

Assignment -2

Q1 Write an R programming code for Clustering?

Ans. Clustering is a technique used to group similar data points together. One of the most popular clustering algorithms is K-Means Clustering.

Code:

library(ggplot2)

# Generate sample data

set.seed(123)

data <- data.frame(

x = rnorm(100, mean = 5, sd = 2),

y = rnorm(100, mean = 5, sd = 2)

)

# Apply K-Means clustering with 3 clusters

kmeans\_result <- kmeans(data, centers = 3)

# Add cluster labels to the dataset

data$cluster <- as.factor(kmeans\_result$cluster)

# Visualize the clusters

ggplot(data, aes(x, y, color = cluster)) +

geom\_point(size = 3) +

ggtitle("K-Means Clustering in R")

Q2. Write an R programming code for Frequent Pattern Mining Algorithm?

Ans: Frequent Pattern Mining is used in market basket analysis to find commonly occurring patterns in data. The Apriori Algorithm is widely used for this purpose.

Code:

library(arules)

# Create a sample dataset (transactions)

transactions <- list(

c("Milk", "Bread", "Butter"),

c("Milk", "Bread"),

c("Milk", "Butter"),

c("Bread", "Butter"),

c("Milk", "Bread", "Butter", "Eggs"),

c("Milk", "Eggs"),

c("Bread", "Eggs")

)

# Convert to transaction format

trans <- as(transactions, "transactions")

# Apply Apriori Algorithm

rules <- apriori(trans, parameter = list(supp = 0.2, conf = 0.6))

# Display the results

inspect(rules)

Q3. Explain any one BI tool used in Data Mining and Business Intelligence

Ans:

Tableau: Tableau is a popular Business Intelligence (BI) tool used for data visualization and analytics. It helps organizations analyze, visualize, and share insights from large datasets.

Features of Tableau:

1. Interactive Dashboards – Allows users to create real-time dashboards.
2. Drag-and-Drop Interface – Easy to use for data exploration.
3. Integration with Databases – Supports multiple databases like MySQL, PostgreSQL, and Excel.
4. Data Blending & Transformation – Enables merging data from different sources.
5. Cloud & On-Premise Deployment – Available for cloud and local use.

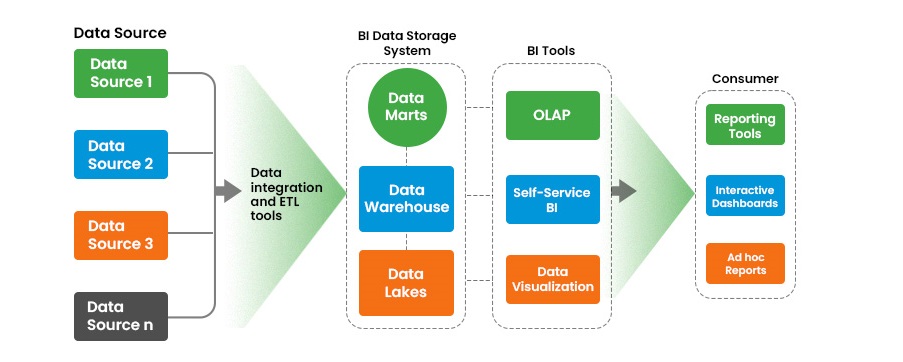
Example Use Case:

* A retail company uses Tableau to analyze sales trends, identify best-selling products, and optimize inventory management.

Q4. Define BI and Give Its Architecture?

Ans: Business Intelligence (BI) refers to technologies, processes, and tools used to collect, store, and analyze business data for decision-making. BI helps organizations gain insights from their data to improve efficiency and strategy.

BI Architecture



BI architecture consists of four main components:

1. Data Source Layer
   * Includes databases (SQL, NoSQL), cloud storage, and data warehouses.
2. ETL (Extract, Transform, Load) Layer
   * Extracts data from multiple sources, transforms it into a standard format, and loads it into a data warehouse.
3. Data Storage Layer
   * Stores structured and unstructured data in data warehouses like Amazon Redshift, Snowflake, or Google BigQuery.
4. Data Analytics & Visualization Layer
   * Uses BI tools (Tableau, Power BI, Google Data Studio) for reporting and analysis.

Q5. Explain Business Intelligence Issues ?

Ans: BI systems face several challenges that affect their implementation and performance.

Key Business Intelligence Issues:

1. Data Quality Issues
   * Inconsistent, duplicate, or incomplete data leads to incorrect insights.
2. Integration Problems
   * Combining data from multiple sources can be complex and time-consuming.
3. Performance & Scalability
   * Large datasets require high processing power, leading to slow performance.
4. Security & Privacy Risks
   * Protecting sensitive business data from breaches is a major challenge.
5. User Adoption & Training
   * Employees need training to effectively use BI tools and interpret data insights.
6. Cost of Implementation
   * Setting up BI infrastructure can be expensive for small businesses.

Q6. Explain Any Data Mining Application Where Business Intelligence Can Be Used

Ans: Application: Customer Churn Prediction

Customer churn prediction helps companies identify customers at risk of leaving and take preventive actions.

How Business Intelligence is Used:

1. Data Collection
   * Collects customer data from CRM systems, transaction history, and social media.
2. Data Processing & Analysis
   * Uses BI tools like Power BI or Tableau to analyze patterns in customer behavior.
3. Machine Learning for Prediction
   * Data Mining algorithms (e.g., Decision Trees, Random Forests) predict which customers might leave.

4.Decision Making

* + Businesses can offer personalized discounts or loyalty programs to retain high-risk customers.

Example:

* Telecom Companies (e.g., Vodafone, Airtel) use BI to predict customer churn and improve customer retention.