Practice Assignment 09

Create a GitHub repository called “st2195_assignment_9” and include a script in R and Python able to perform the following operations:

1. Extend the results on the titanic dataset in the module of machine learning in Python, by using the additional features that are created and considered in the Practice Assignment 7 [5 points]
2. Extend the results further by using other valid classification models listed in the notes on machine learning frameworks. Please try as many as you can [5 points]

Additional Notes:

- Hints
  - Additional feature used in Practice 7 was “family size”, which is calculated using sibsp + parch + 1
    - sibsp: # of siblings / spouses aboard the Titanic
    - parch: # of parents / children aboard the Titanic
  - Start with Python (suggested). You can save the titanic dataset extracted from OpenML in csv format for use in the R version
- Python version
  - Much of the code covered in the lecture notes can be reused for the assignment.
  - Just need to find additional classification models (see https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html)
- R version
  - Useful reference on mlr3 at https://introduction-to-machine-learning.netlify.app/