Deep Learning in Large Astronomical Archives

Bachelor’s Thesis by Ondřej Podsztavek, FIT CTU

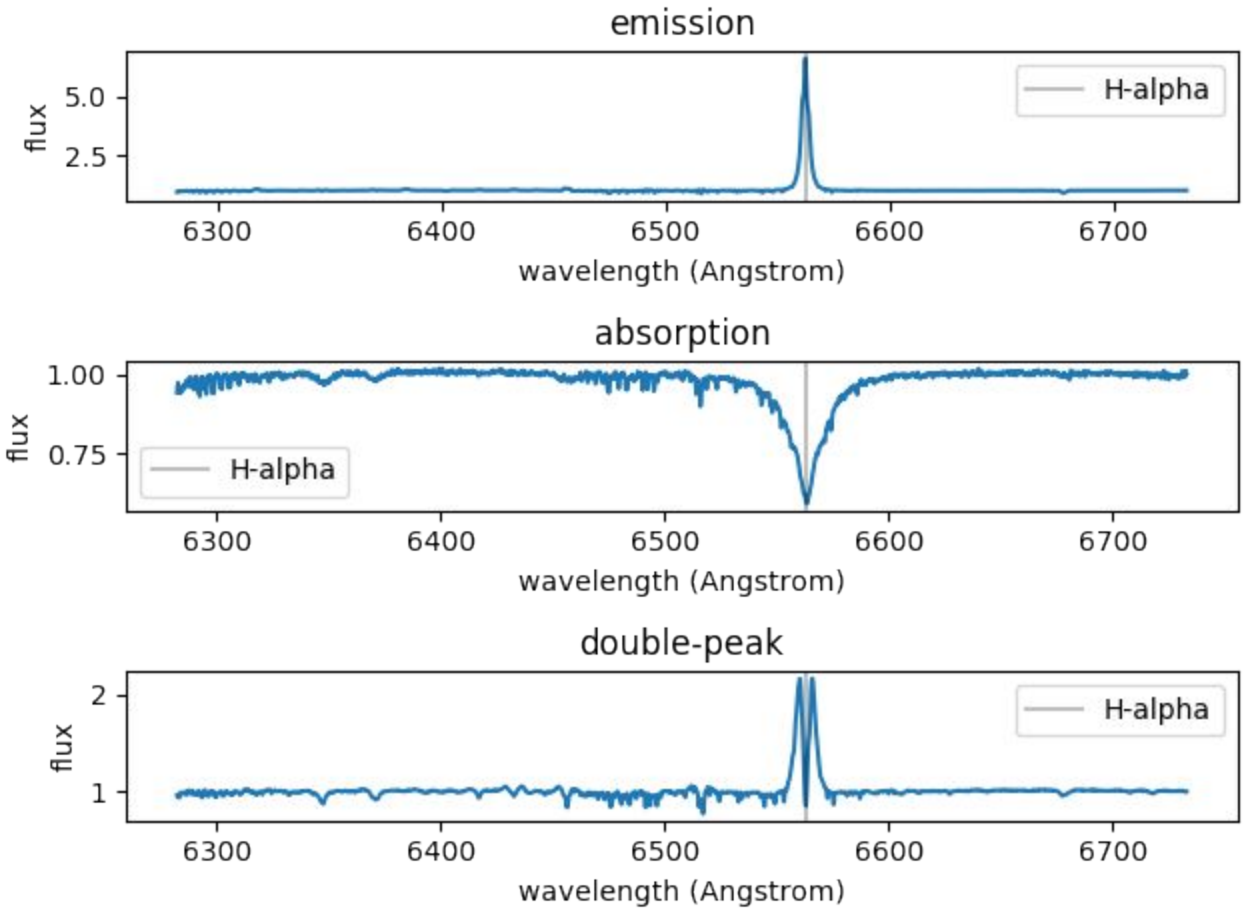
supervisor: Petr Škoda

Motivation

Identify emission-line spectra in the LAMOST spectral survey archive using deep neural network

