Guide to use Azure Machine Learning

Version 1.0

Evelyn Tan

Last updated 23 April 2021

Overview

In this guide, you would need to have an Azure account created in order for you to follow through the steps below. If you do not have an Azure account created, please create one before you continue. After this guide, you would be able to understand the essential components in Azure Machine Learning and run the notebooks:

- ***predict-link-xgboost-part1-github.ipynb***
- ***predict-link-xgboost-part2-github.ipynb***
- ***predict-link-automl-github.ipynb***

Content Page

Overview **1**

Brief Overview of Azure Machine Learning **2**

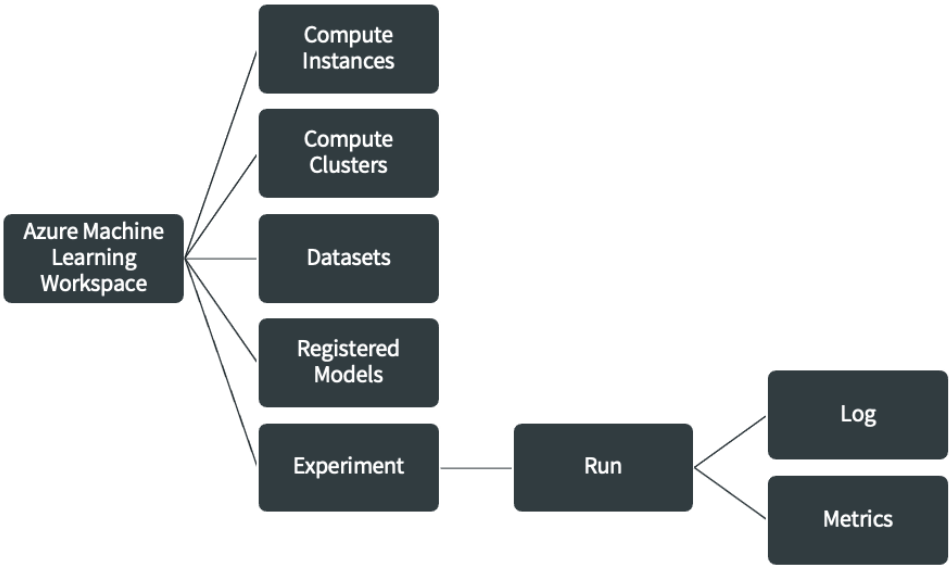
Step 1: Download config.json of Azure Machine Learning **3**

Step 2: Create Compute Instance **3**

Step 3: Upload notebooks **6**

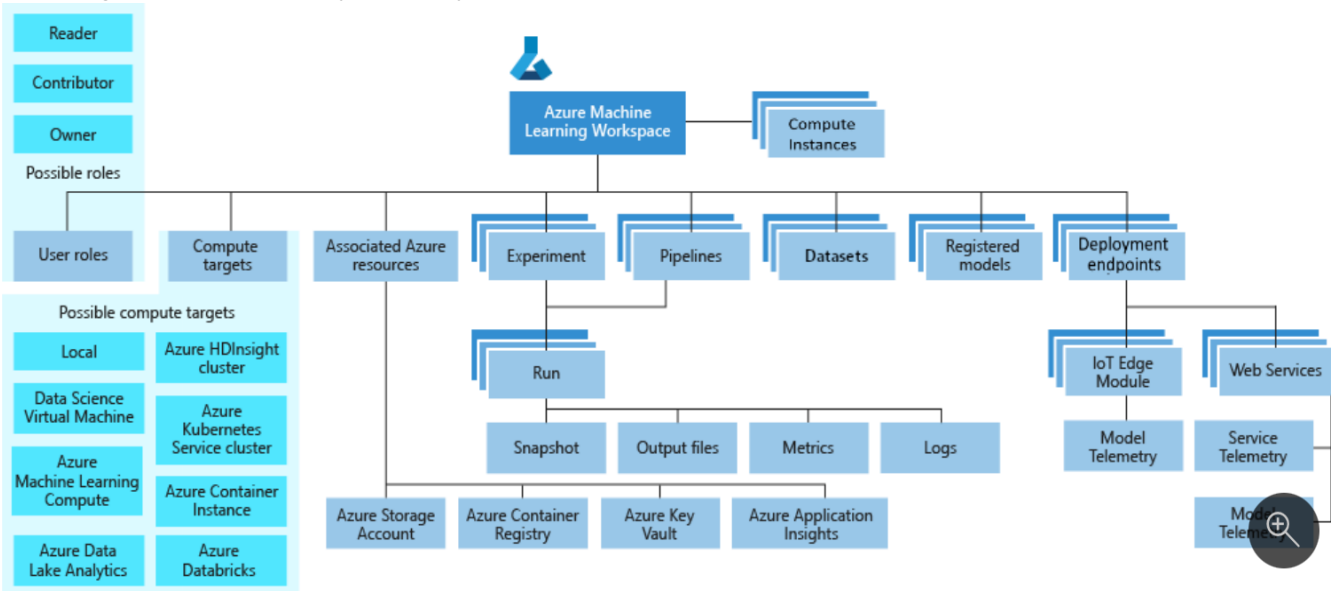
Brief Overview of Azure Machine Learning

