# DEPLOYING MACHINE LEARNING MODELS IN MICROSOFT AZURE ML WORKSPACE

THAPELO LENZI

14 JUNE 2024
METTUS/TESSERAI AI ACADEMY
MLOPS

## Overview

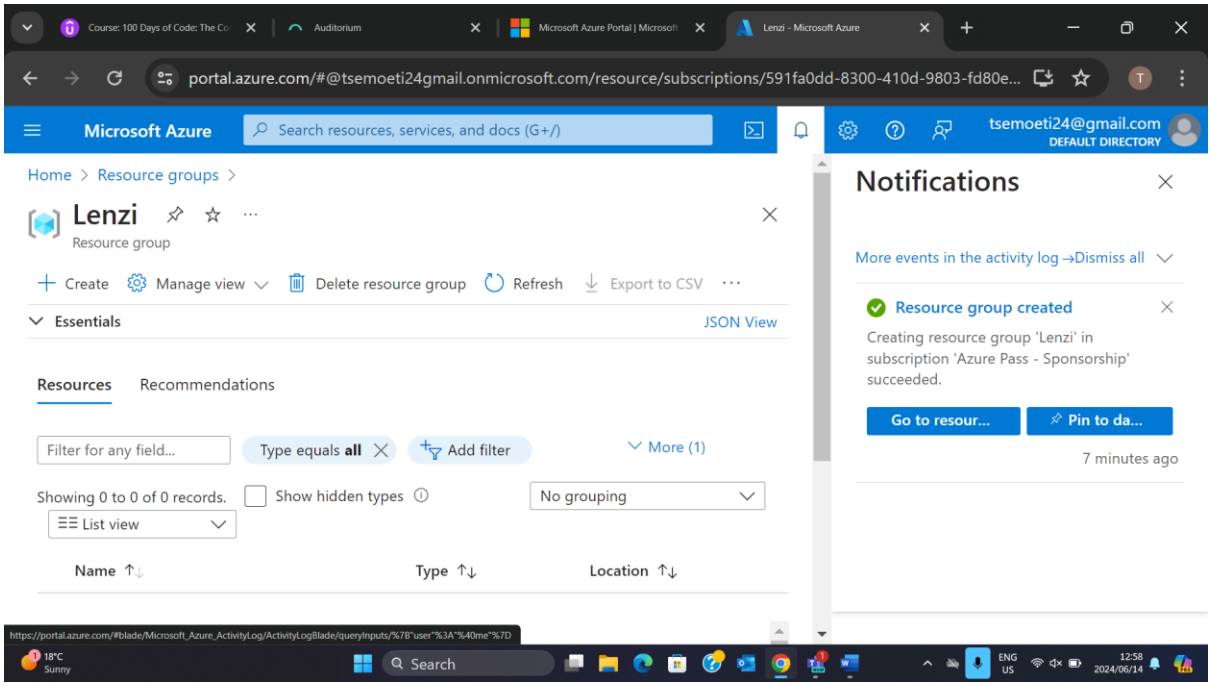
This document provides a brief guide on training and deploying a machine learning model using Microsoft Azure, aimed at data scientists, ml engineers, and/or AI enthusiasts like me. The project makes use of the diabetes dataset from Kaggle, covering steps from data extraction and preparation, through model training and evaluation, to deployment and monitoring. The dataset in use contains several independent (medical predictor variables), with only one target dependent variable. Its main objective is to diagnostically predict whether a patient is diabetic based on the predictor variables. The process involves using Azure ML Studio to build and refine the model and deploying it using Azure's deployment tools.

(https://www.kaggle.com/datasets/akshaydattatraykhare/diabetes-dataset)

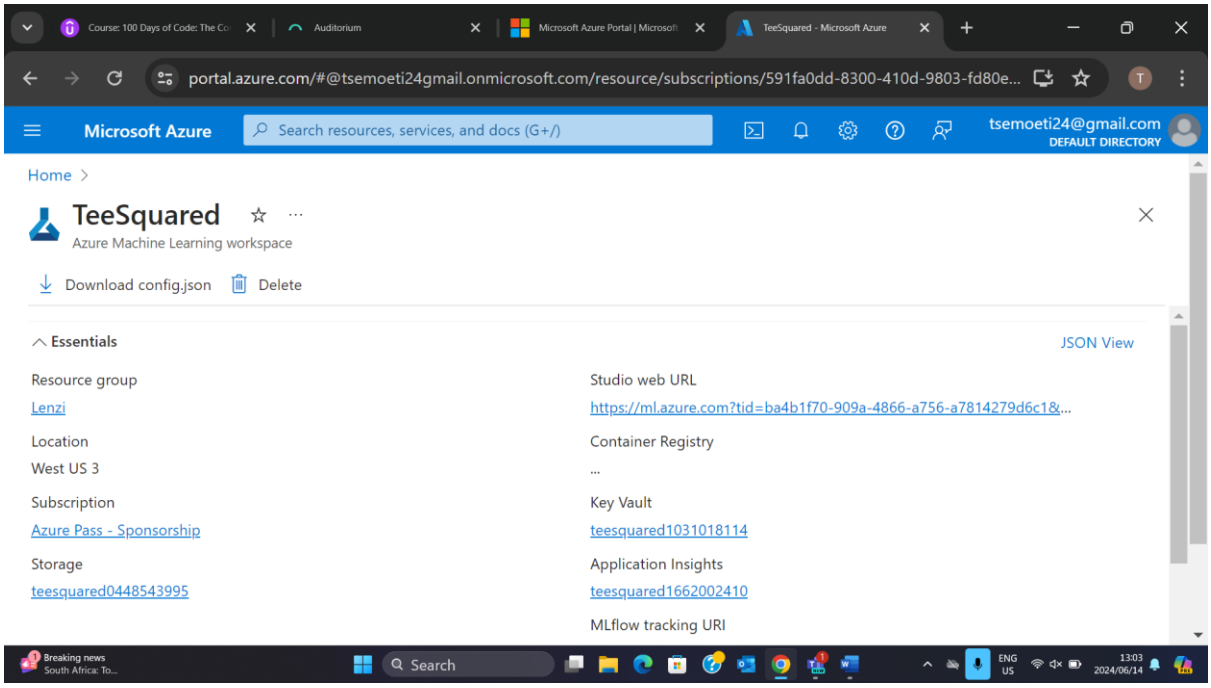
The system architecture for this project uses Microsoft Azure's suite of tools to well-organize the machine learning workflow. The diabetes dataset is stored in Azure Blob Storage, accessed and processed in the Azure ML Workspace (the centre for all the activities). Compute resources, such as Compute Instances, provide scalable processing power for model training and evaluation. Trained models are registered in Azure's Model Registry for version control and easy deployment. The model is deployed using Azure Kubernetes Services or Azure Container Instances (used in this project), creating a web service with RESTful API endpoints for real-time predictions. Performance and health of the deployed model is monitored using Azure Monitor and Application Insights, which ensures maintenance and updates.

## Azure Workspace

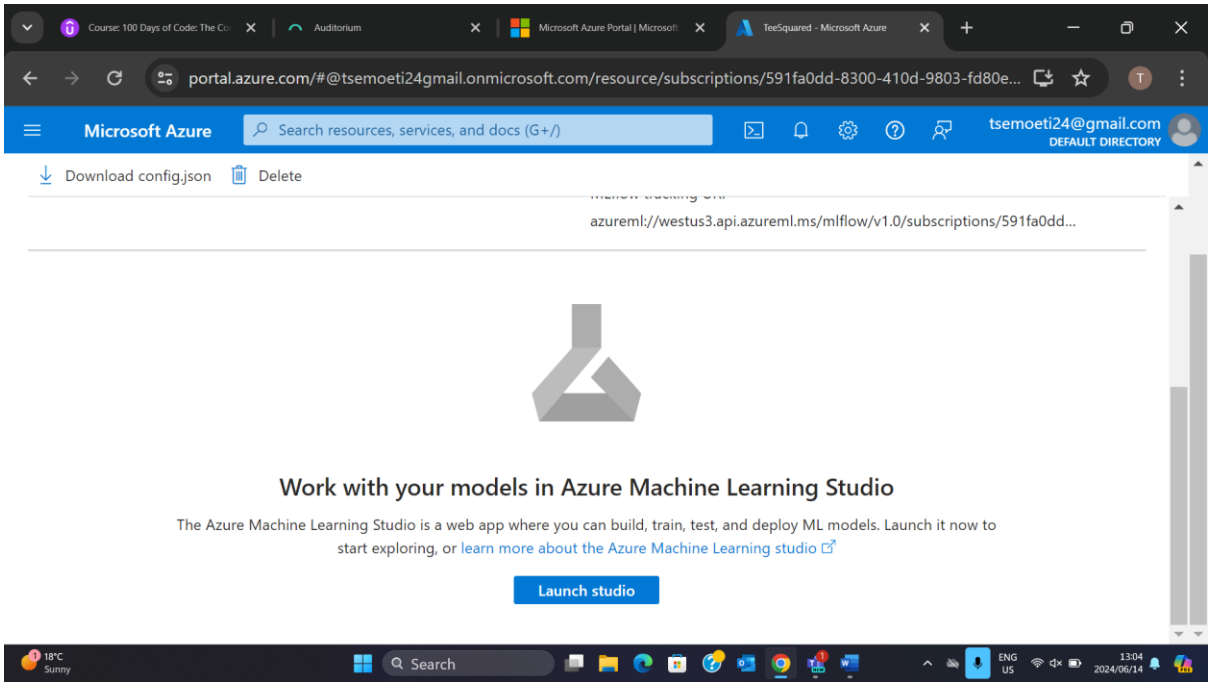
Setting up an Azure Machine Learning Workspace to organize and manage all the projects. Firstly, creating a resource group that will manage all my workspaces.

