

Exercises Chapter 2

Exercise 1

- Sample size n is extremely large with small Predictors P
 - Flexible methods will perform better cause they can capture complex patterns without overfitting , the high n provide enough data to estimate complex relationships
 - Sample size n is small while the predictors P is extremely large
 - Flexible methods will perform worse cause they have high variance and can easily over fit, Linear regression is more stable in high dimensional settings
 - The relationship between the predictors and the response is highly non-linear
 - inflexible models like linear regression assume a simple relationship between the predictions and the response
 - The variance of the error terms $\sigma^2 = Var(\varepsilon)$ is extremely high
 - Flexible is worse cause they amplifies the variance unlike the more Linear approaches
-

Exercise 2

- We collect a set of data on the top 500 firms in the US. For each firm we record profit, number of employees, industry and the CEO salary. We are interested in understanding which factors affect CEO salary.
 - This is a Regression Problem \rightarrow Salary is a quantitative
 - We interested in [Inference](#) \rightarrow Relationship (How all these factors they affect) CEO salary
 - $n = 500$ $p = 3$
 - We are considering launching a new product and wish to know whether it will be a success or a failure. We collect data on 20 similar products that were previously launched. For each product we have recorded whether it was a success or failure, price charged for the product, marketing budget, competition price, and ten other variables.
 - This is a Classification Problem \rightarrow Two Classes Success or Failure and they are qualitative value
 - We interested in Prediction \rightarrow either success or failure
 - $n = 20$, $p = 13$
 - We are interested in predicting the % change in the USD/Euro exchange rate in relation to the weekly changes in the world stock markets. Hence we collect weekly data for all of 2012. For each week we record the % change in the USD/Euro, the % change in the US market, the % change in the British market, and the % change in the German market
 - This is a Regression Problem \rightarrow % is a quantitative value
 - We interested in Prediction
 - $n = 53$, $p = 3 \rightarrow n$ is the number of weeks in a year
-

Exercise 3

- This Exercise challenged my understanding of Variance and the curve related to it and also clarified some concepts

