**Clustering vs Classification**

1. **Definition**:
   * **Clustering**: An unsupervised learning technique that groups a set of objects in such a way that objects in the same group (called a cluster) are more similar to each other than to those in other groups.
     + **Example**: Grouping customers based on their purchasing behavior.
   * **Classification**: A supervised learning technique that assigns a label to an input based on learned patterns from labeled training data.
     + **Example**: Predicting whether an email is spam or not based on labeled examples.
2. **Purpose**:
   * **Clustering**: Primarily used to find hidden patterns or intrinsic structures in data without predefined labels.
     + **Example**: Identifying different species of plants based on their characteristics without knowing the species labels beforehand.
   * **Classification**: Used to predict the category or class of new data points based on training data with predefined labels.
     + **Example**: Classifying handwritten digits into numbers 0-9 using a trained model on labeled digit images.
3. **Labels**:
   * **Clustering**: No predefined labels are given; the algorithm discovers the inherent groupings in the data.
     + **Example**: Segmenting social media posts into topics without knowing the topic beforehand.
   * **Classification**: Predefined labels are required for training the model; each training example is associated with a label.
     + **Example**: Diagnosing diseases based on patient symptoms where the model is trained on a dataset with known diagnoses.
4. **Algorithms**:
   * **Clustering**: Common algorithms include K-means, hierarchical clustering, and DBSCAN.
     + **Example**: Using K-means clustering to segment customers into different groups based on purchasing history.
   * **Classification**: Common algorithms include Decision Trees, Support Vector Machines (SVM), and Neural Networks.
     + **Example**: Using a Support Vector Machine to classify emails as spam or not spam based on a labeled dataset.
5. **Evaluation**:
   * **Clustering**: Evaluation is often more challenging and subjective, relying on metrics like silhouette score, Davies-Bouldin index, or visual inspection.
     + **Example**: Evaluating the quality of customer segments created by clustering by visualizing them and using silhouette score.
   * **Classification**: Evaluation is more straightforward and can be quantified using metrics like accuracy, precision, recall, and F1 score.
     + **Example**: Measuring the accuracy of a spam detection model by comparing predicted labels against actual labels.

**Regression vs Classification**

1. **Definition**:
   * **Regression**: A supervised learning technique used to predict a continuous output or value based on input features.
     + **Example**: Predicting house prices based on features like size, location, and number of bedrooms.
   * **Classification**: A supervised learning technique used to assign a discrete label to an input based on learned patterns from labeled training data.
     + **Example**: Predicting whether a tumor is benign or malignant based on medical imaging data.
2. **Output Type**:
   * **Regression**: Produces a continuous numeric value as output.
     + **Example**: Estimating the annual income of an individual based on their age, education, and experience.
   * **Classification**: Produces a discrete category or class label as output.
     + **Example**: Classifying emails as "spam" or "not spam".
3. **Algorithms**:
   * **Regression**: Common algorithms include Linear Regression, Polynomial Regression, Ridge Regression, and Lasso Regression.
     + **Example**: Using Linear Regression to predict a student's exam score based on their study hours.
   * **Classification**: Common algorithms include Logistic Regression, Decision Trees, Support Vector Machines (SVM), and Neural Networks.
     + **Example**: Using Logistic Regression to predict whether a customer will buy a product or not (binary classification).
4. **Evaluation Metrics**:
   * **Regression**: Evaluated using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R²).
     + **Example**: Measuring the performance of a model predicting house prices using RMSE.
   * **Classification**: Evaluated using metrics such as accuracy, precision, recall, F1 score, and ROC-AUC.
     + **Example**: Measuring the performance of a spam detection model using accuracy and F1 score.
5. **Use Cases**:
   * **Regression**: Used when the goal is to predict a numeric value.
     + **Example**: Forecasting stock prices based on historical data and market indicators.
   * **Classification**: Used when the goal is to categorize data into predefined classes.
     + **Example**: Diagnosing patients with a particular disease based on their medical history and test results.