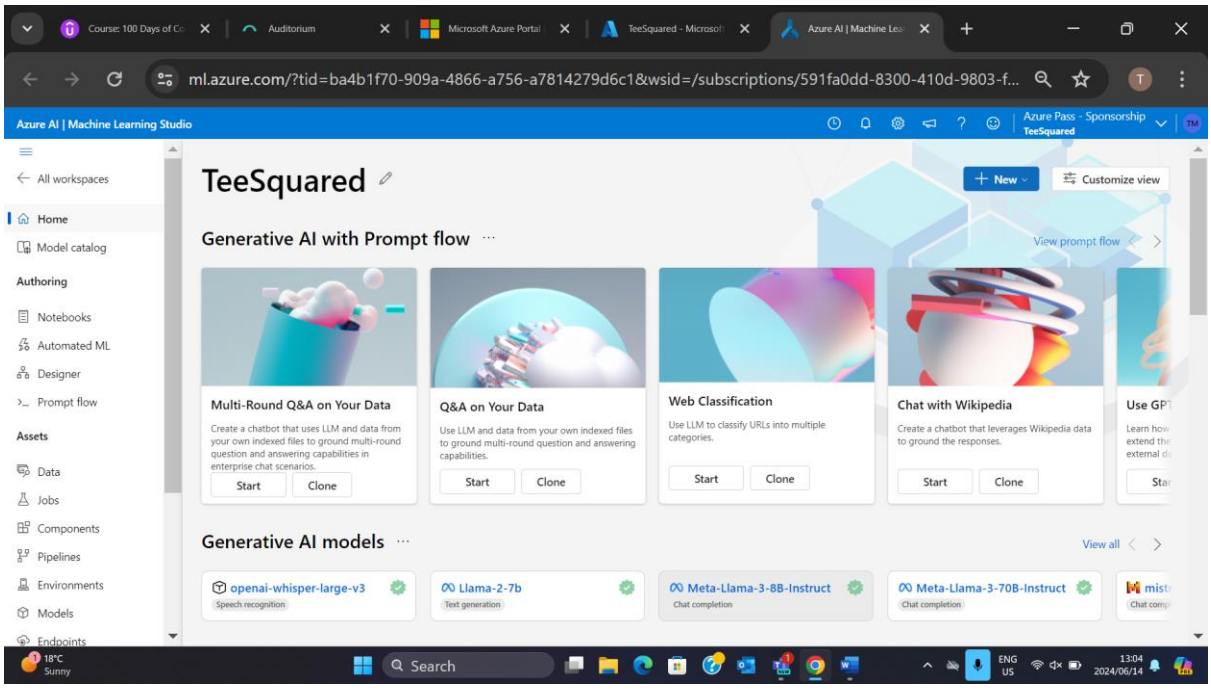



The created Resource Group is Lenzi and in it, a workspace named TeeSquared is created.



Navigating to Azure ML studio.

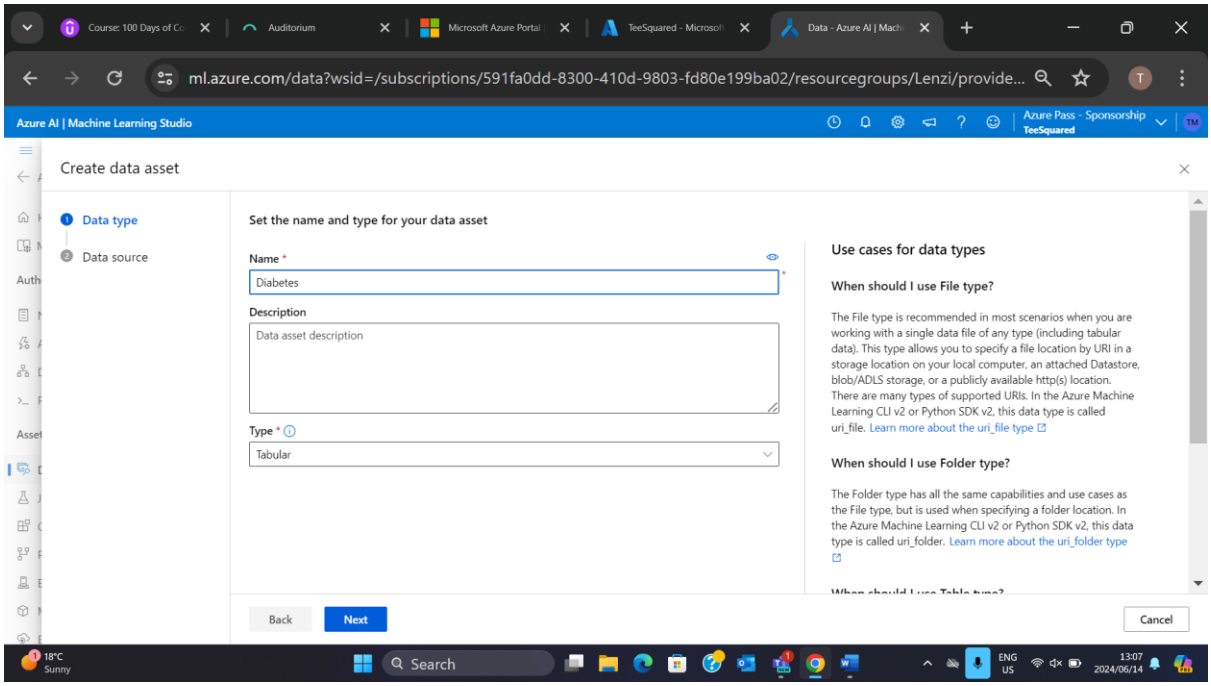


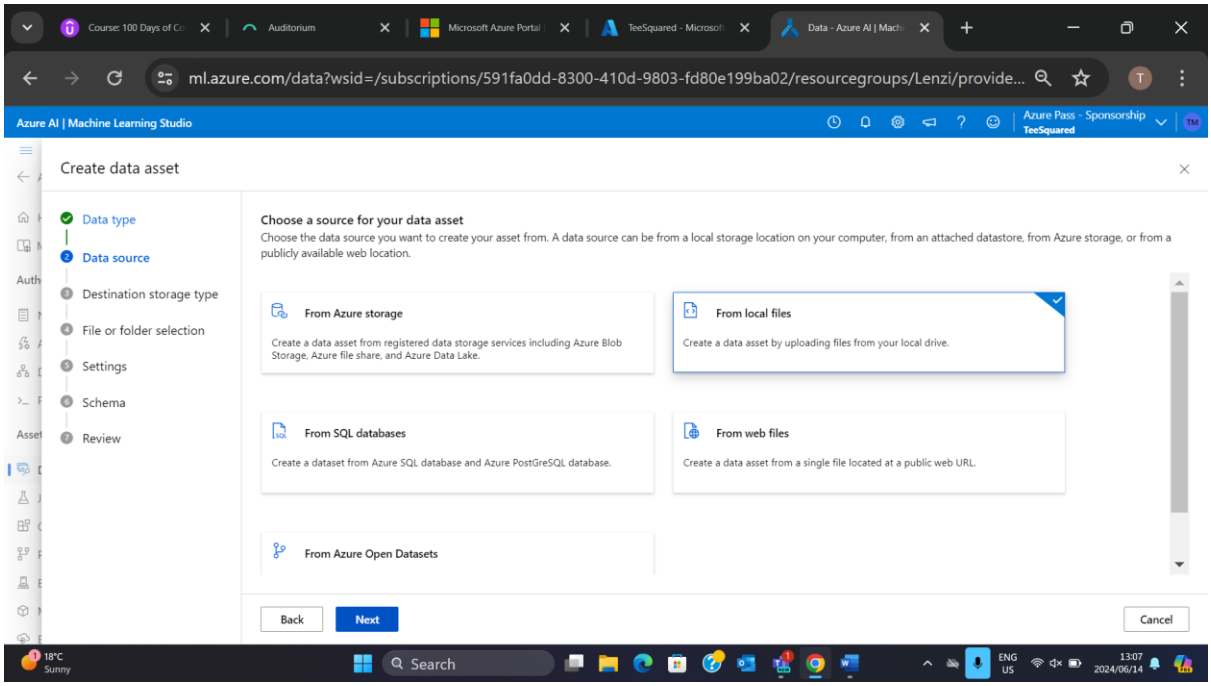



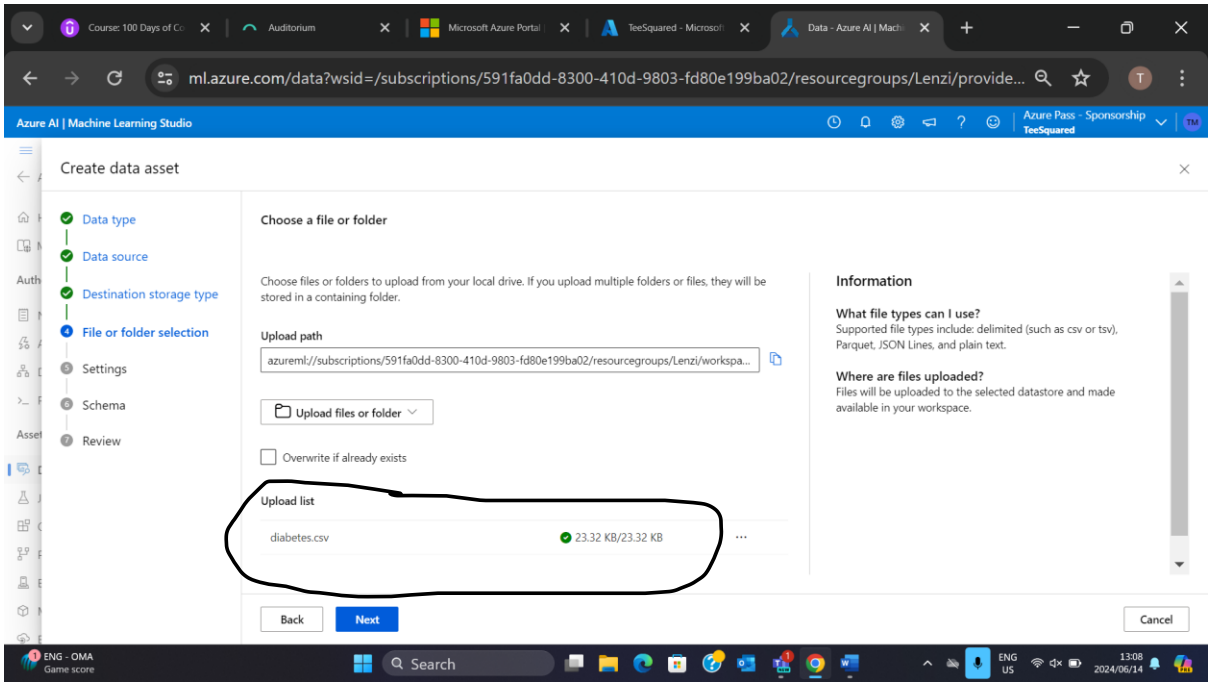

Now, one can access all the tools within the workspace.

