**Final Assessment (35 Marks)**

Data Mining Task Assignment for Cluster Analysis and Classification using Tableau and R

In this assignment, students will engage in a comprehensive data mining task that involves cluster analysis and classification using Tableau and R.

The assignment will cover various aspects of data preprocessing, exploratory data analysis, modeling, and evaluation to provide students with hands-on experience in utilizing these powerful tools for data analysis.

**Task Overview:**

Data Preprocessing:

* Clean and preprocess a dataset (Collected by Students) using R to ensure data quality and consistency.

Exploratory Data Analysis (EDA):

* Perform EDA to gain insights into the dataset, identify patterns, and understand the distribution of data.

Cluster Analysis:

* Utilize Tableau's clustering feature to automatically cluster similar data points based on specific criteria.
* Implement cluster analysis using R to further explore and analyze data clusters.

Classification:

* Build predictive models for classification using techniques such as Decision Trees, Naïve Bayes Classification, and Support Vector Machine in R.
* Evaluate the performance of the classification models using metrics like accuracy or precision.

Integration of Tableau and R:

* Integrate Tableau and R to leverage the strengths of both tools for data visualization and statistical analysis.
* Create calculated fields in Tableau that interact with R scripts to enhance data analysis capabilities.

**Deliverables:**

* Data Mining Report: (5 marks))

Include a detailed report documenting the data preprocessing steps, EDA findings, cluster analysis results, and classification model performance.

* Visualization Dashboard: (5 marks)

Develop an interactive dashboard in Tableau that showcases the clustered data points and classification results.

* Code Implementation: (5 marks)

Provide the R scripts used for cluster analysis and classification, along with explanations of the code logic.

* Presentation/Viva (20 marks), **findings with meaningful, comprehensive analysis and good data visualization will get better marks.**