Motivation

Identify emission-line spectra in the LAMOST spectral survey archive using deep neural network **trained on spectra from Ondřejov archive.**



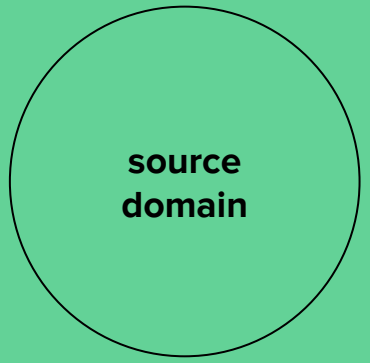
The three classes

The first challenge

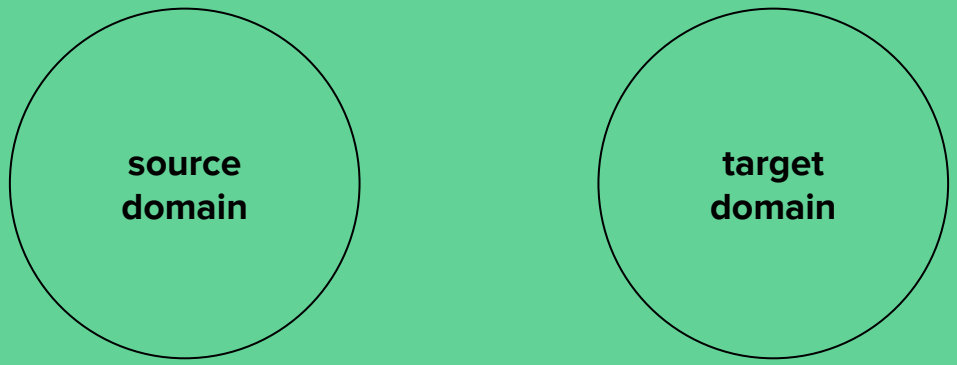
Domain adaptation of Ondřejov training data to LAMOST resolution.

