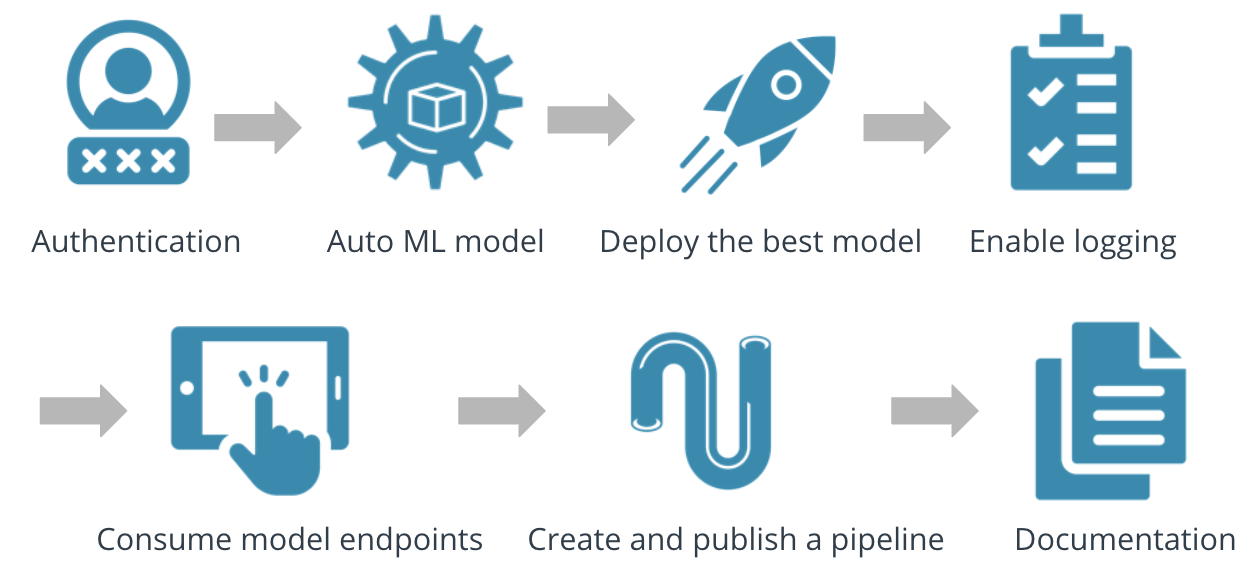
# Operationalizing Machine Learning

This is Udacity Azure Machine Learning Nanodegree project 2, I've worked with the Bank Marketing dataset which contains 21 features and the goal is to predict if a person will take a loan or not. The target column is y.

In this project using Microsoft Azure we will configure a cloud-based Machine Learning model, deploy it, and consume it. We will also create, publish, and consume a pipeline using Azure Python SDK.



**Project main steps:**

In this project, we will be following the below steps:

Authentication:

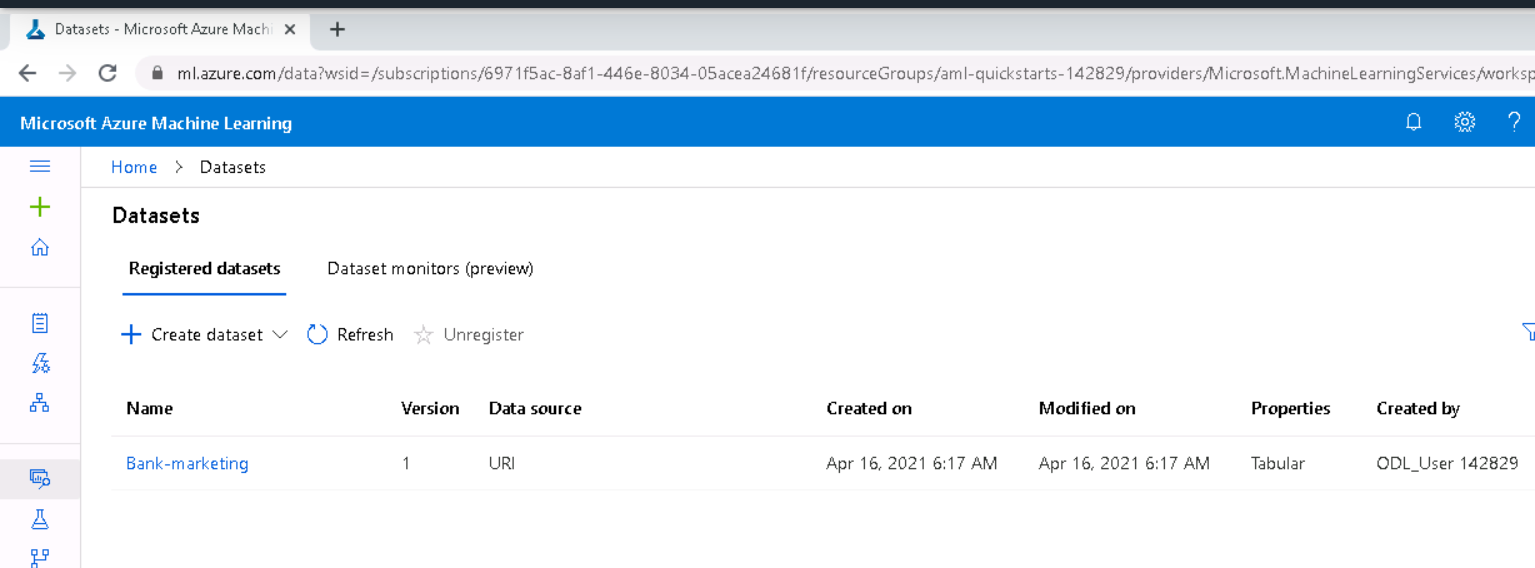
In this step, we need to create a Security Principal (SP) to interact with the Azure Workspace.Here I’m using the lab Udacity provided to me, so I have skipped this step as I’m are not authorized to create a security principal.

