XGBoost - READMe

-----------------------

Description: This folder contains a Google Colab notebook "XGBoost.ipynb", with which several XGBoost related experiments were conducted.

-----------------------

Prerequisites:

- Logged in with a Google account, with access to Google Colab, Google Colab's CPUs and GPUs (subscription needed to have prioritised access GPU's for a significant longer time) (for this notebook, you only need CPUs)

- Granting permission to run each of our Google Colab notebooks and python files.

- Mount drive at /content/drive/ [This is done when running the notebook, see below]

- The folders/files regarding Preprocessed training and validation datasets are at the described location.

-----------------------

How to run:

Before running, the user can specify which preprocessed datasets they want to work with.

In the first section, there is a variable

PREPROCESSING\_CHOICE defeined by PREPROCESSING\_OPTIONS[INDEX], where INDEX is an integer [0;8]. The user can choose INDEX from that range, which corresponds to the preprocessed dataset.

After selecting INDEX:

Run the first section "Init and load data" to setup everything needed for this notebook.

Then run the second section "XGBoost Classifier", is divided into 2 subsections, named after the model, from top to bottom

-----------------------

Output:

In the directory /content/drive/CIL22022, the trained models are stored as file and the performance of the model is logged.