Covariance, correlation and decision trees

Tutorial/Lab work: Week 2

Instructions: This is a coding exercise. Kindly stage+commit+push the `MATLAB` code to the "`week02`" directory on your INM431 Github repository. Grading: All attempts will get 1 point. **Please make sure that the document is pushed to Github by 20/10/2021.**

Problem: 1 Covariance and Correlation

In `MATLAB`, load the hospital.mat data set and:

1. Calculate the covariance between Weight and Blood pressure
2. Calculate the correlation coefficient between Age and Blood pressure

Hints:

- Familiarize yourself with the hospital.mat dataset, find more details here: `https://uk.mathworks.com/help/stats/dataset-array-columns.html`
- Notice that `BloodPressure` is composed of `SysPressure` and `DiaPressure`.
- Find out how to calculate covariances in `MATLAB` here: `https://uk.mathworks.com/help/matlab/ref/cov.html`
- Find out how to calculate correlations in `MATLAB` here: `https://uk.mathworks.com/help/matlab/ref/corrcoef.html`
- Optional: check out the following list of `MATLAB` sample datasets and consider calculating more covariance matrices: `https://uk.mathworks.com/help/stats/sample-data-sets.html`

Problem 2: Decision trees

Apply `MATLAB` function: `fitctree` and predict to the `fisheriris.mat` dataset

Hints:

- Find out what this data set is about at: `https://en.wikipedia.org/wiki/Iris_flower_data_set`
- Follow `MATLAB` examples of Decision Trees here: `https://www.mathworks.com/help/stats/decision-trees.html`
- Learn more about displaying trained Decision Trees at: `https://uk.mathworks.com/help/stats/view-decision-tree.html`
- Learn more about making predictions using a Decision Tree: https://uk.mathworks.com/help/stats/prediction-using-classification-and-regression-trees.html
- Finally, learn implementing cross-validation for Decision Trees (`crossval`): `https://www.mathworks.com/help/stats/classificationtree.crossval.html`