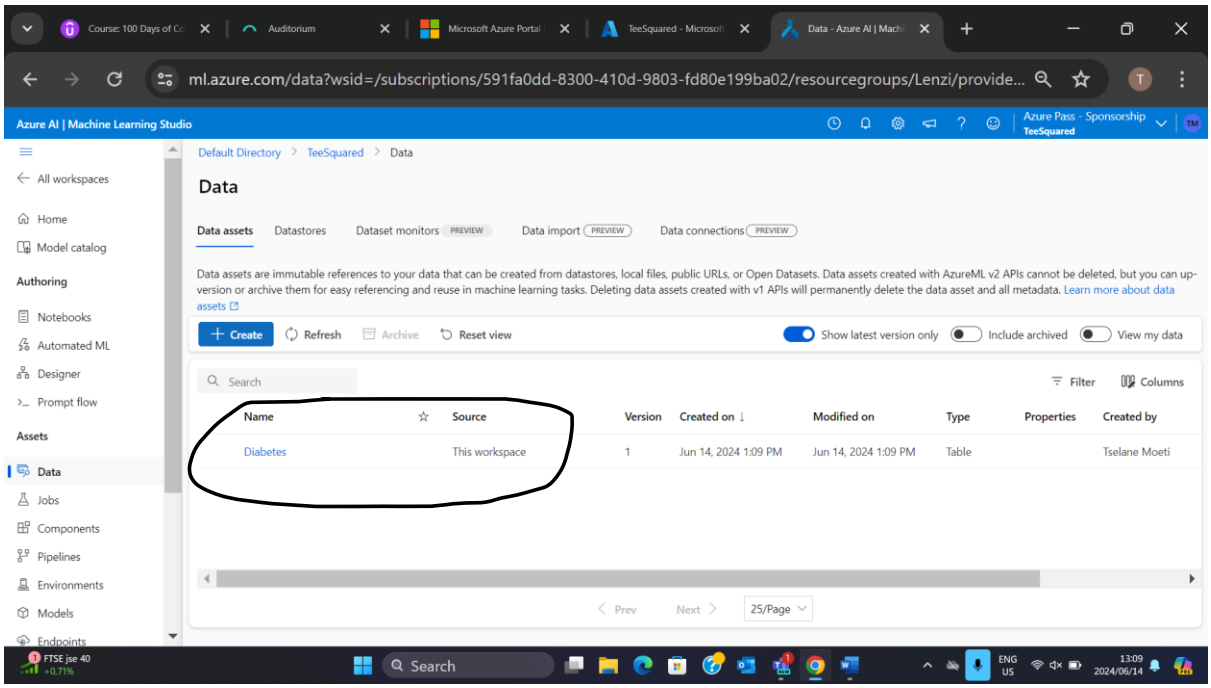
## Data Preparation

The dataset is loaded into Azure Blob Storage for easy access using the Data tab in Azure Workspace.


Course: 100 Days of Co
Auditorium
Microsoft Azure Portal
TeeSquared - Microsoft
Data - Azure AI | Machi
ml.azure.com/data?wsid=/subscriptions/591fa0dd-8300-410d-9803-fd80e199ba02/resourcegroups/Lenzi/provide...
Azure AI | Machine Learning Studio
Azure Pass - Sponsorship
TeeSquared
Create data asset
Data type
Data source
Set the name and type for your data asset
Name *
Diabetes
Description
Data asset description
Type *
Tabular
Use cases for data types
When should I use File type?
The File type is recommended in most scenarios when you are working with a single data file of any type (including tabular data). This type allows you to specify a file location by URI in a storage location on your local computer, an attached Datastore, blob/ADLS storage, or a publicly available http(s) location. There are many types of supported URIs. In the Azure Machine Learning CLI v2 or Python SDK v2, this data type is called uri_file. Learn more about the uri_file type
When should I use Folder type?
The Folder type has all the same capabilities and use cases as the File type, but is used when specifying a folder location. In the Azure Machine Learning CLI v2 or Python SDK v2, this data type is called uri_folder. Learn more about the uri_folder type
Back
Next
Cancel
18°C
Sunny
Search
ENG
US
13:07
2024/06/14


Course: 100 Days of Co
Auditorium
Microsoft Azure Portal
TeeSquared - Microsoft
Data - Azure AI | Machi
ml.azure.com/data?wsid=/subscriptions/591fa0dd-8300-410d-9803-fd80e199ba02/resourcegroups/Lenzi/provide...
Azure AI | Machine Learning Studio
Azure Pass - Sponsorship
TeeSquared
Create data asset
Data type
Data source
Destination storage type
File or folder selection
Settings
Schema
Review
Choose a source for your data asset
Choose the data source you want to create your asset from. A data source can be from a local storage location on your computer, from an attached datastore, from Azure storage, or from a publicly available web location.
From Azure storage
Create a data asset from registered data storage services including Azure Blob Storage, Azure file share, and Azure Data Lake.
From local files
Create a data asset by uploading files from your local drive.
From SQL databases
Create a dataset from Azure SQL database and Azure PostGreSQL database.
From web files
Create a data asset from a single file located at a public web URL.
From Azure Open Datasets
Back
Next
Cancel
18°C
Sunny
Search
ENG
US
13:07
2024/06/14

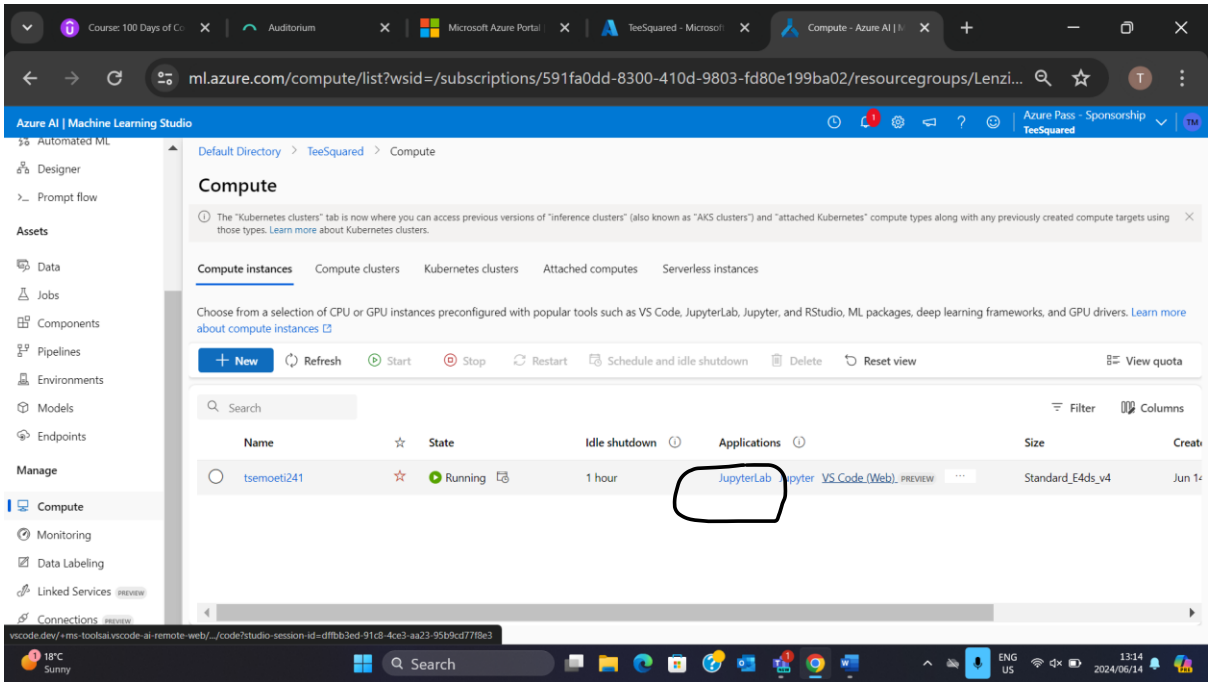



Now, the data is in storage.

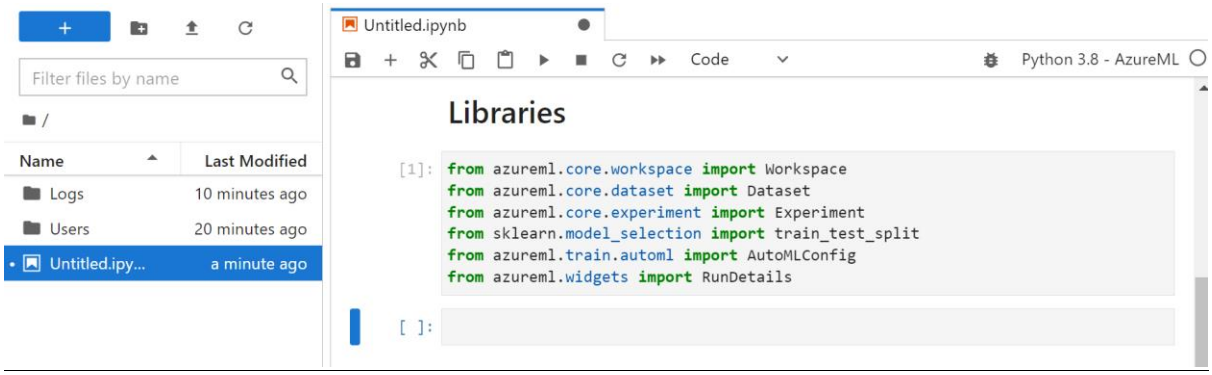
## Compute Instance

For efficient model training, provision a Compute Instance. In the Compute Instance applications, the JupyterLab application is used to trained the model using the Python 3.8-AzureML kernel.

