

CLASSIFICATION ERROR METRICS

Intro Metrics • Plots • Multi-class Recap

Learning Objectives & Agenda

Learning objectives



Be able to

- Choose an appropriate error metric
- Evaluate a classifier's performance
- Generalize error metrics to multiclass classification

Agenda



- Motivating Example
- Metrics
- Plots
- Multiclass problems

Motivating Example

Choosing the Right Error Measurement



You are asked to build a classifier for leukemia.

Your data set is *imbalanced*: 1% patients with leukemia, 99% healthy

You're asked to create a classifier with high *accuracy:* total % of predictions that are correct.

Choosing the Right Error Measurement



You are asked to build a classifier for leukemia.

Your data set is imbalanced: 1% patients with leukemia, 99% healthy

You're asked to create a classifier with high *accuracy*: total % of predictions that are correct.

Solution:

Build a simple model that always predicts "healthy". Accuracy will be 99%...

METRICS

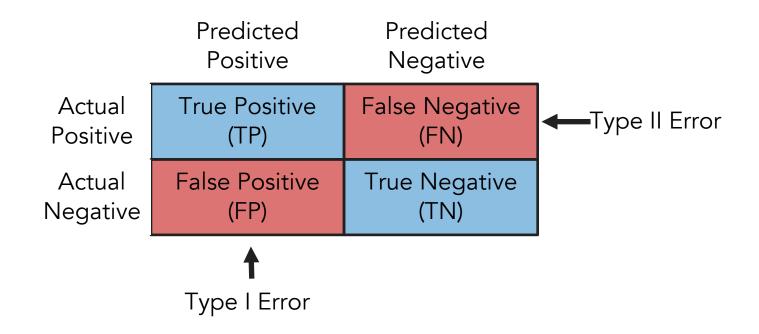
Confusion Matrix



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

Confusion Matrix





Accuracy: Predicting Correctly

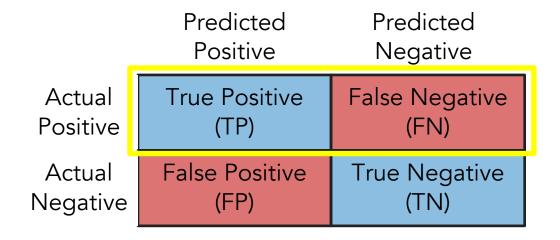


	Predicted Positive	Predicted Negative	
Actual	True Positive	False Negative	
Positive	(TP)	(FN)	
Actual	False Positive	True Negative	
Negative	(FP)	(TN)	

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

Recall: Identifying All Positive Instances

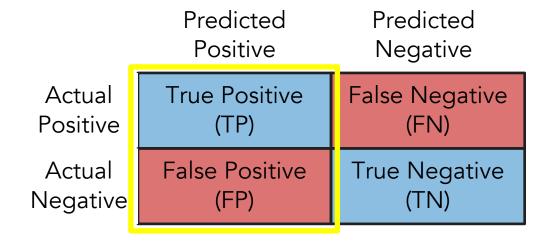




Recall or
$$=$$
 $\frac{TP}{TP + FN}$

Precision: Identifying Only Positive Instances

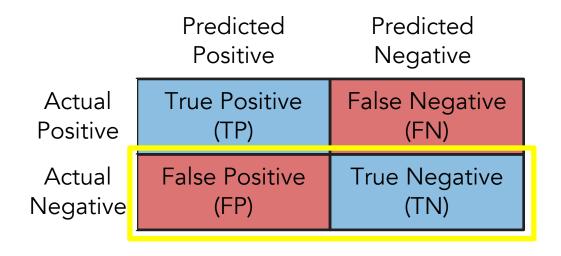




$$Precision = \frac{TP}{TP + FP}$$

Specificity: Avoiding False Alarms





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



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

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



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

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

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



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

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

Recall or Sensitivity =
$$\frac{TP}{TP + FN}$$
Specificity = $\frac{TN}{FP + TN}$
F1 = 2 $\frac{Precision * Recall}{Precision + Recall}$