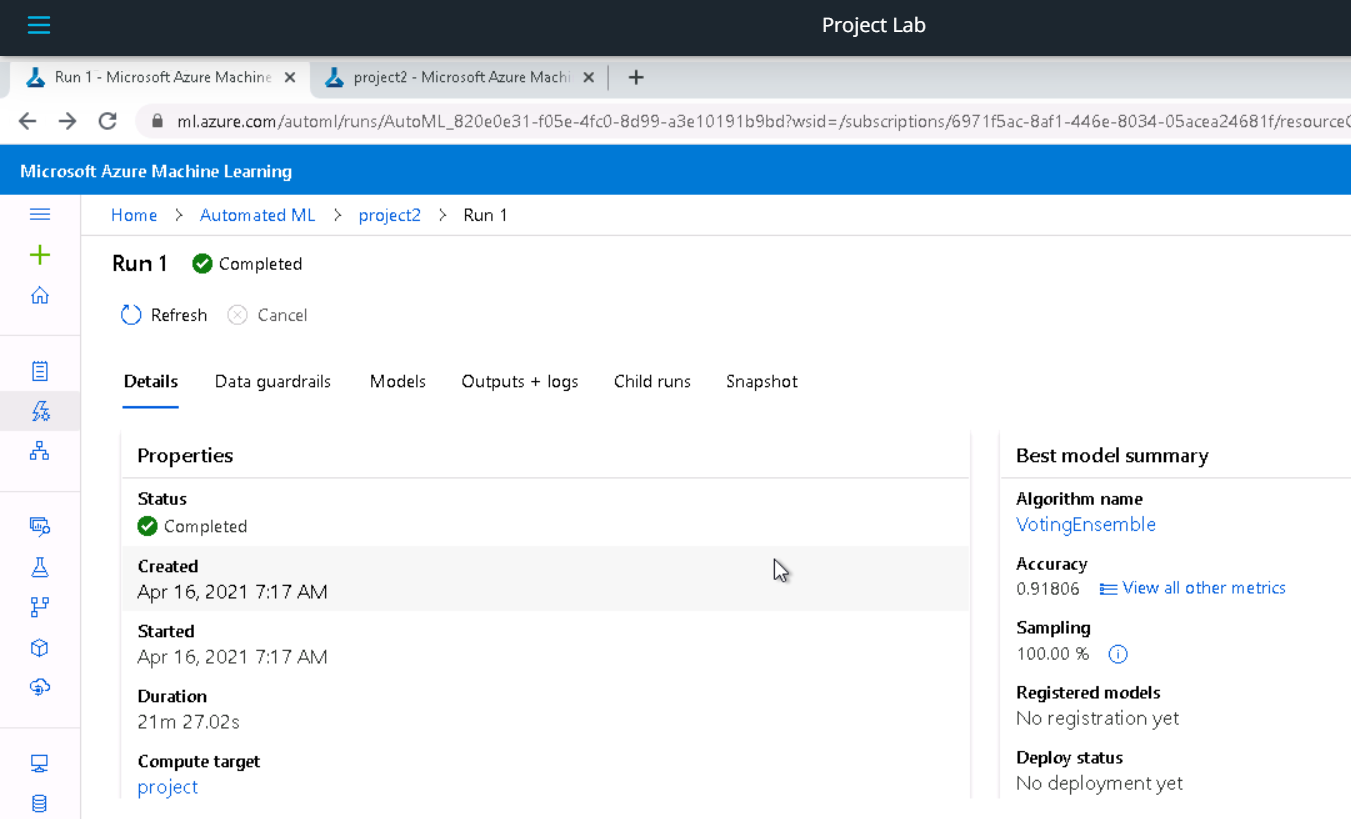
Automated ML Experiment

In this step, we create an experiment using Automated ML, configure a compute cluster, and use that cluster to run the experiment.

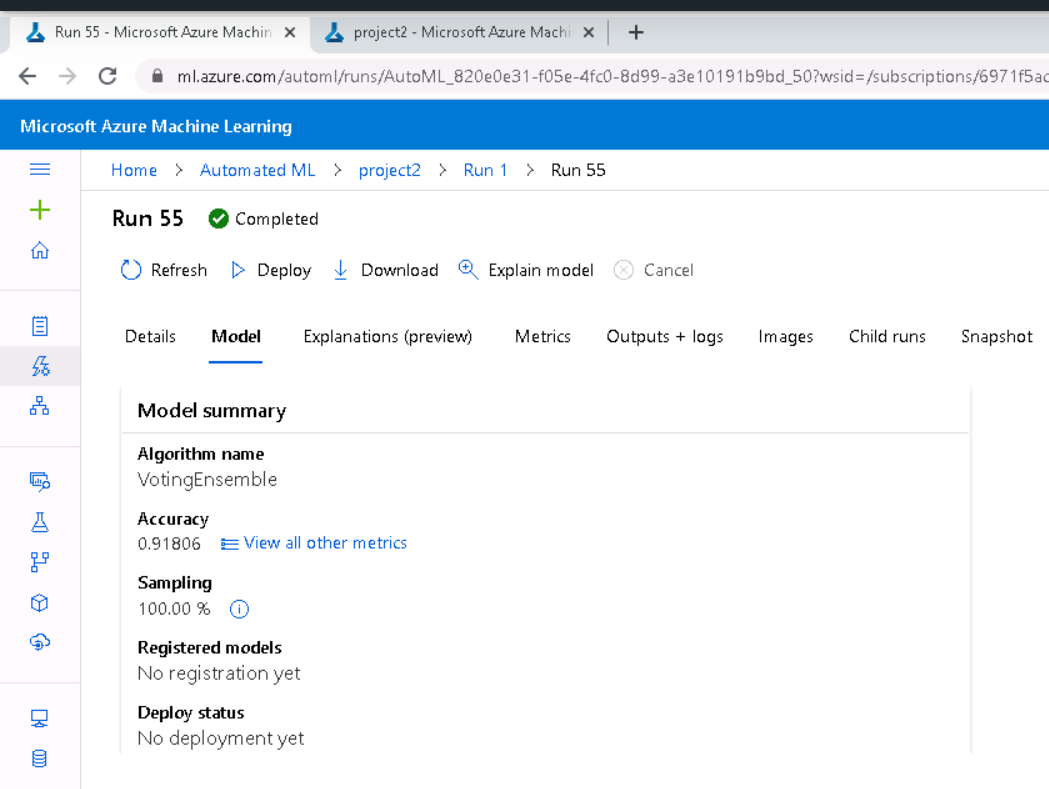
First, we need to use the Bank Marketing dataset (bankmarketing\_train.csv) that is provided to Azure Machine Learning Studio.

