For each image representation, a different notebook has been used which makes the tfrecords. Tfrecords are compressed files, which are convenient when used as training data set in deep learning models. Before the exported images from google earth engine, the tiff files, are written to tfrecords, some postprocessing is done.

Firstly

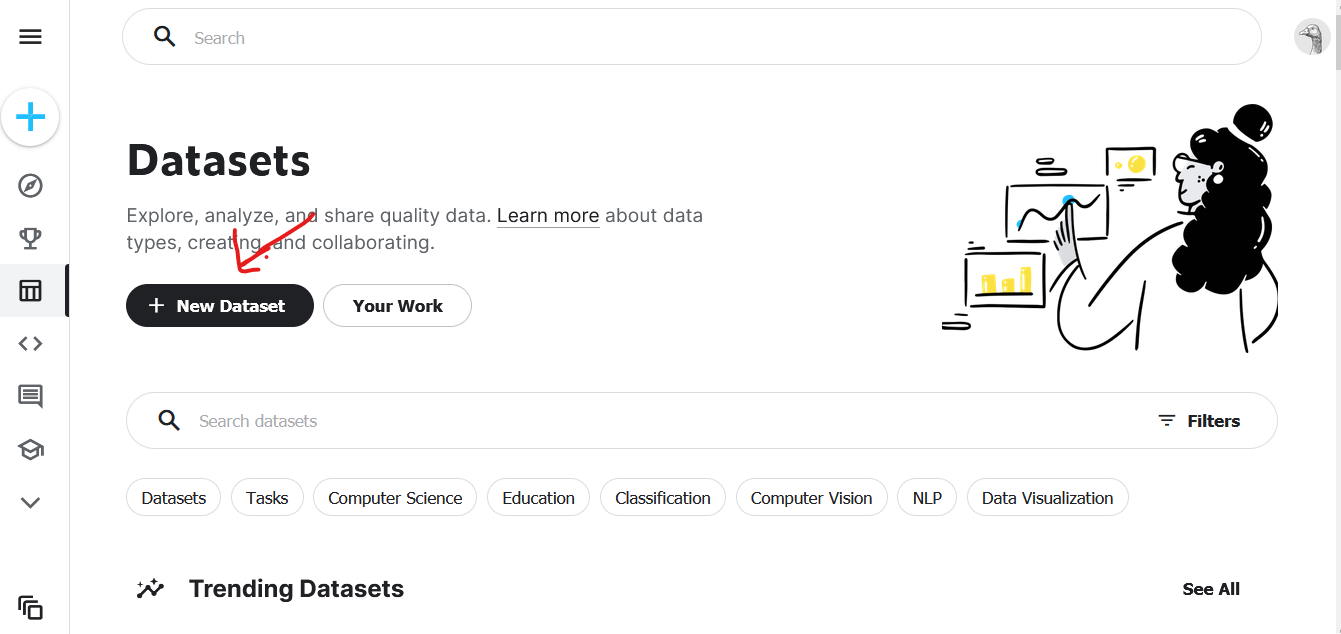
* the images are divided into training,testing, validation dataset
* the minimum and maximum values of the training dataset are calculated
* the pixel values smaller than or equal to -30 decibel are filtered out, they are replaced by -29 decibel
* the images’ pixel values are scaled to the range of [-1,1]\* wrt to the minimum and maximum values of the training dataset

\*This scaling is necessary since the images within the deep learning models must be within this scale. For the cGAN this is [-1,1], for the U-Net [0,1]. In order to ensure that the testing and validation datasets scaled values do not overstep this range, since they are scaled based on the minimum and maximum of the training dataset, sometimes it is chosen to scale to a slightly smaller value like e.g. 0.8 instead of 1.

