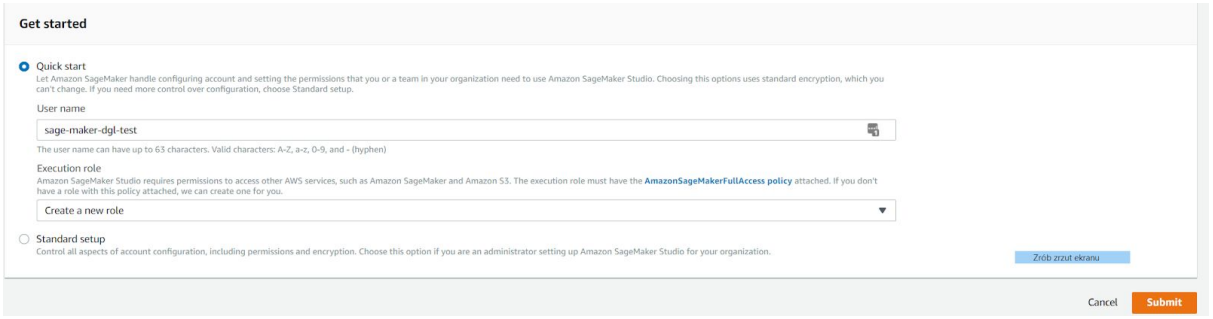
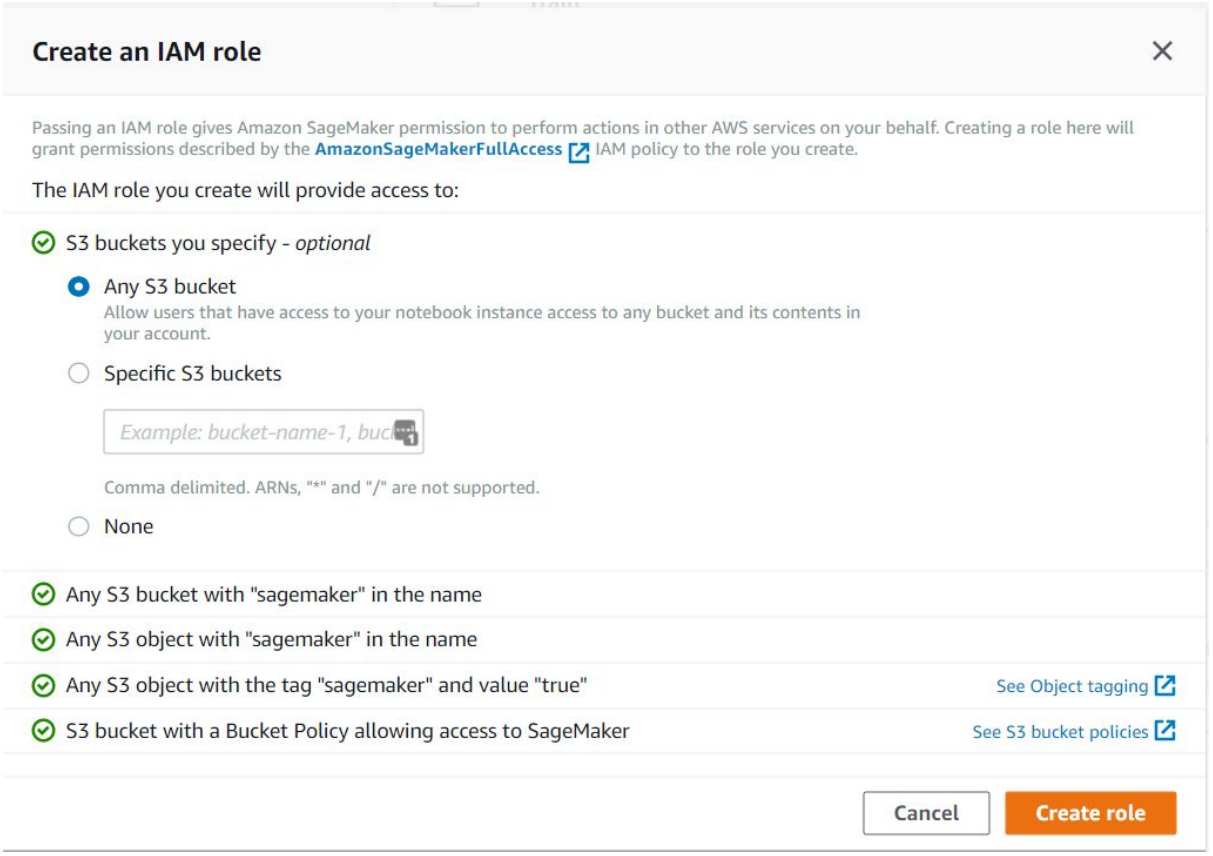
SageMaker with DGL integration