Transfer Learning
Domain Adaptation

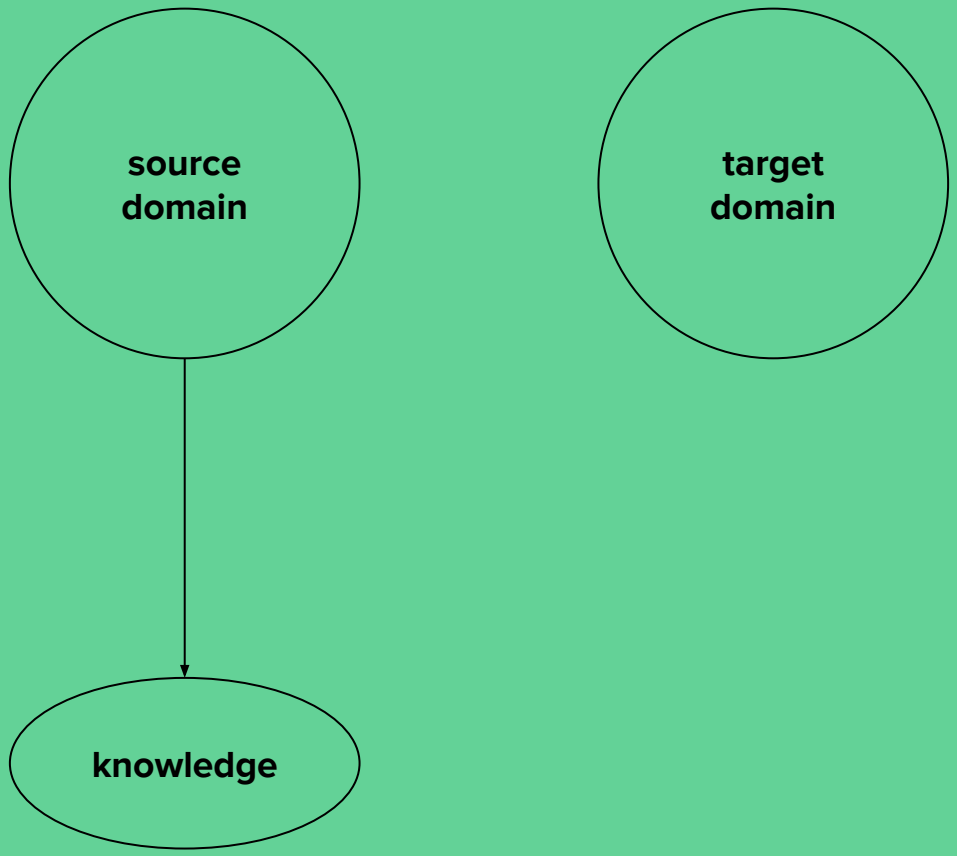
Transfer Learning
Domain Adaptation



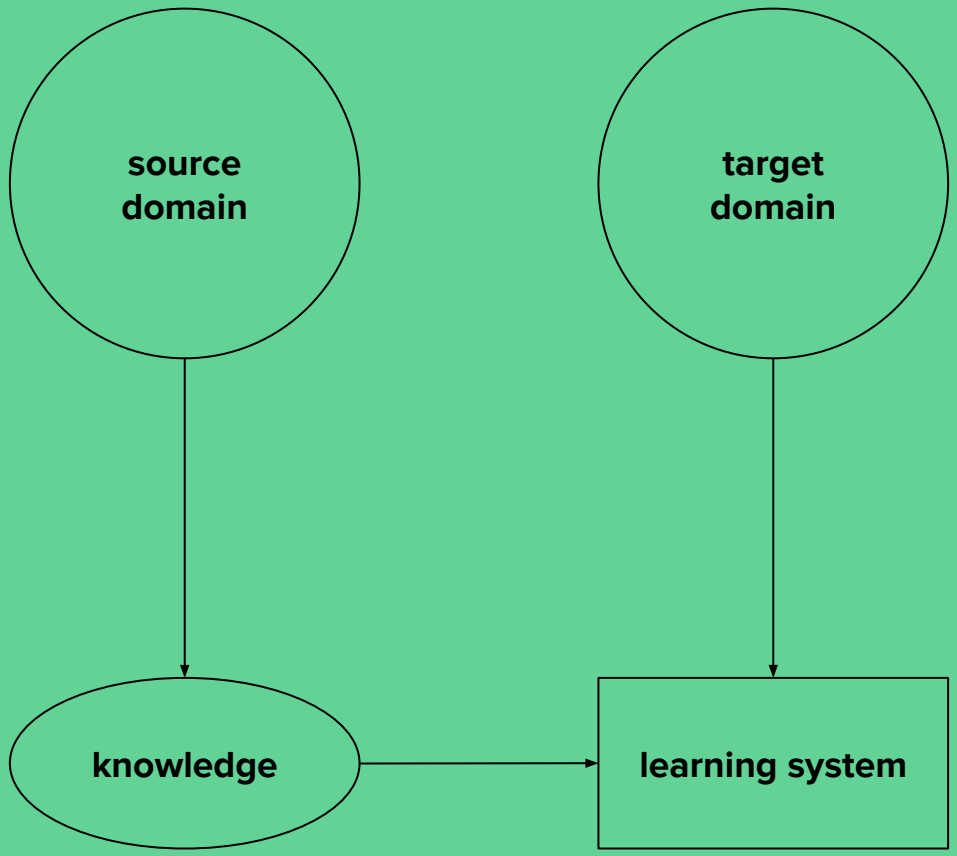
Transfer Learning
Domain Adaptation