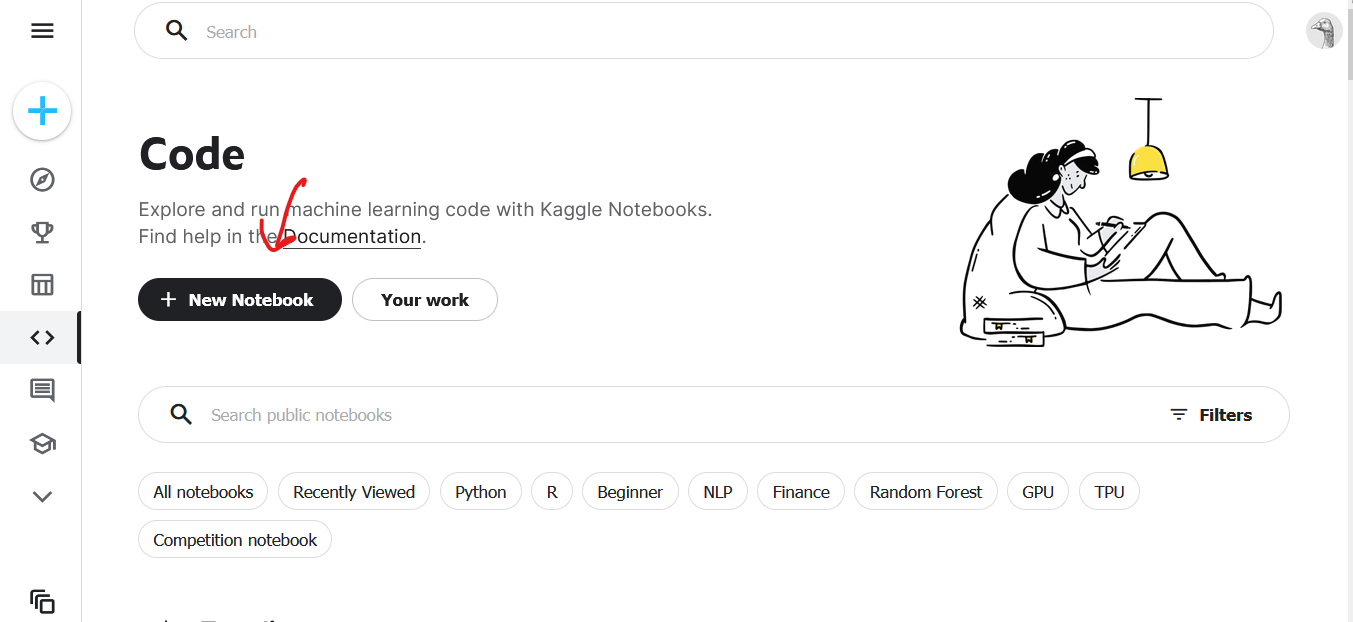
If these tfrecords are uploaded to Kaggle as a dataset, and then linked back to the ipynb Colab Notebook then the TPU can be used for free. This works as follows:

-Make an account on Kaggle

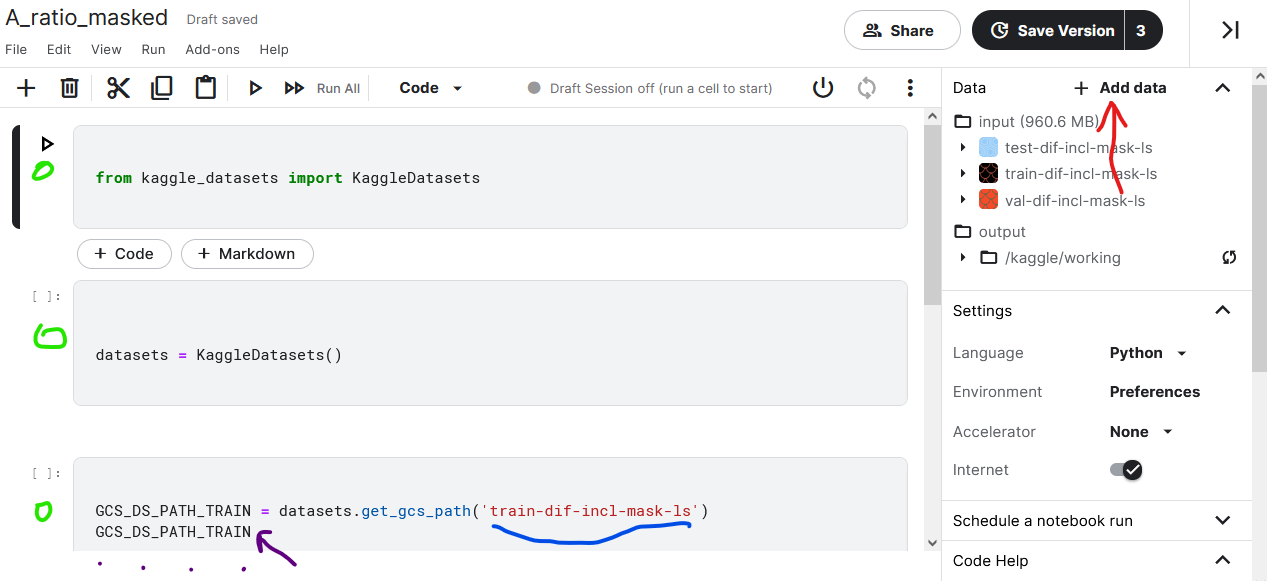
-Create a new dataset on Kaggle, upload your downloaded tfrecords from your local files



-Create a new code on Kaggle



-In this code window/viewer, click on upload datasets in the upper right corner. Here you can upload your Kaggle Datasets. Upload the Kaggle dataset which contains your tfrecords. This is indicated by the red arrow in the figure below



* Then add some code, as shown in the figure above. The input to datasets.get\_gcs\_path(‘input’) is the name of your kaggle dataset. This only works when this same dataset is added to your code, as shown in previous step
* If you run the last part of code, on the purple points there will appear a link. This link contains your tfrecords in a cloudfile in Kaggle. This link can thus be used in, e.g. Google Colab, to access your tfrecords while in the cloud in another code.
* Run the following line in Google Colab, your files will then appear.
* How to get a training dataset from these files is explained in the ipynb’s concerning the DL models (cGAN and Unet)