ENGSCI 712 – Computational Techniques for Signal Processing

Module I: Feature Extraction, Assignment 1.1 & 1.2

Andreas W. Kempa-Liehr

Department of Engineering Science

Due date 18.08.2020 10pm– CSP-assignment-1.pdf, Rev. 795a84d

Outline

Assignment instructions

- The scope of the assignment has been reduced, because the 9th lecture was cancelled due to lockdown restrictions.

Overview

1. Setup your virtual environment (instructions follow)
2. Write both assignment tasks 1.1 and 1.2 using the same Jupyter notebook
3. Export your assignment as HTML

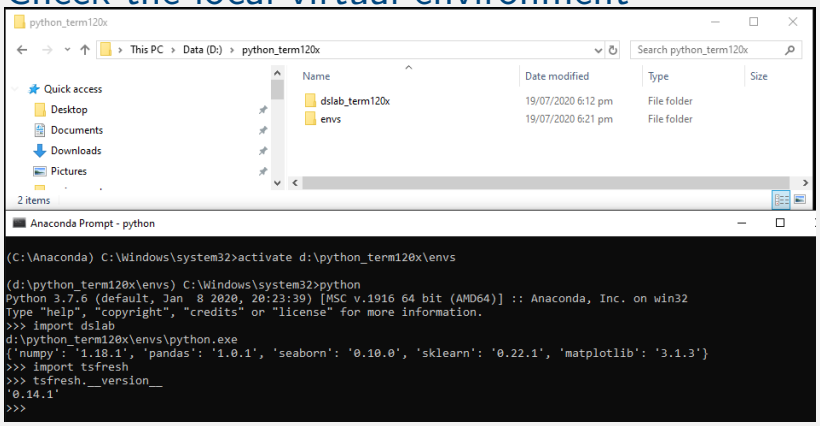
Due date 18.08.2020 10pm

- Upload the HTML version of your notebook to Canvas

Setting up your development environment

Lab Computer in 439-427

Check the local virtual environment



Private Computer

- Follow the instruction at

`https://stash.auckland.ac.nz/projects/DSEA/repos/dslab/browse?at=term1205`

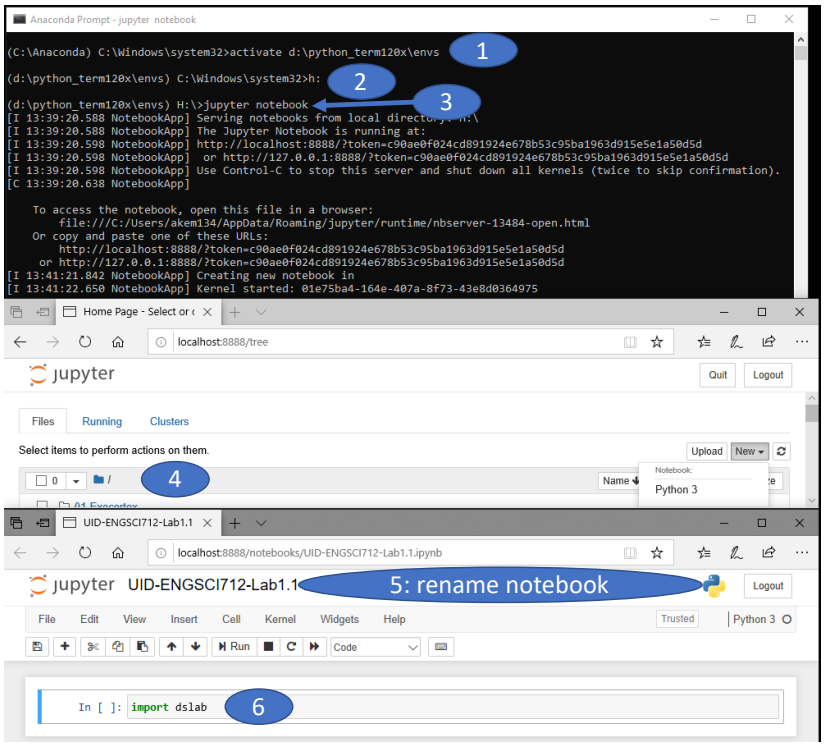
Download Jupyter notebook with assignment tasks 1.1

- Download `FE03-CSP-assignment_1.1.ipynb` from Canvas → files → assignment 1

`https://canvas.auckland.ac.nz/files/5168204/download?download_frd=1`

- Your browser might add the ending `.txt` to the downloaded file. In this case, you need to rename the file such that its filename ends with `.ipynb`.

