k-fold cross validation

Ravi Kumar Tiwari
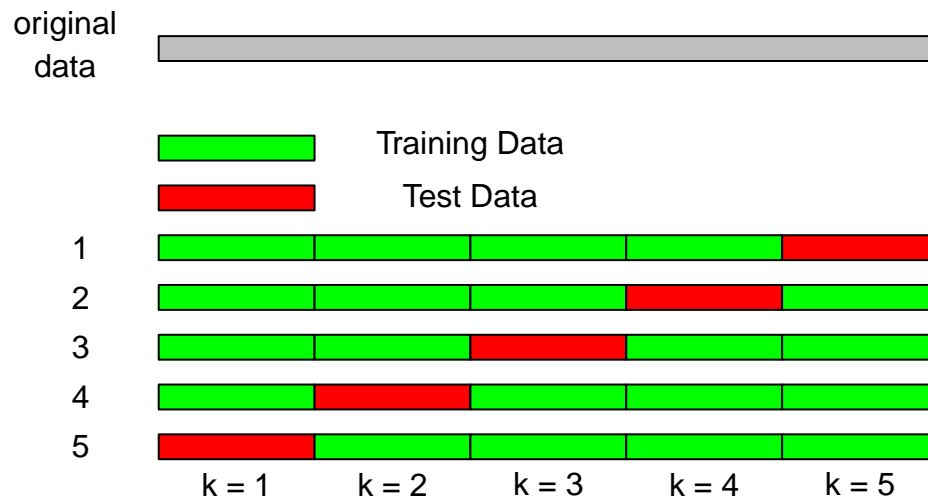
14 July 2016

k-fold Cross-validation

This method involves randomly splitting dataset into k-folds of equal size. Out of k-fold, one group of observation is held-out and the remaining k-1 groups of observations are used to train the model.

This method is very useful to estimate 1. test-error associated with a given learning method in order to evaluate its performance (model assessment) 2. choose appropriate level of flexibility (model selection)

5–fold cross–validation illustration

Leave-one-out cross-validation (LOOCV) is a special case of k-fold cross-validation where k = n, where n is the number of observations.