Yeabin Moon, Ph.D.
BUS 212A-2 Spring 2023
January 5, 2023

Homework 1

Submission Instructions

1) You have to submit three PDF files on Latte. Do not compress the files.
2) The first PDF should be your answers for the questions below.
3) The second PDF is your most recent resume. I want to see the quality one.
4) The last one is a cold email to your dream job.

Question 1. For the following scenarios, explain whether a flexible learning model performs better (or worse) than an inflexible model. Supply your intuition.

1) The sample size (the number of rows) is extremely large, and the number of features (the number of columns) is small.
2) The relationship between the features and target is highly non-linear.
3) The variance of the error terms ($\sigma^2 = Var(\epsilon)$) is extremely high.

Question 2. Explain when a less flexible model might be preferred against a more flexible one. You have to supply your intuition using real-life applications.

Question 3. Provide three different figures:

- Each plot includes five different curves: bias, variance, training error, test error, Bayes error curves
- The first plot is from a very flexible learning model.
- The second plot is from a very inflexible learning model.
- The third plot is from a model between the two models above.
- x-axis represents the model's flexibility, and the y-axis represents the values for each curve.
- Make sure to label everything clearly.
- You can use a software simulation or handwriting.

Question 4. Two teams play a best of 7 baseball game. That is, whoever wins 4 games first. Given that each team has a 50 % chance of winning in any given round (no draws), what is the probability that the series goes to 7 games?

Question 5. One in a thousand people have a particular disease, and the test for the disease is 98 % correct in testing for the disease. On the other hand, the test has a 1 % error rate if the person being tested does not have the disease. If someone tests positive, what are the odds they have the disease?

Question 6. You are given an unfair coin having an unknown bias towards heads or tails. How can you generate fair odds using this coin?

Question 7. Say you have a deck of 50 cards made up of cards in 5 different colors, with 10 cards of each color, numbered 1 through 10. What is the probability that two cards you pick at random do not have the same color and are also not the same number?

Question 8. Explain a confidence interval to a non-technical audience.