Spectroscopy AI application: identification of absorption line systems

Lecture 12

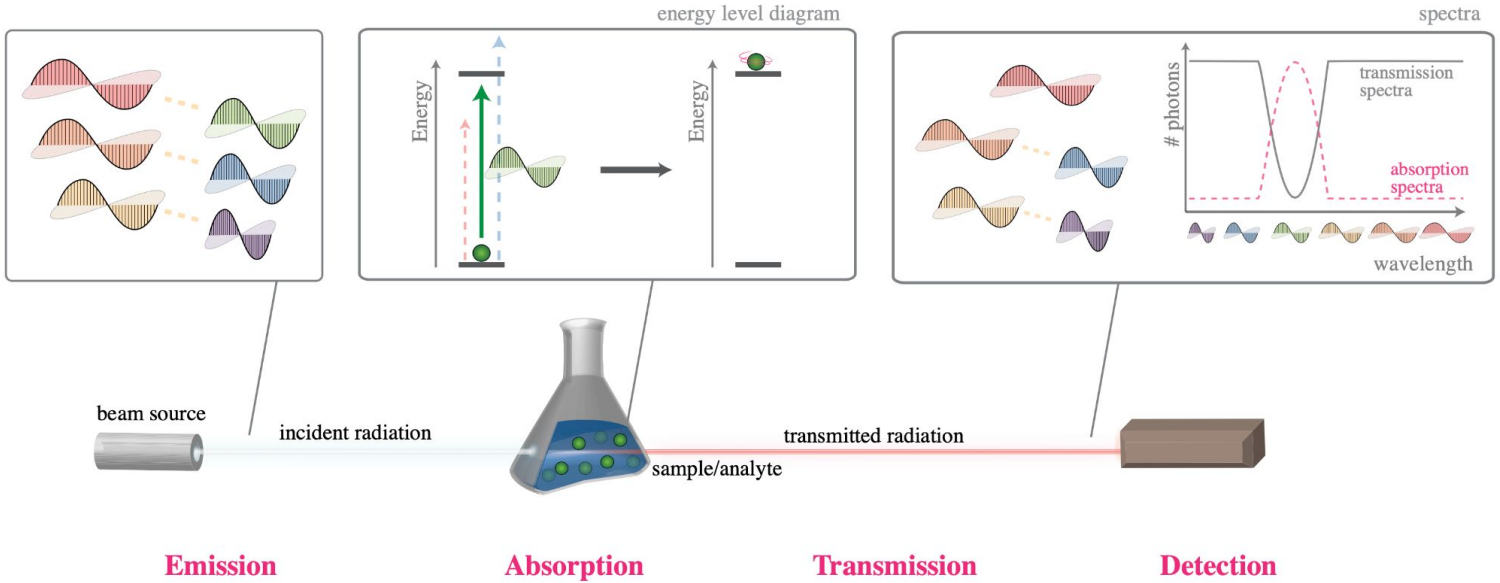
Course of:

