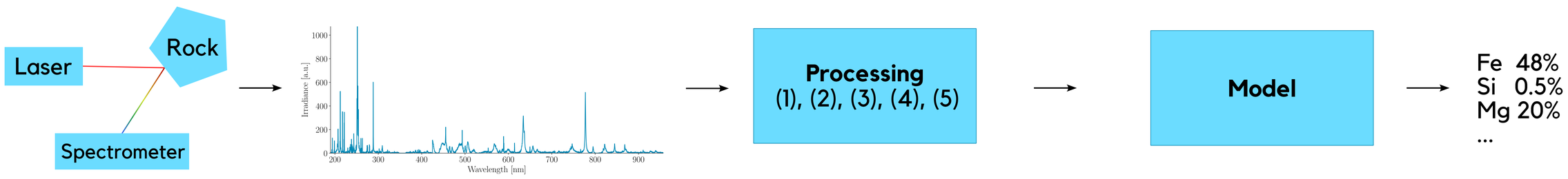
How to feed emission spectra into machine learning models

C. H. Egerland, A. Lomashvili, E. Clave, K. Rammelkamp, S. Schröder, H.-W. Hübers
Deutsches Zentrum für Luft- und Raumfahrt e.V., Institut für Optische Sensorsysteme, Berlin, Germany

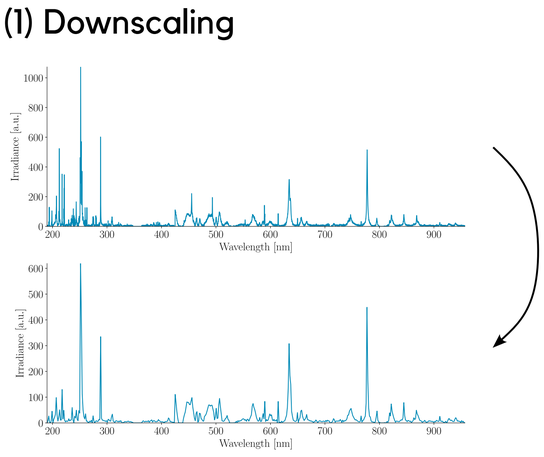
Emission spectra are their **own** kind of **data type**

From experiment to machine learning

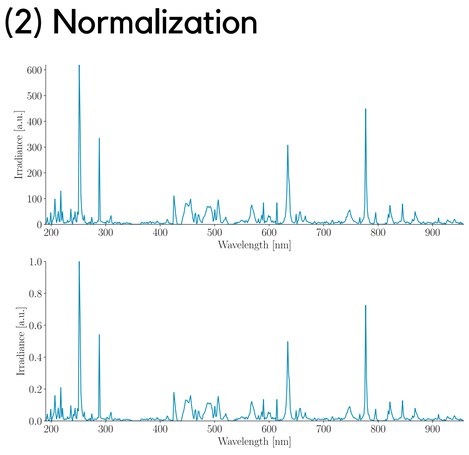


Preprocessing

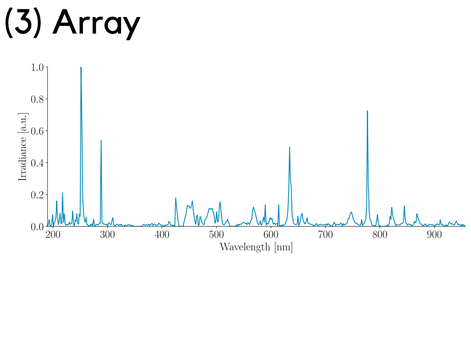
(1) Downscaling



(2) Normalization



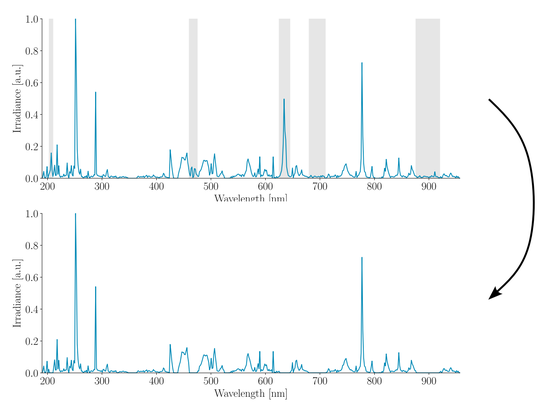
(3) Array



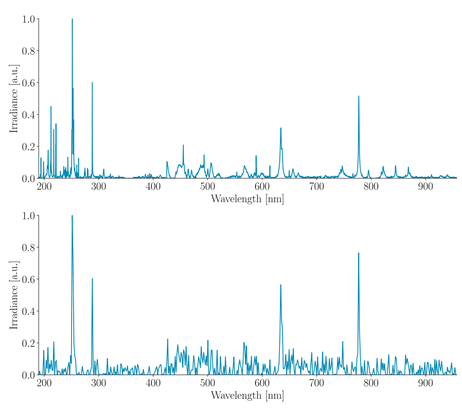
[0.1, 0, 0, 0.195, ... , 0.65, 0]

Data Augmentation

(4) Masking



(5) Noise addition



NO shifting
NO inverting
NO cropping

Machine Learning Models

Machine learning for laser-induced breakdown spectroscopy involves classification as well as quantitative analysis (regression) of targets.
Different methods are used from classical (partial) least squares to neural networks [1].
Labelled spectral data is limited, yet known data augmentation techniques of time series [2] are restricted
because of the inadvertent information loss or label change [3,4].

[1] Boucher et al., "A Study of Machine Learning Regression Methods for Major Elemental Analysis of Rocks Using Laser-Induced Breakdown Spectroscopy", https://doi.org/10.1016/j.sab.2015.02.003
[2] Wen et al., "Time Series Data Augmentation for Deep Learning: A Survey", https://doi.org/10.24963/ijcai.2021/631
[3] Anderson et al., "Post-Landing Major Element Quantification Using SuperCam Laser Induced Breakdown Spectroscopy", https://doi.org/10.1016/j.sab.2021.106347
[4] Zorov et al., "A Review of Normalization Techniques in Analytical Atomic Spectrometry with Laser Sampling: From Single to Multivariate Correction", https://doi.org/10.1016/j.sab.2010.04.009.

Contact me: