# Evaluation metrics

# Evaluation metrics for classification algorithm

- Evaluating your machine learning algorithm is an essential part of any project. There are different types of evaluation metrics available:
- Confusion matrix
- Accuracy
- Precision
- Recall
- F Score

# Confusion matrix

- A confusion matrix is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.
- A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix.
- The confusion matrix shows the ways in which your classification model is confused when it makes predictions. It gives insight not only into the errors being made by your classifier but more importantly the types of errors that are being made.

# Confusion Matrix

| | | | |
|---|---|---|---|
| | Predicted class | | |
| Actual Class | | Class = Yes | Class = No |
| | Class = Yes | True Positive | False Negative |
| | Class = No | False Positive | True Negative |

Where,
True Positive (TP): The number of correct classifications of the positive examples
False Negative (FN): The number of incorrect classifications of the positive examples
False Positive (FP): The number of incorrect classifications of the negative examples
True Negative (TN): The number of correct classifications of the negative examples

# Accuracy

- It is the ratio of number of correct predictions to the total number of input samples.

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made}$$

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

# Precision

- Precision p is the number of correctly classified positive examples divided by the total number of examples that are classified as positive.

$$p = \frac{TP}{TP + FP}.$$

# Recall

- Recall r is the number of correctly classified positive examples divided by the total number of actual positive examples in the test set.

$$r = \frac{TP}{TP + FN}.$$

# F Score

- F1 score combines precision and recall into one measure.
- The F score is used to measure a test's accuracy

$$F_1 = \frac{2pr}{p+r}$$

# The F score

The F score reaches the best value, meaning perfect precision and recall, at a value of 1.

The worst F score, which means lowest precision and lowest recall, would be a value of 0.

F1 Score is the weighted average (harmonic mean) of Precision and Recall.

Therefore, this score takes both false positives and false negatives into account.

# Evaluation metrics for regression algorithms

- Mean Absolute Error (MAE)
- Mean Square Error (MSE)
- Root Mean Square Error (RMSE)
- R square

# Mean absolute error (MAE)

MAE is the sum of **absolute** differences between actual and predicted values. It doesn't consider the direction, that is, positive or negative.

$$\text{MAE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_1|}{n}$$

| Actual Value (y) | Predicted Value (y hat) | Error (difference) | Absolute Error | |
|---|---|---|---|---|
| 100 | 130 | -30 | 30 | |
| 150 | 170 | -20 | 20 | |
| 200 | 220 | -20 | 20 | |
| 250 | 260 | -10 | 10 | |
| 300 | 325 | -25 | 25 | |
| | | | 21 | **Mean** |

# Mean Square Error (MSE)

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

| Actual Value (y) | Predicted Value (y hat) | Error (difference) | Squared Error | |
|---|---|---|---|---|
| 100 | 130 | -30 | 900 | |
| 150 | 170 | -20 | 400 | |
| 200 | 220 | -20 | 400 | |
| 250 | 260 | -10 | 100 | |
| 300 | 325 | -25 | 625 | |
| | | | 485 | Mean |

# Root mean square error (RMSE)

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

| Actual Value (y) | Predicted Value (y hat) | Error (difference) | Squared Error | |
|---|---|---|---|---|
| 100 | 130 | -30 | 900 | |
| 150 | 170 | -20 | 400 | |
| 200 | 220 | -20 | 400 | |
| 250 | 260 | -10 | 100 | |
| 300 | 325 | -25 | 625 | |
| | | | **485** | **Mean** |
| | | | 22.02271555 | **Square root of mean** |

R square

It is the ratio of distances between predicted & mean and actual & mean

$$R^2 = \frac{\sum (y_p - \bar{y})^2}{\sum (y - \bar{y})^2}$$