Part 1: Metrics for Evaluating Machine Learning Models

1. Definitions

- **True Positives (TP):** The model correctly predicted a positive outcome (the actual outcome was positive).
  (Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)
- **True Negatives (TN):** The model correctly predicted a negative outcome (the actual outcome was negative).
  (Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)
- **False Positives (FP):** The model incorrectly predicted a positive outcome (the actual outcome was negative). Also known as a Type I error.
  (Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)
- **False Negatives (FN):** The model incorrectly predicted a negative outcome (the actual outcome was positive). Also known as a Type II error.
  (Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)
- **Confusion Matrix:** A confusion matrix is a matrix that summarizes the performance of a machine learning model on a set of test data. It is a means of displaying the number of accurate and inaccurate instances based on the model's predictions. It is often used to measure the performance of classification models, which aim to predict a categorical label for each input instance. It is structured as follows:

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

  (Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)

2. Metrics

- **Accuracy:** Accuracy is used to measure the performance of the model. It is the ratio of Total correct instances to the total instances.

  $$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

  (Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)

- **Precision:** Precision is a measure of how accurate a model's positive predictions are. It is defined as the ratio of true positive predictions to the total number of positive predictions made by the model.

$$\text{Precision} = \frac{TP}{TP + FP}$$

(Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)

- **Recall (Sensitivity):** Recall measures the effectiveness of a classification model in identifying all relevant instances from a dataset. It is the ratio of the number of true positive (TP) instances to the sum of true positive and false negative (FN) instances.

$$\text{Recall} = \frac{TP}{TP + FN}$$

(Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)

- **F1-Score:** F1-score is used to evaluate the overall performance of a classification model. It is the harmonic mean of precision and recall.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

(Source: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/)

- **Mean Squared Error (MSE):** Mean Squared Error represents the average of the squared difference between the original and predicted values in the data set. It measures the variance of the residuals.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

(Source: https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e)

- **Root Mean Squared Error (RMSE):** Root Mean Squared Error is the square root of Mean Squared error. It measures the standard deviation of residuals.

$$\text{RMSE} = \sqrt{\text{MSE}}$$

(Source: https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e)

- **Mean Absolute Error (MAE):** The Mean absolute error represents the average of the absolute difference between the actual and predicted values in the dataset. It measures the average of the residuals in the dataset.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

(Source: https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e)

- **Coefficient of Determination (R^2):** The coefficient of determination or R-squared represents the proportion of the variance in the dependent variable which is explained by the linear regression model. It is a scale-free score i.e. irrespective of the values being small or large, the value of R square will be less than one.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where $\bar{y}$ is the mean of the actual values. (Source: https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e)