



MODEL EVALUATION METHODS

SYRACUSE UNIVERSITY
School of Information Studies

MODEL EVALUATION METHODS

What methods can measure model fitness before using it in real predictions?

Some evaluation methods have been designed to test the model on training data while controlling model overfitting.

Hold-out test

Cross-validation

HOLD-OUT TEST

Split the training data to two subsets, using one subset for training and the other for testing.

The splitting ratio is determined by the training set size in that both subsets cannot be too small.

50/50 or 2:1 are common splitting ratios.

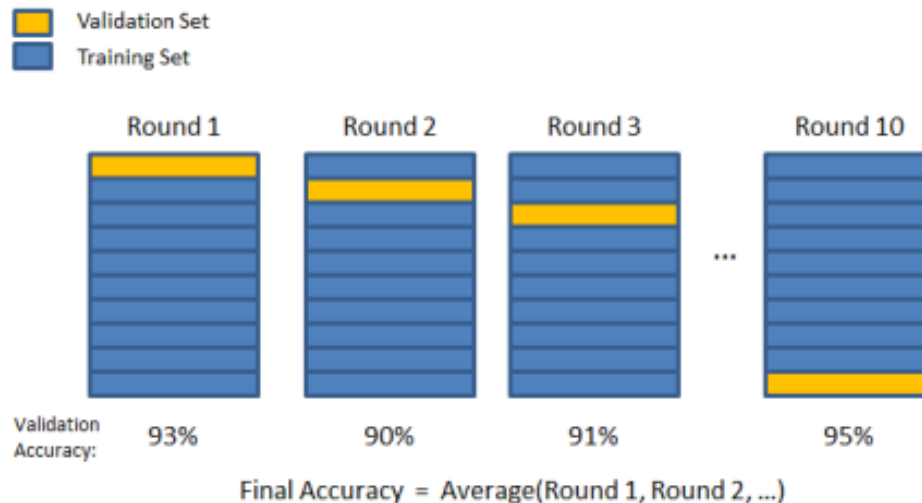
Advantage: Fast

Shortcoming: When the split changes, the test result changes too
High variability in the test result

CROSS-VALIDATION (CV)

N is determined by the training set size. The larger the N, the longer it takes to run the experiment.

Five and 10 are common choices for N.



<http://chrisjmcormick.wordpress.com/2013/07/31/k-fold-cross-validation-with-matlab-code/>

LEAVE ONE OUT

An extreme case of cross-validation

N equals the training set size S

Advantage

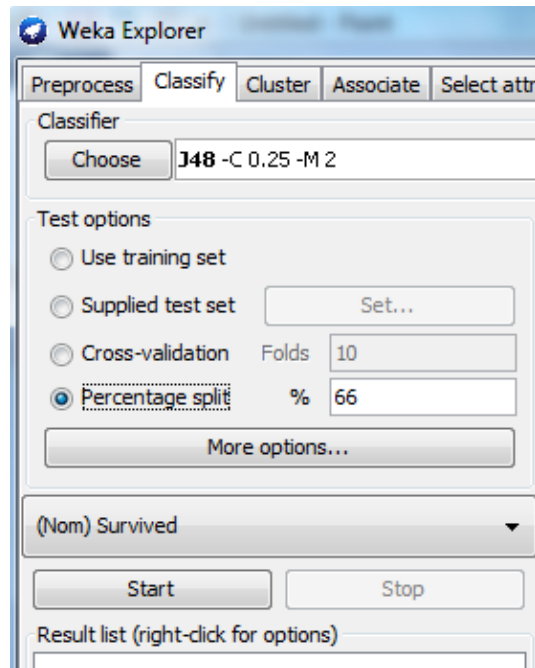
No variability in the test result (always get the same result)

Problems

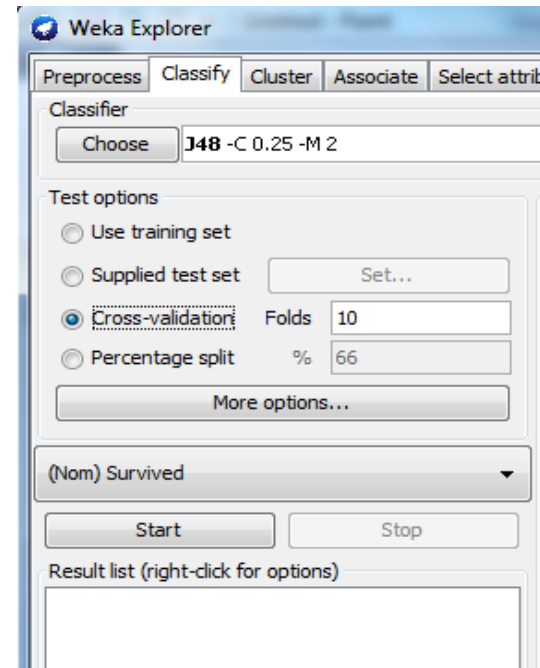
The most time-consuming method

Usually used on very small data sets

HOLD-OUT TEST VS. CROSS-VALIDATION



Weka test option for **hold-out test**



Weka test option for **cross-validation**

HOLD-OUT TEST VS. CROSS-VALIDATION

Hold-out test

Pro: Fast

Con: High variability in the result, depending on the split

Cross-validation

Pro: Less variability and thus more reliable error estimation

Con: Takes longer time

WHICH MODEL EVALUATION METHODS TO CHOOSE?

CV is the standard method.

When data set is huge, hold-out test can save time.

When data set is small, leave one out can be considered.