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



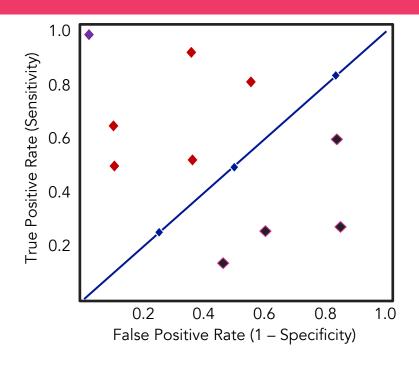
Accuracy	% correct	I reat both classes equally
Precision	% correct of those guessed positive	Only positive instances
Recall (Sensitivity)	Recall (Sensitivity) % correct of those that are positive	
Specificity	% correct of those that are negative	All negative instances
F1	Balances precision ar	nd recall

Accuracy =
$$\frac{TP + TN}{TP + FN + FP + TN}$$
 Recall or $\frac{TP}{Sensitivity} = \frac{TP}{TP + FN}$ Precision = $\frac{TP}{TP + FN}$ Specificity = $\frac{TP}{TN}$ F1 = 2 $\frac{Precision * Recall}{Precision + Recall}$

PLOTS

Receiver Operating Characteristic (ROC)

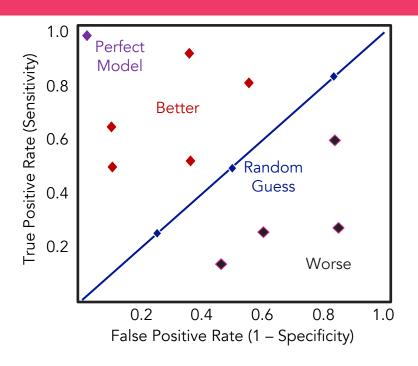




Evaluation of model at all possible thresholds

Receiver Operating Characteristic (ROC)

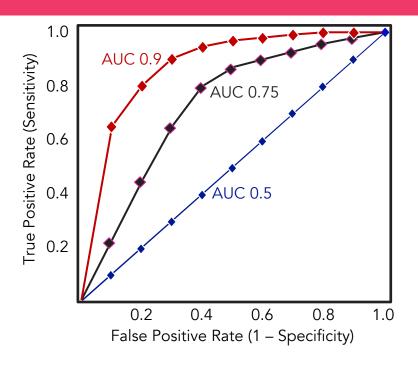




Evaluation of model at all possible thresholds

Area Under Curve (AUC)

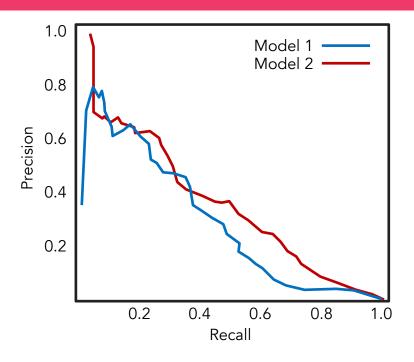




Measures total area under ROC curve

Precision Recall Curve (PR Curve)





Measures trade-off between precision and recall

MULTI-CLASS

Multiple Class Error Metrics



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

Multiple Class Error Metrics



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

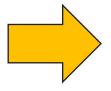
Accuracy =
$$\frac{TP1 + TP2 + TP3}{Total}$$

Multiple Class Error Metrics



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

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



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

Recap

Learning objectives



Be able to

- Choose an appropriate error metric
- Evaluate a classifier's performance
- Generalize error metrics to multiclass classification

Agenda



- Motivating Example
- Metrics
- Plots
- Multiclass problems

Takeaways



Choose an appropriate error metric for the business problem (not for the results)

Accuracy, precision, recall, specificity, and F1 are tailored for different needs

AUC measures how well two classes are being separated

Error metrics generalize to multiclass problems

QUESTIONS?