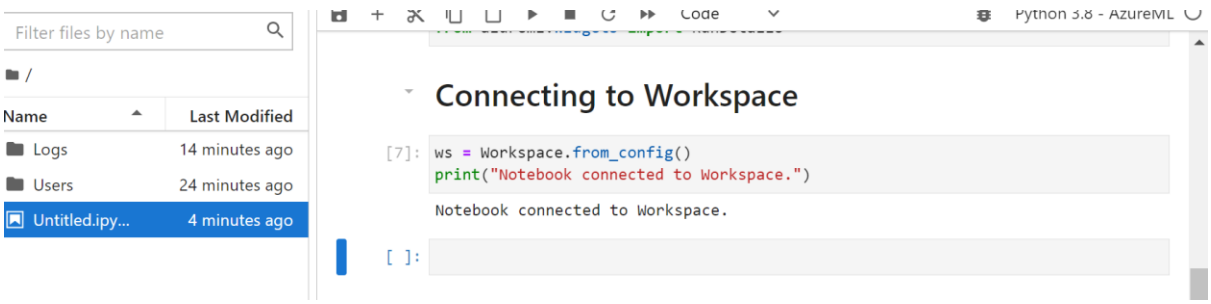



## JupyterLab Environment

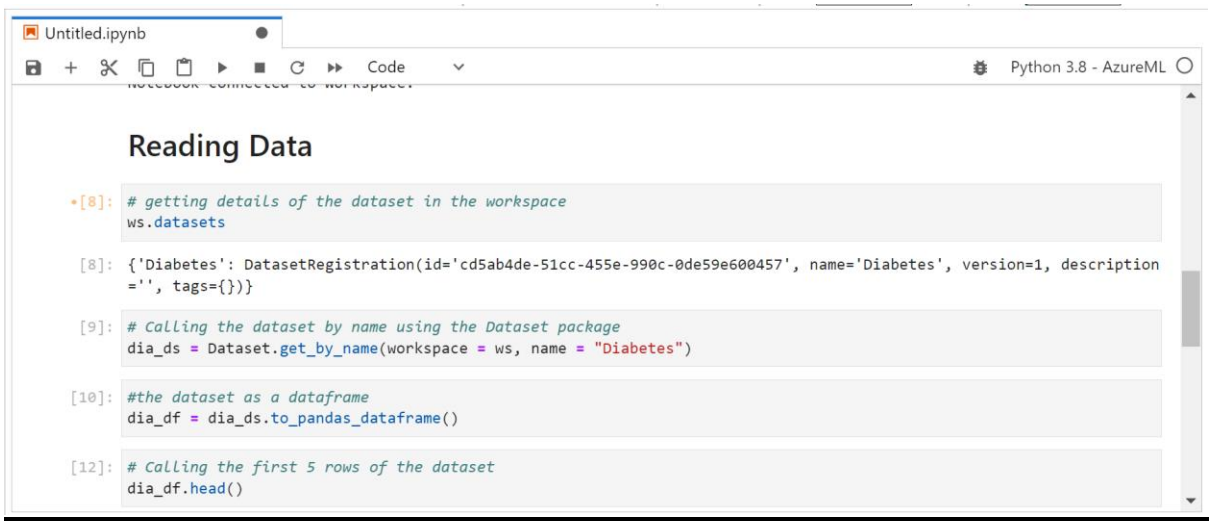
Necessary libraries

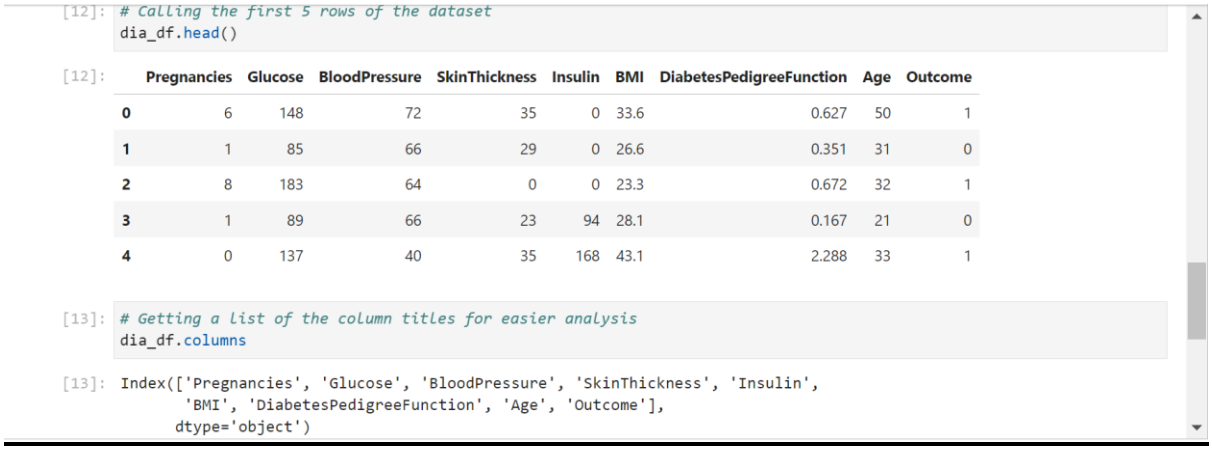



In order to read diabetes dataset in the JupyterLab environment, the workspace is called in the JupyterLab environment.



Extracting the dataset from the workspace.



[12]:
```
# Calling the first 5 rows of the dataset
dia_df.head()
```

| | Pregnancies | Glucose | BloodPressure | SkinThickness | Insulin | BMI | DiabetesPedigreeFunction | Age | Outcome |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 6 | 148 | 72 | 35 | 0 | 33.6 | 0.627 | 50 | 1 |
| 1 | 1 | 85 | 66 | 29 | 0 | 26.6 | 0.351 | 31 | 0 |
| 2 | 8 | 183 | 64 | 0 | 0 | 23.3 | 0.672 | 32 | 1 |
| 3 | 1 | 89 | 66 | 23 | 94 | 28.1 | 0.167 | 21 | 0 |
| 4 | 0 | 137 | 40 | 35 | 168 | 43.1 | 2.288 | 33 | 1 |

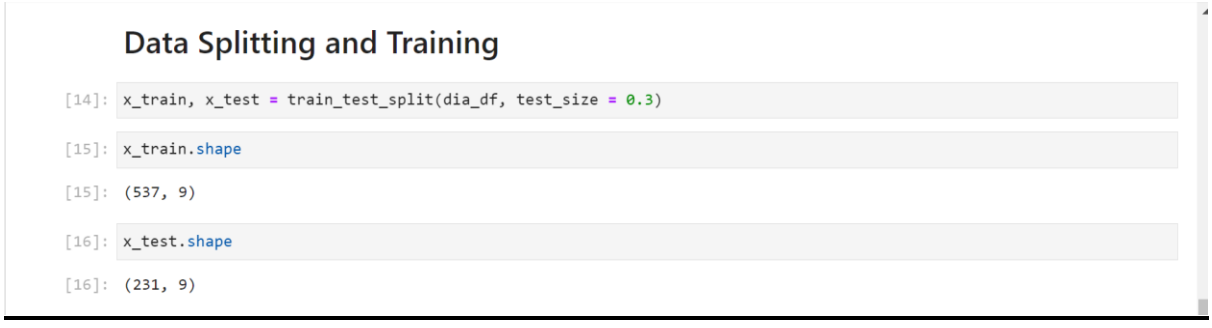
[13]:
```
# Getting a list of the column titles for easier analysis
dia_df.columns
```

[13]:
```
Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin',
       'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],
      dtype='object')
```

The above dataset is the clean version of the diabetes dataset, thus there is no need for the cleaning stage. So one can go ahead to split and train the data.

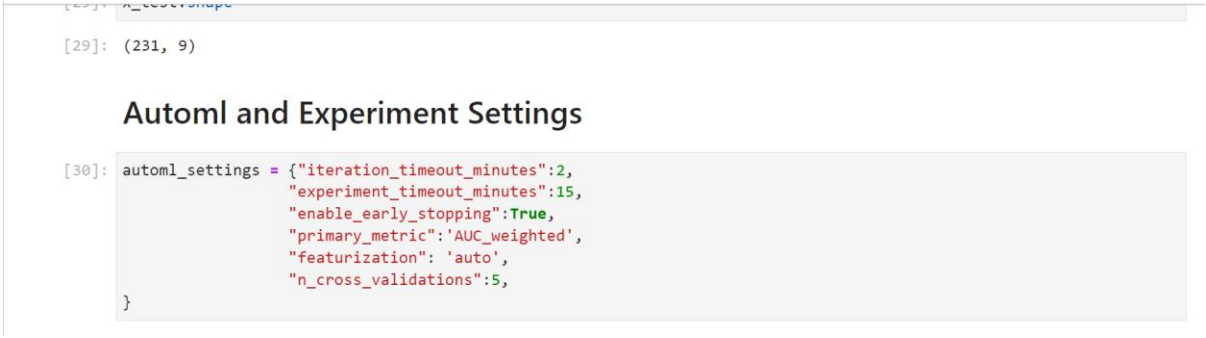
### Data Training

70% of the dataset will be used to train the model and 30% will be used to test it.



## AutoML and Algorithm specifications

Automl automates several important tasks in the machine learning workflow, which include feature engineering and algorithm selection. In the following snippet, by defining these parameters, automl efficiently manages the duration of training iterations and experiments, stopping the iterations early when the performance metrics are stable. The metric used for algorithm selection is “AUC_weighted”. With the n_cross_validations parameter set to 5, ensures a thorough validation of models during training, improving their generalization capabilities.



```
[29]: (231, 9)
```

Automl and Experiment Settings

```
[30]: automl_settings = {"iteration_timeout_minutes":2,
                        "experiment_timeout_minutes":15,
                        "enable_early_stopping":True,
                        "primary_metric":'AUC_weighted',
                        "featurization": 'auto',
                        "n_cross_validations":5,
}
```

These configurations emphasise the role of AutoML in optimizing machine learning pipelines, making it easier to deploy accurate predictive models. In the case of this project, accurately predicting diabetes outcomes.

Specifying task and algorithm

```
[32]: automl_config = AutoMLConfig(task='classification',
                                  debug_log='automl_errors.log',
                                  training_data = x_train,
                                  label_column_name = "Outcome",
                                  **automl_settings)
```

An experiment named Diabetes-Experiment is created in the workspace to track and manage all the model training runs

## Experiment for Deployment

[33]:
```
experiment = Experiment(ws, "Diabetes-Experiment")
```

[*]:
```
run = experiment.submit(automl_config, show_output=True)
```

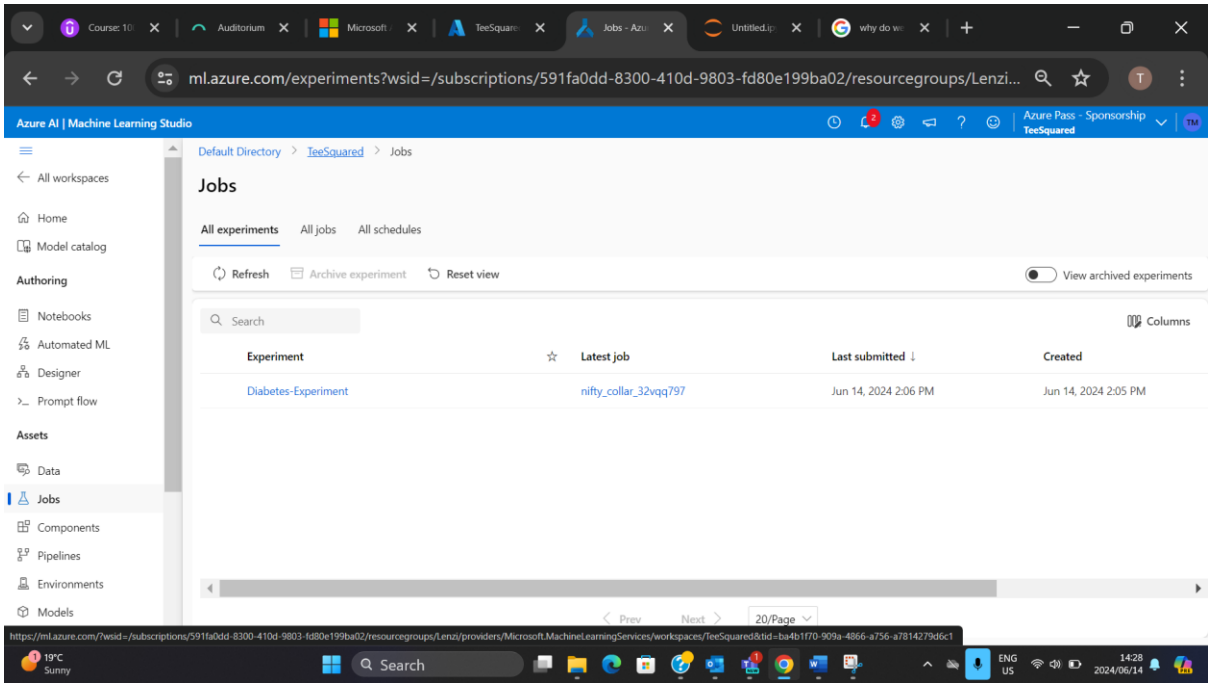
```
No run_configuration provided, running on local with default configuration
Running in the active local environment.
```

| Experiment | Id | Type | Status | Details Page | Docs Page |
|---|---|---|---|---|---|
| Diabetes-Experiment | AutoML_f595901b-e4a2-42c2-bda7-6e8469ec21e1 | automl | Preparing | Link to Azure Machine Learning studio | Link to Documentation |

