Text

Description automatically generated

Text, letter

Description automatically generated

Text

Description automatically generated

Text

Description automatically generatedGraphical user interface

Description automatically generated

Text

Description automatically generated with medium confidence

1. Using *n* for observations and *k* for observations randomly split up in the data set we use the equation *n* / *k*. We can now use the average of *k* to obtain the MSE estimate. The observations we used to split up is used as a validation set.
2. i.
   * K-Fold Validation can be applied to almost all learning algorithms
   * The computation time for K-Fold Validation is approximately K times of the computation time of “Validation Set Approach” that is acceptable for most learning algorithms
   * 5-Fold or 10-fold validation have been shown to yield the test error rate that suffer neither from excessively high bias nor from very high variance.
   * Disadvantages are that there may be overestimation when based on a few observations

ii. The LOOCV

* The LOOCV method can have a shorter computation time but in special cases