Other Aspects of Classifier Evaluation

More attributes might degrade performance if they are noisy. Better performance might be acheived by selecting a subset of the attributes.

More complex classifiers might degrede performance because of overfitting.

An unrepresentative dataset can adversely affect future performance. Nearly all algorithms assume that the dataset comes from the same distribution as the application environment.

ROC Curves

ROC (Receiver Operating Characteristic) curves show the tradeoffs between false positives and false negatives.

False Positive Rate = number of incorrect positive predictions / number of negative examples

False Negative Rate = number of incorrect negative predictions / number of positive examples

Changing the bias weight of an SVM or output neuron can affect this tradeoff.

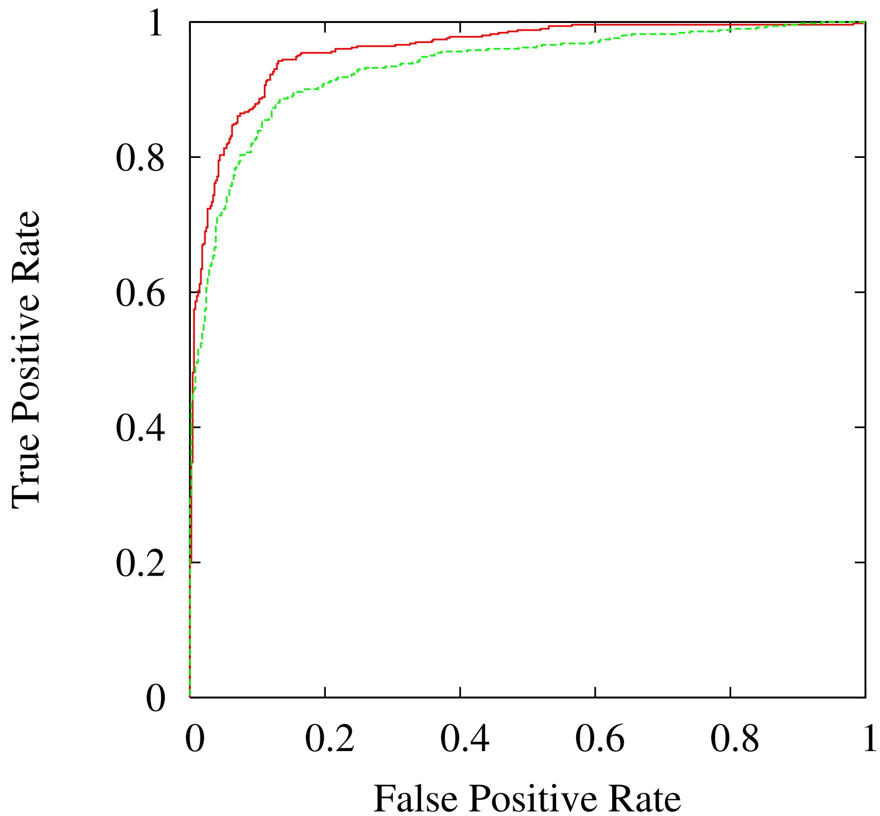
Change in Bias Weight	Effect	False Positives	False Negatives
increase	increase positive predictions	increase	decrease
decrease	increase negative predictions	decrease	increase

An ROC curve plots the the false positive rate vs. the true positive rate ($= 1 -$ false negative rate) for different bias weights.

We want a high true positive rate and a low false positive rate.

For a given false positive rate, we prefer a higher true positive rate, so prefer higher ROC curves.

The next figure shows the ROC curves for two SVMs. They were generated using the SVM values from 10-fold CV.

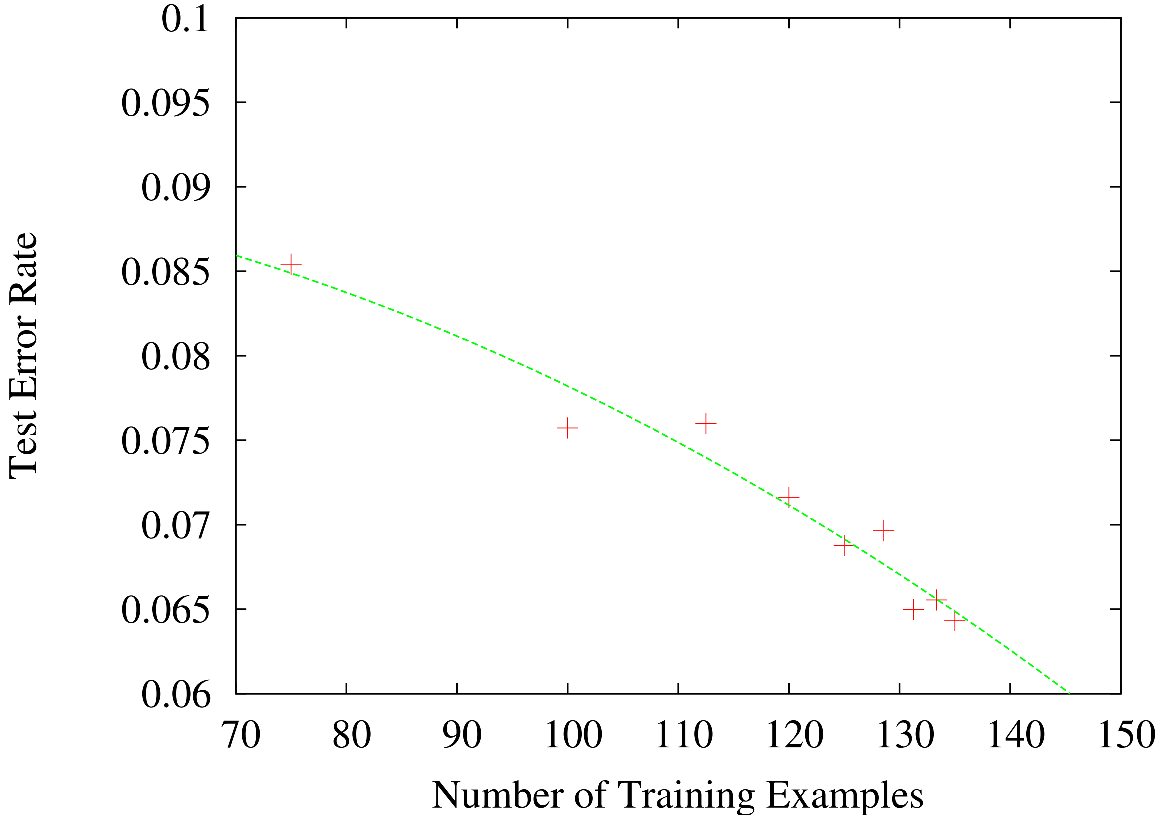


Generalization Curves

A generalization curve shows how the classifier improves with additional data.

If you have k-fold CV, then obtain the error rate from 2-fold CV (50% used for training) to 10-fold CV (90% used for training), and then fit a quadratic curve to the data to extrapolate.

The minimum of the quadratic curve provides an rough idea of whether more data will lead to much lower error rates.


0.1
0.095
0.09
0.085
0.08
0.075
0.07
0.065
0.06
Test Error Rate
70
80
90
100
110
120
130
140
150
Number of Training Examples