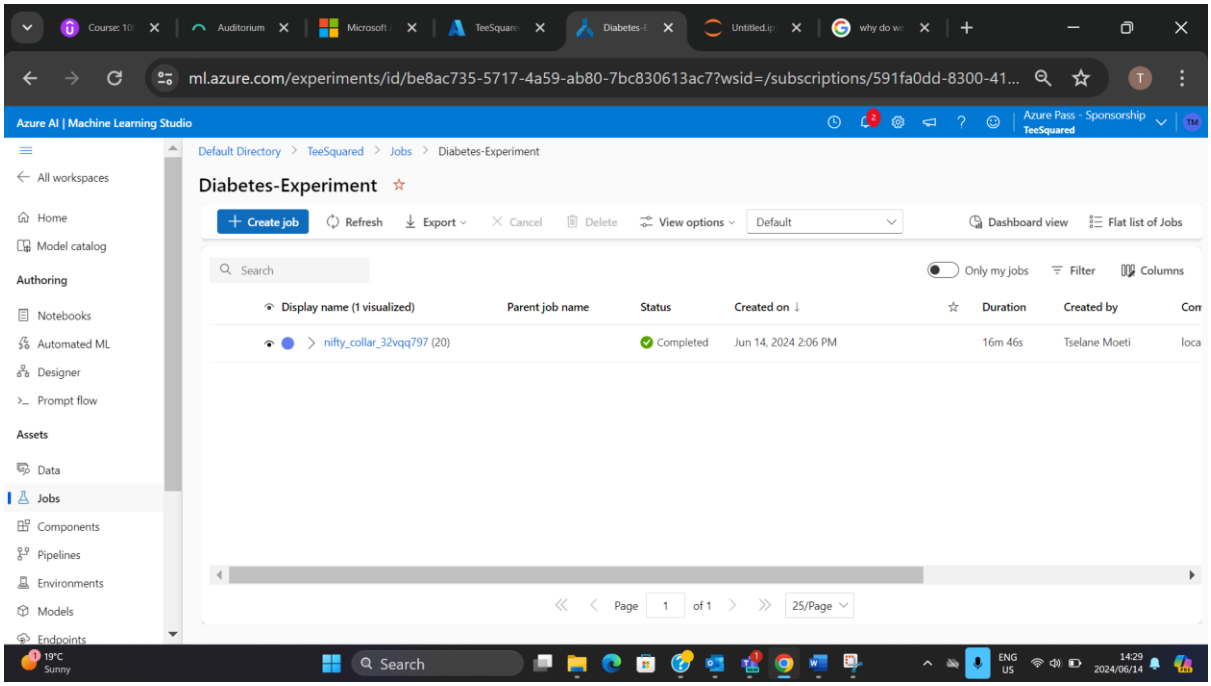
```
Current status: DatasetEvaluation. Gathering dataset statistics.
Current status: FeaturesGeneration. Generating features for the dataset.
Current status: DatasetFeaturization. Beginning to fit featurizers and featurize the dataset.
Current status: DatasetFeaturizationCompleted. Completed fit featurizers and featurizing the dataset.
```

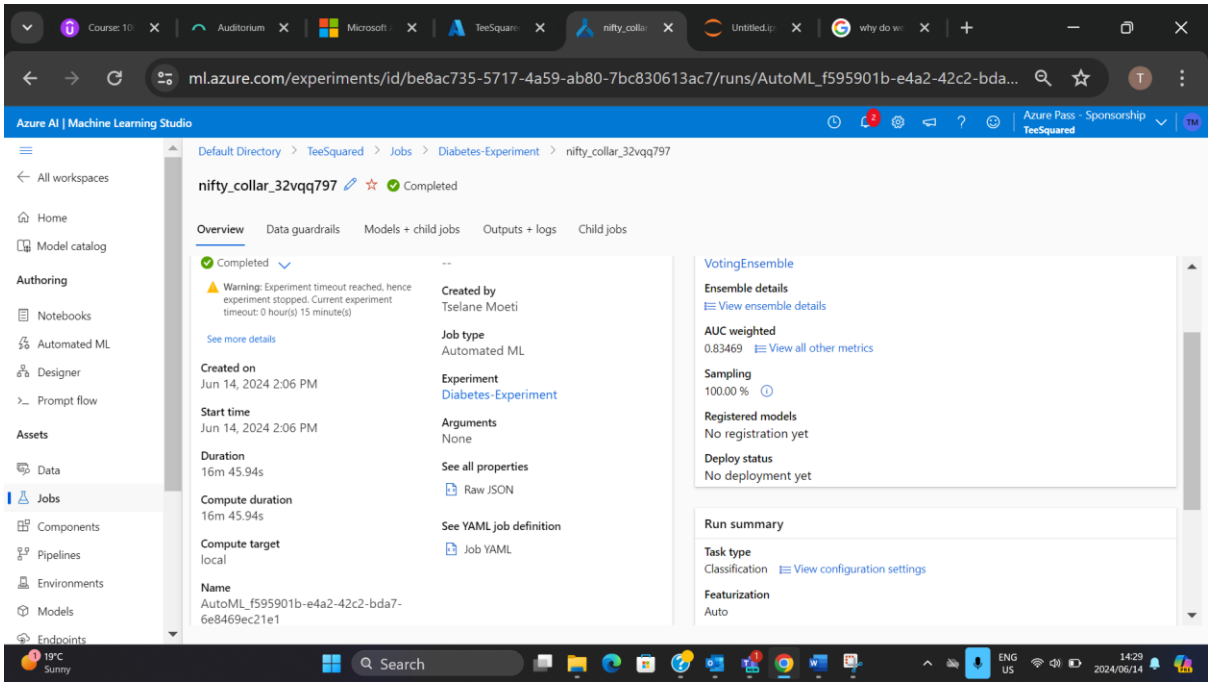
```
Current status: DatasetEvaluation. Gathering dataset statistics.
Current status: FeaturesGeneration. Generating features for the dataset.
Current status: DatasetFeaturization. Beginning to fit featurizers and featurize the dataset.
Current status: DatasetFeaturizationCompleted. Completed fit featurizers and featurizing the dataset.
Current status: DatasetCrossValidationSplit. Generating individually featurized CV splits.
2024/06/14 12:07:02 WARNING mlflow.sklearn: Model was missing function: predict. Not logging python_function flavo
r!

*****************************************************************************************
DATA GUARDRAILS:

TYPE:         Class balancing detection
STATUS:       PASSED
DESCRIPTION:  Your inputs were analyzed, and all classes are balanced in your training data.
              Learn more about imbalanced data: https://aka.ms/AutomatedMLImbalancedData

*****************************************************************************************

TYPE:         Missing feature values imputation
STATUS:       PASSED
```

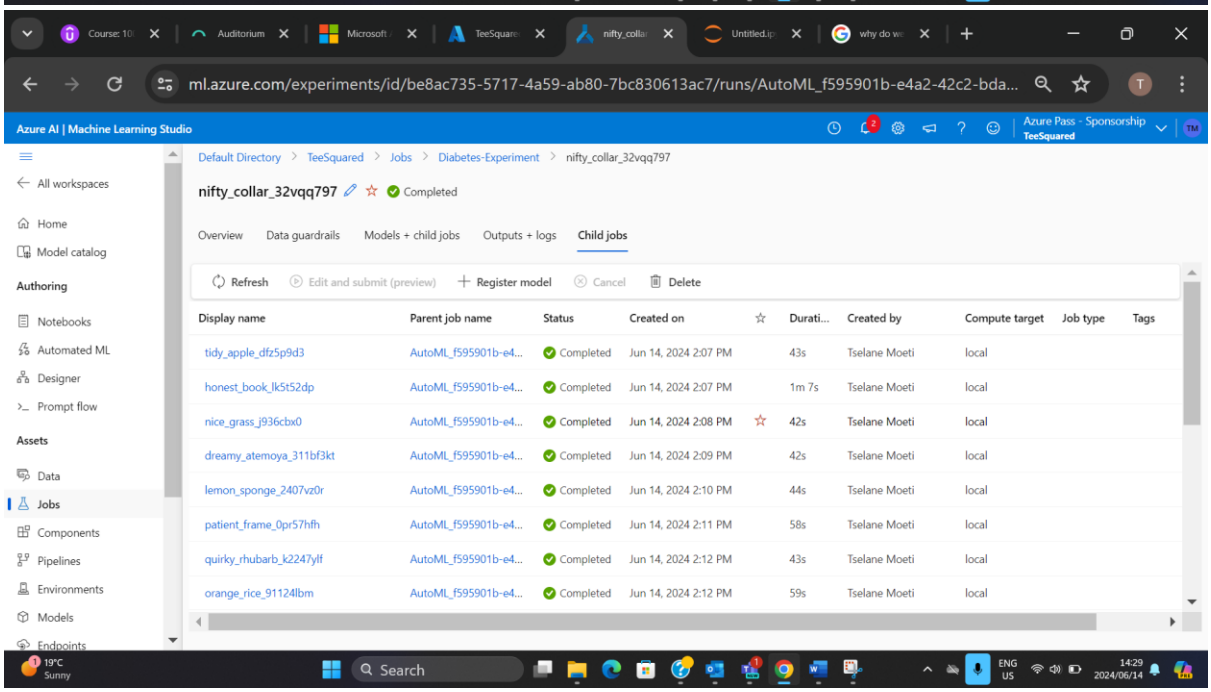
```
    13   StandardScalerWrapper ExtremeRandomTrees       0:00:42             0.7554    0.8192
    14   StandardScalerWrapper RandomForest             0:00:44             0.7983    0.8192
    15   MaxAbsScaler LightGBM                          0:00:41             0.8195    0.8195
    16   MaxAbsScaler LogisticRegression                0:00:41             0.7928    0.8195
    17   StandardScalerWrapper ExtremeRandomTrees       0:00:46             0.8197    0.8197
    18   VotingEnsemble                                 0:00:46             0.8347    0.8347
    19   StackEnsemble                                  0:00:46             0.8303    0.8347
Stopping criteria reached at iteration 20. Ending experiment.
********************************************************************************************
Current status: BestRunExplainModel. Best run model explanations started
Current status: ModelExplanationDataSetSetup. Model explanations data setup completed
Current status: PickSurrogateModel. Choosing LightGBM as the surrogate model for explanations
Current status: EngineeredFeatureExplanations. Computation of engineered features started
2024-06-14:12:24:27,38 INFO     [explanation_client.py:334] Using default datastore for uploads
Current status: EngineeredFeatureExplanations. Computation of engineered features completed
Current status: RawFeaturesExplanations. Computation of raw features started
Current status: RawFeaturesExplanations. Computation of raw features completed
Current status: BestRunExplainModel. Best run model explanations completed
********************************************************************************************
```

