



Training and Test Datasets

- Partition the original data (randomly or stratified) into a training set and a test set. (e.g. 70/30)

- It can be useful to evaluate the training error, but you should not look at training error alone.
- Training error is not an estimate of **generalization error** (on a test set or cross-validated), which is what you should care more about.
- Training error vs test error over time is an useful thing to calculate. It can tell you when you start to overfit your model, so it is a useful metric in supervised machine learning.

Training Set vs.
Test Set

Training Error vs.
Test Error



- Regression: R^2 , MSE, RMSE
- Classification: Accuracy, F1, H-measure, Log-loss
- Ranking (Binary Outcome): AUC, Partial AUC

Performance Metrics



Training and Test Data Sets

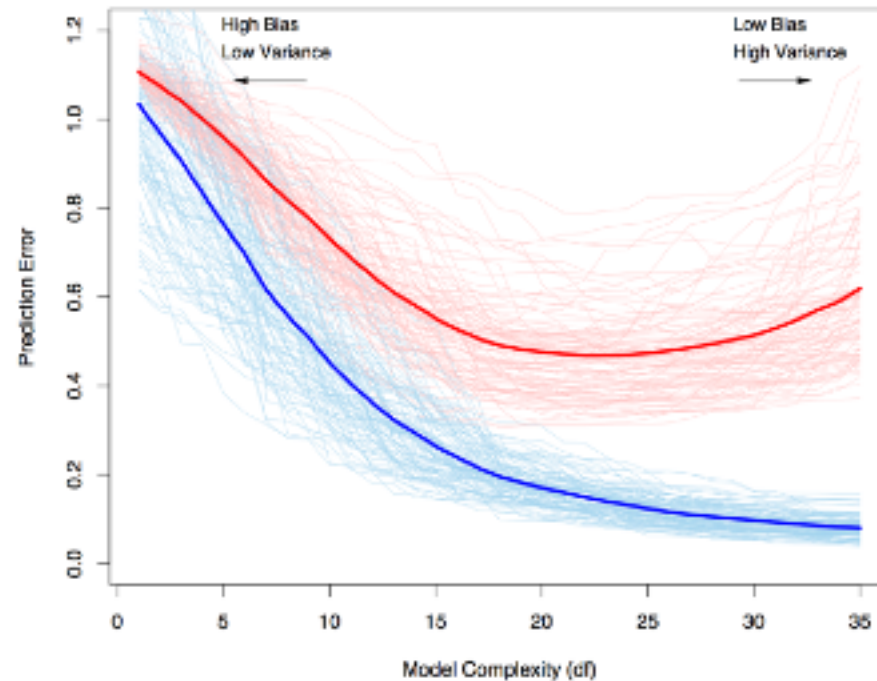


FIGURE 7.1. Behavior of test sample and training sample error as the model complexity is varied. The light blue curves show the training error \overline{err} , while the light red curves show the conditional test error Err_T for 100 training sets of size 50 each, as the model complexity is increased. The solid curves show the expected test error Err and the expected training error $E[\overline{err}]$.

Source: Elements of
Statistical Learning

Training and Test Data Sets

Training Set vs.
Test Set

Training Error vs.
Test Error

Performance Metrics

- Partition the original data (randomly or stratified) into a **training** set and a **test** set. (e.g. 70/30)
-
- It can be useful to evaluate the training error, but you should not look at training error alone.
 - Training error is not an estimate of **generalization error** (on a test set or cross-validated), which is what you should care more about.
 - Training error vs test error over time is an useful thing to calculate. It can tell you when you start to overfit your model, so it is a useful metric in supervised machine learning.
-
- Regression: R^2 , MSE, RMSE
 - Classification: Accuracy, F1, H-measure, Log-loss
 - Ranking (Binary Outcome): AUC, Partial AUC