

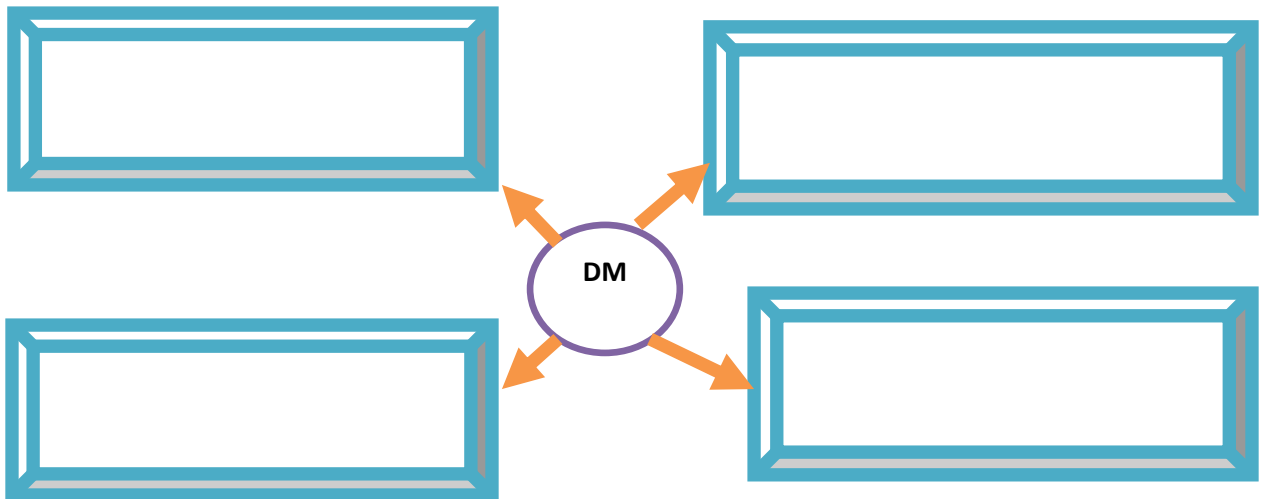


WORKBOOK

UNIT III

TOPIC 1: INTRODUCTION TO DATA MINING

1. Write the other names for Data Mining



2. Which is data/information?

- Each student's test score.
- The average score of a class or of the entire school



3. Which leads to wisdom? Write the hierarchy.



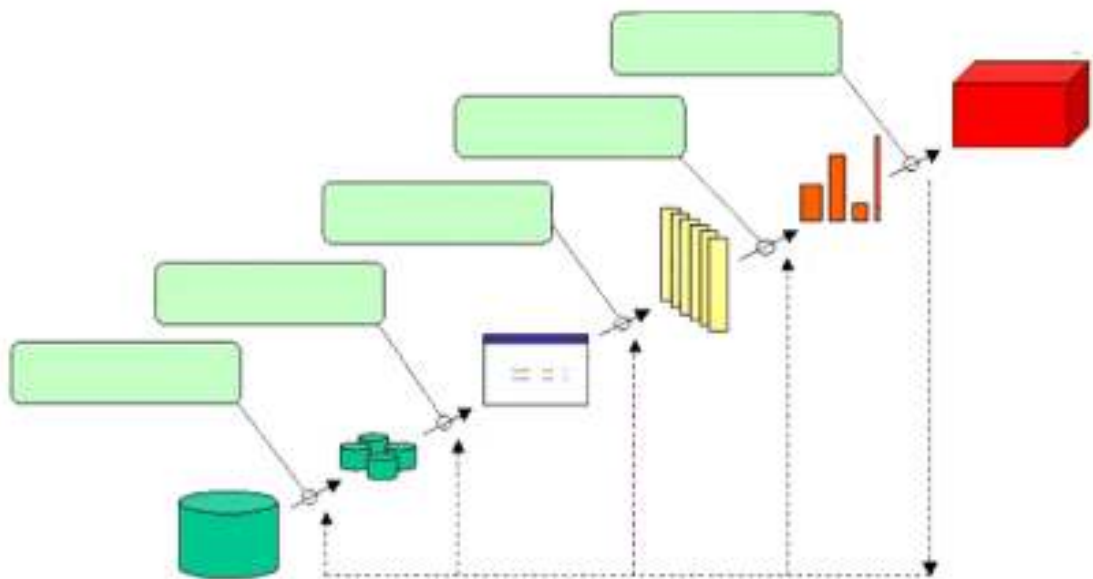
data

TOPIC 2: TYPES OF DATA

1. Write relational tables for customer, item, employee, and branch.

[illegible]

