1. (3 points) Bagging can only be applied to decision trees.
   (a) True
   (b) False
   (c) It depends
   (d) All of the above
   (e) None of the above
2. (3 points) Random forest can be applied on regression and classification tasks.
   (a) True
   (b) False
   (c) It depends
   (d) All of the above
   (e) None of the above
3. (3 points) To apply bagging to regression trees which of the following **is/are** true in such case?
   (a) We build the N regression with N bootstrap sample.
   (b) We take the average the of N regression trees.
   (c) Each tree has a high variance with low bias.
   (d) (a) and (b)
   (e) (a) and (c)
   (f) (b) and (c)
   (g) (a), (b) and (c)

Consider the `Customer_Churn.csv` datafile. Each row represents a customer, each column contains customer's attributes described on the column Metadata. The data set includes information about:

- Customers who left within the last month, the column is called Churn.
- Services that each customer has signed up for phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies.
- Customer account information: how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges.
- Demographic info about customers: gender, age range, and if they have partners and dependent

4. **In Python**, answer the following:
   (a) (3 points) Using the `pandas` library, read the csv file and create a data-frame called `churn`.
   (b) (3 points) Using the `numpy` library, create a variable called `Churn_numb` that takes the value of 1 when `Churn = Yes` and 0 when `Churn = No`.
   (c) (5 points) Using the `sklearn.ensemble` library, build a random forest classifier, in which `Churn_numb` is the target variable, and `tenure` and `MonthlyCharges` are the input variables. Predict the likelihood of churn of a customer with `tenure` equal to 20, and `MonthlyCharges` equal to 80.

5. **In R**, answer the following:

   (a) (3 points) Using the `read.csv` function, read the csv file and create a data-frame called `churn`.

   (b) (3 points) Using the `ifelse` function, create a variable called `Churn_numb` that takes the value of 1 when `Churn = Yes` and 0 when `Churn = No`.

   (c) (5 points) Using the `randomForest` package, build a random forest classifier, in which `Churn_numb` is the target variable, and `tenure` and `MonthlyCharges` are the input variables. Predict the likelihood of churn of a customer with `tenure` equal to 12, and `MonthlyCharges` equal to 250.

The End.