Signal and imaging acquisition and modelling in environment

17/04/2023

Federico De Guio - Matteo Fossati

Absorption spectroscopy is a powerful tool in many fields of science

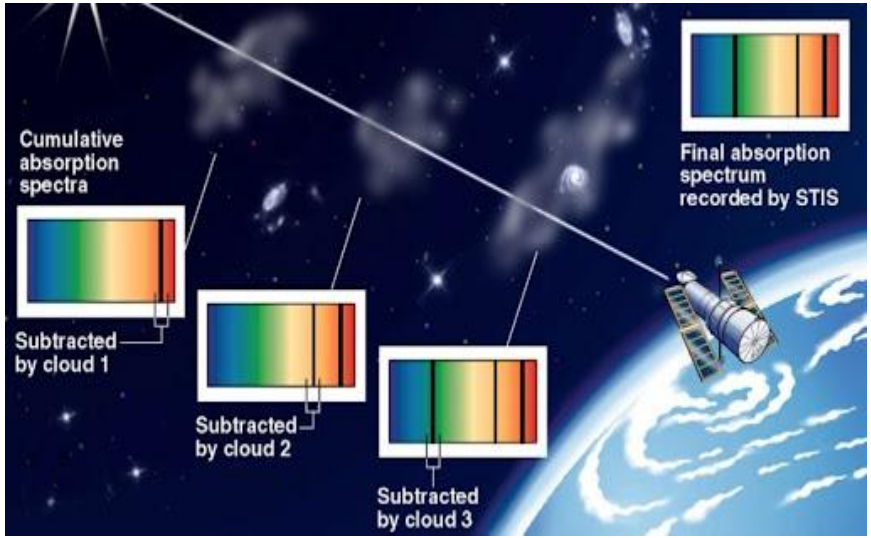


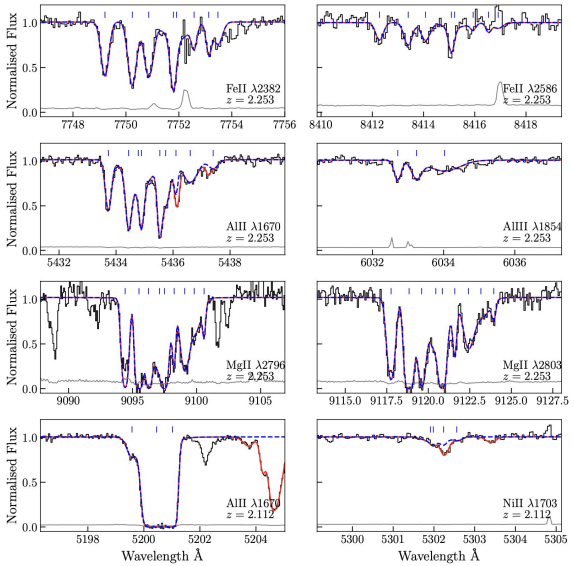
This technique can reveal the composition and density of a sample and has many applications in: Medicine, Astronomy, Environmental monitoring, Chemistry, ….

AI is fundamental in Astronomical absorption spectroscopy

An absorption line system can occur anywhere in a recorded spectrum of a distant source.

AI is fundamental in Astronomical absorption spectroscopy





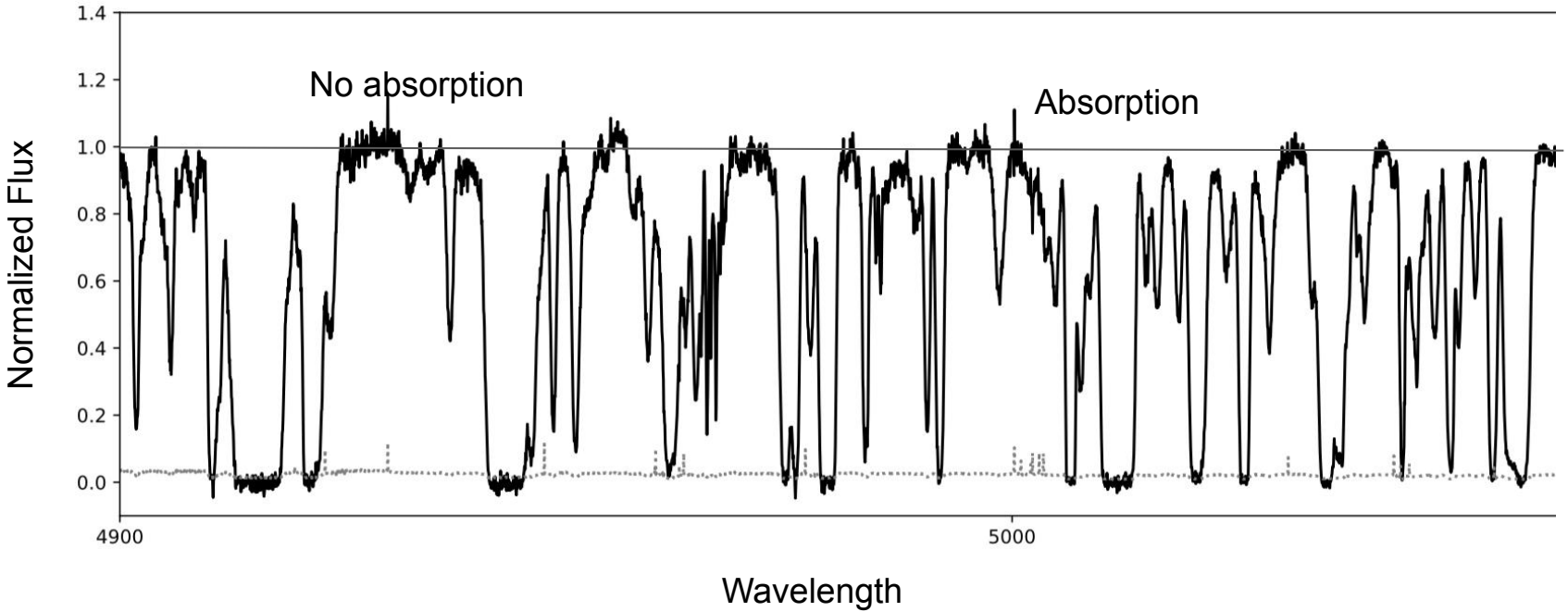
The number, position, and properties of absorption lines can vary from spectrum to spectrum.
Human inspection is impractical beyond few tens of spectra.

