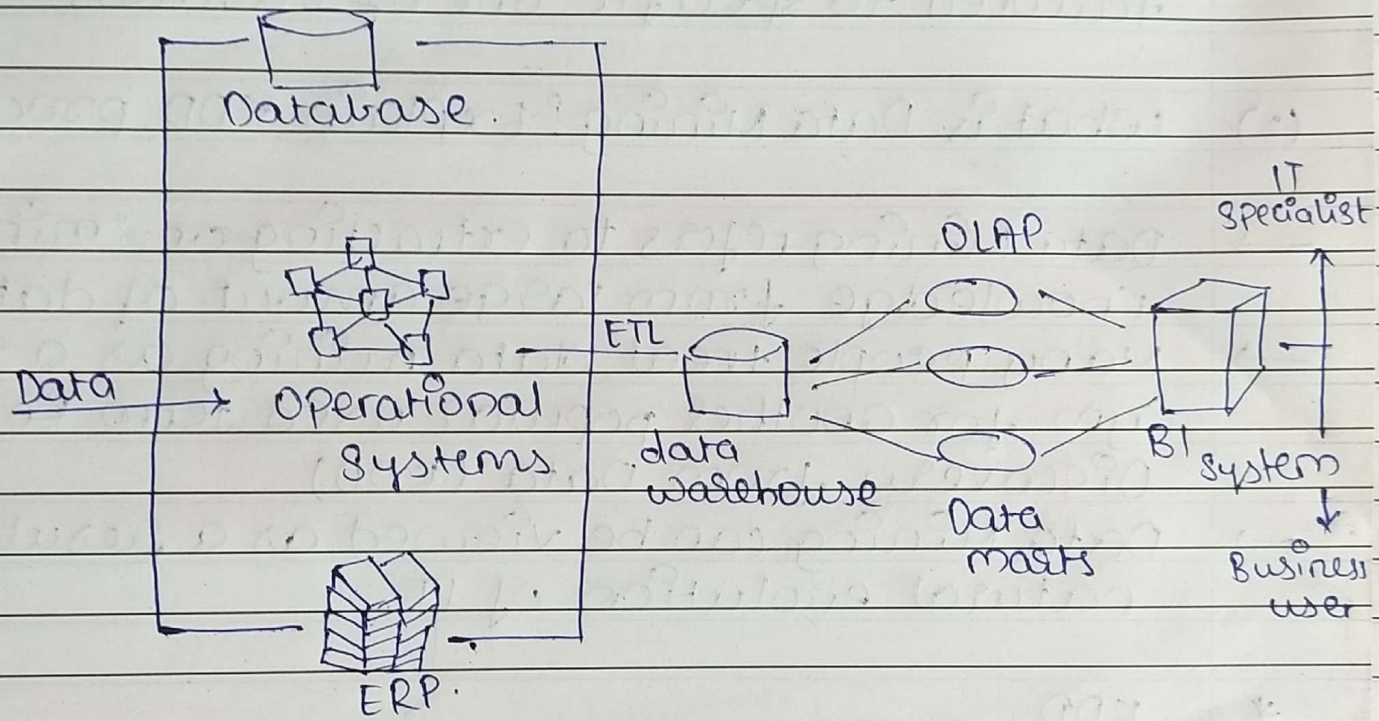
Meghpara Jeet. J
160410116056

* ASSIGNMENT-1 *

Date: L4 IT-1

(1) Explain BI architecture with its components.

→



→ Forrester defines BI broadly as a "set of methodologies, processes, architecture, & technologies that transform raw data into meaningful & useful info that allow user to make informed business decision & decisions with real-time data".

→ BI traditionally focused on reports, dasboards, and answering predefined questions.

→ BI system evolved by adding layer of data staging to increase accessibility of business data to business user.

→ Data from warehouse were loaded in OLAP cubes, as well as data marts stored in data warehouse.

YASH

Royal ECO

Date:

→ OLAP cubes facilitated analysis of data over several dimensions.

→ Data marts present subset of data in warehouse, tailored to specific line of business.

(2) What is Data Mining? Explain KDD process.

→ Data Mining refers to extracting or "mining" knowledge from large amount of data.

→ Many people treat data mining as a synonym for another popular used term Knowledge Discovery from Data (KDD)

→ Data Mining can be viewed as a result of natural evolution of IT.

* KDD.

— There are some pre-processing operations which are required to make pure data in data warehouse before use that data for Data Mining processes.

→ Iterative sequence of follo steps are:

(i) Data cleaning:
To remove noise & inconsistent data.

(ii) Data integration:
where multiple data sources may be combined.

(iii) Data selection:
data retrieved from database.