To start off, Azure Machine Learning Workspace is a centralized place to work with all the artifacts you create when you use Azure Machine Learning Service. A small overview of the workspace of key components is illustrated in the following diagram.



A workspace can contain compute instances which is virtual machines loaded with essential libraries for performing data science and machine learning algorithms. A compute cluster can be created to increase the number of nodes running in the virtual machines to have the ability of using parallel computations. Datasets are required as it is essential for model training. Once you have a model trained and you want the model to be deployed, the model should be registered in the workspace. Experiments are required to be created as it will track the training runs, we use to build our models.

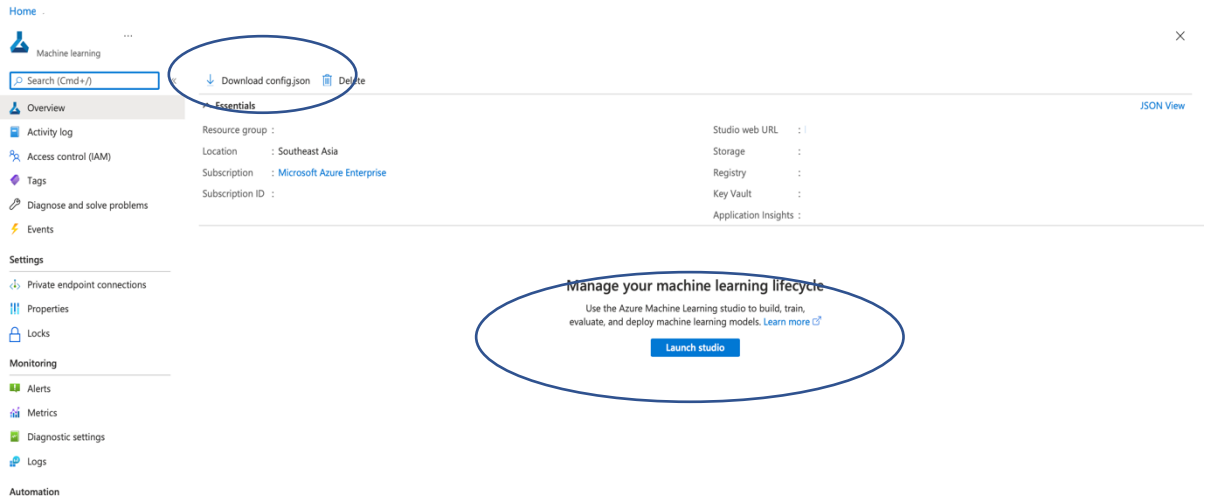
This is the full view of the architecture of azure machine learning in azure documentation[1]. The diagram above displays the key components that is required in the workspace.