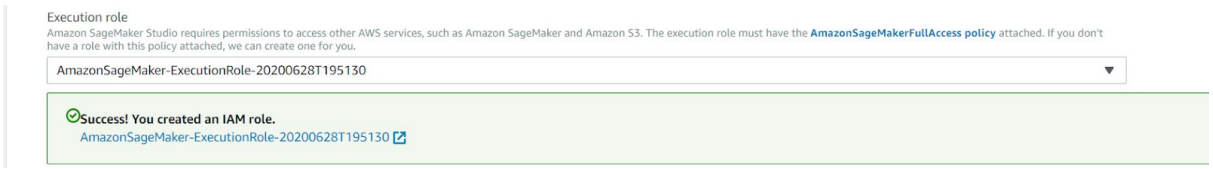
Example 1 - Semi-supervised nodes classification

At first, new SageMaker service account needs to be created. SageMaker must be searched within AWS services.

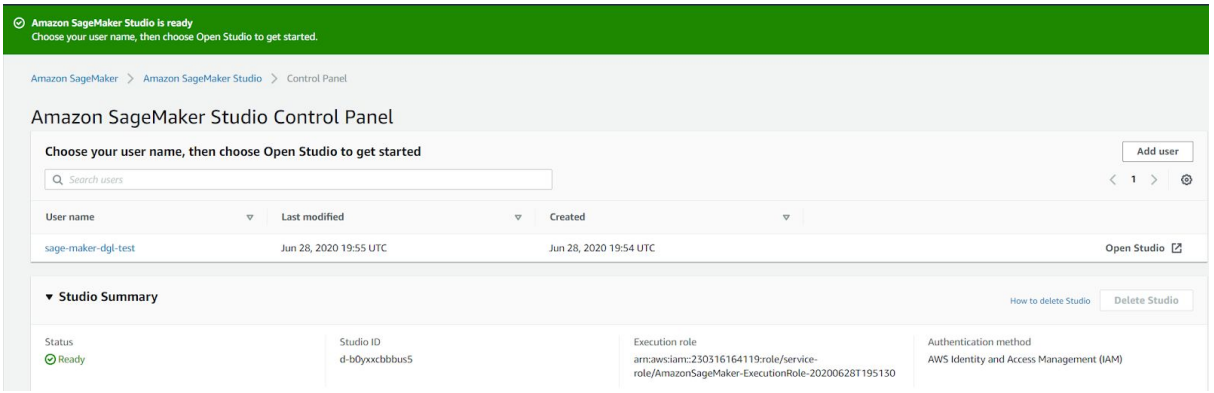


To make use of majority of SageMaker features, S3 bucket is needed. The UI creator helps to create one on demand:

