410255:LP-VI BI	
Experiment No: 5	Perform the data classification using classification algorithm. Or Perform the data clustering using a clustering algorithm.

Objective of the Assignment : To introduce the concepts and components of Business Intelligence (BI)

Prerequisite:

1. Basics of Tableau.

Contents for Theory :

1. What is Clustering and classification?
2. Clustering in Tableau:
3. Classification in Tableau:

Theory:

1. What is Clustering and classification?

Clustering and classification are two important techniques used in bioinformatics to analyze biological data.

Clustering is the process of grouping similar objects or data points together based on their similarity or distance from each other. In bioinformatics, clustering is often used to group genes or proteins based on their expression patterns or sequences. Clustering can help identify patterns and relationships between different genes or proteins, which can provide insights into their biological function and interactions.

Classification, on the other hand, is the process of assigning a label or category to a new observation based on its features or characteristics. In bioinformatics, classification is often used to predict the function or activity of a new gene or protein based on its sequence or structure.

Classification can help identify new drug targets or biomarkers for disease diagnosis and treatment.

Both clustering and classification are important tools for analyzing large and complex biological datasets and can provide valuable insights into the underlying biological processes.

Clustering in Tableau:

1. Connect to the data: Connect to the data set that you want to cluster in Tableau.

2. Drag and drop the data fields: Drag and drop the data fields into the view, and select the data points that you want to cluster.

3. Choose a clustering algorithm: Select a clustering algorithm from the analytics pane in Tableau. Tableau provides several built-in clustering algorithms, such as K-Means and Hierarchical Clustering.

4. Define the number of clusters: Define the number of clusters that you want to create. You can do this manually or let Tableau automatically determine the optimal number of clusters.

5. Analyze the clusters: Visualize the clusters and analyze them using Tableau's built-in visualizations and tools.

Classification in Tableau:

1. Connect to the data: Connect to the data set that you want to classify in Tableau.

2. Drag and drop the data fields: Drag and drop the data fields into the view, and select the target variable that you want to predict.

3. Choose a classification algorithm: Select a classification algorithm from the analytics pane in Tableau. Tableau provides several built-in classification algorithms, such as Decision Trees and Random Forest.

4. Define the model parameters: Define the model parameters, such as the maximum tree depth or the number of trees to use in the forest.

5. Train the model: Train the model on a subset of the data using Tableau's built-in cross-validation functionality.

6. Evaluate the model: Evaluate the accuracy of the model using Tableau's built-in metrics, such as confusion matrix, precision, recall, and F1 score.

7. Predict the target variable: Use the trained model to predict the target variable for new data.

8. Visualize the results: Create visualizations to communicate the results of the classification analysis using Tableau's built-in visualization tools.

Conclusion : In this way we implement classification and clustering using Tableau.

Assignment Questions :
1. What is Classification and Clustering?
2. List the Classification Algorithms.
3. List the Clustering Algorithms.