THE UNIVERSITY OF TENNESSEE KNOXVILLE

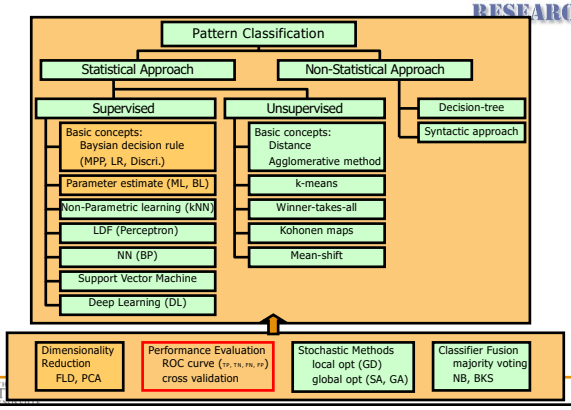
AICIP RESEARCH

ECE471-571 – Pattern Recognition

Lecture 8 – Performance Evaluation

Hairong Qi, Gonzalez Family Professor
Electrical Engineering and Computer Science
University of Tennessee, Knoxville
http://www.eecs.utk.edu/faculty/qi
Email: hqi@utk.edu

AICIP RESEARCH



AICIP RESEARCH

Some Terminology

- Often used in automatic target recognition and medical diagnosis
- True positive
  - The object is there and our classifier says it is there
- True negative
  - The object is not there and our classifier says it is not there
- False negative (false misses)
  - The object is there and our classifier says it is not there
- False positive (false hits)
  - The object is not there and our classifier says it is there

Some Terminology (cont')

AICIP RESEARCH

- Sensitivity
  - Probability of a true-positive = TP/(TP+FN)
- Specificity
  - Probability of a true-negative = TN/(TN+FP)
- The probability of a correct decision = (TP+TN)/N, where N is the total number of samples

THE UNIVERSITY OF TENNESSEE

4

Some Terminology (cont')

AICIP RESEARCH

- Precision = TP/(TP+FP)
- Recall = TP/(TP+FN)
- Accuracy = (TP+TN)/(TP+TN+FP+FN)

THE UNIVERSITY OF TENNESSEE

5

Parameters vs. Performance

AICIP RESEARCH

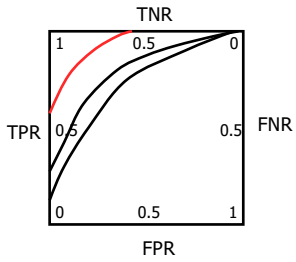
- Once we have designed our classifier, we invariably have some parameters we'd like to adjust. E.g.,
  - Prior probability
- The optimal classifier is one with sensitivity as close to 100% as possible, and at the same time with specificity as close to 100% as possible

THE UNIVERSITY OF TENNESSEE

6

ROC – Receiver Operating Characteristic

AICIP RESEARCH



- Each curve represents the performance of a particular classifier as some parameter is varied over its range
- Of the three curves, the one with the sharpest bend, which passes closest to the upper left corner, is the best
- Calculate the area above the curve, the one with the smallest area is the best

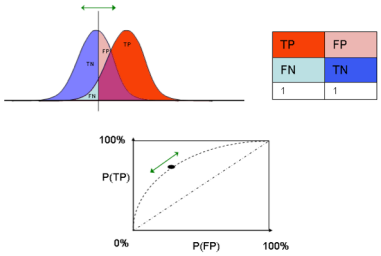
- TPR: TP out of the total actual positives (Sensitivity or Recall)
- FPR: FP out of the total actual negatives (1-Specificity)

THE UNIVERSITY OF TENNESSEE KNOXVILLE

7

ROC (cont'd)

AICIP RESEARCH



http://en.wikipedia.org/wiki/Receiver_operating_characteristic

THE UNIVERSITY OF TENNESSEE KNOXVILLE

8

Determine ROC from Test Data

AICIP RESEARCH

- Apparent error rate vs. true error rate
  - Apparent error rate: counting the number of elements in the training set that are misclassified. However, this error rate leads to an optimistic result since the classifier has been designed to minimize the number of misclassifications of the training set
- Solutions???

THE UNIVERSITY OF TENNESSEE KNOXVILLE

9

Solution 1 – Separating the training set and the test set

AICIP RESEARCH

- Divide the training set into two parts (randomly) and build the classifier using half of the data, then test the classifier on the other half. This other half is called the validation set.
- Clarification: training set, test set, validation set

Solution 2 – The Leave-One-Out Approach

AICIP RESEARCH

- Assume there are n points in the training set.
- Remove point 1 from the set and design the classifier (determine the pdf) using the other n-1 points. Then test the classifier on point 1.
- Repeat for all points.
- The resulting error rate can be shown to be an almost unbiased estimate of the expected true error rate
- This requires we design n classifiers. However, we only need to do it once.

m-Fold Cross Validation

AICIP RESEARCH

- A generalization to both solution 1 and 2.
- The training set is randomly divided into m disjoint sets of equal size. The classifier is trained m times, each time with a different set held out as a validation set
- For example, when m = 3,
  - m1+m2 as training, test on m3
  - m1+m3 as training, test on m2
  - m2+m3 as training, test on m1
- When m=2, it is solution 1
- When m=n, it is solution 2 (n is the number of samples in the original training set)