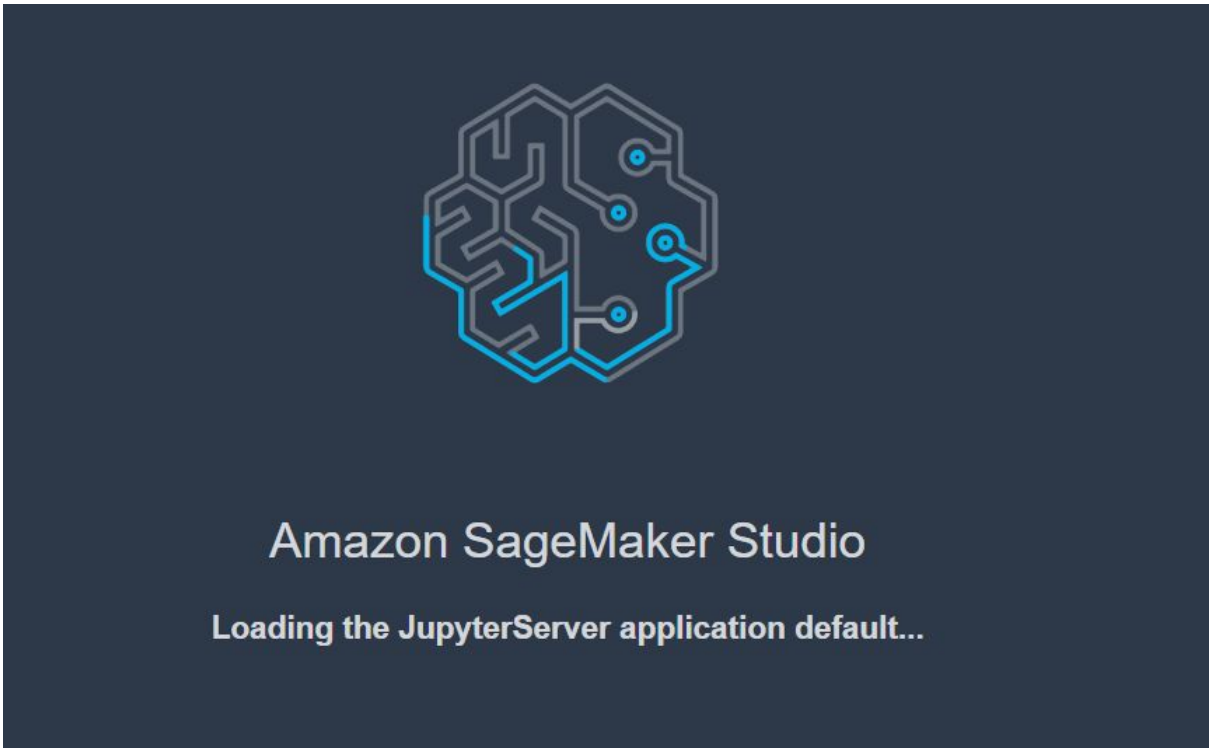


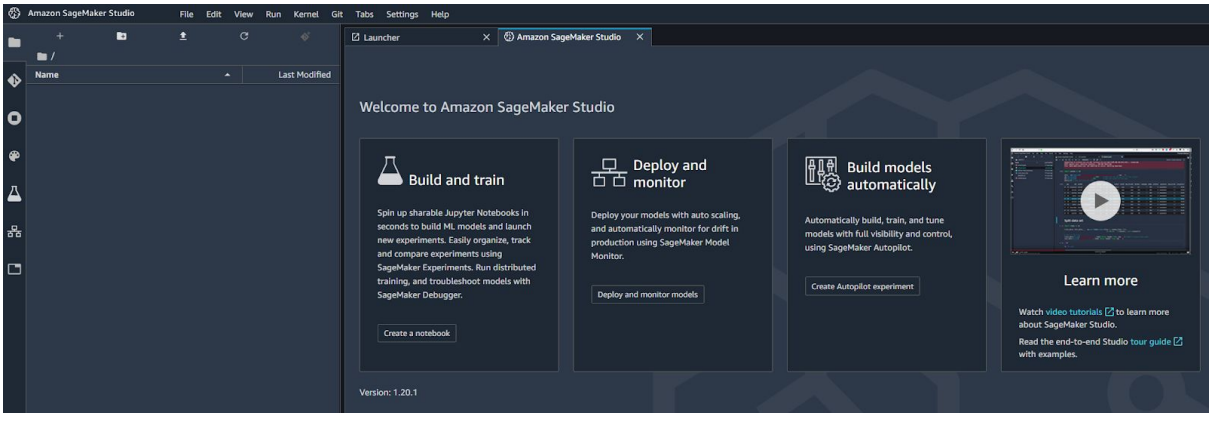


After initialization is completed - one can access SageMaker Studio Control Panel

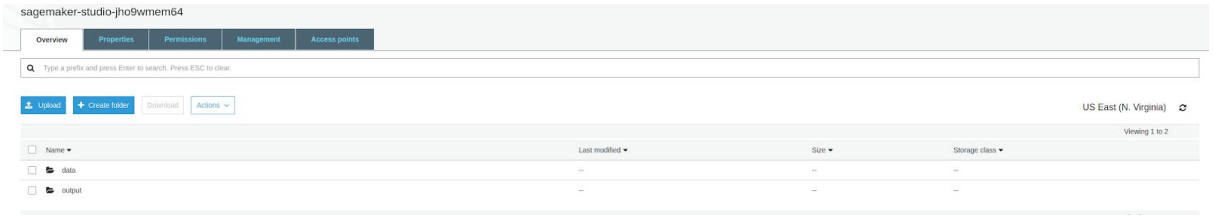


With use of **Open Studio** button one can initialize SageMaker Studio which is kind of a IDE based on Jupyter to train and deploy ML models.