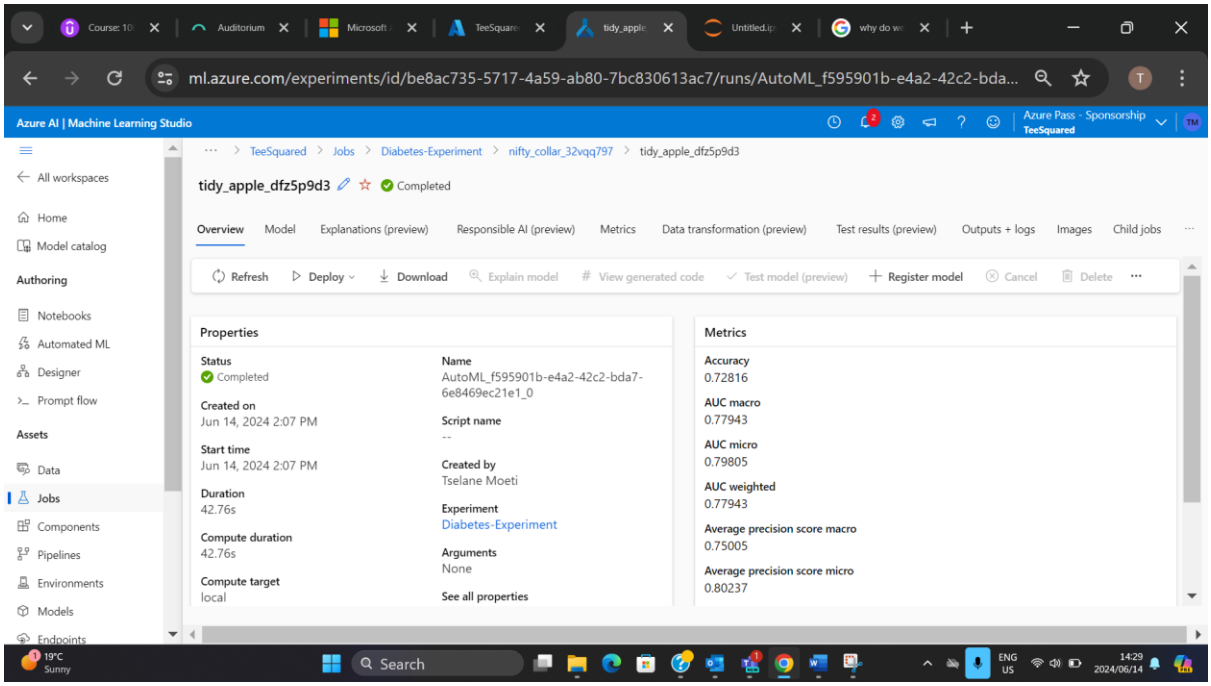



The following shows all the jobs and child jobs under the created experiment job.


ml.azure.com/experiments/id/be8ac735-5717-4a59-ab80-7bc830613ac7?wsid=/subscriptions/591fa0dd-8300-41...
Azure AI | Machine Learning Studio
Azure Pass - Sponsorship
TeeSquared
All workspaces
Home
Model catalog
Authoring
Notebooks
Automated ML
Designer
Prompt flow
Assets
Data
Jobs
Components
Pipelines
Environments
Models
Endpoints
Default Directory > TeeSquared > Jobs > Diabetes-Experiment
Diabetes-Experiment
Create job
Refresh
Export
Cancel
Delete
View options
Default
Dashboard view
Flat list of Jobs
Search
Only my jobs
Filter
Columns
Display name (1 visualized)
Parent job name
Status
Created on
Duration
Created by
nifty_collar_32vqq797 (20)
Completed
Jun 14, 2024 2:06 PM
16m 46s
Tselane Moeti
Page 1 of 1
25/Page
19°C
Sunny
Search
ENG
US
14:29
2024/06/14


ml.azure.com/experiments/id/be8ac735-5717-4a59-ab80-7bc830613ac7/runs/AutoML_f595901b-e4a2-42c2-bda...
Azure AI | Machine Learning Studio
Azure Pass - Sponsorship
TeeSquared
Default Directory > TeeSquared > Jobs > Diabetes-Experiment > nifty_collar_32vqq797
nifty_collar_32vqq797
Completed
All workspaces
Home
Model catalog
Authoring
Notebooks
Automated ML
Designer
Prompt flow
Assets
Data
Jobs
Components
Pipelines
Environments
Models
Endpoints
Overview
Data guardrails
Models + child jobs
Outputs + logs
Child jobs
Warning: Experiment timeout reached, hence experiment stopped. Current experiment timeout: 0 hour(s) 15 minute(s)
See more details
Created on
Jun 14, 2024 2:06 PM
Start time
Jun 14, 2024 2:06 PM
Duration
16m 45.94s
Compute duration
16m 45.94s
Compute target
local
Name
AutoML_f595901b-e4a2-42c2-bda7-6e8469ec21e1
Created by
Tselane Moeti
Job type
Automated ML
Experiment
Diabetes-Experiment
Arguments
None
See all properties
Raw JSON
See YAML job definition
Job YAML
VotingEnsemble
Ensemble details
View ensemble details
AUC weighted
0.83469
View all other metrics
Sampling
100.00 %
Registered models
No registration yet
Deploy status
No deployment yet
Run summary
Task type
Classification
View configuration settings
Featurization
Auto
19°C
Sunny
Search
ENG
US
14:29
2024/06/14


ml.azure.com/experiments/id/be8ac735-5717-4a59-ab80-7bc830613ac7/runs/AutoML_f595901b-e4a2-42c2-bda...
Azure AI | Machine Learning Studio
Azure Pass - Sponsorship
TeeSquared
Default Directory > TeeSquared > Jobs > Diabetes-Experiment > nifty_collar_32vqq797
nifty_collar_32vqq797
Completed
Overview
Data guardrails
Models + child jobs
Outputs + logs
Child jobs
Refresh
Edit and submit (preview)
Register model
Cancel
Delete
Display name
Parent job name
Status
Created on
Durati...
Created by
Compute target
Job type
Tags
tidy_apple_dfz5p9d3 AutoML_f595901b-e4... Completed Jun 14, 2024 2:07 PM 43s Tselane Moeti local
honest_book_lk5t52dp AutoML_f595901b-e4... Completed Jun 14, 2024 2:07 PM 1m 7s Tselane Moeti local
nice_grass_j936cbx0 AutoML_f595901b-e4... Completed Jun 14, 2024 2:08 PM 42s Tselane Moeti local
dreamy_atemoya_311bf3kt AutoML_f595901b-e4... Completed Jun 14, 2024 2:09 PM 42s Tselane Moeti local
lemon_sponge_2407vz0r AutoML_f595901b-e4... Completed Jun 14, 2024 2:10 PM 44s Tselane Moeti local
patient_frame_0pr57hfh AutoML_f595901b-e4... Completed Jun 14, 2024 2:11 PM 58s Tselane Moeti local
quirky_rhubarb_k2247ylf AutoML_f595901b-e4... Completed Jun 14, 2024 2:12 PM 43s Tselane Moeti local
orange_rice_91124lbm AutoML_f595901b-e4... Completed Jun 14, 2024 2:12 PM 59s Tselane Moeti local
19°C
Sunny
Search
ENG
US
14:29
2024/06/14




The outputs can be retrieved and reviewed from all the model training runs to identify the best-performing model. The artifacts that can be examined to evaluate model performance include the score files. The outputs can be graphically displayed as well.

## Getting the Run output

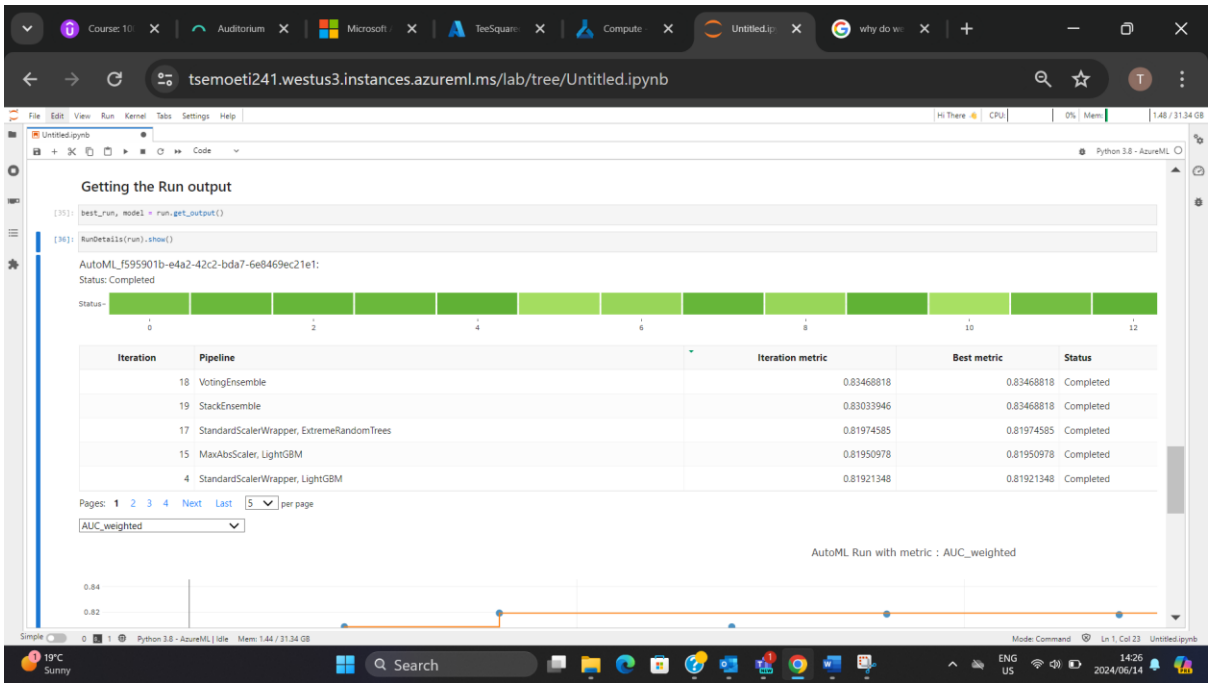
[35]:
```
best_run, model = run.get_output()
```

