Ensemble of top 3 models - READMe

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

Description: This folder contains various Google Colab notebook to the ensemble of the top 3 chosen models

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

Content:

- Least Squares Ensemble Validation and Test.ipynb: Learns a ridge classifier on the training data (.csv files) that were produced by the specified modeles. Furthermore it generates the prediction on the validation set, as well as the kaggle submission file

- Majority Voting Ensemble Test submission for Kaggle.ipynb: Generates prediction on testset set, the kaggle submission file

- Majority Voting Ensemble Validation.ipynb: Generates prediction on validation set

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

Prerequisites:

- Logged in with a Google account, with access to Google Colab, Google Colab's 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 files containing the intermediary prediction (for each chosen model) are at the described location.

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

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

How to run: Majority Voting Ensemble Test submission for Kaggle.ipynb

Before running, the user can specify

- which 3 files (predictions of the chosen 3 models) to perform a majority vote on (specify the path relative to /content/drive/CIL 2022/

The ensemble collects previously generated .csv files, that contain the prediction of the model.

Run all sections, top to bottom.

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

Output:

The final prediction is in the directory "/content/drive/CIL 2022/data/test data/majority\_voting\_ensemble\_submission.csv"

This is the submission file for Kaggle

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

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

How to run: Majority Voting Ensemble Validation.ipynb and Least Squares Ensemble Validation and Test.ipynb

- Similar as described above; the files produced are NOT the submission file for Kaggle