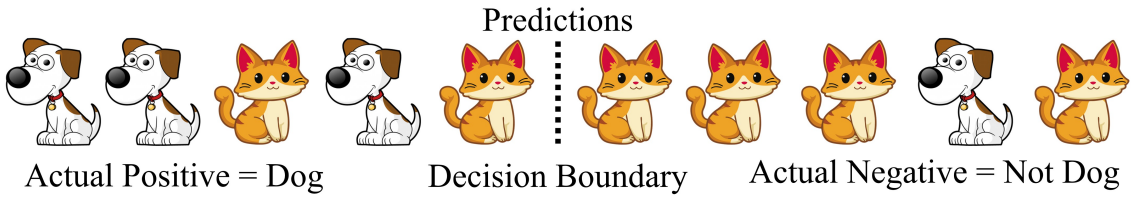
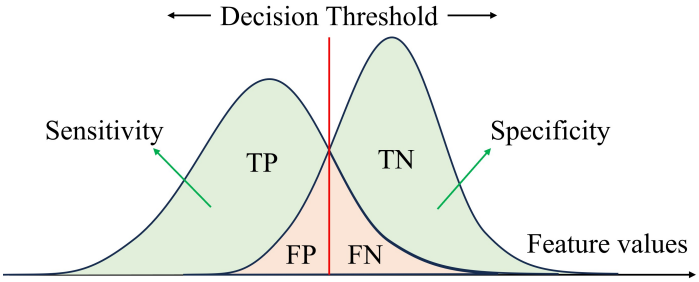
Evaluation Metrics for Classification

Imbalance Classification Problem: When there is an unequal distribution of classes in the training dataset (biased/skewed distribution). It will be harder to predict the minority class. If we have a test set with 100 patients (95 healthy and 5 cancer) and the model wrongly predicts all cancer patients still accuracy would be 95%. Therefore, accuracy is not a suitable evaluation metric for imbalanced classification.



		Actual Class: Positive (Dog)	Actual Class: Negative (Not Dog)	
Predicted Class	Positive (Dog)	True Positive (TP) TP = 3	False Positive (FP) FP = 2 Type I Error	Precision $\frac{TP}{(TP+FP)}$
Predicted Class	Negative (Not Dog)	False Negative (FN) FN = 1 Type II Error	True Negative(TN) TN = 4	
		Recall (Sensitivity) True +ve rate (TPR) $\frac{TP}{(TP+FN)}$	Specificity $\frac{TN}{(TN+FP)}$	Accuracy $\frac{TP+TN}{(TP+TN+FP+FN)}$



Accuracy: It is the ratio of correct predictions to the total number of predictions.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{7}{10} = 0.7$$

False +ve rate/False Acceptance Rate: It is the ratio of wrong dog predictions to total dog samples.

$$\text{FPR} = \frac{FP}{FP+TN} = 1 - specificity = \frac{2}{6} = 0.333$$

Precision: How accurate model predicts dog as dog in total dog predictions.

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{3}{5} = 0.6$$

Recall/Sensitivity/True +ve Rate: How accurate model predicts dog in total dog samples.

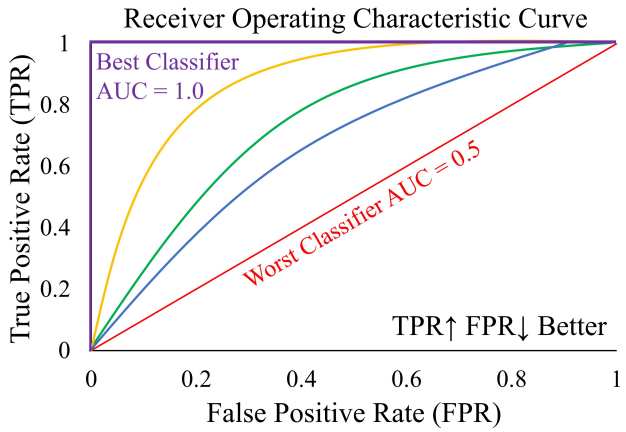
$$\text{Recall} = \frac{TP}{TP+FN} = \frac{3}{4} = 0.75$$

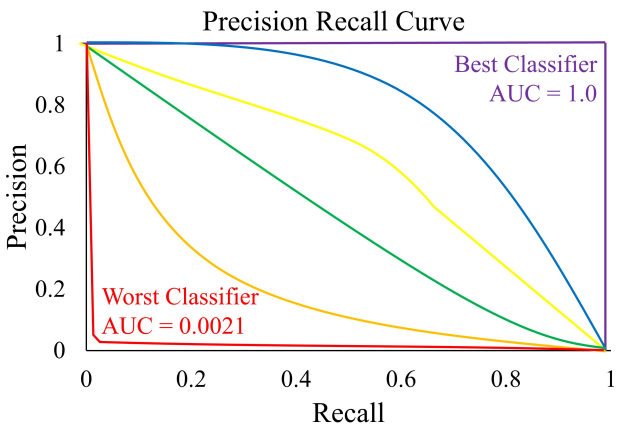
Specificity: How accurate model predicts not-dog in total not-dog samples.

$$\text{Specificity:} = \frac{TN}{TN+FP} = \frac{4}{6} = 0.667$$

F1 Score Harmonic mean of Precision and Recall.

$$\text{F1 score:} = 2 \times \frac{Prec \times Rec}{Prec + Rec} = \frac{0.9}{1.35} = 0.67$$





ROC curve and PR curve are developed by varying the decision threshold for a given distribution. These curves tell about how good the classifier is and help to choose the decision threshold value. ROC curve is used for balanced data and PR curve is useful for imbalanced data.

For any mistake/query: alishafique3@gmail.com For references and updates: Website, Github