

June 7, 2018

5.7 (b) The training error goes down as we increase D , but this does not necessarily mean the model with a large D is a good representation of the data. As we can see in the first figure, when $D > 5$, the model is quite complicated and the fitting curve seeks to pass through every data point, which indicates the overfitting problem. Also, we can see the model underfits the data if $D = 1$, so we should choose $D = 3$ or $D = 5$.

5.10

Eric should choose $D = 5$ as it produces minimal testing error.

Stanley could choose either $D = 3$ or $D = 8$ but $D = 3$ produces a simpler model.

Even though Kyle could choose $D = 10$, the decreasing trend in both the training and testing error curves suggests that we are still in the underfitting zone! So Kyle should try even larger values for D until the testing error ceases to decrease or goes back up.

In Kennys plot, the training error does not consistently go down which is indicative of a problem in his implementation. Therefore he should look into his implementation for debugging!