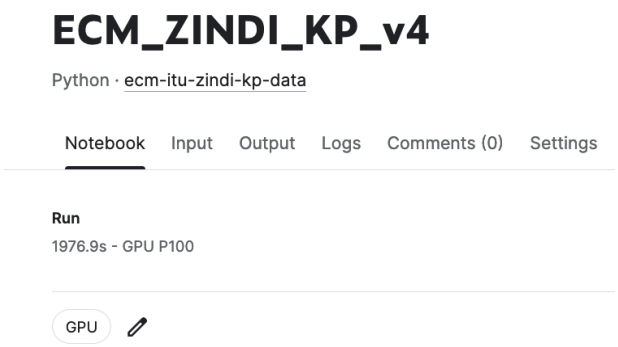
Environment:
Make sure to run all the codes in the **kaggle GPU environment** for reproducibility and RAM out of memory error issue as I have used kaggle to code because of more RAM allotment.

Steps:

1. Open a kaggle notebook
2. Add data (your data will remain private, so no worries)
    1. Data folder is attached in the solution folder (you can directly upload the folder or the zipped version, also attached in the solution folder)
        i. Name: ecm-itu-zindi-kp-data
            1. This folder contains training data downloaded from ITU website
                a. Imgs_202307101549519358
                b. Imgs_2023071012123392536
                c. Imgs_2023071012130978799
                d. imgs_2023071012133740345
            2. Test data downloaded from zindi
                a. SampleSubmission
        ii. If you decide to download and create your own folder, make sure to place both training and test in same folder and change path in the notebook.
3. Go to notebook options and make sure the following values are in following fields:
    1. Accelerator: GPU P100
    2. Language: Python
    3. Persistence: No Persistence
    4. Environment: Pin to original environment (2023-09-07)
    5. Internet: on
    6. The above values are **NOT** all default so please make sure by checking once. Please **TURN ON the GPU.**
4. Import notebook, upload and Run **ecm-zindi-kp-v4-training-and-prediction-notebook.ipynb**
    1. Check GPU is on, save and commit this notebook (top right corner)
    2. The output of this notebook contains models+ final submission.
        i. Models (these are fast ai models for each fold)
            1. Nn_models (for each fold)
        ii. Keras_NN models for each fold
        iii. submission_fastai.csv
        iv. submission_keras.csv
        v. submission_ensemble.csv
        vi. Submission_ensemble_hm.csv
5. Final submission of leaderboard: **submission_keras.csv**
    1. Please use this csv to match the scores.
        i. There should be no to very very minor difference in score as almost everything is seeded.



The notebook should take less than 2500 seconds to run.

The other **shap-analysis-model-explanation.ipynb** notebook is for the organizers to understand the models trained. You **NEEDNOT** run it to match the score on LB. There is a separate notebook for this because shap analysis for ANN models take a lot of time to run. The input to this notebook is again the input data and the output models of above training notebook. If you decide to run this. Please change the paths in the code accordingly.

Finally, thanks a lot for this competition and do let me know if you just need the committed version of all the above notebooks and I can share with the appropriate person.

NOTE: please **DONOT** run this on any other platform or environment as the version of
`Scipy`
`Numpy`
`Pandas`
`fastai`
`sklearn`
`torch`
`tensorflow`
`keras`
`shap`

wont match and the results wont be reproducible. Also there could be RAM OOM errors.

Following are the versions in the current kaggle environment:

```
import shap
import fastinference
from platform import python_version
print("Python Version:",python_version())
print("Numpy Version:",np.__version__)
print("Pandas Version:",pd.__version__)
print("fastai Version:",fastai.__version__)
print("fastinference Version:",fastinference.__version__)
print("scipy Version:",scipy.__version__)
print("sklearn Version:",sklearn.__version__)
print("torch Version:",torch.__version__)
print("tensorflow Version:",tf.__version__)
print("keras Version:",keras.__version__)
print("shap Version:",shap.__version__)
```

```
Python Version: 3.10.12
Numpy Version: 1.23.5
Pandas Version: 2.0.3
fastai Version: 2.7.12
fastinference Version: 0.0.36
scipy Version: 1.7.3
sklearn Version: 1.2.2
torch Version: 2.0.0
tensorflow Version: 2.12.0
keras Version: 2.12.0
shap Version: 0.42.1
```

!pip install numpy==1.23.5
!pip install pandas==2.0.3
!pip install fastai==2.7.12
!pip install plotly==5.17.0
!pip install plotnine==0.12.3
!pip install fastinference==0.0.36
!pip install scipy==1.7.3
!pip install sklearn==1.2.2
!pip install torch==2.0.0
!pip install tensorflow==2.12.0
!pip install shap==0.42.1

THANKS
Farzi_Data_Scientists
Krishna Priya
Rajat Ranjan