[36]:
```
RunDetails(run).show()
```

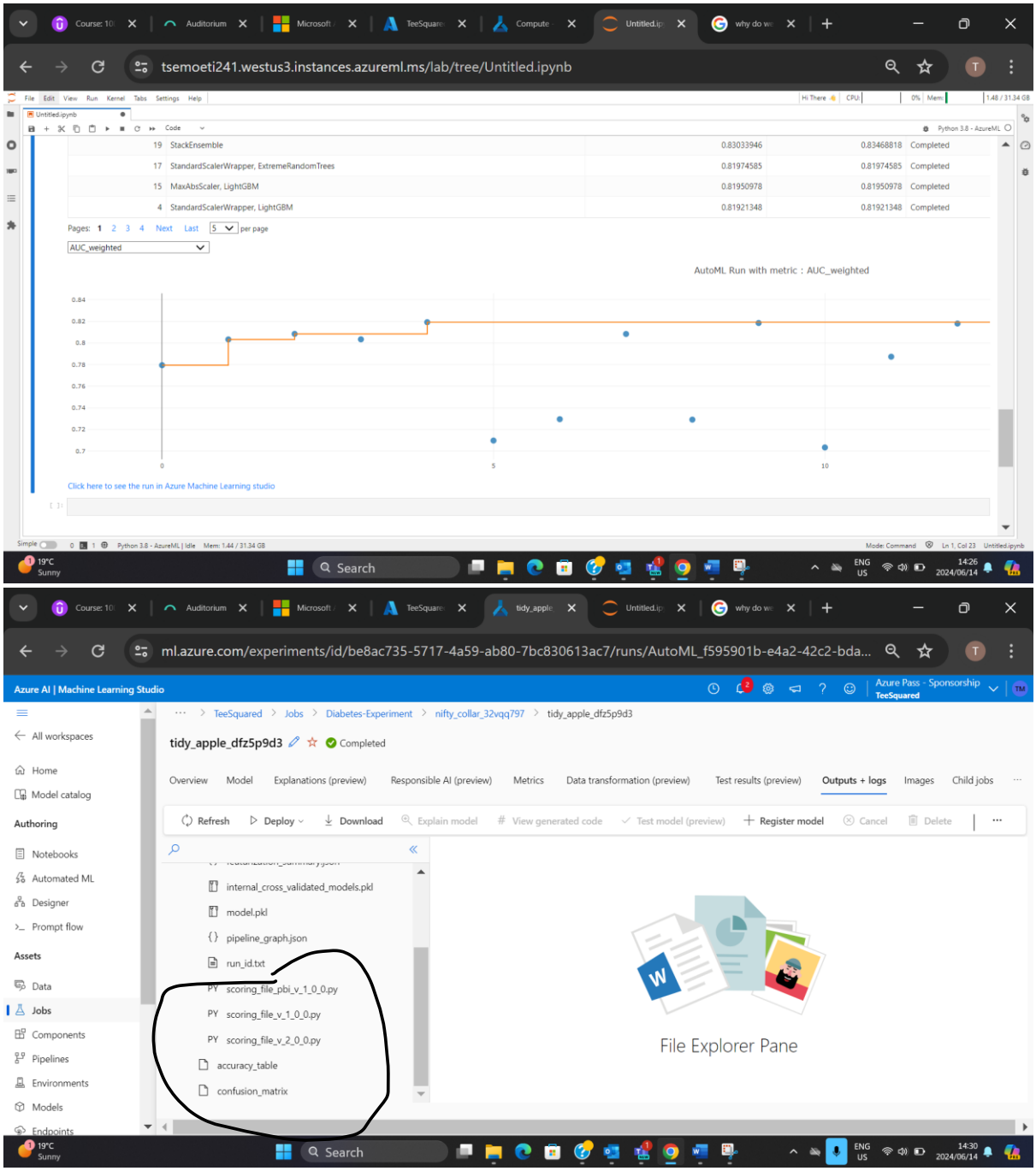
AutoML_f595901b-e4a2-42c2-bda7-6e8469ec21e1:

Status: Completed


| Iteration | Pipeline | Iteration metric | Best metric | Status | Duration | Start |
|---|---|---|---|---|---|---|




| Iteration | Pipeline | Iteration metric | Best metric | Status |
|---|---|---|---|---|
| 18 | VotingEnsemble | 0.83468818 | 0.83468818 | Completed |
| 19 | StackEnsemble | 0.83033946 | 0.83468818 | Completed |
| 17 | StandardScalerWrapper, ExtremeRandomTrees | 0.81974585 | 0.81974585 | Completed |
| 15 | MaxAbsScaler, LightGBM | 0.81950978 | 0.81950978 | Completed |
| 4 | StandardScalerWrapper, LightGBM | 0.81921348 | 0.81921348 | Completed |

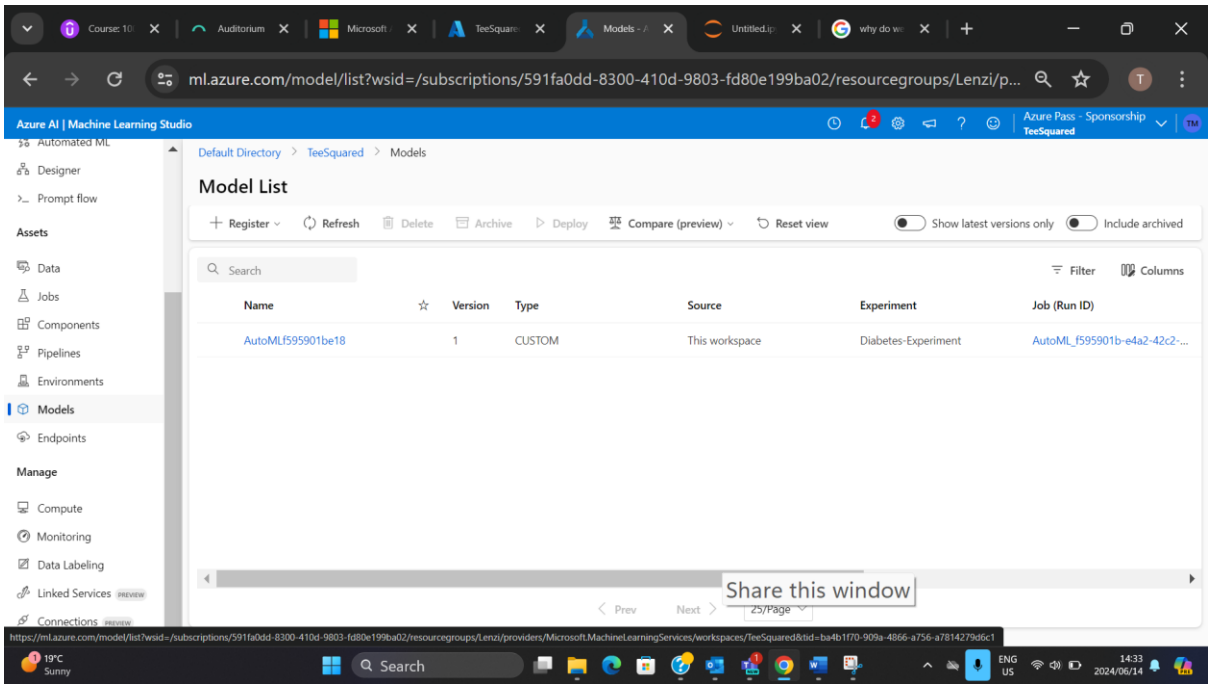



For version control and deployment, the trained model is registered in the Azure Machine Learning Model Registry.

Click here to see the run in Azure Machine Learning studio

### Registering the Model

[37]:
```
model_name = best_run.properties["model_name"]
registered_name = run.register_model(model_name = model_name, description= "AutoML Diabetes",tags = None)
```




## Deployment

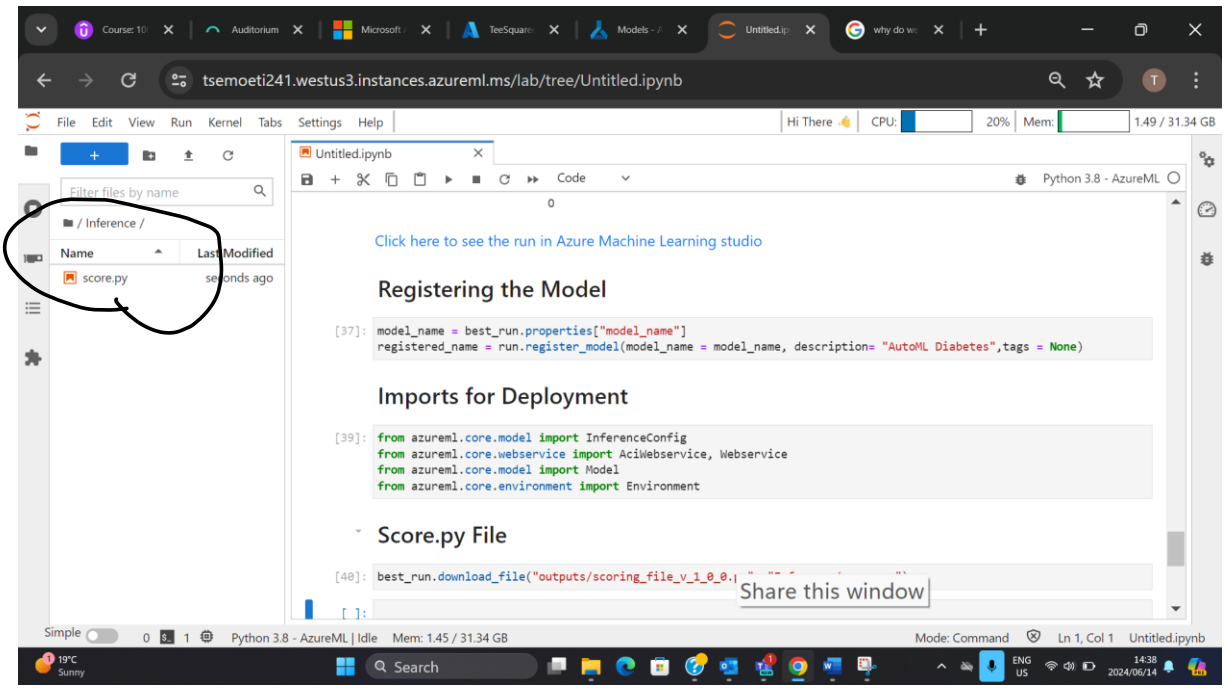
### Imports for Deployment

[39]:
```
from azureml.core.model import InferenceConfig
from azureml.core.webservice import AciWebservice, Webservice
from azureml.core.model import Model
from azureml.core.environment import Environment
```

Creating and downloading the score file which contains the scoring script for handling predictions.

