Problem 4.25

1. No, we should not select the learner with minimum validation error. First, the VC bound is:

Because each learner is trained on different size of validation set, is different for each model. Therefore selecting the minimum does not guarantee the minimum due to the fact that is different.

First, the validation error is an unbiased estimate of because the final hypothesis was obtained independently of the data points in the validation set.

Second, the validation process is equivalent to learning a hypothesis from using the data in the validation set. Therefore, we can apply the VC bound.

Using Hoeffding inequality to obtain a bound on the in-sample and out-of-sample errors:

The validation errors are 'in-sample' errors for this learning process. That is

Since we have evaluatations

Because

Let , we then have