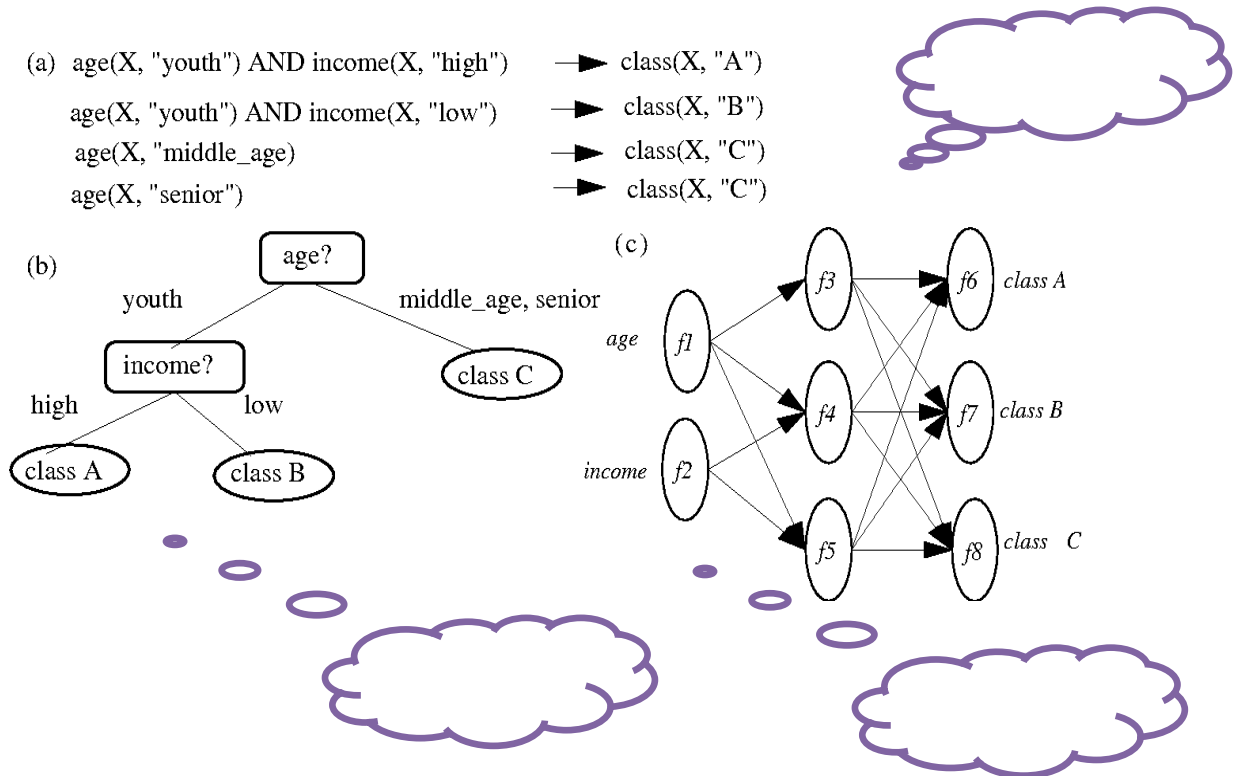
2. Fill the diagram to make as the step for KDD process



3. Mention other kinds of data except Relational database, Data warehouse and Transactional Database.

TOPIC 3: DATA MINING FUNCTIONALITIES

1. Mention the various classification models described here.



2. Match the following

S.NO	PARTICULARS	DESCRIPTION
1	Frequent item sets	refers different structural forms (eg graphs, trees) that are combined with itemsets or subsequences.
2	Frequent Subsequences or sequential pattern	refers to a set of items that frequently appear together in a transaction data set, such as milk and bread.

