Evaluation Baselines

School of Information Studies
Syracuse University

Metrics of Model Performance

Most common measure

- Accuracy

Accuracy

Accuracy: total correct/total

- So you get 90% accuracy; isn't that great?!
- Maybe not, it depends on your baseline.

Random Guess Baseline

50% for binary classification

1/n for n-class classification

Majority Vote Baseline

- What is majority vote baseline?
  - It's a trivial classifier that classifies all examples to the majority class.
  - Example: A spam data set includes 90% spams and 10% regular mails, so the majority vote baseline is 90%.
  - It means this trivial classifier predicts everything as spam and gets 90% accuracy.
  - A good classifier has to beat this baseline to claim effectiveness.

Which Baseline to Choose?

Accuracy: total correct/total

- So you get 90% accuracy; isn't that great?!
- If your data set is balanced with equal number of examples in each category, then **yes**!
- But maybe not, if you have a very skewed data set, in which case you need to compare your classifier's accuracy against the majority vote baseline.
- Or compare your classifier to the state-of-the-art classifier.