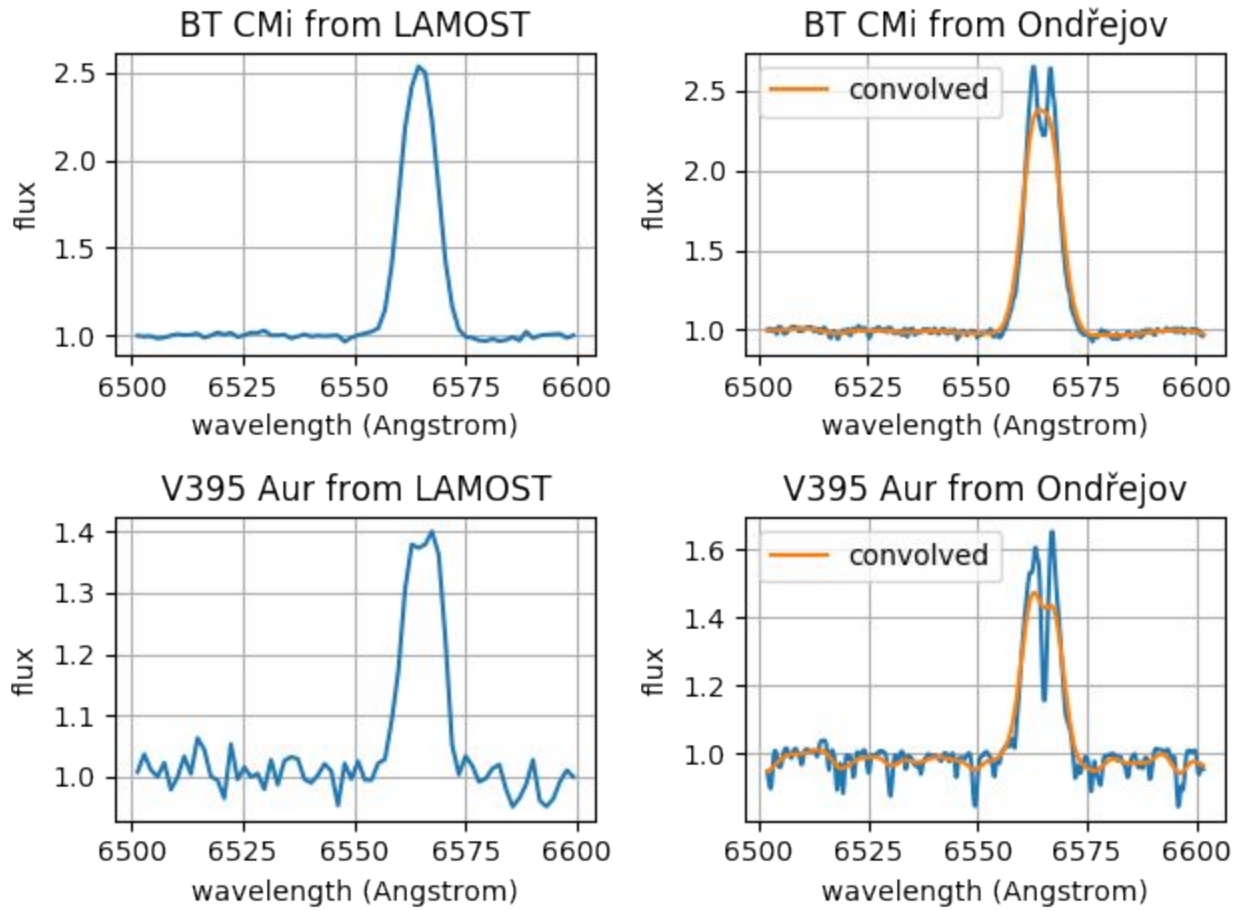


Transfer Learning
Domain Adaptation



Transfer Learning
Domain Adaptation

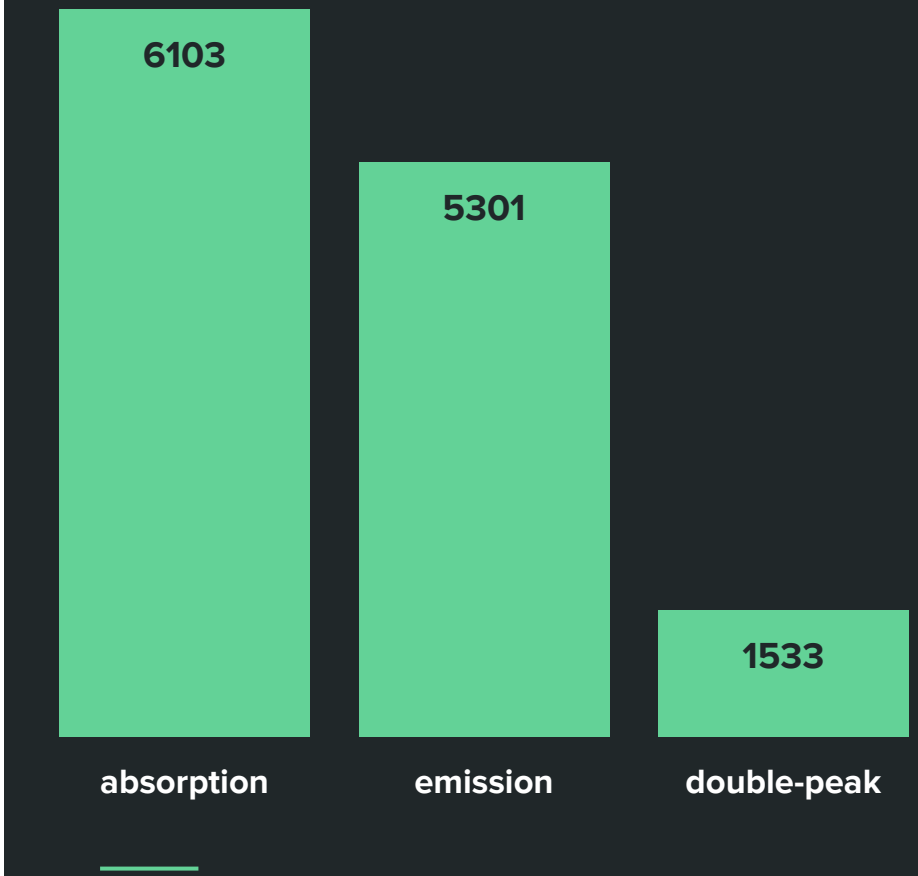




Gaussian blur

The second challenge

