Statistical Modelling & Machine Learning HW3
(Due: 12/05/2021, Sunday)

Instruction:

- There is no correct or unique answer in this homework.
- I will give your HW score based on your results and analysis procedure.
- Submit your analysis report and a `R` code files.
- Your report should include the **description of your analysis procedure**.
- Your `R` code should show the procedure that you obtain the final result (**Do NOT include `R` codes for all procedure that you have tried**).

1. Consider the `train.csv` and `test.csv` data files. Using the `train.csv` dataset, build your predictive model and then apply your predictive model to `test.csv` dataset and report the F-measure value for the test set.

Situation:

- Suppose that A bank performs marketing campaigns through phone calls to clients and the bank wants to predict clients who buy financial products or services by the campaigns. Also, suppose that the bank wants to know what factors are important for the prediction.

Description of variables:

- `X1`: Client's age.
- `X2`: Client's job.
- `X3`: Marital status.
- `X4`: Education level.
- `X5`: Does client have in default?
- `X6`: Average yearly balance.
- `X7`: Does client have housing loan?
- `X8`: Does client have personal loan?
- `X9`: Main communication type.
- `X10`: Last contact day of the month.
- `X11`: Last contact month of year.
- `X12`: Last contact duration.
- `X13`: The number of contacts performed during this campaign.

- `X14`: The number of days that passed by after the client was last contacted from a previous campaign.
- `X15`: The number of contacts performed before this campaign.
- `X16`: Outcome of the previous marketing campaign.
- `Y`: Did client buy financial product or service? (yes: marketing success; no: marketing failure).

Instruction and suggestion for analysis:

- You can use any predictive models including data models and algorithmic models.
- There are 'unknown' category in some variables. You can consider it as a new category or as a miss value.