Model evaluation techniques in machine learning:

By: Loga Aswin

1. Accuracy: The ratio of correctly predicted instances to the total instances in the dataset.

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

2. Precision: The ratio of true positive predictions to the total positive predictions.

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

3. Recall (Sensitivity): The ratio of true positive predictions to the total actual positive instances.

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

4. F1 Score: The harmonic mean of precision and recall, providing a balanced measure of a model's performance.

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall}$$

5. Specificity: The ratio of true negative predictions to the total actual negative instances.

$$\textbf{Specificity} = \frac{\textbf{True Negatives}}{\textbf{True Negatives} + \textbf{False Positives}}$$

6. ROC Curve (Receiver Operating Characteristic): A graphical representation of a classifier's performance at various thresholds.

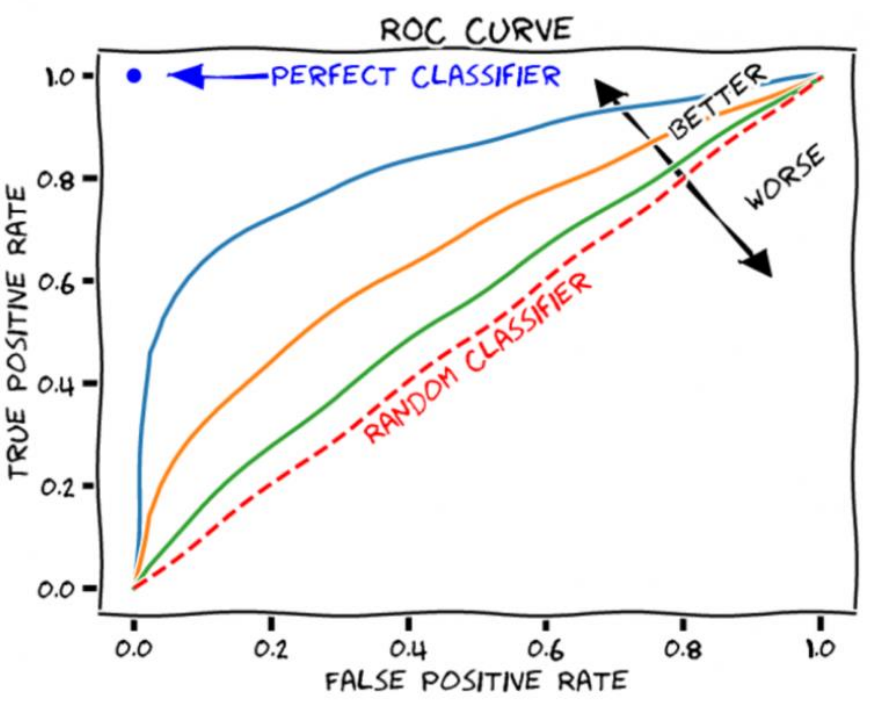
7. AUC-ROC (Area Under the ROC Curve): A metric that quantifies the overall performance of a binary classifier.

The True Positive Rate or Recall is defined as

$$TPR = \frac{TP}{TP + FN}$$

The False Positive Rate is defined as

$$FPR = \frac{FP}{FP + TN}$$



8. Confusion Matrix: A table that describes the performance of a classification model by comparing predicted and actual class labels.

	Class 1 Predicted	Class 2 Predicted
Class 1 Actual	TP	FN
Class 2 Actual	FP	TN

9. Mean Absolute Error (MAE): The average of absolute errors between predicted and actual values in regression tasks.

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

10. Mean Squared Error (MSE): The average of squared errors between predicted and actual values in regression tasks.

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

11. **Root Mean Squared Error (RMSE**): The square root of the MSE, providing a more interpretable error measure.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Predicted_i - Actual_i)^2}{N}}$$

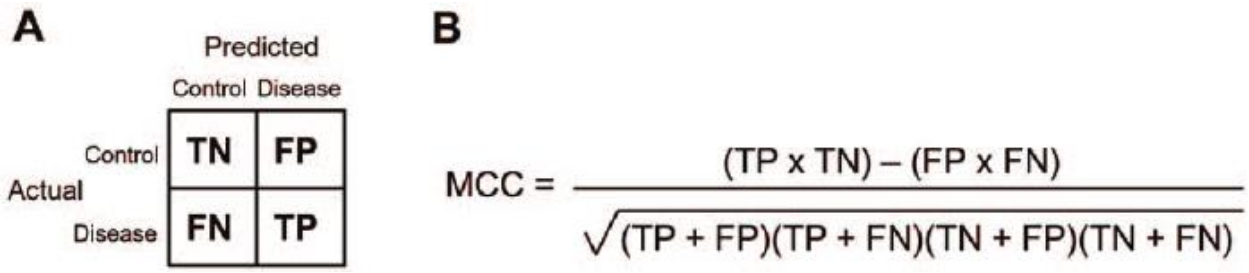
12. R-squared (R^2): A measure of how well a regression model fits the data, indicating the proportion of variance explained.

13. Log-Loss: A measure of the accuracy of a probabilistic classifier's predictions.

14. Cohen's Kappa: A statistic that measures inter-rater agreement for categorical items.

$$\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$$

15. Matthew's Correlation Coefficient (MCC): A measure of the quality of binary classifications.



16. Gini Coefficient: A metric for evaluating the inequality in a distribution.

17. Brier Score: A metric for assessing the accuracy of probabilistic forecasts.

18. Silhouette Score: A measure of cluster quality, indicating how similar an object is to its own cluster compared to others.

$$silhouettescore = (p - q)/max(p, q)$$

p = mean distance to the points in the nearest cluster

q = mean intra-cluster distance to all the points.

19. Adjusted Rand Index (ARI): A measure of the similarity between two data clusterings.

20. Cross-Validation: A technique for assessing a model's performance by splitting data into multiple subsets, training and testing on different subsets.