Samples with 100.000, 1.000.000 spectra NEED AI.

AI is fundamental in Astronomical absorption spectroscopy

We define three main problems:

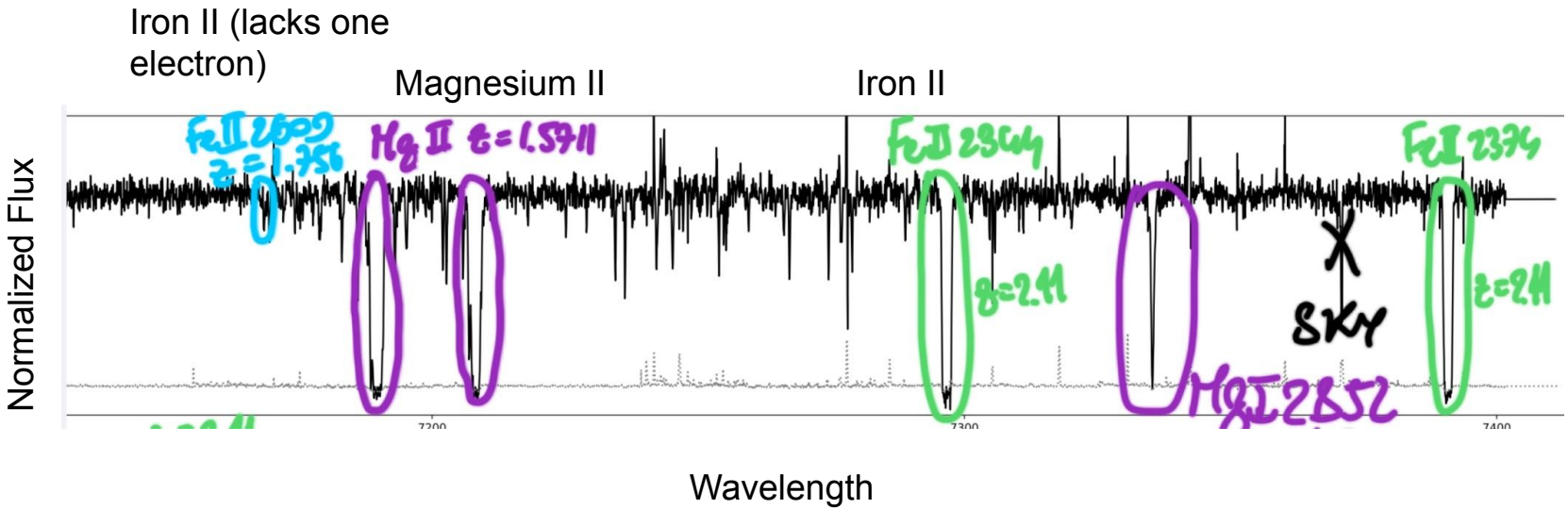
1. Identification (which pixels in a spectrum are absorption systems and which are not)



AI is fundamental in Astronomical absorption spectroscopy

We define three main problems:

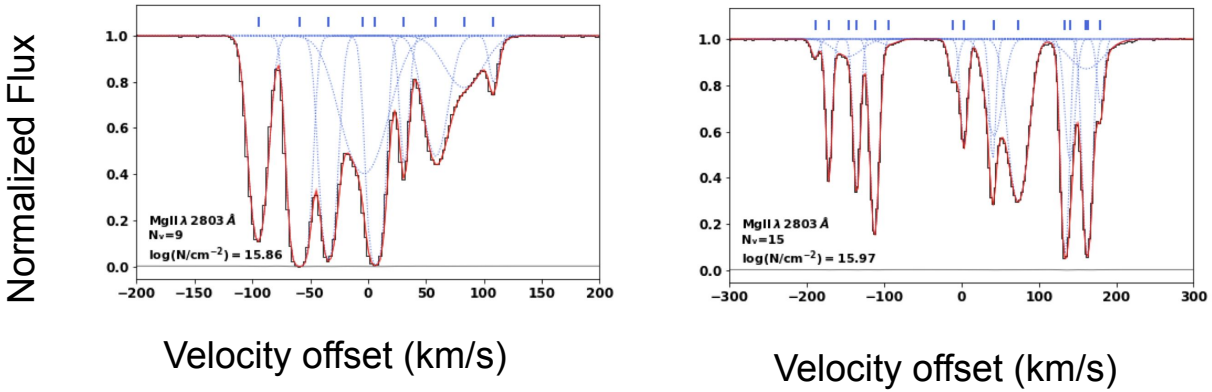
1. Identification (which pixels in a spectrum are absorption systems and which are not)
2. Classification (identify which atomic element is producing an absorption line system)