[1] https://docs.microsoft.com/en-us/azure/machine-learning/concept-workspace

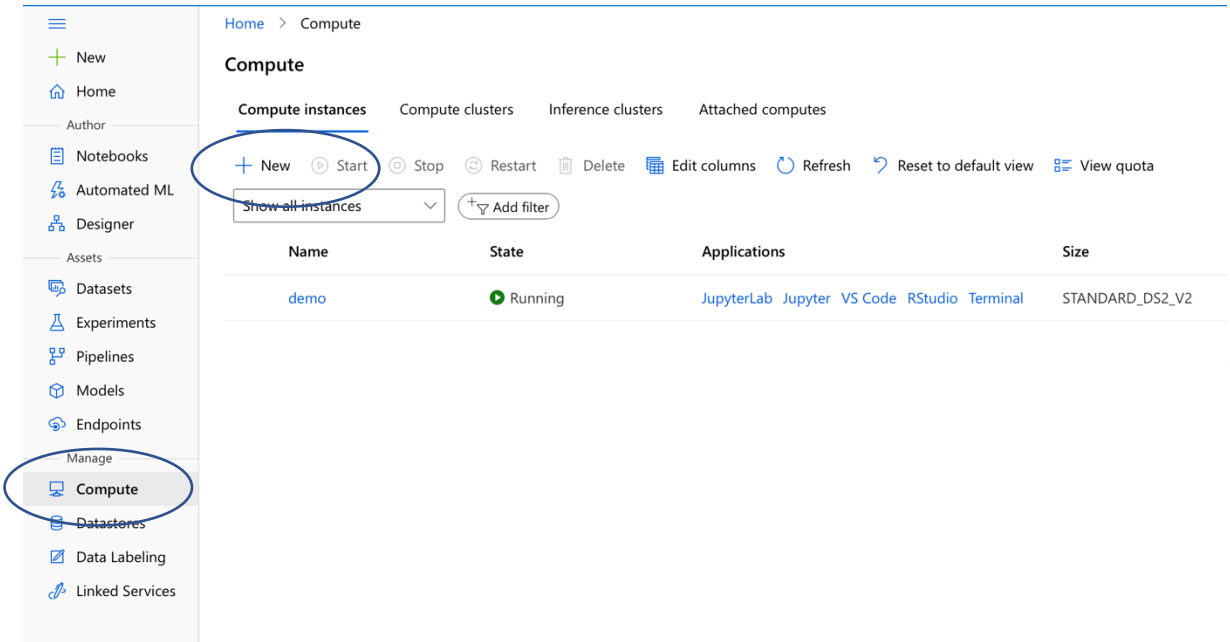
Step 1: Download config.json of Azure Machine Learning

Once your Azure Machine Learning Service has been created, click on the name and you be prompted this page. Click on *Download config.json* to download the credentials of Azure Machine Learning. You would need this file later on when creating notebooks in Azure Machine Learning. Next, click on *Launch Studio* to open Azure Machine Learning Studio. This is the place where you will run model training.

