How do we measure a classifier's performance?

Oliver W. Layton

CS251: Data analysis and visualization

Lecture 28, Spring 2019

Wednesday April 17

Plan

- Confusion matrix
- Binary classification metrics
- Shortcomings of percent correct
- ROC curve and AUC

Measure a classifier's performance

- **Holy grail:** It should predict all the class labels in test set. Do the same on different sets that we present to model.
  - ...almost never happens!
- When we run our classifier on test set, how do we know if it works well?
  - Compute percent correct: $\frac{\text{Num labels correctly predicted}}{\text{Size of test set}}$
  - Analyze *number* and *type of errors* made.

Confusion matrix

- Table of counts of which class labels assigned by model.
- Allows to to quantify how the model may confuse one class or another.
- For **binary classification (2 classes: labeled 0 and 1)**, confusion matrix is 2×2
- Generally $N \times N$.

Dog data example confusion matrix

	Predict Pug	**Predict Lab**
Pug	100	0
Lab	30	70

- How large test set?
- How would perfect classification look like?
- Which test class is being confused for what?
- What would a symmetric confusion matrix indicate?

Fruit data 3x3 example confusion matrix

	Predict Apple	**Predict Pear**	**Predict Orange**
Apple	100	0	0
Pear	0	100	0
Orange	0	97	3

Quantifying confusion matrix results

- Are the results from one confusion matrix better/worse than other? Good in general?
- Binary (2 class) classification are so common that there are metrics computed to answer these questions.
- Let's draw a diagram and define the confusion metrics.

Summary of binary classification metrics (1/2)

- Class A: Sick classification
- Class B: Healthy classification
- **True positive (TP):** Number of data points in A classified as A.
  - Sick patents classified as sick.
- **False positive (FP):** Number of data points in B classified as A.
  - Healthy patents classified as sick.
- **False negative (FN):** Number of data points in A classified as B.
  - Sick patents classified as healthy.
- **True negative (TN):** Number of data points in B classified as B.
  - Healthy patents classified as healthy.

Summary of binary classification metrics (2/2)

- **TP rate (sensitivity):** $\frac{TP}{TP+FN}$. Proportion correctly classified as sick.
- **FP rate:** $\frac{FP}{FP+TN}$. Proportion of healthy classified as sick.
- **Precision:** $\frac{TP}{TP+FP}$. Of the people classified as sick, what fraction actually are?