

Name: Vijay Gurung

ID Number: MST03-0070

Date: 25-June-24

Task: - Difference between Classification Vs Clustering & Regression VS Classification.

Question1:-

Q.1 Different between Classification Vs Clustering with Examples?

The key differences between Classification and Clustering with examples are as below:-

1. **Classification** is a supervised learning technique, while **Clustering** is an unsupervised learning technique.

For Example: Classifying emails as spam or not spam Supervised, but grouping customers based on their purchasing behavior Unsupervised.

2. **Classification** requires labeled training data, while **Clustering** does not.

For Example: Classifying images as dogs or cats requires labeled images. However, clustering images based on visual similarity does not need labels.

3. **Classification** aims to approximate a mapping function from inputs to discrete outputs, while **Clustering** aims to discover hidden patterns and group similar instances together.

For Example: Classifying loan applicants as high-risk or low-risk. However, grouping customers based on their credit scores and income levels.

4. **Common classification algorithms** include Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines.

Common clustering algorithms include K-Means, Hierarchical Clustering, DBSCAN, and Gaussian Mixture Models.

For Example: Using Logistic Regression to classify customers as likely to churn or not. However, using K-Means to group customers based on their demographics and purchase history.

5. Classification models are typically evaluated using metrics like accuracy, precision, recall and F1-score.

Clustering models are evaluated using metrics like silhouette score, Calinski-Harabasz index, and Davies-Bouldin index.

For Example: Evaluating a spam detection classifier based on its accuracy vs. evaluating a customer segmentation model based on the silhouette score of the resulting clusters.

6. Classification is commonly used in applications like spam filtering, sentiment analysis and credit risk assessment.

Clustering is commonly used in applications like market segmentation, image segmentation and anomaly detection.

For Example: Using classification to detect fraudulent credit card transactions. But, using clustering to identify unusual network traffic patterns.

Conclusion, while both **Classification** and **Clustering** aim to group data instances, they differ in their learning type, need for labeled data, purpose, algorithms, evaluation metrics and applications. Classification is a supervised technique that predicts discrete labels, while Clustering is an unsupervised technique that discovers hidden patterns in data.

Question2:-

Q.2 Different between Regression VS Classification with Examples?

The key differences between Regression VS Classification with examples are as below:-

1. Regression: Predicts continuous numerical values (e.g., house prices, stock prices, etc.).

Classification: Predicts discrete class labels (e.g., spam/not spam, disease/no disease, etc.).

For Example:

Predicting a student's exam score (regression) vs. classifying a student as passing or failing (classification).

2. Regression: Aims to find the best-fit line or function that minimizes the error between predicted and actual values.

Classification: Aims to find the decision boundary that best separates the classes.

For Example:

Predicting the price of a house (regression) vs. classifying a loan applicant as high-risk or low-risk (classification).

3. Regression: Evaluated using metrics like Mean Squared Error (MSE), R-squared and Mean Absolute Error (MAE).

Classification: Evaluated using metrics like Accuracy, Precision, Recall and F1-score.

For Example:

Evaluating the accuracy of a model predicting house prices (regression). However, evaluating the precision of a model classifying email as spam or not spam.

4. Regression: Linear Regression, Polynomial Regression, Support Vector Regression, Decision Tree Regression.

Classification: Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, K-Nearest Neighbors.

For Example:

Using Linear Regression to predict sales (regression). However, using Logistic Regression to classify customers as likely to churn or not (classification).

5. Regression: Predicting continuous values like house prices, stock prices, weather forecasts, and sales figures.

Classification: Classifying emails as spam or not, identifying fraudulent transactions, diagnosing medical conditions, and categorizing images.

For Example:

Using regression to predict the price of a used car (regression) vs. using classification to identify whether an email is spam or not (classification).

6. Regression: Assumes a linear or non-linear relationship between the independent and dependent variables.

Classification: Does not make any assumptions about the underlying distribution of the data.

For Example:

Linear Regression assumes a linear relationship between the features and the target variable, while Logistic Regression does not make any assumptions about the distribution of the data.

Conclusion:- The key differences between **Regression and Classification** lie in the output type, learning objective, evaluation metrics, algorithms, applications and assumptions.

Regression is used to predict continuous numerical values, while Classification is used to predict discrete class labels.