

## **Machine Learning and Fintech Hw 2**

### **1. Theoretical problem**

Introduction to statistical learning with applications in R. Page 414: 3, 4.

### **2. Computing problem.**

In this homework assignment, you will download the dataset on NewE3 “seeds\_dataset” to solve the following problems.

The explanation of variables can find in the resource :

<http://archive.ics.uci.edu/ml/datasets/seeds>

- (a) Analysis the data by K-means clustering. Explain how you choose the number of cluster and illustrate the answer clearly.
- (b) Analysis the data by the hierarchical clustering. Under which similarity metric and linkage method does it produce clusters closest to the original classes?