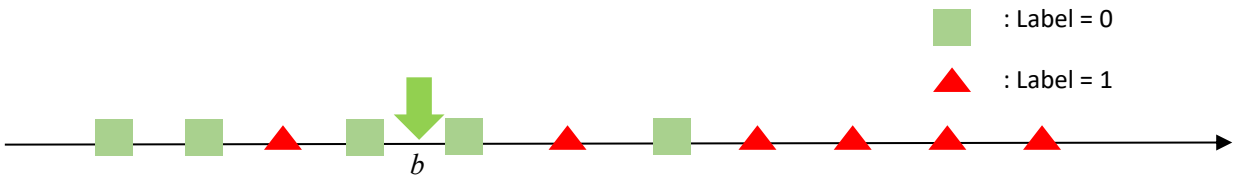
ECE/CS/ME 539 – Fall 2024 — Activity 7

1 Performance Metrics

Consider the same set of posterior probabilities estimated by a classifier for a 2-class pattern classification problem. Assume now the threshold b is at a different value. Manually evaluate the following quantities (represent fractional values in the form of fractions, not decimal numbers):



TP = _____ FN = _____ FP = _____ TN = _____

Sensitivity = _____ Specificity = _____

Pr. Miss = _____ Pr. False Alarm = _____

Precision = _____ Recall = _____

2 Calculating AUROC by Hand

In this exercise, you will calculate the Area Under the Receiver Operating Characteristic Curve (AUROC) for a simple binary classification model. AUROC is a common metric used to evaluate the performance of a classifier, which measures the ability of the model to distinguish between positive and negative classes.

Dataset and Model

Consider a 2-dimensional dataset with 6 points, as shown in the table below:

Point	x_1	x_2	True Label
1	0.5	0.5	Negative
2	1.0	0.5	Negative
3	1.0	2.0	Negative
4	1.0	1.0	Positive
5	2.5	1.0	Positive
6	3.0	2.0	Positive

The points are classified using a simple linear classifier with the decision boundary defined by:

$$f(x_1, x_2) = x_1 + x_2$$

The decision rule is: if $f(x_1, x_2) \geq$ threshold, classify the point as **Positive**; otherwise, classify it as **Negative**.

Step 1: Compute the Classifier Scores

For each point, compute the classifier score $f(x_1, x_2)$.

Step 2: Determine the TPR and FPR at Different Thresholds

To create the ROC curve, you need to calculate the True Positive Rate (TPR) and False Positive Rate (FPR) at various thresholds. Use the thresholds 1.0, 1.5, 2.5, 3.0, 4.0, and 5.0.

Step 3: Plot the ROC Curve

Plot the ROC curve using the FPR (x-axis) and TPR (y-axis) values for each threshold.

$$\text{Thresholds: } 1.0, 1.5, 2.5, 3.0, 4.0, 5.0$$

Mark each point on the graph and *connect them with a line* to form the ROC curve.

Step 4: Calculate the AUROC by Hand

To calculate the Area Under the ROC Curve (AUROC), divide the area under the curve into simple geometric shapes (e.g., rectangles, triangles). Compute the area of each shape and sum them up to find the total area.

Step 5: Interpolation

In Step 3, we *connected* dots with a straight line. How can we justify this? Is there a classifier that can actually operate on a point on the straight line?

Conclusion

The AUROC is a measure of how well your classifier distinguishes between positive and negative classes. A higher AUROC indicates better performance.