Tree Learning
Exercises

Exercise 1.

1. Describe in your own words the learning and classification process of decision trees.
2. Which form should a decision tree have? As broad as possible, or as deep as possible? Why?

Exercise 2.
A hospital introduces a software to support diagnosing patients. For that purpose, data on ill and healthy patients have been collected. Use the data to create a decision tree that predicts a patients health status based on their heart rate and blood pressure.

Patient	Heart Rate	Blood Pressure	class
1	irregular	normal	ill
2	regular	normal	healthy
3	irregular	abnormal	ill
4	irregular	normal	ill
5	regular	normal	healthy
6	regular	abnormal	ill
7	regular	normal	healthy
8	regular	normal	healthy

Use information gain to compute the tree.

Exercise 3.
Create a Jupyter Notebook and

1. load the IRIS dataset
2. select only features 2 and 3 (counting from 0) as features, thus creating a 2-dimensional feature space
3. use the random seed 1 if you use random
4. create a classification experiment, preprocess the data, split into training (60%) and test data (40%), use random sampling
5. think about feature scaling – is it necessary?

6. use a diagram to indicate the relationship between tree depth and the classification quality in terms of accuracy

7. plot the decision regions of the best classifier (use the python module `classification_viz` for that purpose).