Data Science Lab 2024: Worksheet 7

May 10, 2024

1 Convolutional neural network

Exoplanet classification: Replace the fully connected neural network model in Worksheet 6 with a convolutional neural network to classify exoplanet host stars in the Kepler dataset.

(a) Load the data and split them into input and output sets. Plot an example each of the two classes of stars (both for the train and test set).

(b) Create a CNN model.

(c) Train the model.

(d) Evaluate the model on a validation and test dataset. What is the precision and recall of the model?

(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. Use the performance metrics discussed in the lecture to show how your model performance changes with choice of model architecture, model hyperparameters etc.

(g) Apply data augmentation. Does it help with the model performance?