Link:   
<https://automlsamplenotebookdata.blob.core.windows.net/automl-sample-notebook-data/bankmarketing_train.csv>



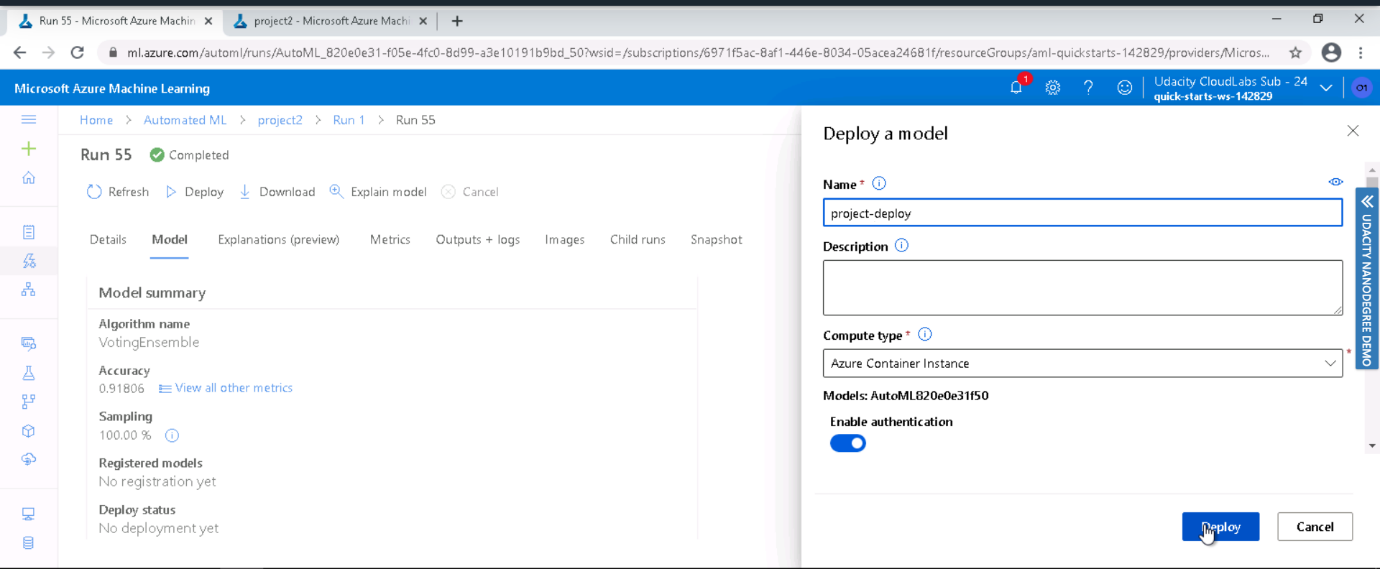


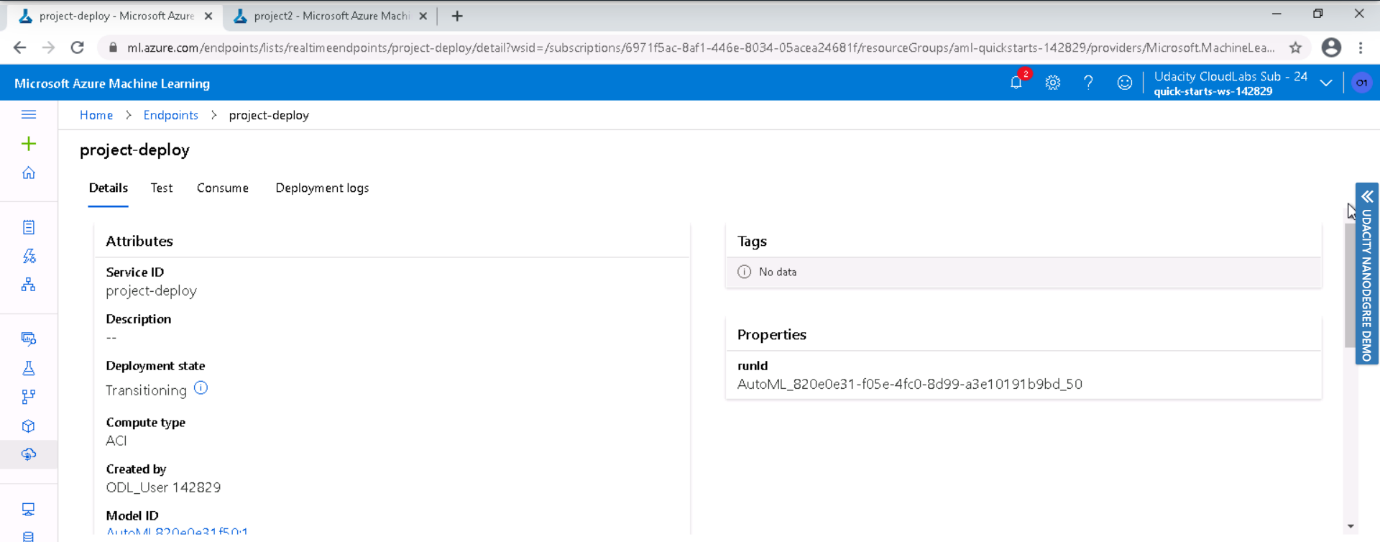
The best algorithm found is VotingEnsemble with an accuracy of 91.8%.

