~~> DMDV Imp ques. - Refer my notes & questions given in that.

* Ch-4 Classification and Clustering.

- Classification vs Clustering
- Classification vs Prediction
- Supervised vs Unsupervised learning.
- Decision Tree Induction
  - ↳ MIMP (numerical)
    - Decision tree using ID3 - entropy & info gain
      - ↳ weather/football numerical to find root node

Classificⁿ

- Naive Bayes Classification IMP
  - ↳ theory & theorem
  - ↳ numerical (fruit example) (numerical may come)
- Partitioning method
  - ↳ K-Means - MIMP numerical } distance formulas
  - ↳ K-medoid
- Hierarchical method
  - ↳ Agglomerative } theory and diagram
  - ↳ Divisive

Clustering

- DBSCAN algorithm
- DBSCAN Reachability. - 3 methods
- Evaluation of clustering - 3 methods
- Outlier Detection.

* Ch-5 Statistical Representation of Data.

- Numerical - mean, mode, median, variance, standard deviaⁿ
- Dispersion of data - dispersion, range, quartile, IQR, 5-no. summary, outlier.
- Dispersion by boxplot analysis.
- Histogram vs Bar graph.

* Data visualization = Quantile Plot, Quantile-Quantile Plot, Scatter Plot, Loess Curve

* Ch-6 = Theoretical as per session on Power BI

* Ch-1 Intro to Data mining.

→ Theoretical

- Data mining defi, importance ✓
- Data mining architecture ✓✓
- KDD ✓✓
- Major issues in data mining ✓
- Applications ✓✓

* Ch-2 Data pre-processing

- Tasks in data pre-processing – 4 tasks ✓
- Data cleaning ✓✓✓
  - → fill missing data
  - → noisy data & it's handling.
  - → Clustering
  - → regression
  - → Binning method numerical ✓✓✓
- Data Transformation.
- Data Reduction – Attribute subset selection
- Data Discretization
- Concept Hierarchy generation

* Ch-3 Mining Frequent Patterns

- Frequent pattern mining
- Association Rule Mining ✓✓✓
  - ↳ with example ✓✓
- Apriori algorithm
- FP-growth algorithm