Predictive Accuracy Evaluation

Dr Muhammad Atif Tahir

Professor

School of Computer Science

National University of Computing & Emerging Sciences

Karachi Campus

Predictive Accuracy Evaluation

The main methods of predictive accuracy evaluations are:

- Resubstitution (N ; N);
- Holdout (2N/3 ; N/3)
- k-fold cross-validation (N-N/k ; N/k)
- Leave-one-out (N-1 ; 1)

where *N* is the number of instances in the dataset

Predictive Accuracy Evaluation

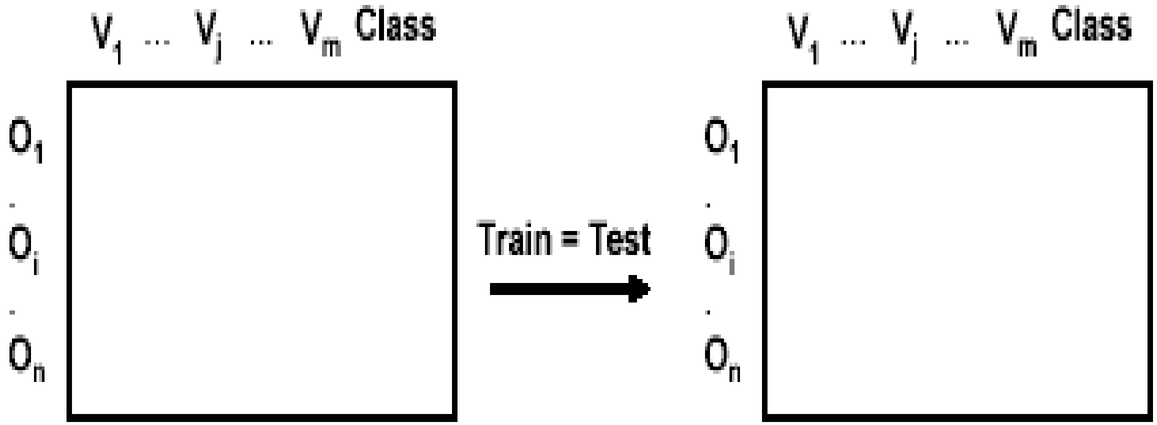
N = 100

Resubstitution (100 ; 100);
Holdout (70 ; 30)
k-fold cross-validation (90 ; 10);
Leave-one-out (99 ; 1);

where N is the number of instances in the dataset

Resubstitution (N ; N)

Testing the classification model by using the given data set (already used for „training“)



Resubstitution Error Rate

- Error rate is obtained from training data
- NOT always 0% error rate, but usually (and hopefully) very low!
- Resubstitution error rate indicates only how good (bad) are our results (rules, patterns, NN) on the TRAINING data; expresses some knowledge about the algorithm used

Holdout (2N/3 ; N/3)

- The holdout method reserves a certain amount for testing and uses the remainder for training – so they are disjoint!
- Usually, one third for testing, and the rest for training
- Train-and-test; repeat

Holdout (2N/3 ; N/3)

- Generally, the larger is the training the better is the classifier
- The larger the test data the more accurate the error estimate

- ***Holdout* procedure:** a method of splitting original data into training and test set

- Dilemma: ideally both training and test set should be large! What to do if the amount of data is limited?
- How to split?

k-fold cross-validation (N-N/k ; N/k)

- cross-validation is used to prevent the overlap!
- cross-validation avoids overlapping test sets:

  - first step: split data into k subsets of equal size

  - second step: use each subset in turn for testing, the remainder for training
  - The error estimates are averaged to yield an overall error estimate

Cross-validation

- Standard cross-validation: 10-fold cross-validation
- Why 10?

Extensive experiments have shown that this is the best choice to get an accurate estimate. There is also some theoretical evidence for this. *So interesting!*

Improve cross-validation

- Even better: *repeated* cross-validation

Example:

10-fold cross-validation is repeated 10 times and results are averaged (reduce the variance)

A particular form of cross-validation

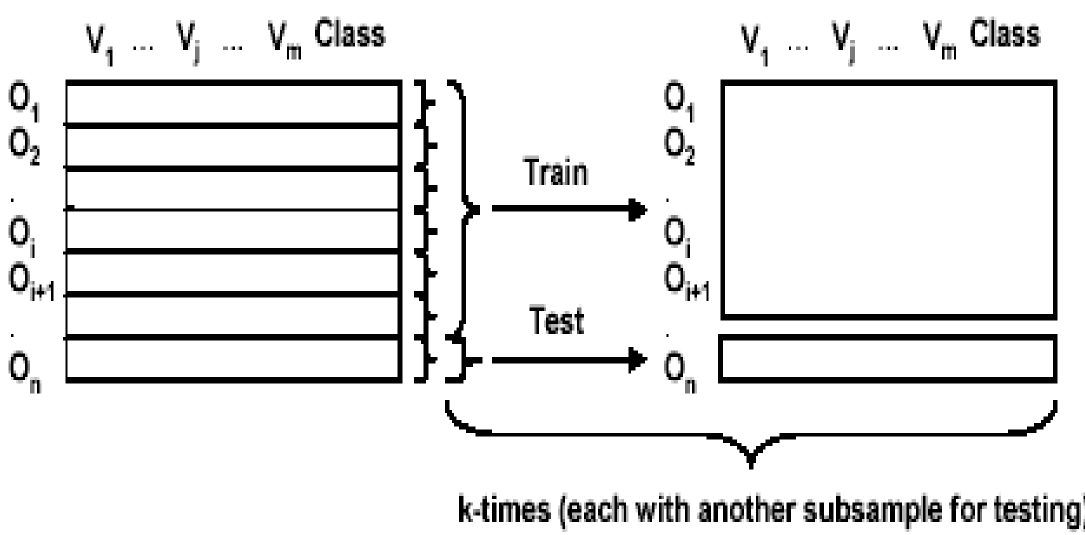
- k-fold cross-validation: (N-N/k ; N/k)
- If k= N, what happens?
- We get:
- (N-1; 1)
- **It is called “leave –one –out”**

Leave-one-out (N-1 ; 1)

Cross-Validation (for moderated sample sizes) → Sampling without replacemen

- Dividing the given data set into ***m* subsamples of equal size**
- Each subsample is tested by using a model generated from the remaining *(m-1)* subsamples

→ **Leave-One-Out**: m = Number of objects



Leave-one-out (N-1 ; 1)

- Make best use of the data
- Involves no random subsampling
- Stratification is not possible
- Very computationally expensive