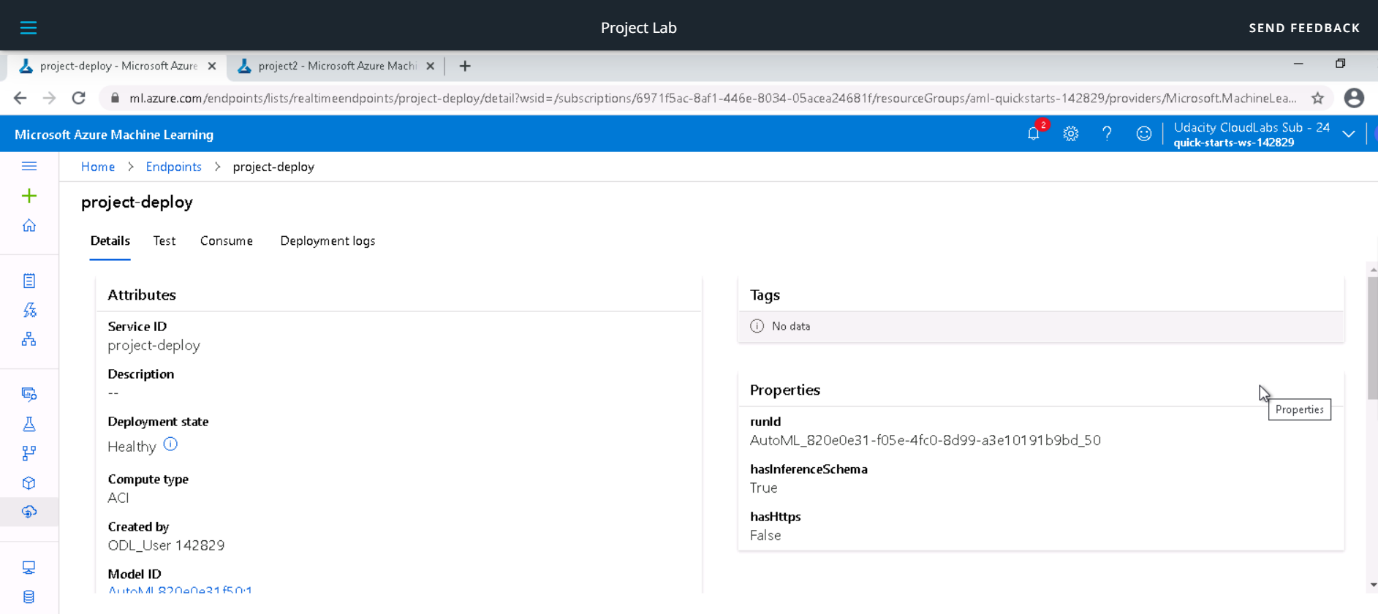


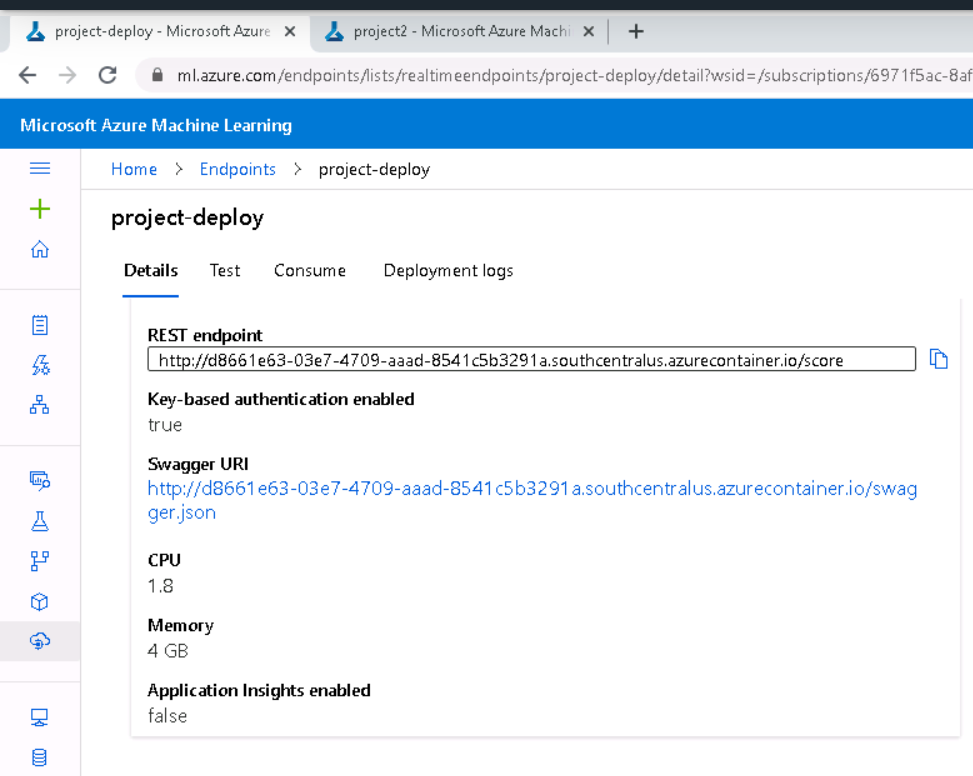
Deploy the best model

To interact with the best model obtained we have to deploy it using Azure Container Instance (ACI) and authentication is enabled.