3	Frequent Substructures	pattern that customers tend to purchase : first a PC , followed by camera, then a memory card.
---	------------------------	--

3. Define Support and Confidence

TOPIC 4: INTERESTINGNESS OF PATTERNS

1. What makes pattern interesting?



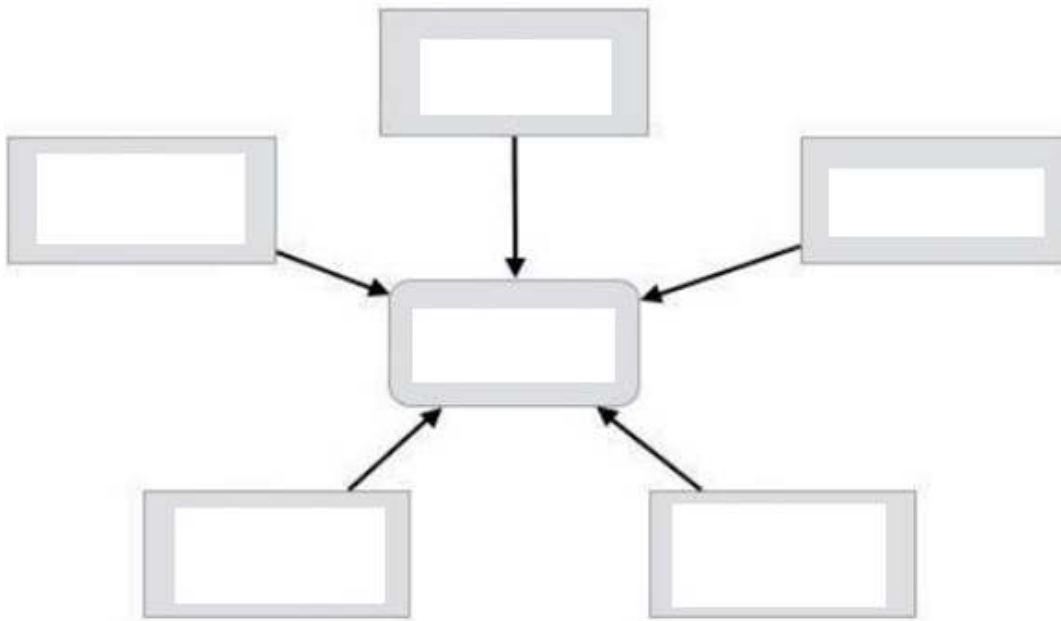


2. Write the definition for Conditional Probability $P(Y|X)$ and $P(X|Y)$

3. Can a DM system generate only interesting patterns?

TOPIC 5: CLASSIFICATION OF DATA MINING SYSTEMS

1. Data mining incorporated many techniques from other domains such as,



2. Calculate the mean for the following.
30,36,47,50,52,52,56,60,63,70,70

3. Compare supervised, unsupervised and semi-supervised learning.

TOPIC 6: DATA MINING TASK PRIMITIVES

1. List the DM Task Primitives

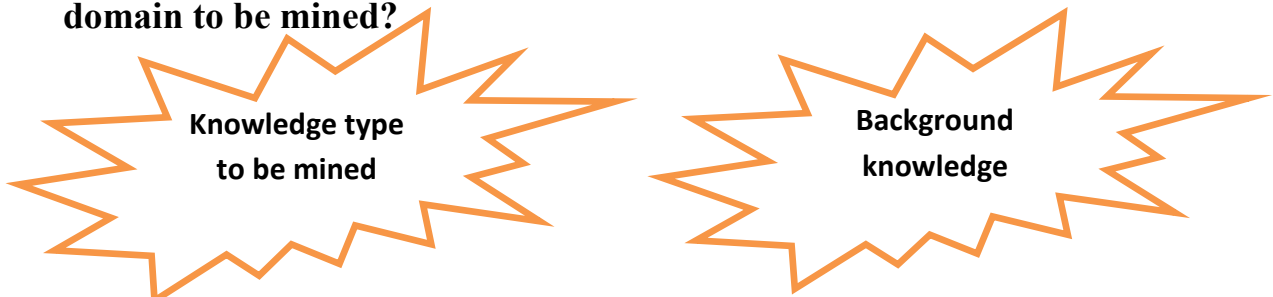
2. Which primitive specifies the data upon which mining is to be performed?



