Data Science Lab 2024: Worksheet 6

May 2, 2024

1 Neural Networks

In the lecture we discussed a classical feed forward network and went through an example classifer for exoplanet detection. Using the network description in the lecture train a model to classify Kepler transit light curves.

The accompanying files exoTrain.csv and exoTest.csv contain the preprocessed training and test dataset respectively. Each row is an example (observation of a star). The label column marks whether the star has an exoplanet or not (2 if has an exoplanet and 1 if no exoplanet has been detected).

(a) Load the data and apply any pre-processing step if needed. Plot an example each of the two classes of stars (both for the train and test set).

(b) Create a model.

(c) Train the model.

(d) Evaluate the model on a validation and test dataset.

(e) Can your network classify the stars correctly? Compare the predicted class with the observed class.

(f) Tune the different hyperparameters and note down how the model predictions improve/worsen with the changes. Visualsize the changes (e.g., plot the test loss and accuracy as a function of the hyperparameters).

NOTE: pre-processing steps may include: handling missing values, data cleaning, feature scaling, normalization or standardization of the dataset, to name a few.