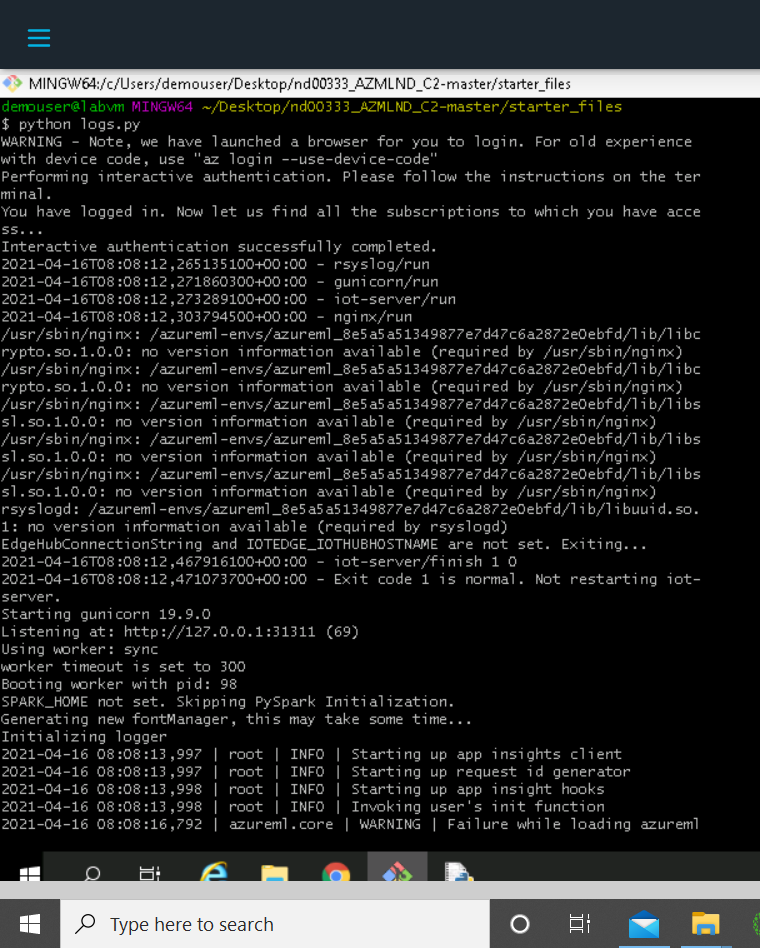


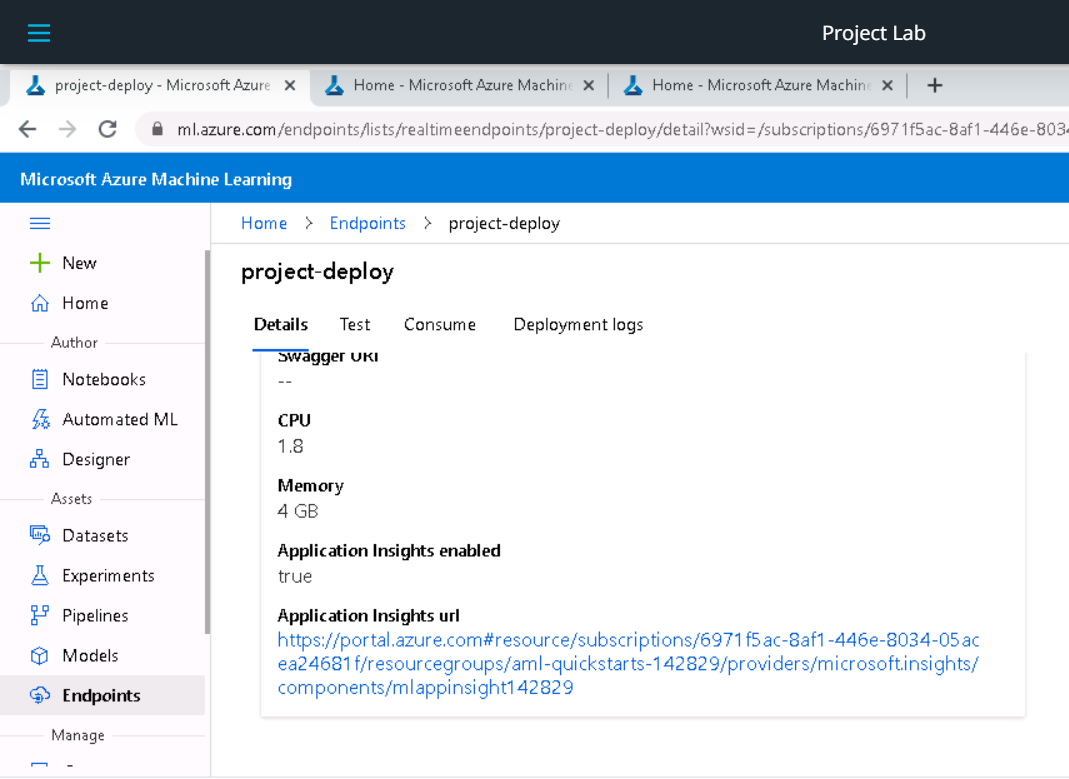




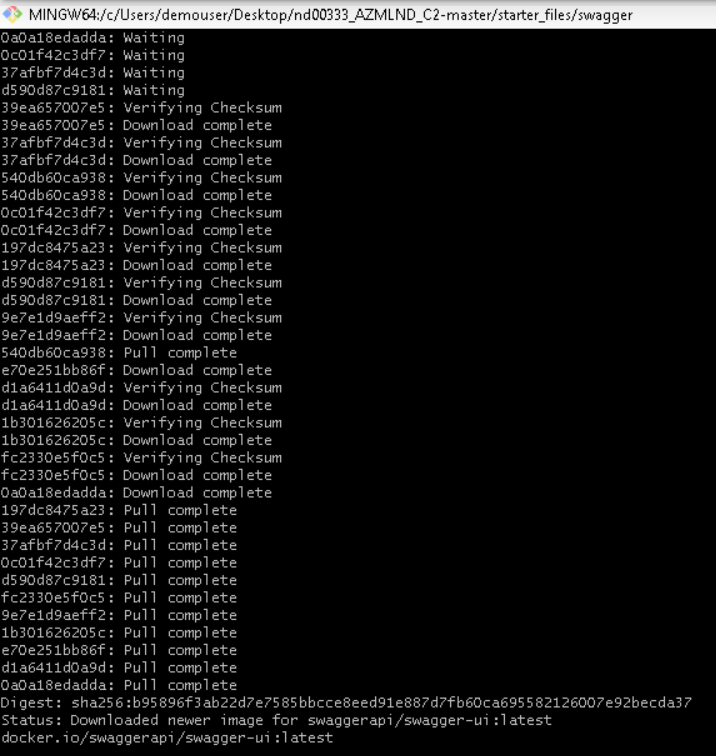
Enable logging

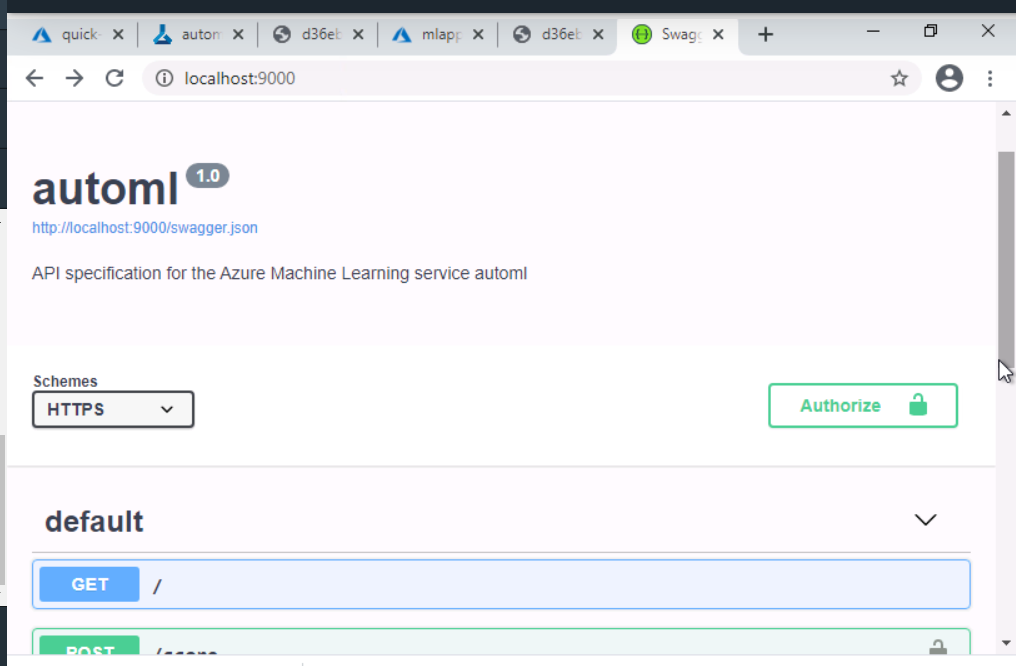
We enable application insights in logs.py and then run the script to view the logs generated.