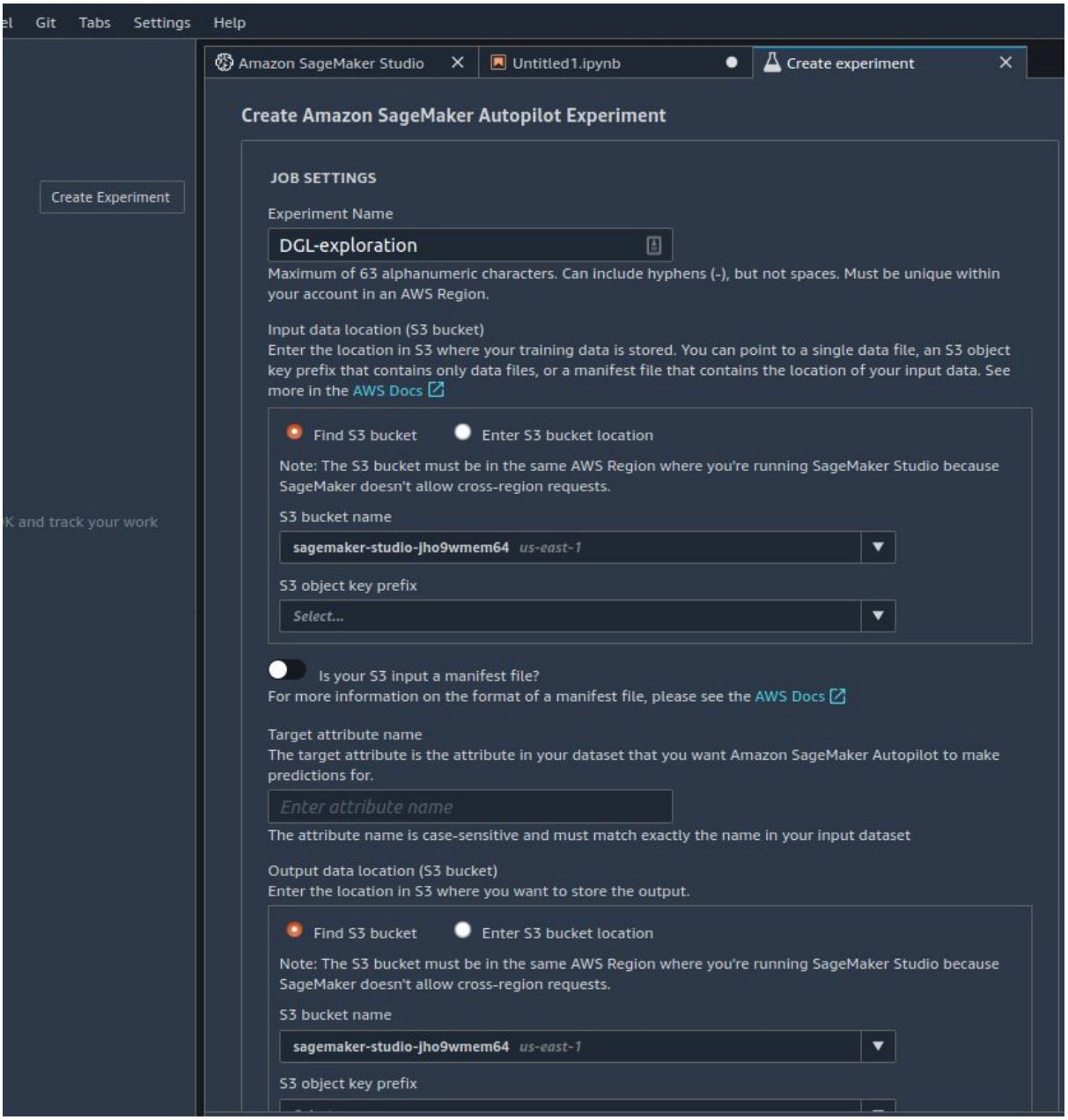




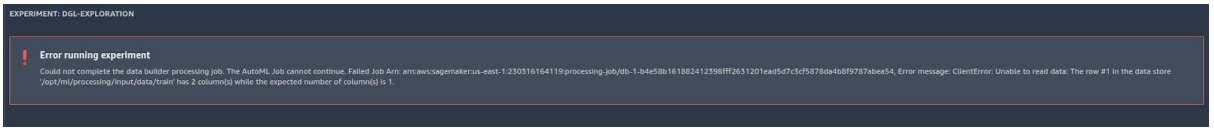
To create an experiment (**Create Experiment** button) one needs some structure on S3 bucket (to store input and output data in controllable manner). Here the data *directory* contains training dataset.