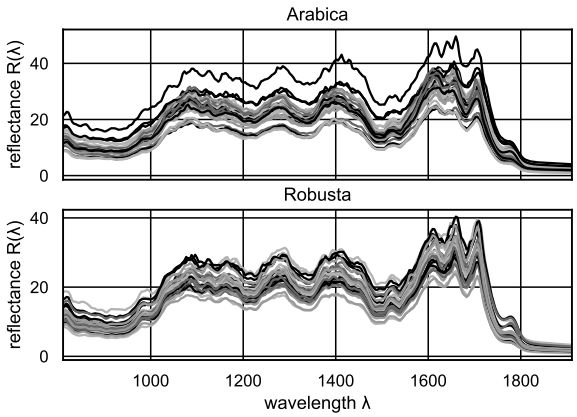
Start working with Jupyter Notebook



1. Activate virtual environment (this command will be different on your private computer). Check the location of your virtual environment using `conda env list`
2. Go to the folder, which has the Jupyter notebook `FE03-CSP-assignment_1.1.ipynb` from Canvas
3. Start Jupyter notebook server. If command `jupyter` is missing: `conda install jupyter`
4. Select `FE03-CSP-assignment_1.1.ipynb`
5. Rename notebook using your UID
6. Check correct installation of `dslab`

Welcome to CoffeeAI

You have decided to join the Auckland based technology startup CoffeeAI, which develops an *intelligent* coffee machine named *eBarista*. This coffee machine is supposed to automatically tune its grinding and brewing process according to the coffee beans provided.
The heart of each *eBarista* is a Fourier-Transform Infrared Spectrometer, which measures the infrared spectrum of absorption of the coffee beans.
Your task is to use computational signal processing to generates features, which can be used for discriminating Arabica and Robusta beans.



Data source: Romain Briandet, E. Katherine Kemsley, and Reginald H. Wilson. "Discrimination of Arabica and Robusta in Instant Coffee by Fourier Transform Infrared Spectroscopy and Chemometrics". In: *Journal of Agriculture and Food Chemistry* 44.1 (1996), pp. 170–174. DOI: 10.1021/jf950305a

In order to evaluate your features, you're going to fit a logistic regression classifier and perform an in-sample analysis.

Fourier Transform Infrared Spectrometer

FTIR spectrometers (Fourier Transform Infrared Spectrometer) are widely used in organic synthesis, polymer science, petrochemical engineering, pharmaceutical industry and food analysis. In addition, since FTIR spectrometers can be hyphenated to chromatography, the mechanism of chemical reactions and the detection of unstable substances can be investigated with such instruments.
Source: `https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Spectroscopy/Vibrational_Spectroscopy/Infrared_Spectroscopy/How_an_FTIR_Spectrometer_Operates`

FTIR_Spectra_instant_coffee.csv

- Download the FTIR_Spectra_instant_coffee.csv from Canvas.
- Load the data set into your Jupyter notebook using the read_csv function from pandas

```
import pandas as pd
data = pd.read_csv(...)
```

- Visualize the data set using functions from seaborn

```
import seaborn as sns
```

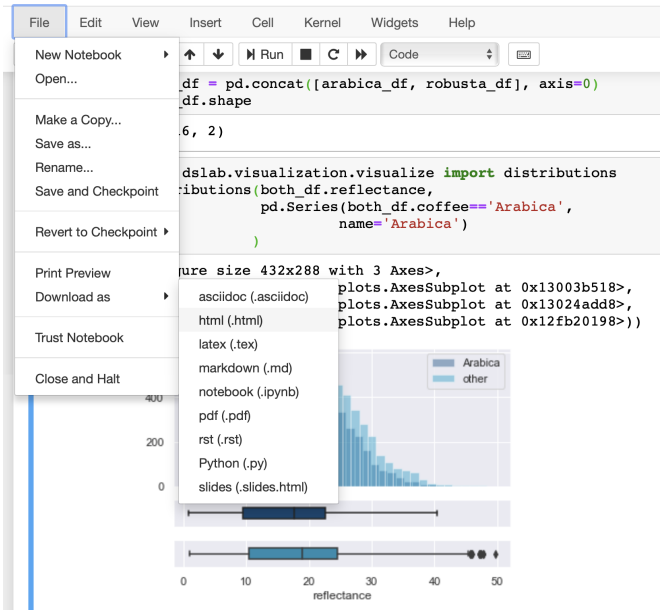
seaborn: statistical data visualization — seaborn 0.10.1 documentation

Design signal processors, which extract information about the statistics of the signals

- Extract the following features from the spectra: minimum, mean, standard deviation, variance, and maximum
- ~~Fit a logistic regression classifier using the L1 norm~~
- ~~Inspect the most important features~~
- Visualize the extracted features
- For each of the extracted features
  - Evaluate the accuracy and confusion of a signal operator, which aims to distinguish the different types of coffee beans
- Design a signal operator, which uses two of the extracted features for distinguishing the different types of coffee beans

Write and submit your report

- Export your Jupyter notebook as html



- Upload the html file to Canvas