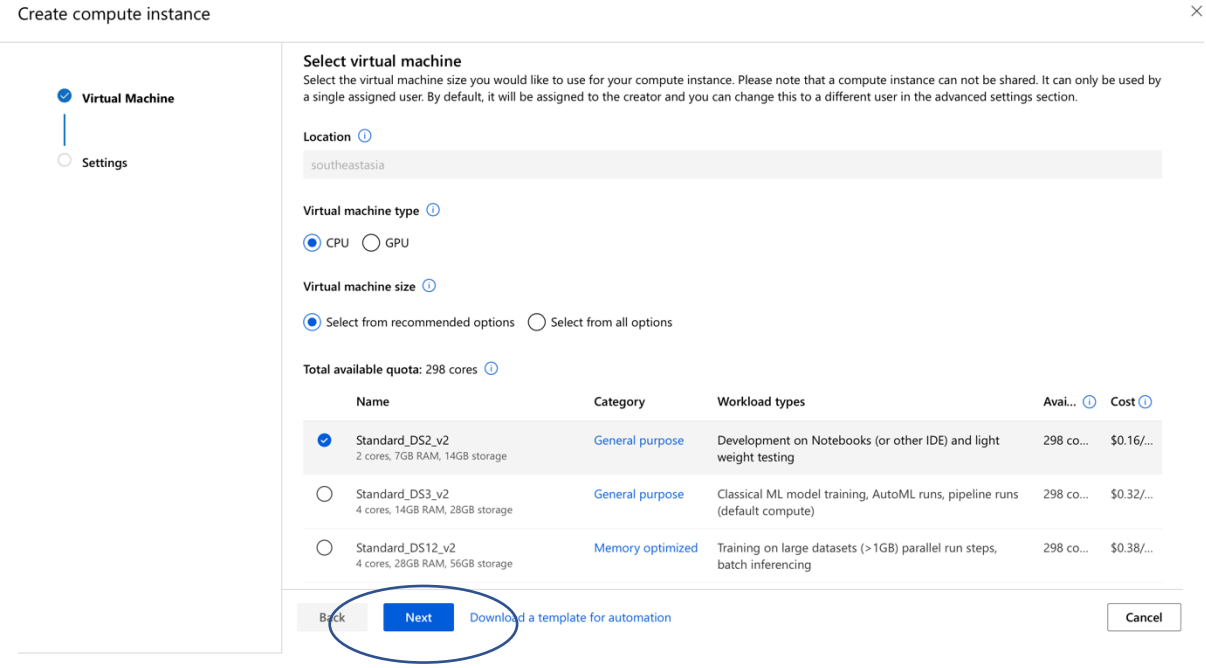


Step 2: Create Compute Instance

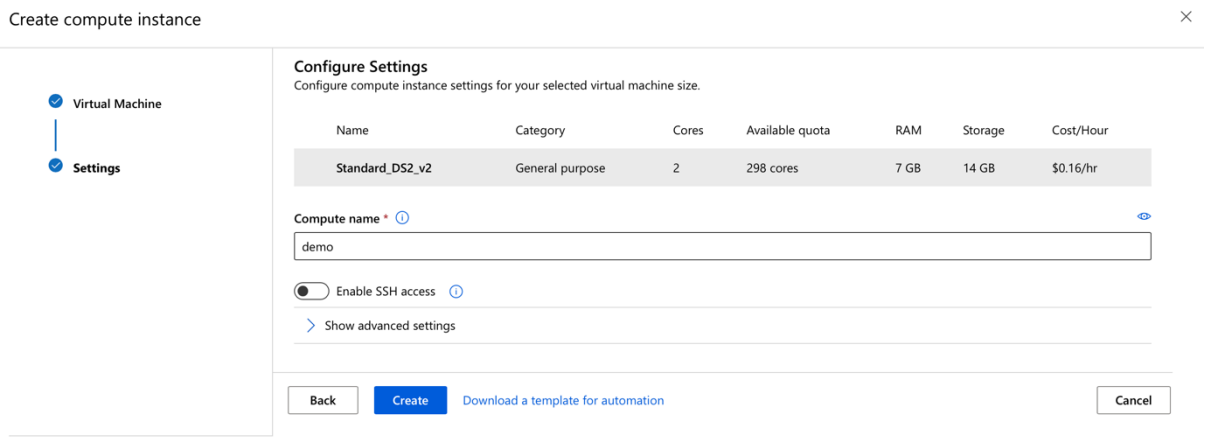
To start creating notebooks, you would need *compute instance* created and running. Click on *compute* on the left sidebar and *new* to create the compute.



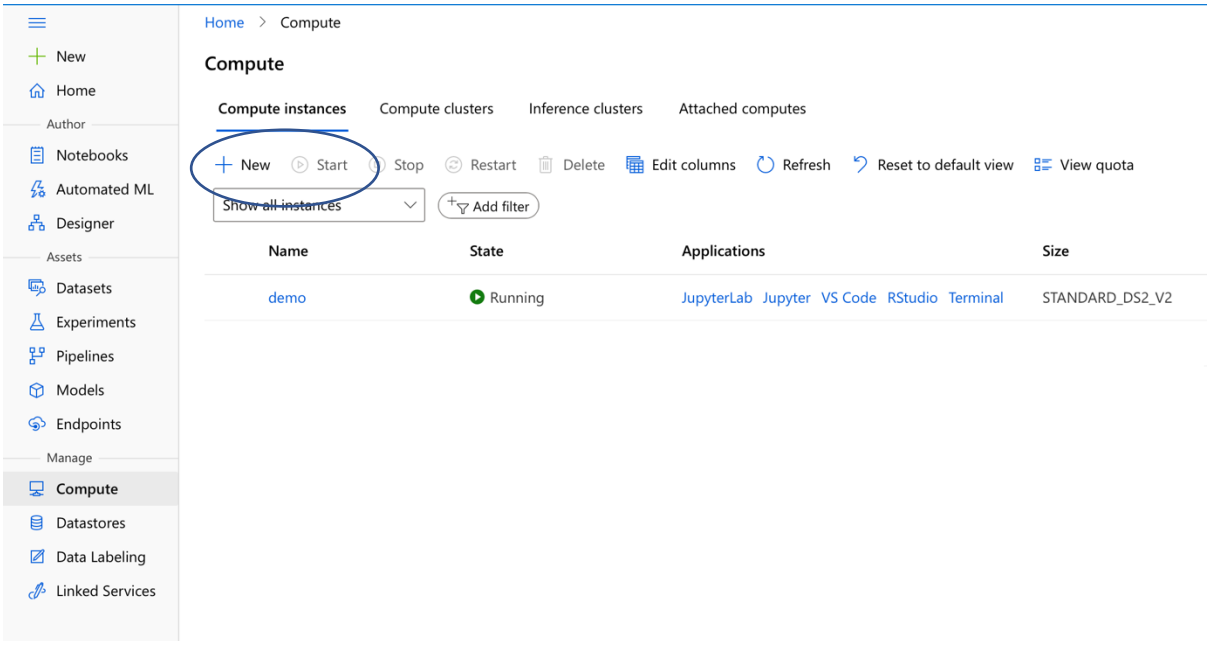
After clicking *new*, you will be prompted to this page. I selected the configurations shown below. If you would like your code to run faster, you can choose the configuration to have more cores, more rams and more storage. Do note that once your configuration has been set, you will not be able to change them. In order to make changes, you will need to create a new compute instance. Click *next* when you have selected your settings.