AI is fundamental in Astronomical absorption spectroscopy

We define three main problems:

1. Identification (which pixels in a spectrum are absorption systems and which are not)
2. Classification (identify which atomic element is producing an absorption line system)
3. Physical measurements (Column density of the gas, temperature, turbulence, number of Gaussian components)



AI is fundamental in Astronomical absorption spectroscopy

We define three main problems:

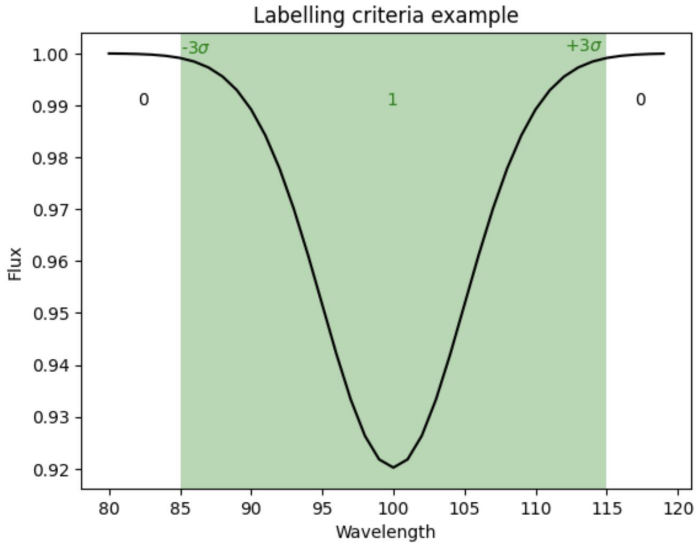
1. Identification (which pixels in a spectrum are absorption systems and which are not)
2. Classification (identify which atomic element is producing an absorption line system)
3. Physical measurements (Column density of the gas, temperature, turbulence, number of Gaussian components)

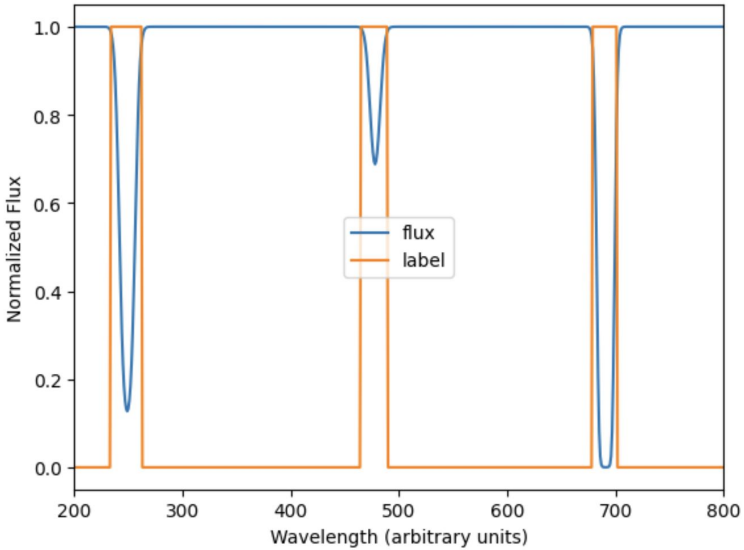
Today we address the first problem: *Identification*, you will generate a sample of random Gaussian absorption profiles and you will use AI to identify them.
This problem is sometimes called *segmentation (i.e. assign a label to each pixel in an image/spectrum)*

My group actively works on all these three problems and we offer AI4ST stages and thesis projects on these problems.