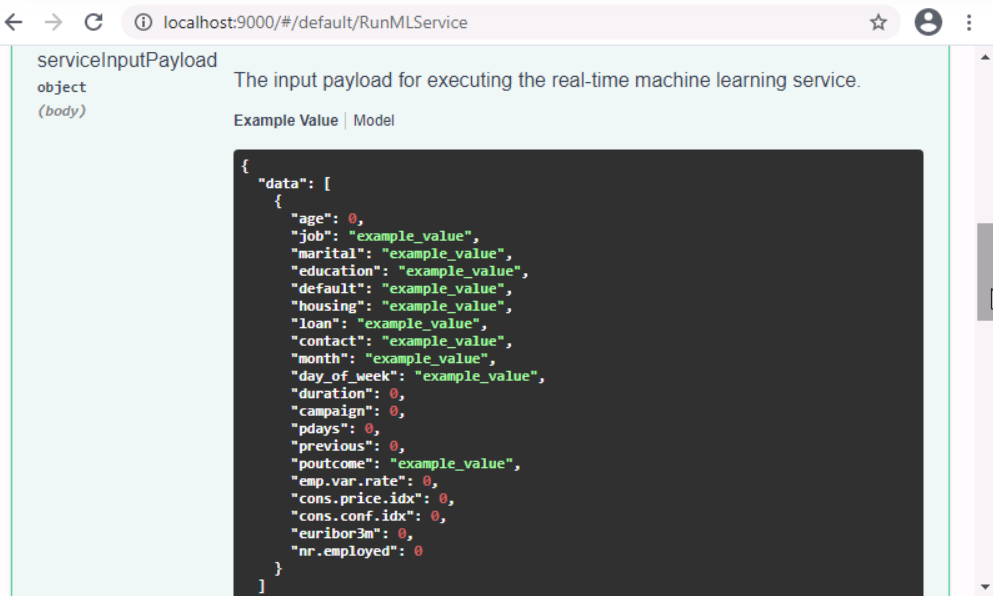




Swagger Documentation

Swagger is a tool that eases the documentation efforts of HTTP APIs. It makes it easier to explain what type of GET and POST requests. For consuming our best AutoML model using Swagger, we first need to download the swagger.json file provided to us in the Endpoints section of Azure Machine Learning Studio. 

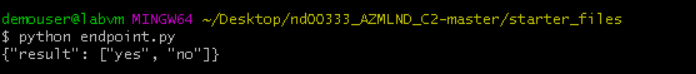




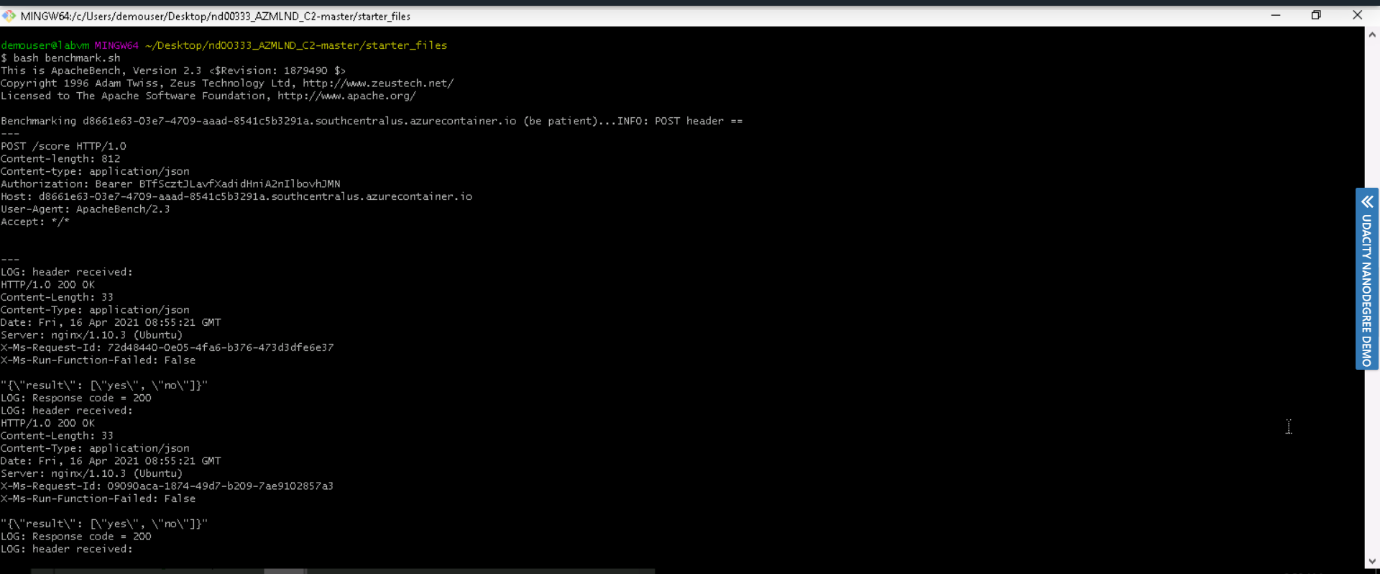
After running the script, we can find our best model's documentation instead of the default Swagger page.

Consume model endpoints

Finally, we can now interact with the model and feed some test data to it by adding the scoring\_uri and the key to the endpoint.py script and running it. So once a model has been deployed, an endpoint will be available which allows user to send inputs to the trained  
model and get a response back. This is called as consuming deployed service.