Imbalanced training dataset balanced with SMOTE.

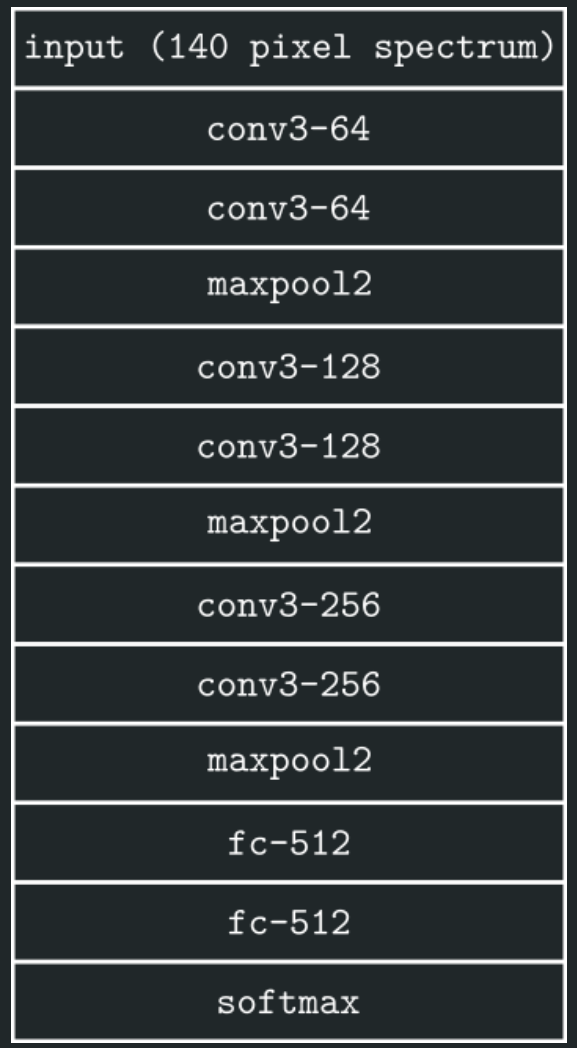


Convolutional Network

Rather deeper network to have representation power and **dropout** to reduce overfitting.