(iv) Data transformation:
data transformed in required form

YASH

Royal ECO

Date:

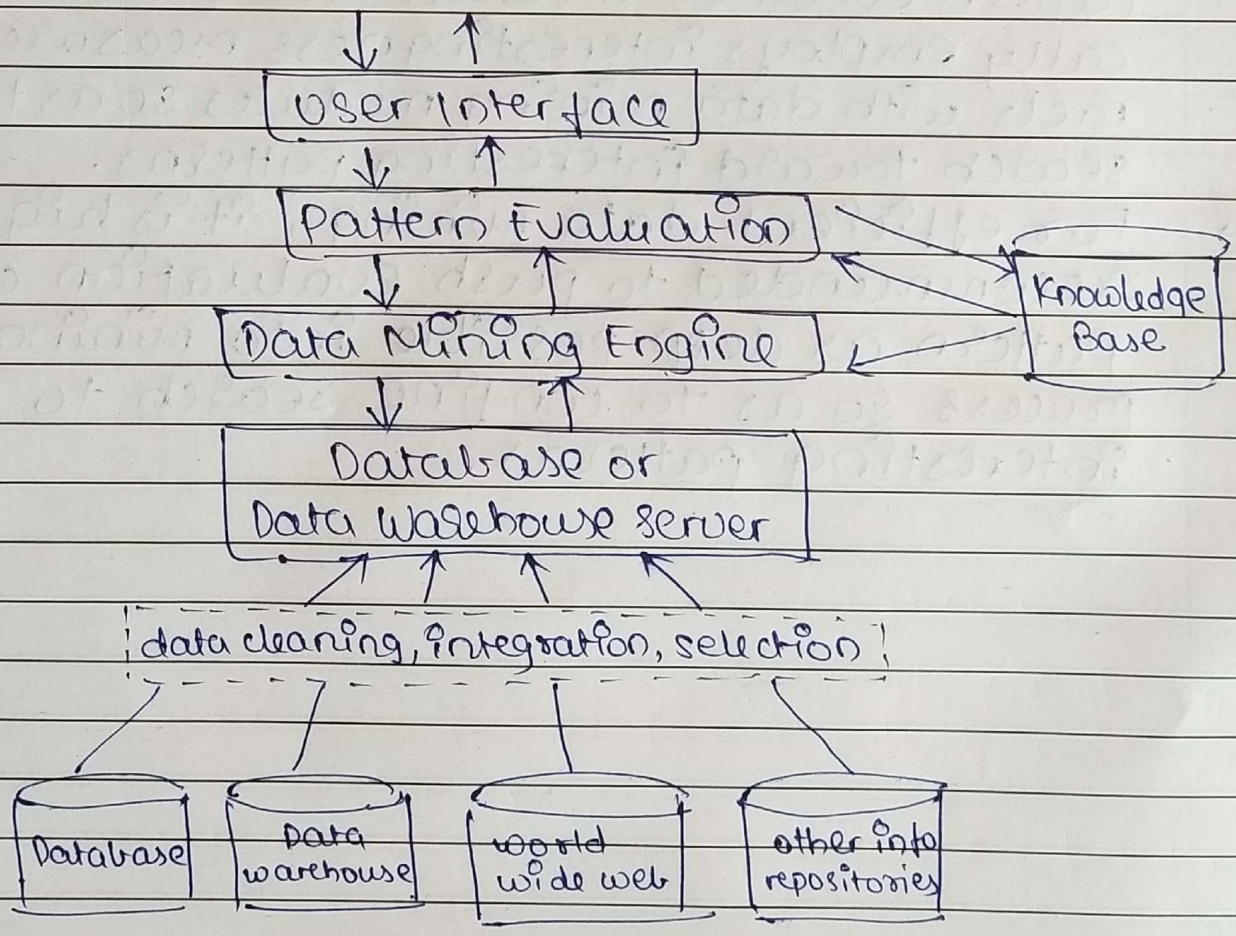
(v) Data mining:
Intelligent methods are applied in order to extract data patterns.

(vi) Pattern evaluation:

(vii) Knowledge presentation:
where visualization & knowledge repr technique are used to present mined knowledge to user.

(3) Explain Architecture of Data mining system with major compo involved.

→



YASH

Royal ECO

Date:

→ Knowledge base: This is domain knowledge that is used to guide search or evaluate interestingness of resulting patterns.

→ Data warehouse typically provide simple & concise view around particular subject issues by excluding data that are not useful decision support process.

→ Data mining engine: It consists of a set of functional modules for tasks such as characterization, association & correln, classification, prediction, and evolution analysis.

→ Pattern evolution module: This component typically employs interestingness measures & interacts with data mining modules so as to focus search toward interesting patterns.

→ For efficient data mining, it is highly recommended to push evaluation of pattern as deep possible into mining process so as to confine search to only interesting patterns.