

Exercise 1:

- a) Take a look at the `spam` dataset (`?mlr3::mlr_tasks_spam`). Shortly describe what kind of classification problem this is and access the corresponding task predefined in `mlr3`.
- b) Use a decision tree to predict spam. Try refitting with different samples. How stable are the trees?
Hint: Use `rpart.plot()` from the package `rpart.plot` to visualize the trees. (You can access the model of a learner by its class attribute `model`)
- c) Use the random forest learner `classif.ranger` to fit the model and state the oob-error.
- d) Your boss wants to know which variables have the biggest influence on the prediction quality. Explain your approach in words as well as code.
Hint: use an adequate variable importance filter as described in <https://mlr3filters.ml-org.com/#variable-importance-filters>.

Exercise 2:

Generate an artificial dataset with the function call `mlbench.spirals(n = 500, sd = 0.1)`. (The function `mlbench.spirals` is part of the `mlbench` package.) Visualize the decision boundaries of a random forest using the `classif.ranger` learner from `mlr3learners`. Create plots with `plot_learner_prediction` from `mlr3viz` for an increasing number of trees. (Start with `num.trees = 1`) Explain what you see.