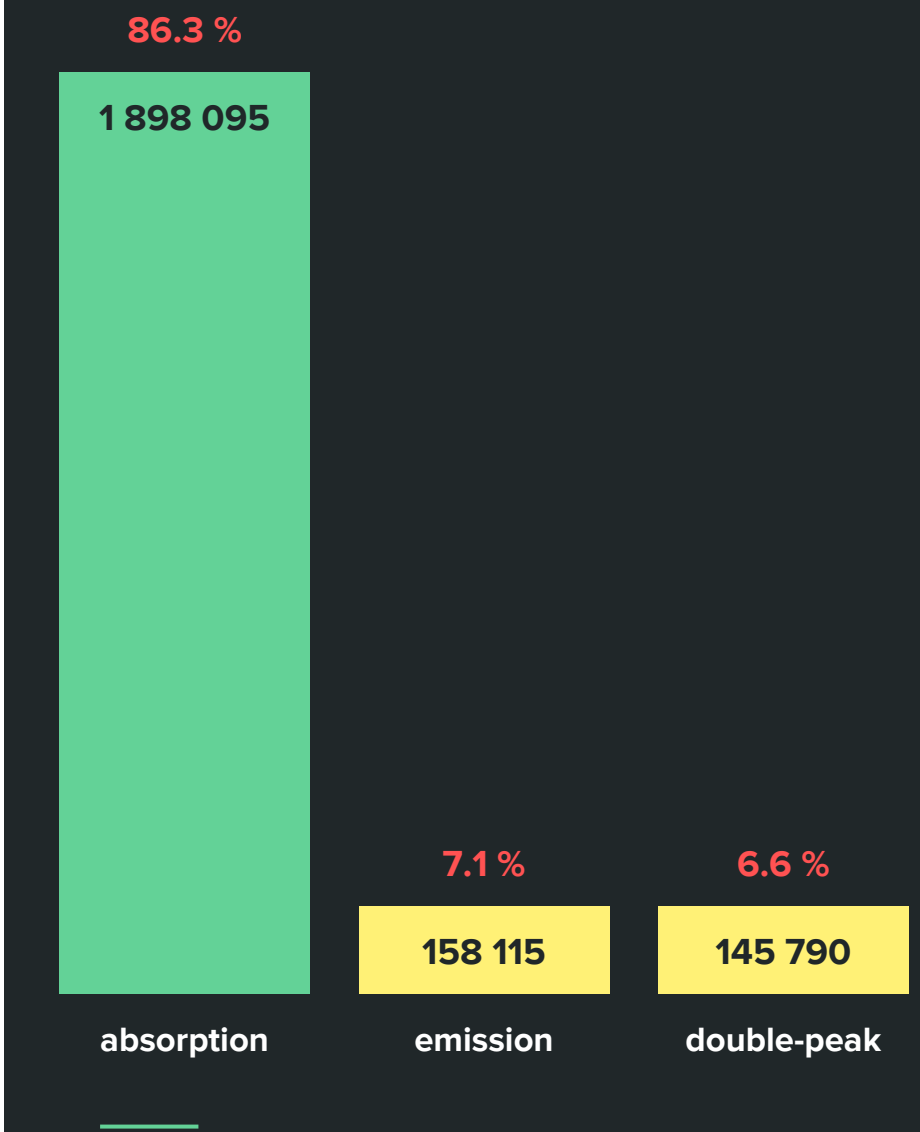
Inspired by **VGGNet** and suited to 1D spectrum.

