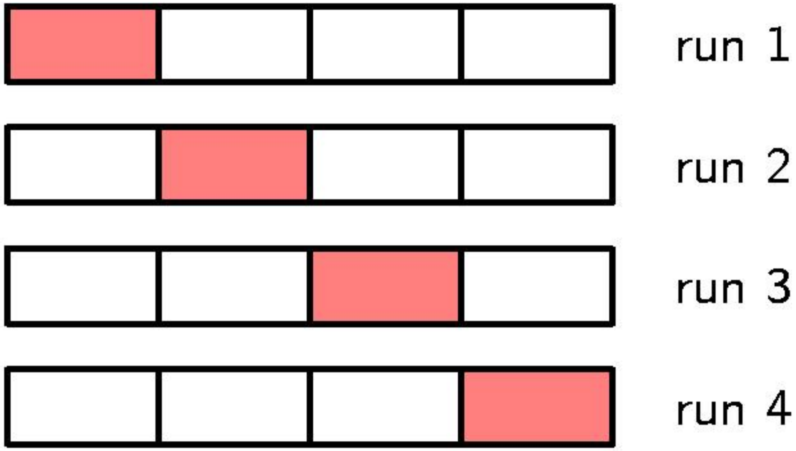
CPE/EE 695: Applied Machine Learning

Lecture 4 - 1: Evaluating Models

Dr. Shucheng Yu, Associate Professor
Department of Electrical and Computer Engineering
Stevens Institute of Technology

Model Evaluation

Cross-validation



Evaluation methods

- **Leave-one-out cross-validation**: This method is used when the data set is very small.
    - It is a special case of cross-validation
- Each fold of the cross validation has only **a single test example** and all the rest of the data is used in training.
- If the original data has *m* examples, this is ***m*-fold cross-validation**

Classification measures

- Accuracy is only one measure (error = 1-accuracy).

  Accuracy is not always suitable

  - In text mining, we may only be interested in the documents of a particular topic, which are only a small portion of a big document collection.
  - In classification involving skewed or highly imbalanced data, e.g., network intrusion and financial fraud detections, we are interested only in the minority class.
    - High accuracy does not mean any intrusion is detected.
    - E.g., 1% intrusion. Achieve 99% accuracy by doing nothing.
- The class of interest is commonly called the **positive class**, and the rest **negative classes**.

Precision and recall measures

confusion matrix

	Classified Positive	Classified Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

where

TP: the number of correct classifications of the positive examples (**true positive**),
FN: the number of incorrect classifications of positive examples (**false negative**),
FP: the number of incorrect classifications of negative examples (**false positive**), and
TN: the number of correct classifications of negative examples (**true negative**).

FP: Type I error , false alarm
FN: Type II error

Precision and recall measures (cont…)

	Classified Positive	Classified Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

$$p = \frac{TP}{TP + FP}. \qquad r = \frac{TP}{TP + FN}.$$

Precision *p* is the number of correctly classified positive examples divided by the total number of examples that are classified as positive.

Recall *r* is the number of correctly classified positive examples divided by the total number of actual positive examples in the test set.

An example

	Classified Positive	Classified Negative
Actual Positive	1	99
Actual Negative	0	1000

- This confusion matrix gives
  - precision p = 100% and
  - recall r = 1%

  because we only classified one positive example correctly and no negative examples wrongly.
- Note: precision and recall only measure classification on the positive class.

F_1-value (also called F_1-score)

- It is hard to compare two classifiers using two measures. F_1 score combines precision and recall into one measure

$$F_1 = \frac{2pr}{p+r}$$

F_1-score is the harmonic mean of precision and recall.

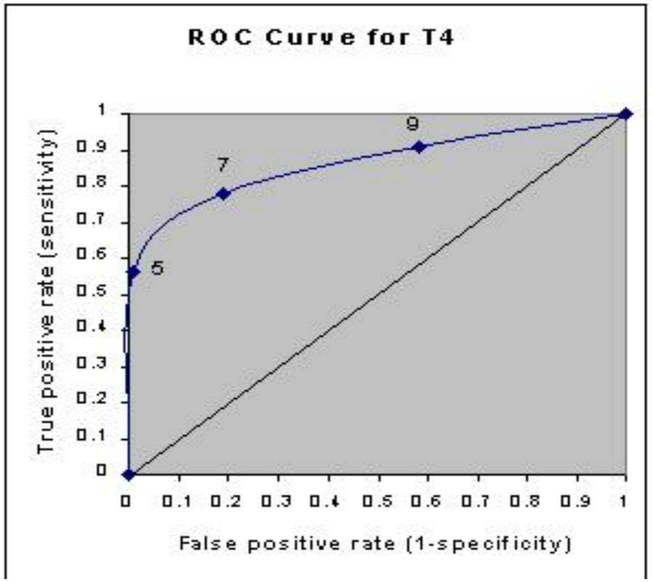
$$F_1 = \frac{2}{\frac{1}{p}+\frac{1}{r}}$$

- The harmonic mean of two numbers tends to be closer to the smaller of the two.
- For F_1-value to be large, both *p* and *r* must be large.

ROC(Receiver Operating Characteristic) curve

Sensitivity: the proportion of positives which are correctly identified . recall rate, True Positive Rate

specificity : the proportion of negatives which are correctly identified . True Negative Rate



- It shows the tradeoff between sensitivity and specificity (any increase in sensitivity will be accompanied by a decrease in specificity).
- The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test.
- The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test.
- The area under the curve (AUC) is a measure of accuracy.

Acknowledgement

Significant part of the slides were based on Dr. Christopher M. Bishop's material at https://www.microsoft.com/en-us/research/people/cmbishop/#!prml-book?from=http%3A%2F%2Fresearch.microsoft.com%2F%7Ecmbishop%2Fprml;

Part of the slide materials were based on Dr. Rong Duan's Fall 2016 course CPE/EE 695A Applied Machine Learning at Stevens Institute of Technology.

STEVENS

INSTITUTE of TECHNOLOGY

THE INNOVATION UNIVERSITY®

1870

stevens.edu