

Classification Error Metrics

Choosing the Right Error Measurement

- You are asked to build a classifier for leukemia
- **Training data:** 1% patients with leukemia, 99% healthy
- **Measure accuracy:** total % of predictions that are correct

Choosing the Right Error Measurement

- You are asked to build a classifier for leukemia
- **Training data:** 1% patients with leukemia, 99% healthy
- **Measure accuracy:** total % of predictions that are correct
- Build a simple model that always predicts "healthy"
- Accuracy will be 99%...

Confusion Matrix

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Confusion Matrix

	Predicted Positive	Predicted Negative	
Actual Positive	True Positive (TP)	False Negative (FN)	← Type II Error
Actual Negative	False Positive (FP)	True Negative (TN)	

↑
Type I Error

The diagram illustrates a 2x2 confusion matrix. The columns are labeled 'Predicted Positive' and 'Predicted Negative'. The rows are labeled 'Actual Positive' and 'Actual Negative'. The cells contain 'True Positive (TP)', 'False Negative (FN)', 'False Positive (FP)', and 'True Negative (TN)' respectively. A red arrow points from the 'Type II Error' label to the 'False Negative (FN)' cell. A black arrow points from the 'Type I Error' label to the 'False Positive (FP)' cell.

Accuracy: Predicting Correctly

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

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

Recall: Identifying All Positive Instances

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

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

Precision: Identifying Only Positive Instances

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

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

Specificity: Avoiding False Alarms

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}$$

Error Measurements

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

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

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

Error Measurements

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

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

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

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

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}$$

Error Measurements

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

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

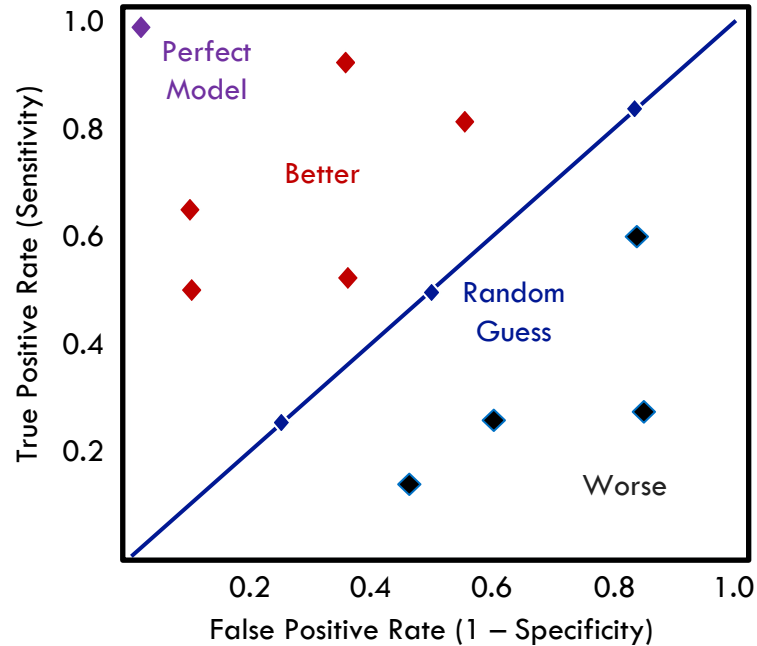
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

