

# The Spectra dataset

This dataset contains **stellar spectra** obtained with the 0.9m Coudé Feed telescope at Kitt Peak National Observatory, classified into various spectral types.

The provided dataset here is a simplification that only keeps 1115 homogeneous spectra with half the resolution (3753 “pixels” remains).

The target only provides the 7 regular spectral types for classification to increase the number of examples per class. Still, the dataset remains highly imbalanced!

Infos about the dataset and the reference paper (*F. Valdes et al. 2004*) are in the info file.

*Input dimension : 3753 (+1)*

*Output dimension : 7*

*Number of example : 1115*

*Class distribution : strong imbalance*

