Rohan Nyati 50007594O (148) Batch-5

Dynamic Paradigm

Sec-B

1) Knowledge discovery ~~from~~ in databases
2) iterative
3) set of values
4) Regression analysis
5) mean square error

Sec-C

1) False
2) True
3) False
4) False
5) True

Sec-A

Q.1) Data mining is a process used ~~for~~ by companies to turn raw data into useful information. By using software to look for patterns in large batches of data, businesses can learn more about their customers to develop more effective marketing strategies, increase sales and decrease costs. Data mining depends on effective data collection, warehousing and computer processing.

Q.2] Data mining refers to the process or method that extracts or "mines" interesting knowledge or pattern from large amounts of data. Data mining involves an integration, rather than a simple transformation, of techniques from multiple disciplines such as statistics, machine learning, pattern recognition, data visualization, neural networks, etc.

Q.3] Stages of Data mining process :-

a.) Data cleaning → data can have many irrelevant and missing parts. To handle this part, data cleaning is done. It involves handling of missing data, noisy data etc.

b.) Data Integration → It tries to reduce redundancy as much as possible without affecting the reliability of the data

c) Data Selection → This is the process by which data relevant to analysis is retrieved from the database.

d) Data Transformation → it includes normalization, aggregation, generalization etc.

e.) Data Mining
f) Pattern evaluation
g) Knowledge Representation

Q.4] A Data Warehousing (DW) is process for collecting and managing data from varied sources to provide meaningfull business insight. It is the electronic storage of a large amount of information by a business or organization. It is a vital component of business intelligence that employs analytical techniques on business data.

Q.5] Data Understanding: Review the data you have, document it, identify data management and data quality issues. Tasks for this phase include:

1) Gathering data
2) Describing
3) Exploring
4) Verifying Quality

Q.6] Types of Relationships:-

1) one-to-one Relationship
2) one-to-many Relationship
3) Many-to-many Relationship

For this, several approaches are:-

1) Inductive Logic Programming (ILP)
2) Statistical Relational Learning (SRL)
3) Graph Mining
4) Propositionalization
5) Multi-view learning