1. (3 points) Does a 3-NN regressor always outperform a 5-NN regressor?
   (a) True
   (b) False
   (c) It depends
   (d) All of the above
   (e) None of the above

2. (3 points) Which of the following models is preferred? Assume the models are trained in the same dataset and evaluated on the same test dataset.

Model	**RMSE**	**MAE**
Random forest with 100	14.3	18.2
Random forest with 300	12.1	15.0
Random forest with 500	12.3	15.9
Random forest with 1000	13.5	14.9
Random forest with 1500	12.1	14.9
Random forest with 2000	12.4	15.0

   (a) Random forest with 100
   (b) Random forest with 300
   (c) Random forest with 500
   (d) Random forest with 1000
   (e) Random forest with 1500
   (f) Random forest with 2000
   (g) None of the above

Consider the `autos.csv` datafile. Each row represents a car, each column contains information such as horsepower, number of cylinders, etc. The goal is to predict the miles per gallon, `mpg`, using the other car attributes.

3. **In Python**, answer the following:
   (a) (3 points) Using the `pandas`, read the csv file and create a data-frame called `autos`.
   (b) (4 points) Split the data into `train` (80%) and `test` (20%).
   (c) (8 points) Using the `train` dataset and the following number of trees (`n_tree = [100, 300, 500, 800, 1000]`) train random forest models in the `train` dataset, in which `cylinders`, `displacement`, `horsepower`, `weight`, and `acceleration` are the input variables, and `mpg` is the target variable. Then, using those models, predict `mpg` on the `test` dataset. Compute the RMSE and MAE for each of the models. Report the model with the smallest RMSE and MAE.

4. **In R**, answer the following:
   (a) (3 points) Using the `read.csv` function, read the csv file and create a data-frame called `autos`.
   (b) (4 points) Split the data into `train` (80%) and `test` (20%).

(c) (8 points) Using the `train` dataset and the following number of trees and depths (`n_tree = c(100, 300, 500, 800, 1000, 1500, 2000)` and `n_depth = c(2, 3, 4, 5, 6)`) train gradient boosting models in the `train` dataset, in which `cylinders`, `displacement`, `horsepower`, `weight`, and `acceleration` are the input variables, and `mpg` is the target variable. Then, using those models, predict `mpg` on the `test` dataset. Compute the RMSE and MAE for each of the models. Report the model with the smallest RMSE and MAE.

The End.