3. Which primitive specifies the specific data mining function to be performed, such as characterization, discrimination, association, classification, clustering, or evolution analysis?

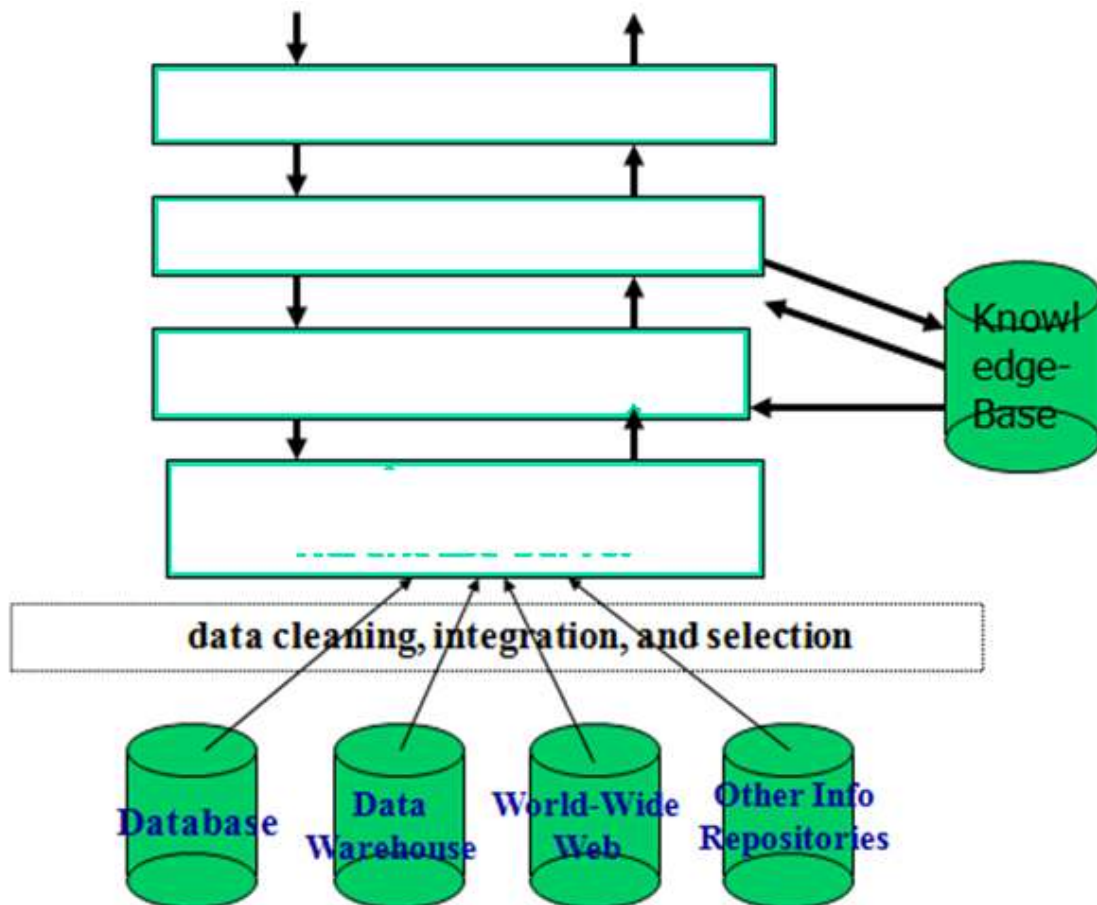


4. Which primitive allows users to specify knowledge they have about the domain to be mined?



TOPIC 7: INTEGRATION OF DATA MINING SYSTEMS WITH A DATA WAREHOUSE

1. Complete the diagram for integration of DM with DW.



2. What are the possible integration schemes?



3.

Means that a DM system will not utilize any function of a DB OR DW system

Means that a system will use some facilities of a DB or DW system, fetching data from a data repository managed by these systems, performing DM and then storing the mining results either