No feature extraction!



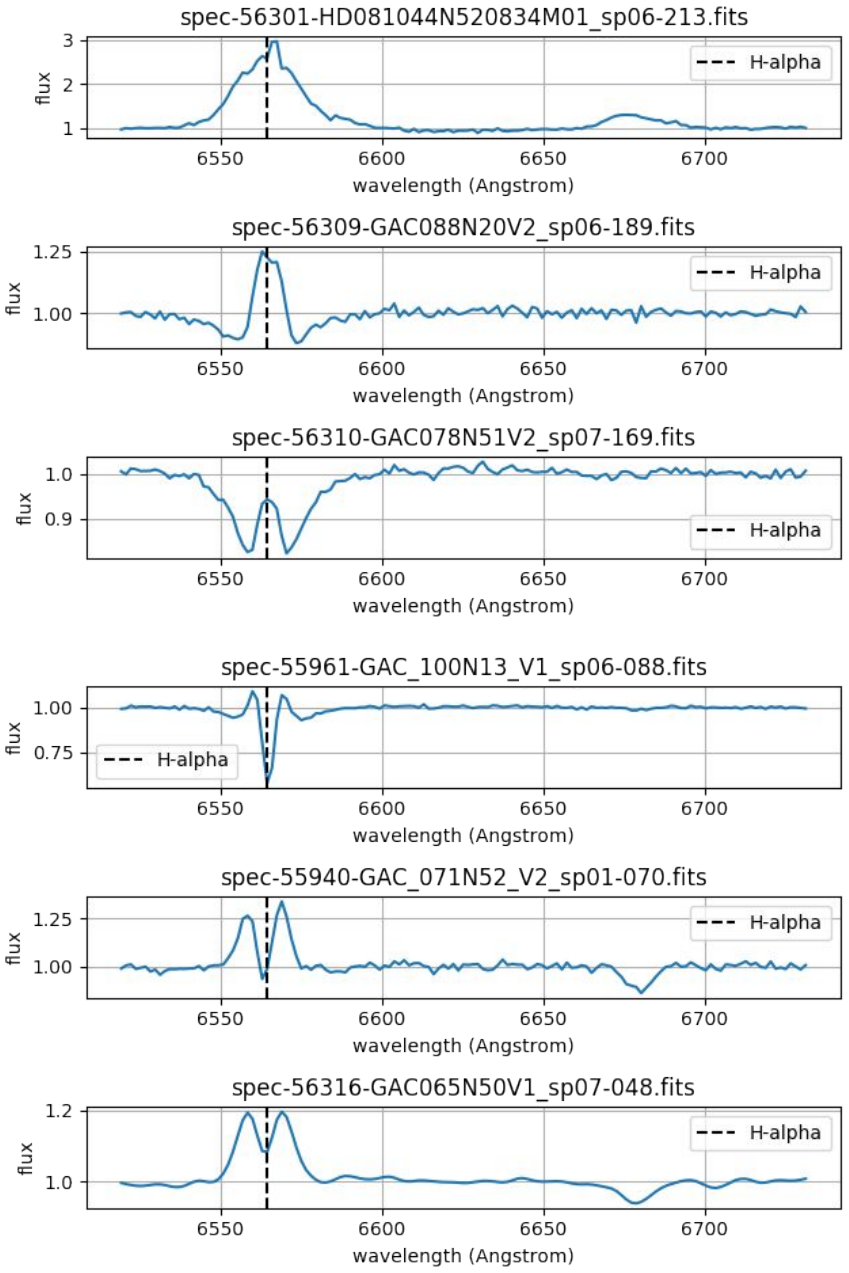
LAMOST data

Classified **2M** of LAMOST spectra **in 4 minutes on GPU**.



Found candidates

No metric to measure performance.



What will I do next?

Improve domain adaptation.

Move it to regression problem.

Thank you.

https://podondra.cz