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

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}$$

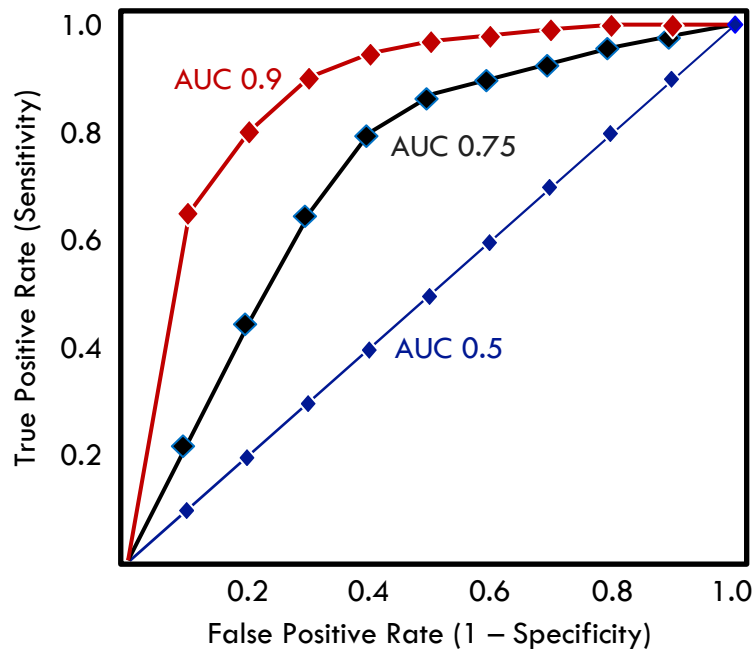
$$F1 = 2 \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Receiver Operating Characteristic (ROC)



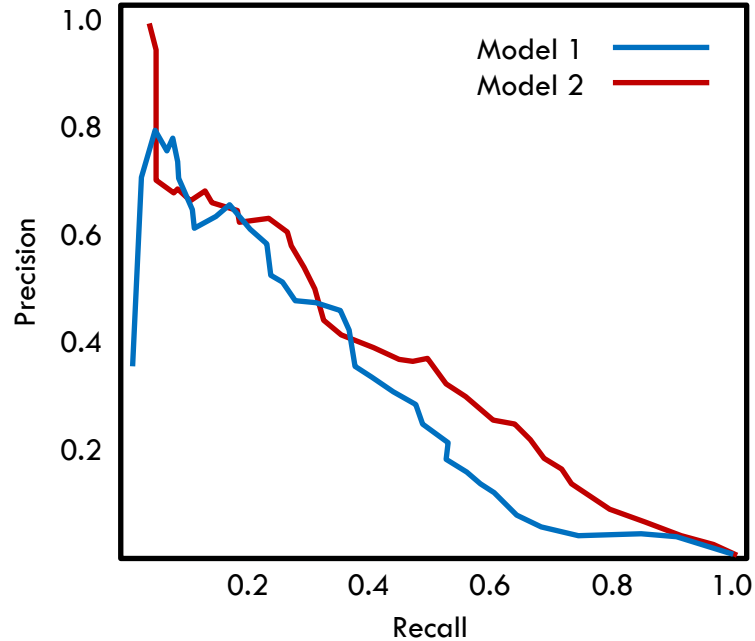
Evaluation of model at all possible thresholds

Area Under Curve (AUC)



Measures total area under ROC curve

Precision Recall Curve (PR Curve)



Measures trade-off between precision and recall

Multiple Class Error Metrics

	Predicted Class 1	Predicted Class 2	Predicted Class 3
Actual Class 1	TP1		
Actual Class 2		TP2	
Actual Class 3			TP3

Multiple Class Error Metrics

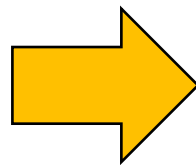
	Predicted Class 1	Predicted Class 2	Predicted Class 3
Actual Class 1	TP1		
Actual Class 2		TP2	
Actual Class 3			TP3

$$\text{Accuracy} = \frac{\text{TP1} + \text{TP2} + \text{TP3}}{\text{Total}}$$

Multiple Class Error Metrics

	Predicted Class 1	Predicted Class 2	Predicted Class 3
Actual Class 1	TP1		
Actual Class 2		TP2	
Actual Class 3			TP3

$$\text{Accuracy} = \frac{\text{TP1} + \text{TP2} + \text{TP3}}{\text{Total}}$$



Most multi-class error metrics are similar to binary versions—just expand elements as a sum