After clicking *next*, enter your compute name and you will be able to create your compute instance.



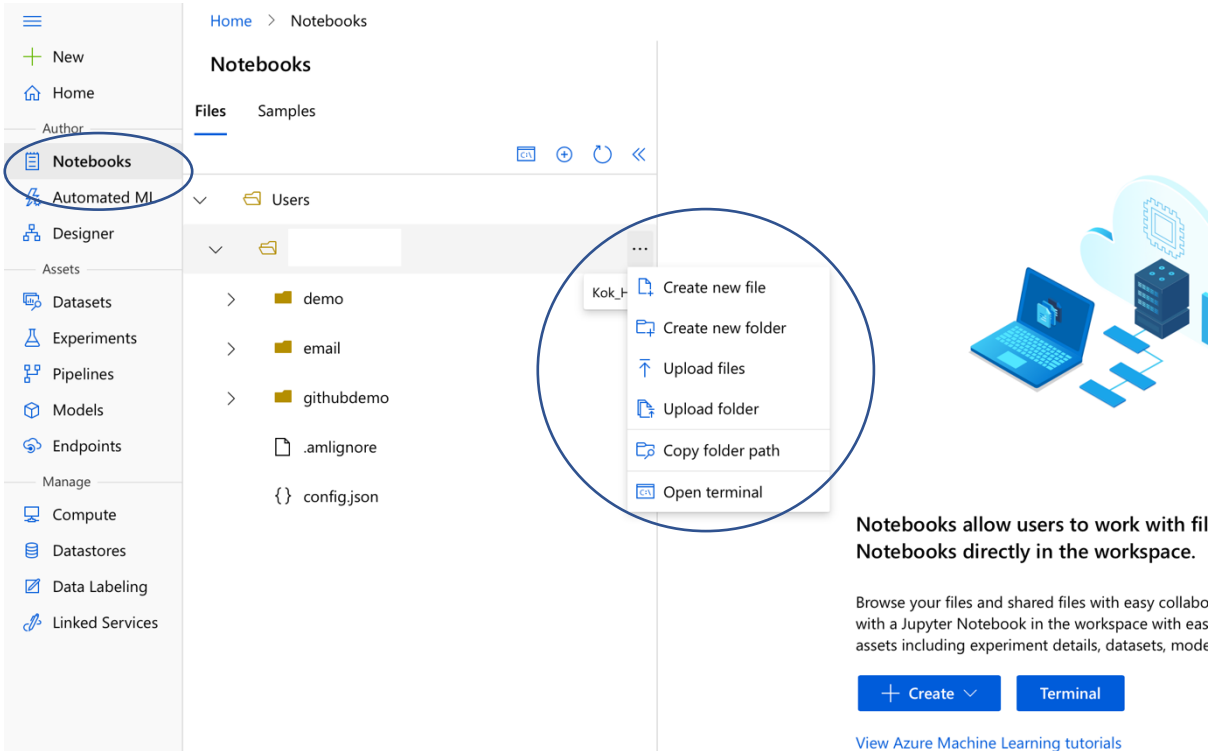
Click *Start* icon to start the compute instance. If you are able to see the state having running shown below, your compute has been created successfully and you are able to start running code in it. You can stop the compute once you have finish running in Azure Machine Learning to reduce costs.



Step 3: Upload notebooks

To start trying Azure Machine Learning, you can upload existing notebooks into the notebooks section. If you are following the tutorial, you can clone the folder in github and upload it into Azure Machine Learning Notebooks.

- amlignore is created by Azure Machine Learning when a file has been uploaded
- config.json is the file downloaded in Step 1.



If you are following the tutorial, you should have at least

- api_config.json
- azure_search_client.py
- config.json

files uploaded. You can also upload the notebooks into the folder g*ithubdemo. config.json* would be the credentials of azure machine learning that we downloaded earlier on in **step 1**. *config.json* can either be inside the folder *githubdemo* or outside the folder. *api_config.json* would be the credentials of Azure Search Service that we created in **Guide for Azure Search Service.pdf**. You would need this *api_config.json* to call the azure search service in *predict-link-xgboost-part1-github.ipynb*.

Click on editors and select *edit in Jupyter* to start coding in python notebooks.