Segmentation of a 1D spectrum

Generate the data and define the labels



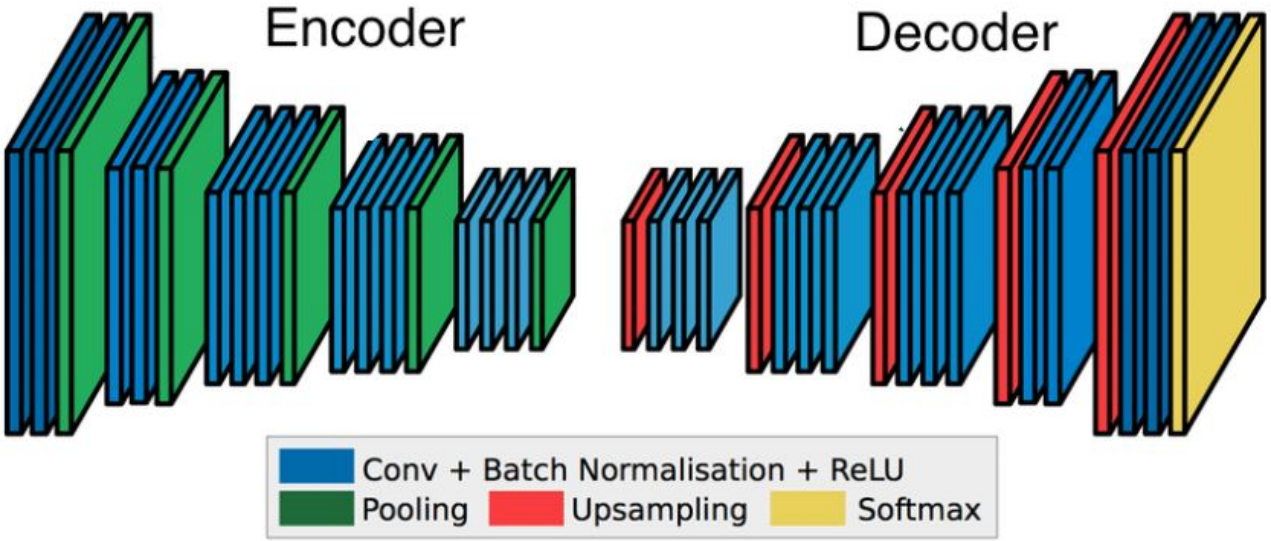


Problem: A Gaussian profile extends indefinitely, where do we cut the label? Cutting at 3 sigma is a reasonable choice.

Segmentation of a 1D spectrum

Which supervised technology would you use to perform such task?

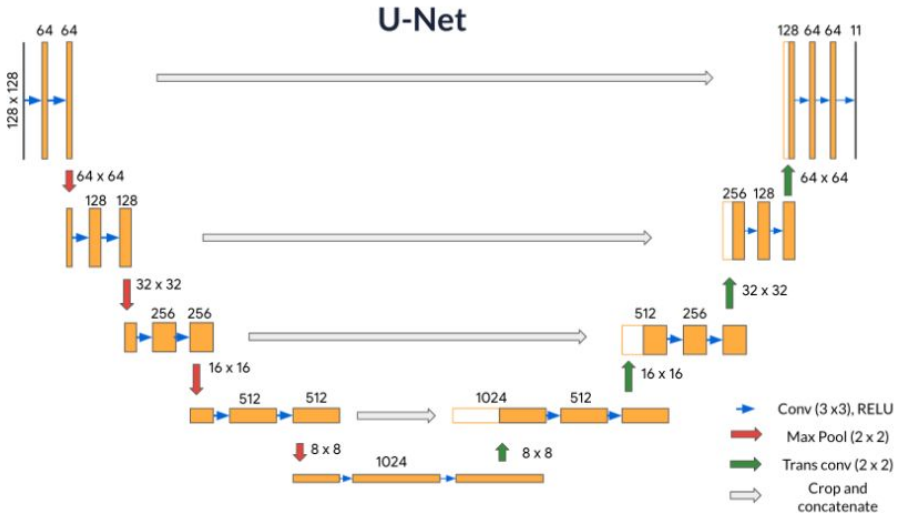
Segmentation of a 1D spectrum



- **Encoder - Decoder structure (similar to the autoencoder)**
- Left half of the net: same structure used to perform classification
- Right half of the net: decode the classification features into an array with pixel based predictions

A popular segmentation model: UNet

- The UNet was designed in 2015 to process biomedical images
  - Classify whether there is a disease
  - Localize the area of abnormality → input and output have the same size
    https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5



Your Task for today

Using the notebook provided you can generate a sample of simulated absorption line spectra, then (today or on Friday) we will move to a realistic dataset.

Inspect the function that generates the absorbers
(why is the absorber applied as flux = cont * exp(-tau)?)

Implement a UNet (or another segmentation network of your choice) to segment the spectra and identify absorption systems.

Study the performance of the network as a function of the SNR of the data, or the number of absorbers in a given spectrum.

Script link:
https://colab.research.google.com/drive/1_SUPb_W5wDeTWMKvL4kaBaLpEcgagOgB?usp=share_link