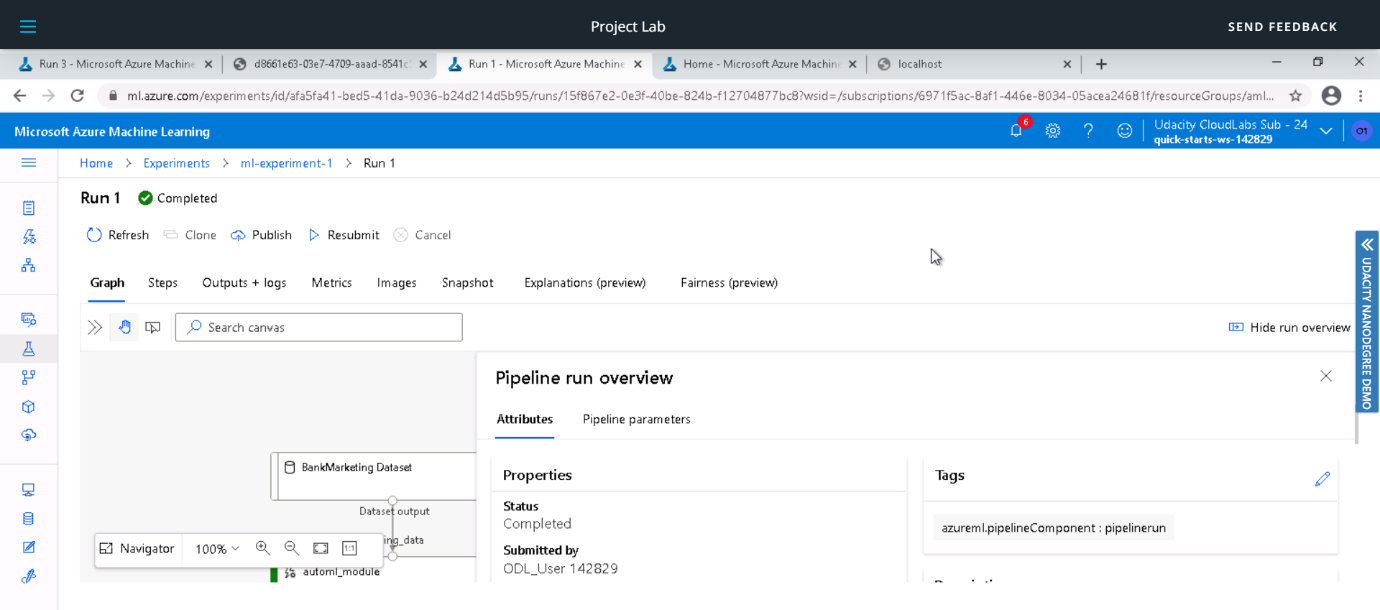


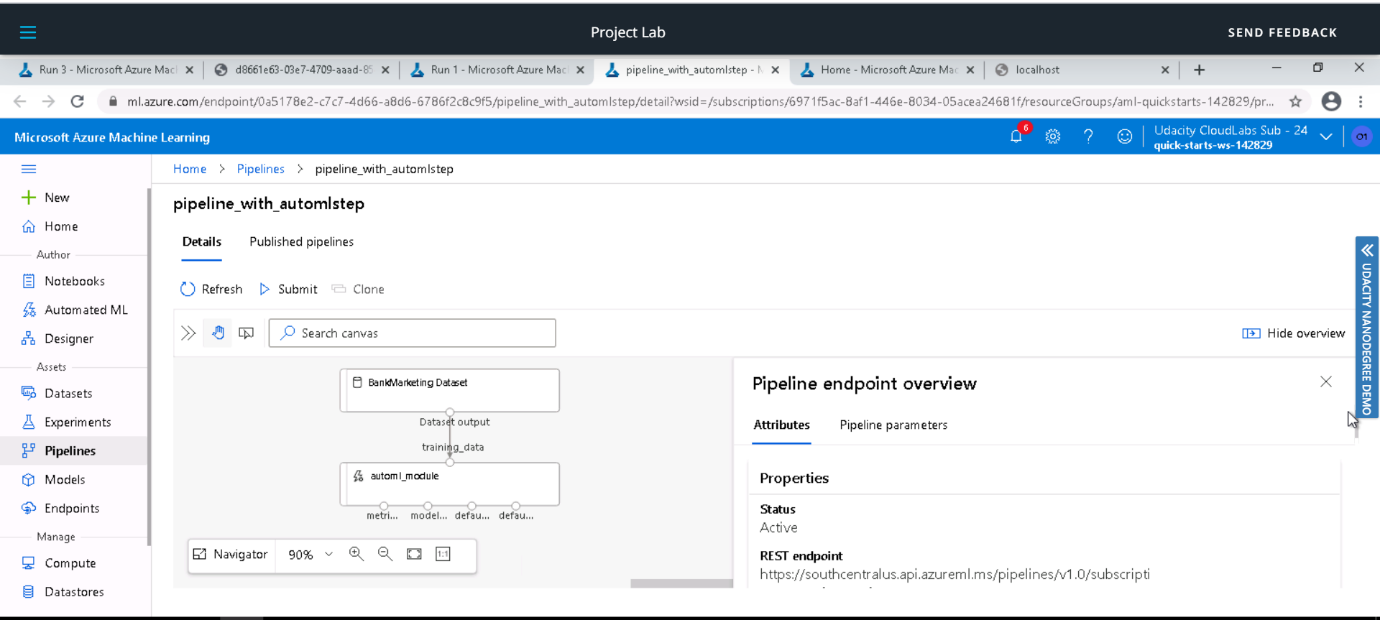
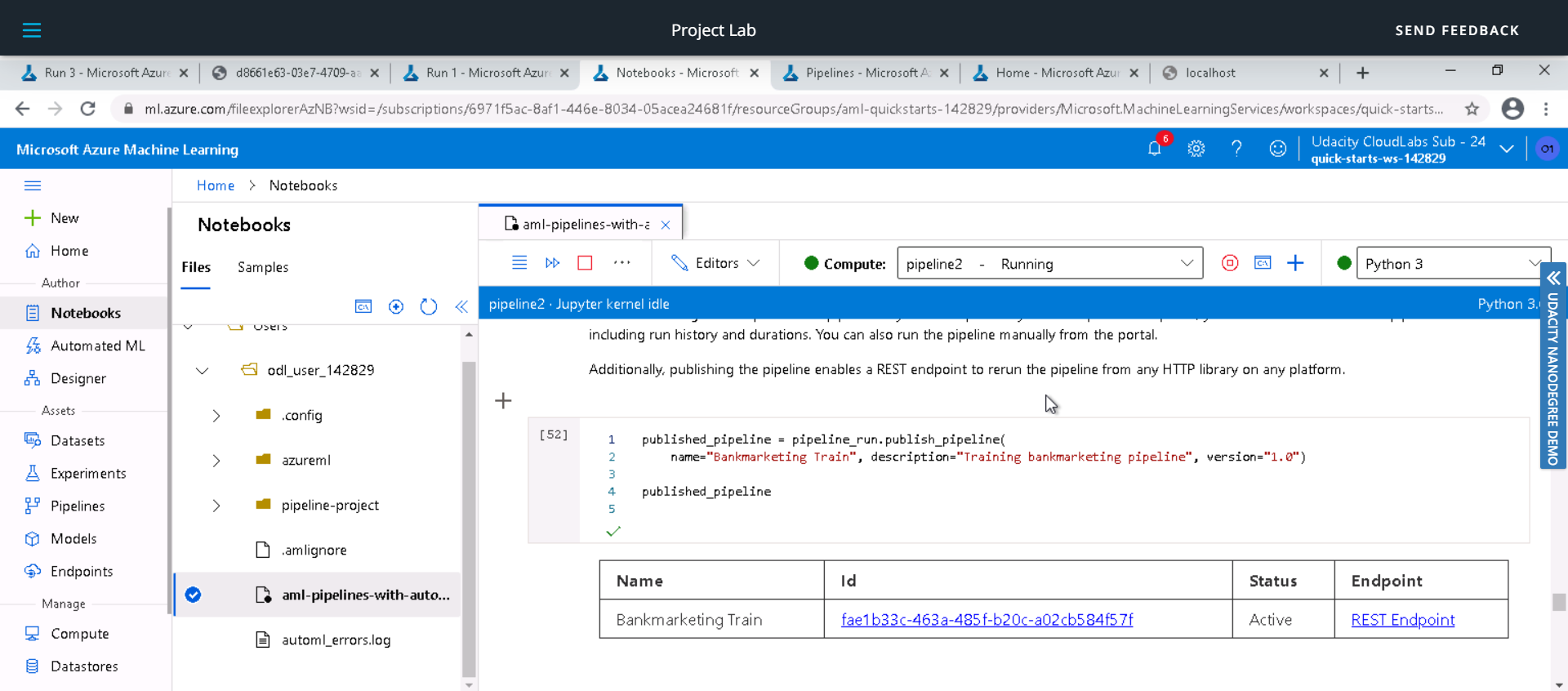
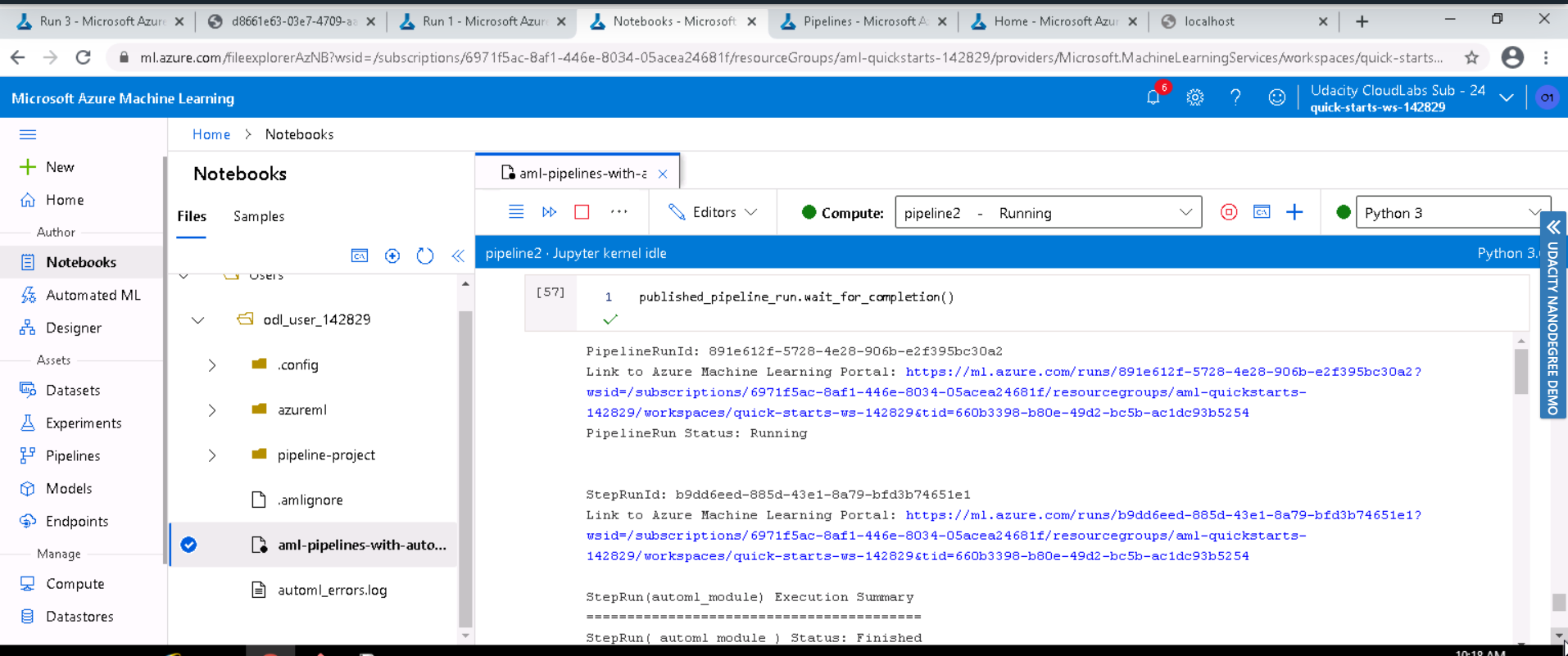
Benchmarking means setting a baseline as an acceptable performance measure. Here we use apache benchmark to see the model performance against our HTTP requests.



Create and publish a pipeline

Publishing a pipeline is the process of making a pipeline publicly available and to automate workflows.





We can also view the graph of pipeline endpoint and also its active status:

Screen Recording:

Screen Recording Link: https://youtu.be/lS\_pDtDVbTs

Future Improvements:

We can try with Deep Learning for better performance.

Resampling or adding more data can resolve class-imbalance issue and improve accuracy.

.