Exercise 5

Arthur Junges Schmidt

08 July 2020

(a) Linear regression This model would have a high bias and a low variance. The high bias occur due to the oversimplification of the regression (underfitting) by using a linear regression for a model with 4 predictors. The variance is lower because it does a better job generalizing for other data sets.

(b) Polynomial regression with degree 3 This model would have a medium value for both bias and variance. While it's more complex and does a better job decreasing the fitting error than linear regression, it still lacks complexity. However, with 3 degrees it would present more variance as it would increase error when utilizing other data sets.

(c) Polynomial regression with degree 10 This model would present a small bias and a high variance. While it would do a good job fitting very well the training data, it may overfit it. By overfitting the data, the variance of the random error would increase.
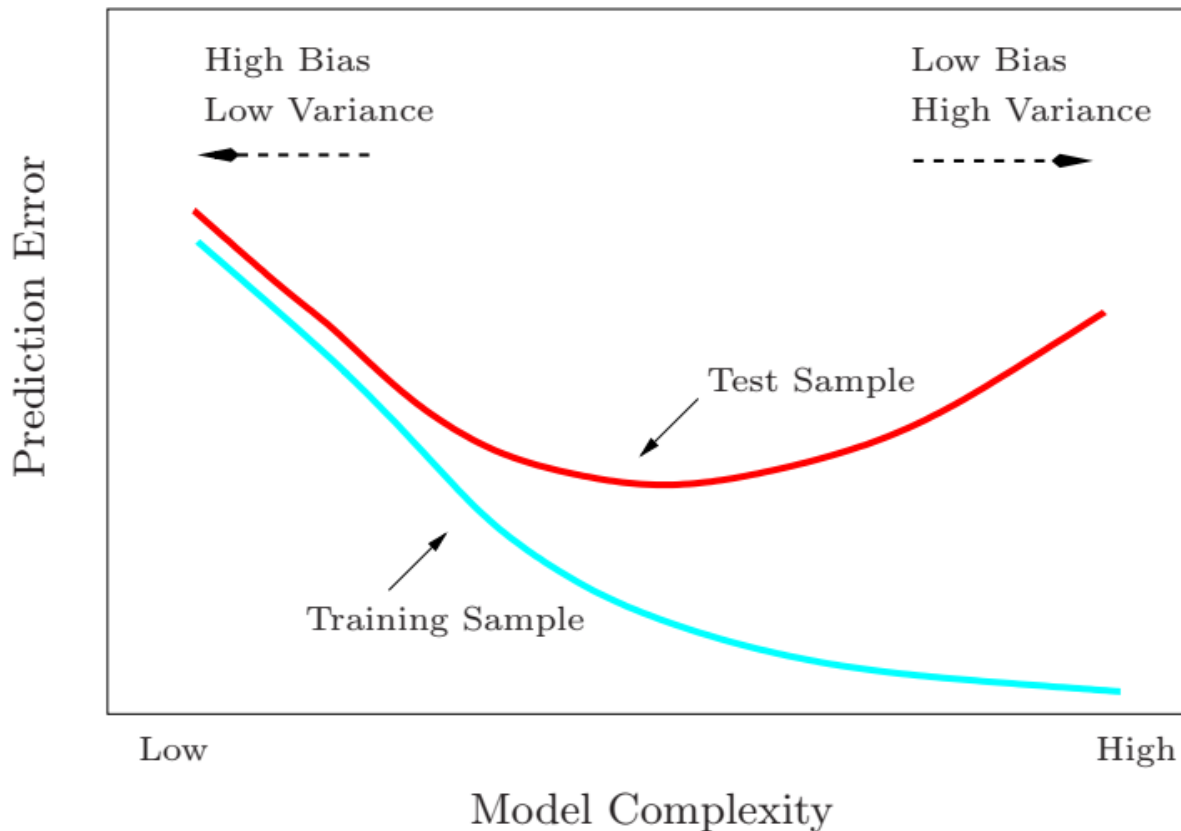

Figure 1: Test and training error as a function of model complexity. Hastie, T., Friedman, J., & Tisbshirani, R. (2018). The Elements of statistical learning: Data mining, inference, and prediction. New York: Springer.