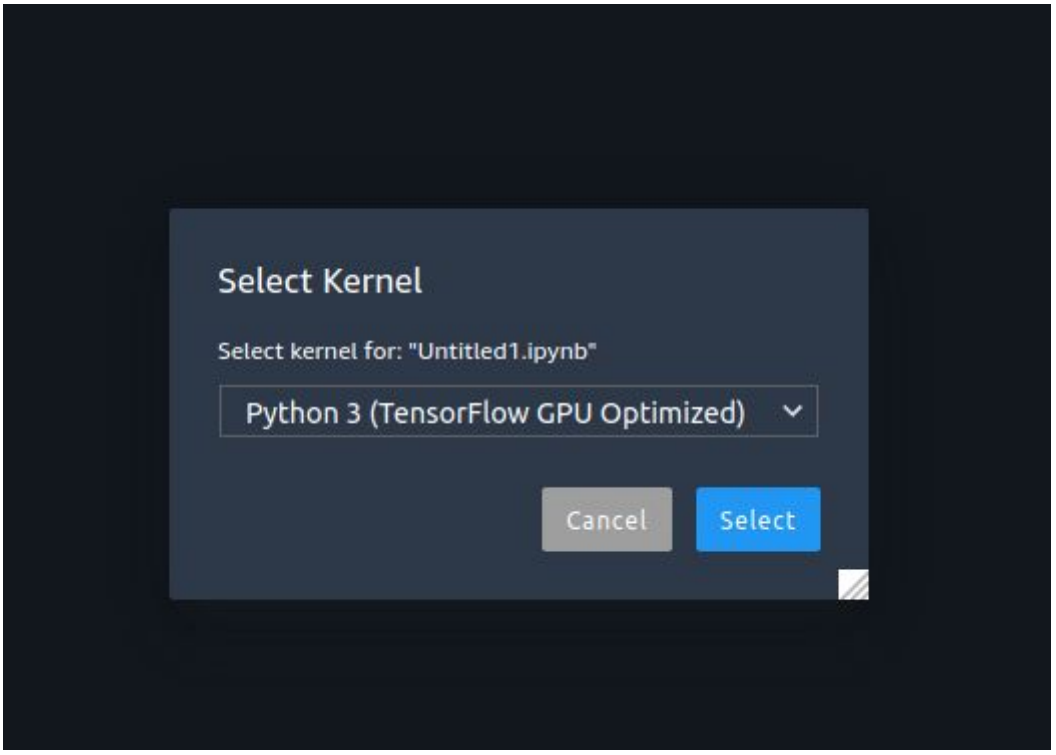
One need to fulfill the form with proper values regarding data structure.



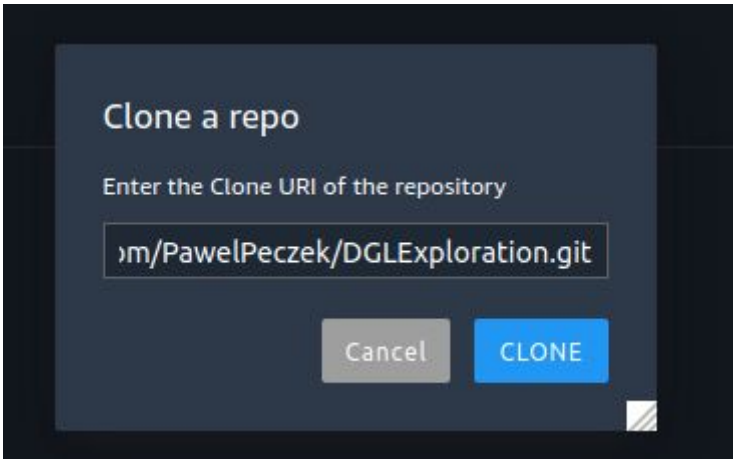
In general case (when handling unstructured data - like in this case), automatic machinery will not be able to help with data indexing.



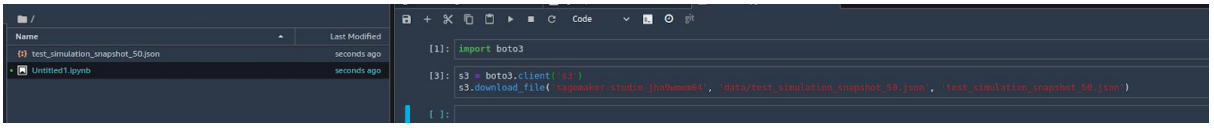
One can trigger new notebook (and choose runtime env).



To have an access to a code one need to clone project repository



As this is done, data must be downloaded from S3:



Then, the code may be executed in the exact form prepared in repository (via Jupyter Notebooks).

When the training is finished - model can be persisted on S3 bucket.