Paul Lintilhac

Foundations of Machine Learning

Homework 2

A: Rademacher Complexity

1)

First, note that if there is only one hypothesis in H, then we can remove the sup{} operator, since sup{const} = const. Then we are starting with the following equation for the Rademacher Complexity:

Since is a convex function, we can apply Jensens inequality to write

Now we expand the squared sum by writing it as a double summation,

Note that since are Rademacher variables, all are independent of each other when (and also independent of ). Therefore . Since , this means that all terms are zero except when . Note that = 1, since is always equal to 1. Then we have

Which is what we wanted to show.

2.

Claim:

Proof:

First, we will derive an expression for the empirical Rademacher complexity,

Because the dot product, supremum, and expectation are all linear operators. Note that we have used the identity . Since the distribution of is the same as the distribution of ,

Therefore,

Claim:

Proof:

Where we have used the sub-additivity of the sup operator. Now by the linearity of expectation we have

Claim:

Proof: