4.8

Greg Matesi

July 5, 2020

Logistic regression

Training error = .20.

Test error = .30

1NN

We know that

$$1/2(\text{Training error}) + 1/2(\text{Test error}) = .18.$$

Using 1 nearest neighbor, the model fits the training sample perfectly. So we can say

$$1/2(0) + 1/2(\text{Test error}) = .18$$
$$(\text{Test error}) = .36$$

Which one to use?

Using this information, we would probably use the one with the lowest training error. So we would use logistic regression.