

How many folds?

• With large number of folds:

+ The bias of the true parameter will be small (estimate very accurate)

The variance of the true error rate estimator will be large

The output will be very large (many experiments)

• With small number of folds:

+The number of experiments and, therefore, computation time are reduced

+ The variance of the estimator will be small

The bias of the estimator will be large (smaller than the true rate)

In practice, the choice of the number of folds depends on the size of the dataset

For large data: even 3-fold cross validation will be quite accurate

For sparse data: leave-one-out in order to train on as many examples as possible

How many folds?

- With large number of folds:
 - + The bias of the true error rate estimator will be small (estimator very accurate)
 - The variance of the true error rate estimator will be large
 - The computational time will be very large as well (many experiments)
- With small number of folds:
 - + The number of experiments and, therefore, computation time are reduced
 - + The variance of the estimator will be small
 - The bias of the estimator will be large (smaller than the true error rate)
- In practice, the choice of the number of folds depends on the size of the dataset
 - For large data: even 3-fold cross validation will be quite accurate
 - For sparse data: leave-one-out in order to train on as many examples as possible

Hyperparameter Tuning