



Evaluation Methods

School of Information Studies
Syracuse University

Evaluation Methods to Avoid Model Overfitting

But... we have to estimate how good a model is 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

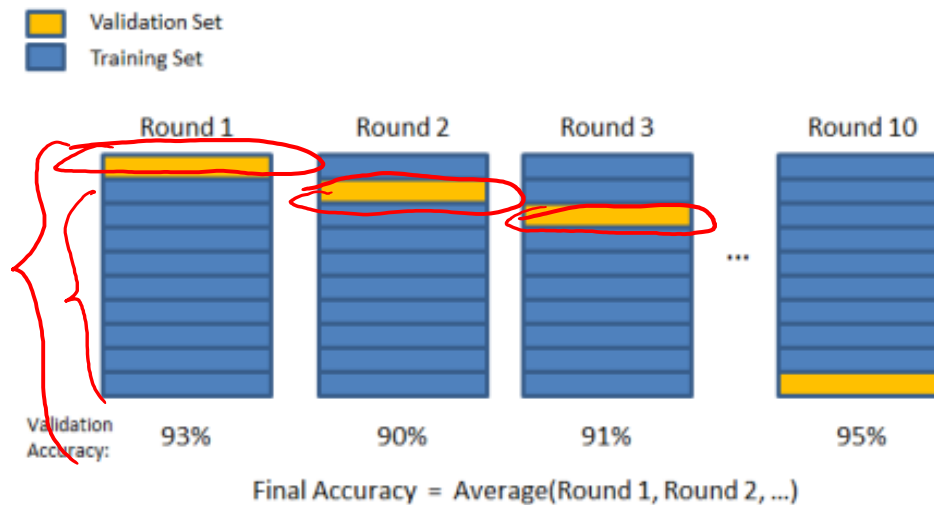
Hold-out test

- 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.

Cross Validation (CV)

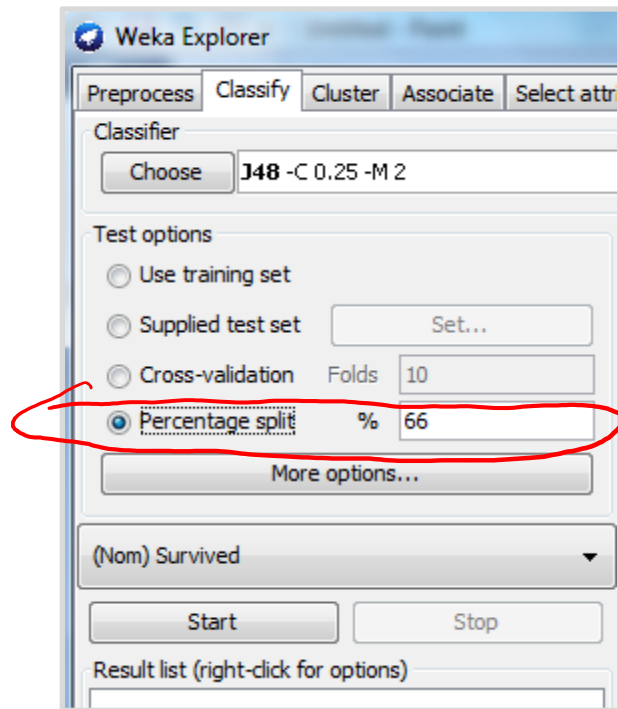
N-fold cross validation (CV)

- N is determined by the training set size. The larger the N, the longer it takes to run the experiment.
- 5 or 10 are common choices for N.

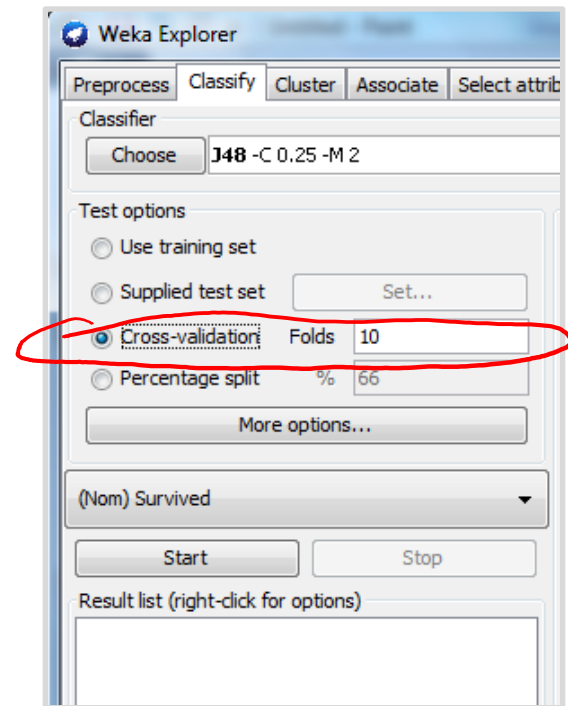


<http://chrisjmcormick.wordpress.com/2013/07/31/k-fold-cross-validation-with-matlab-code/>

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

- Pros: fast
- Cons: high variability in the result depending on the split

Cross validation

- Pros: less variability and thus more reliable error estimation
- Cons: takes longer

| Limitation of Cross Validation

Since cross validation result is still an estimation of the real test error, even if you get good cross validation accuracy on the training data, your final test accuracy may still be different from it, either higher or lower, but the difference should not be large.