1. We firstly notice that the total number of hypothesis is . We also realize that the size of training sample is , and the prediction error of , we have:

For all h, . Then, we know that the total prediction error can be bounded by with probability .

We want to bound this by . Then we have:

Thus, the prediction error could be bounded by .

1. Something
2. Given the formulation in 1, we know that