**UNIT 1**

1. Desigen Guidelines for data warehouse.
2. Multidimension model.
   1. Types:
      1. Star Schema.
      2. Snowflake Schema.
      3. OLAP (Online Analytical Processing)
3. OLAP-Introduction:
4. Characteristics OLAP.
   1. Type of Characteristics:
      1. Multidimensional Data Model.
      2. Interactive Analysis.
      3. Fast Query Performance.
      4. Complex Calculation.
5. Architectures of OLAP.
6. Multidimensional view efficient processing of OLAP queries.
7. ROLAP.
8. MOLAP.
9. HOLAP.
10. Data Cube.
11. Data Cube Operations.
    1. Types:
       1. Roll-up
       2. Drill-Down
       3. Slick
       4. Dice
       5. Pivot
       6. Drill-across
       7. Ranking and Sorting
       8. Calculation and Formulas.
       9. Time-series Analysis.
12. Data Cube Computations.
    1. Types of Computations:
       1. Summation
       2. Average
       3. Minimum and Maximum
       4. Percentages and Ratio.
       5. Growth Rage
       6. Moving Averge.
       7. Correlation and Covariance.
       8. Percentile and Quartile.
       9. Forecasting and predictive analytics.
13. Data Mining.
14. Challenges:
    1. Types:
       1. Data Volume
       2. Data Variety
       3. Data Velocity.
       4. Data Quality.
       5. Privacy and Security.
       6. Scalability.
15. Data Mining Tasks:
    1. Classification.
    2. Regression
    3. Clustering.
    4. Association Rule Mining.
    5. Anomaly Detections
    6. Text Mining.
16. Data.
    1. Types:
       1. Structured Data.\
       2. Semi Structured Data.
       3. Unstructured Data.
17. Data Quality.
18. Data pre-processing.
19. Measures of Similarity and Dissimilarity.

**UNIT 2**

1. **Data Mining.**
2. **Association rule mining.**
   1. **Process.**
      1. **Itemset.**
      2. **Support.**
      3. **Confidence.**
      4. **Association Rules.**
   2. **Application:**
      1. **Market Basked Analysis.**
3. **Naïve Algorithm.**
4. **Apriori Algorithm.**
   1. **Components:**
      1. **Support.**
      2. **Confidence\**
      3. **Lift.**
5. **Direct Hashing and Pruning in Data Mining.**
6. **Dynamic Item set Counting.**
   1. **Steps:**
      1. **Candidate Generation**
      2. **Dynamic Counting.**
      3. **Pruning.**
      4. **Candidate Refinement.**
7. **Mining frequent pattern without candidate generation (FP, growth).**
   1. **Steps:**
      1. **Building the FP Tree.**
      2. **Creating Conditional Pattern Bases.**
      3. **Mining Frequent Pattern.**
8. **Performance evaluation of Algorithm.**
   1. **Methods:**
      1. **Accuracy**
      2. **Precision and Recall.**
      3. **F1 Score.**
      4. **Speed and Efficiency.**
      5. **Robustness and Stability.**
      6. **Generalization Ability.**
      7. **Cross-Validation and Testing.**
      8. **Comparative Analysis.**
9. **Classification.**
10. **Decision Tree.**
    1. **Components:**
       1. **Root Node.**
       2. **Internal Node.**
       3. **Leaf Node.**
       4. **Splitting Criterion.**
    2. **Construction a Decision tree involves:**
       1. **Attributes**
       2. **Splitting**
       3. **Recursion**
       4. **Pruning.**
11. **Tree introduction algorithms -split algorithm based on information theory**
12. **split algorithm based on Gini index split algorithm based on Gini index:**
    1. **Used to Splitting:**
       1. **Calculation of Gini Index.**
       2. **Splitting Criteria Selection.**
       3. **Recursive Partitioning.**
       4. **Stopping Criteria.**
13. **Navis Bayes method.**
    1. **How is works:**
       1. **Baye’s Theorem**
       2. **Naïve Assumption**
       3. **Model Training**
       4. **Classification.**
14. **Estimating predictive accuracy of classification method**
    1. **Methods:**
       1. **Train-Test Split Method.**
       2. **Cross Validation**
       3. **Confusion Matrix.**
       4. **Accuracy.**
       5. **Precision and Recall.**
       6. **F1 Score**

**UNIT 3**

1. **Cluster Analysis.**
2. **Partition Methods.**
   1. **K-means Clustering.**
   2. **Fuzzy C-means.**
3. **Hierarchical Methods.**
   1. **Agglomerative Clustering.**
   2. **Divisive Clustering.**
4. **Density Based Methods.**
   1. **DBSCAN**
   2. **OTICS.**
5. **Dealing with large Database.**
6. **Cluster Software.**
7. **Search Engines.**
8. **Characteristic of Search Engine.**
   1. **Indexing**
   2. **Crawling**
   3. **Querying**
   4. **Ranking Algorithm**
   5. **User Interface.**
   6. **Relevance.**
9. **Search Engine Functionaity.**
   1. **Web Crawling**
   2. **Indexing**
   3. **Query Processing**
   4. **Ranking and Retrieval**
10. **Search Engine Functionality.**
11. **Search Engine Architecture.**
12. **Ranking of Web pages.**
13. **Search Engine History.**
14. **Enterprise Search.**

**UNIT 4**

1. **Web Data Minig.**
2. **Web Terminology and Characteristics.**
3. **Locality and Hierarchy in the Web.**
4. **Web Content Mining.**
5. **Web Usage Mining.**
6. **Web Structure Mining.**
7. **Web Mining Software.**