Hi Zindi

Apologies for the delay in submitting our solution.

The datasets are quite huge, so the models took a while to train.

As features from metadata weren't allowed, we solely focused on generating features from pixels only.

Our solution is a blend of boosting algos and neural networks. Boosting trees were trained on features engineered from vegetation indices and the neural network was trained on raw pixel data.

Blending predictions from these two diverse approaches generalized well and gave us our best score.

Follow these steps to reproduce the solution:

1. Upload the Feature\_Engineering\_&\_CATBOOST.ipynb notebook to colab.

- Enable GPU runtime

- Run all to get the catboost\_models file

2. Upload the Feature\_Engineering\_&\_LGBM.ipynb notebook to colab.

- Enable TPU runtime

- Run all to get the lgbm\_models file

3. Upload the Pixel\_Features-Pytorch.ipynb notebook to colab

- Enable GPU runtime

- Run all to get the pytorch\_models file

4. Finally upload the Ensemble.ipynb notebook to colab

- Upload the lgbm\_models file

- Upload the catboost\_models file

- Upload the pytorch\_models file

- Run all to get the final submission file

Use the below notebooks for data download and manipulation:

- Data Download

- Numpy\_Extraction\_for\_Month\_Start\_Month\_End

- Numpy Extraction\_for\_25\_Periods

- Field\_Aggregation\_Mean

For any queries or clarifications, don't hesitate to contact us.

Thank you for growing DataScience in Africa.

Kind Regards,

Darius.