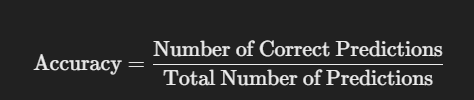
**1. Classification Evaluation Metrics**

When working with classification models (where the output is categorical), we want to measure how well our model predicts the correct class labels. Here are some popular metrics:

**a. Accuracy**

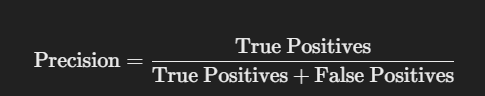
* **Definition**: The proportion of correctly predicted instances out of the total instances.
* **Formula**:



* **Use Case**: Useful when the dataset is balanced (i.e., when each class has a similar number of samples).
* **Limitation**: Can be misleading if the dataset is imbalanced (e.g., 95% of the data is class A, and only 5% is class B).

**b. Precision**

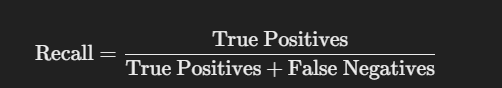
* **Definition**: The proportion of true positive predictions among all predicted positives.
* **Formula**:



* **Use Case**: Important when **false positives** are costly (e.g., in spam detection where mistakenly marking a legitimate email as spam is undesirable).

**c. Recall (Sensitivity or True Positive Rate)**

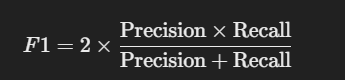
* **Definition**: The proportion of actual positives that were correctly predicted.
* **Formula**:



* **Use Case**: Useful when **false negatives** are costly (e.g., in medical diagnosis where missing a positive case is dangerous).

**d. F1-Score**

* **Definition**: The harmonic mean of precision and recall. It provides a balance between the two.
* **Formula**:



* **Use Case**: Useful when you need a balance between precision and recall, especially with imbalanced datasets.

**e. Confusion Matrix**

* **Definition**: A table that summarizes the number of true positives, true negatives, false positives, and false negatives.
  + **True Positives (TP)**: Correctly predicted positive cases.
  + **True Negatives (TN)**: Correctly predicted negative cases.
  + **False Positives (FP)**: Incorrectly predicted positive cases (Type I error).
  + **False Negatives (FN)**: Incorrectly predicted negative cases (Type II error).
* **Use Case**: Provides a detailed breakdown of classification performance beyond simple accuracy.

**f. ROC-AUC (Receiver Operating Characteristic - Area Under Curve)**

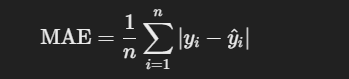
* **Definition**: Measures the model's ability to distinguish between classes. It plots the **True Positive Rate (Recall)** against the **False Positive Rate**.
* **AUC (Area Under the Curve)**: A score of 1.0 indicates perfect classification, while 0.5 indicates random guessing.
* **Use Case**: Good for comparing models and understanding their performance at various classification thresholds.

**2. Regression Evaluation Metrics**

For regression models (where the output is a continuous numerical value), these metrics evaluate how closely the predicted values match the actual values.

**a. Mean Absolute Error (MAE)**

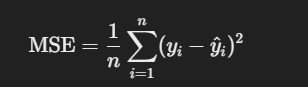
* **Definition**: The average absolute difference between the predicted and actual values.
* **Formula**:



* **Use Case**: Simple to understand and interpret; less sensitive to outliers than other metrics.

**b. Mean Squared Error (MSE)**

* **Definition**: The average squared difference between the predicted and actual values.
* **Formula**:



* **Use Case**: Commonly used because it penalizes larger errors more heavily than MAE.
* **Limitation**: Sensitive to outliers due to squaring the errors.

**c. Root Mean Squared Error (RMSE)**

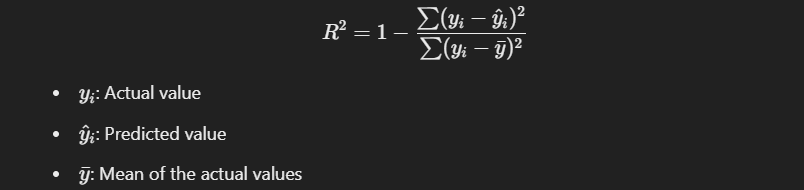
* **Definition**: The square root of the Mean Squared Error, giving an error measure in the same units as the target variable.
* **Formula**:



* **Use Case**: Useful when you want to interpret the error in the original scale of the data.

**d. R² Score (Coefficient of Determination)**

* **Definition**: Measures the proportion of variance in the target variable explained by the model.
* **Formula**:



* **Use Case**: Ranges from 0 to 1, where a value closer to 1 indicates better performance.
* **Limitation**: Can be misleading when used with non-linear models.

**3. Clustering Evaluation Metrics**

For clustering models (used in unsupervised learning where there are no predefined labels), we need metrics that measure the quality of the clusters formed.

**a. Silhouette Score**

* **Definition**: Measures how similar an object is to its own cluster (cohesion) compared to other clusters (separation).
* **Formula**: Ranges from -1 to +1, where a higher value indicates well-separated clusters.
* **Use Case**: Useful to determine the optimal number of clusters.