## Score.py File

```
40]: best_run.download_file("outputs/scoring_file_v_1_0_0.py", "Inference/score.py")
```



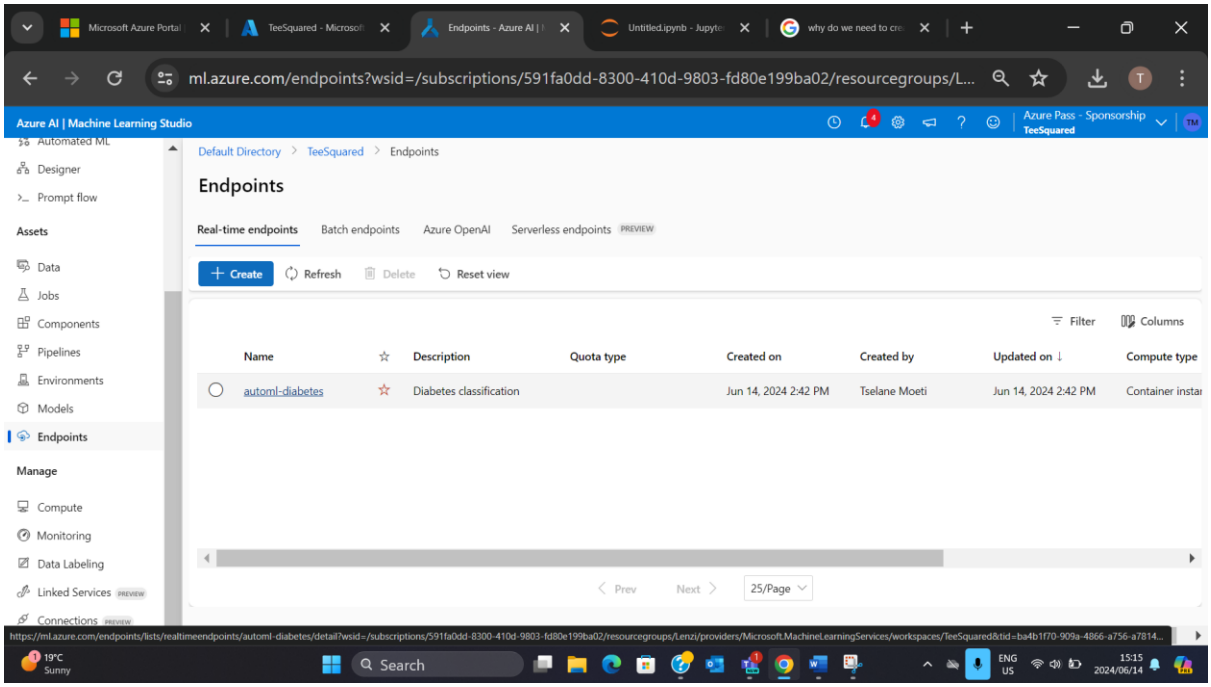

## Deploying

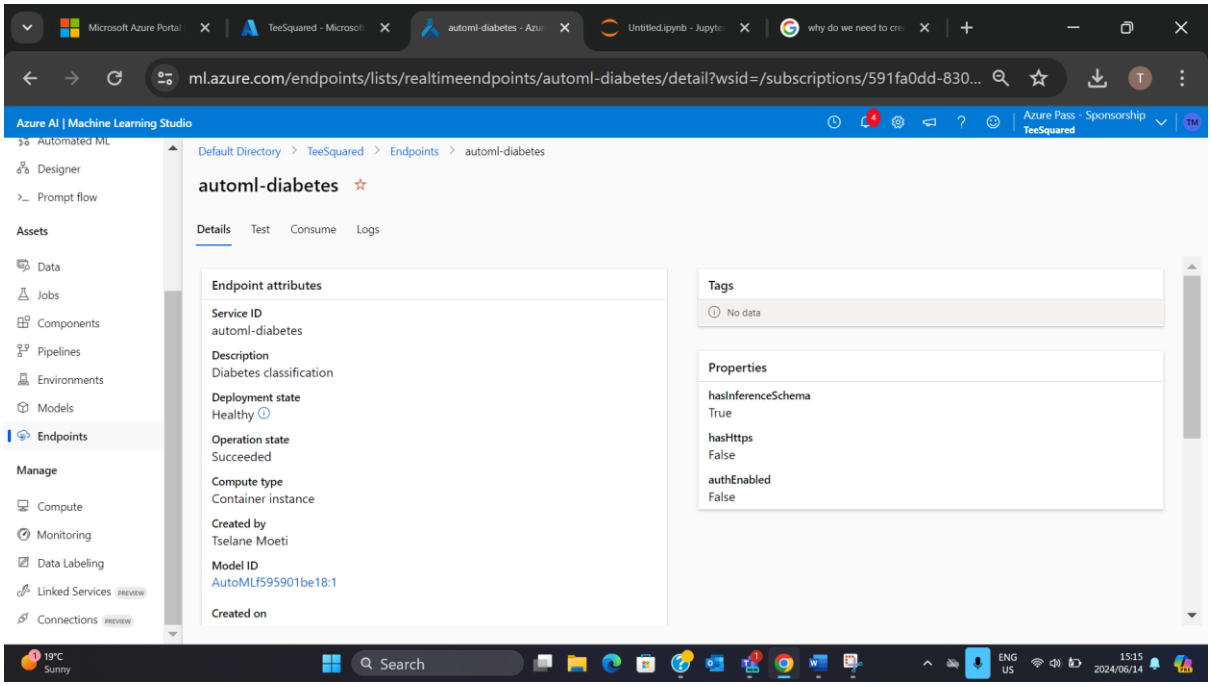
```
[42]: from azureml.automl.core.shared import constants

best_run.download_file(constants.CONDA_ENV_FILE_PATH, "myenv.yml")
env = Environment.from_conda_specification(name="myenv",file_path="myenv.yml")

inference_config = InferenceConfig(entry_script="inference/score.py", environment=env)
aciconfig = AciWebservice.deploy_configuration(cpu_cores=1, memory_gb=1, description="Diabetes classification")
service = Model.deploy(ws, "automl-diabetes", [registered_name], inference_config,aciconfig)
service.wait_for_deployment(True)
```

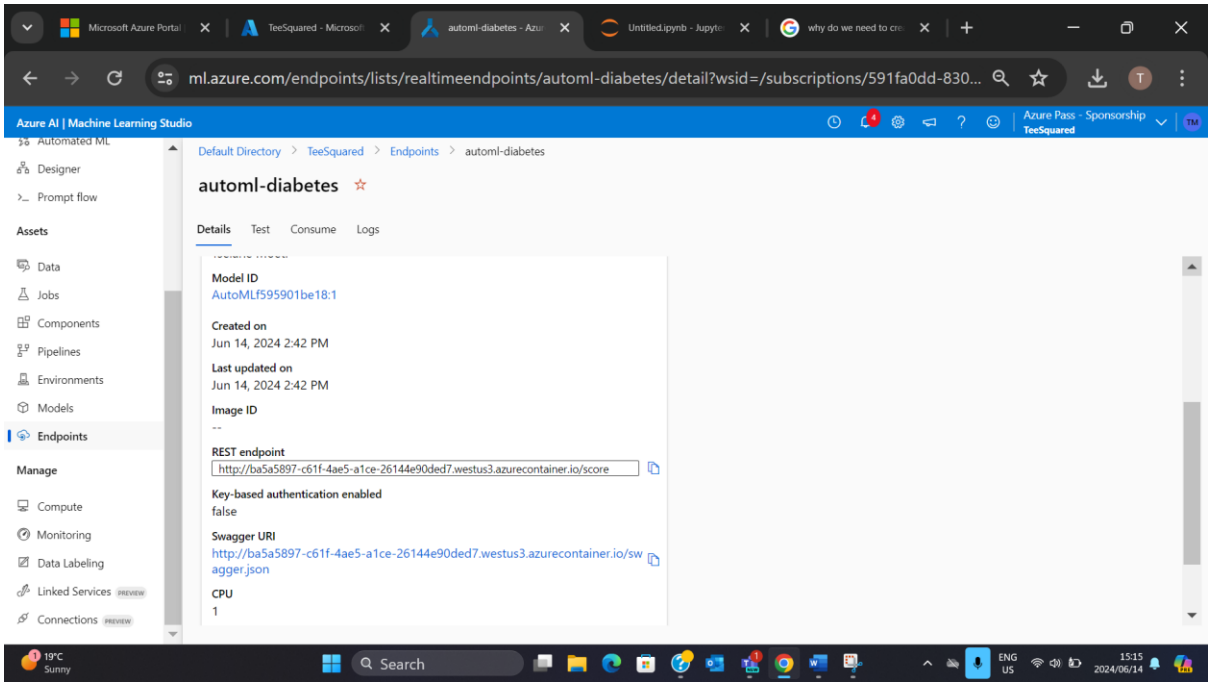
```
Tips: You can try get_logs(): https://aka.ms/debugimage#dockerlog or local deployment: https://aka.ms/debugimage#debug-locally to debug if deployment takes longer than 10 minutes.
Running
2024-06-14 12:42:32+00:00 Creating Container Registry if not exists..
2024-06-14 12:52:33+00:00 Registering the environment..
2024-06-14 12:52:39+00:00 Building image..
2024-06-14 13:11:10+00:00 Generating deployment configuration.
2024-06-14 13:11:11+00:00 Submitting deployment to compute..
2024-06-14 13:11:19+00:00 Checking the status of deployment automl-diabetes..
2024-06-14 13:13:17+00:00 Checking the status of inference endpoint automl-diabetes.
Succeeded
ACI service creation operation finished, operation "Succeeded"
```

```
[ ]:
```

Establishing endpoints for real-time predictions.





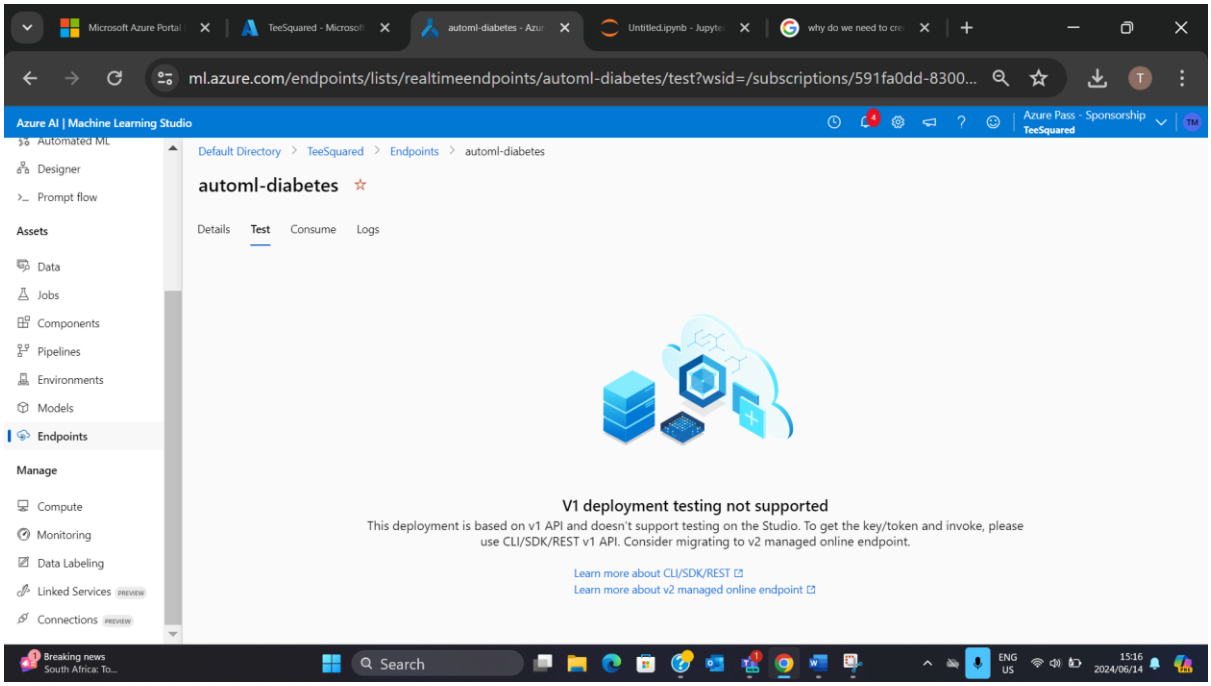




## The Deployment URL:

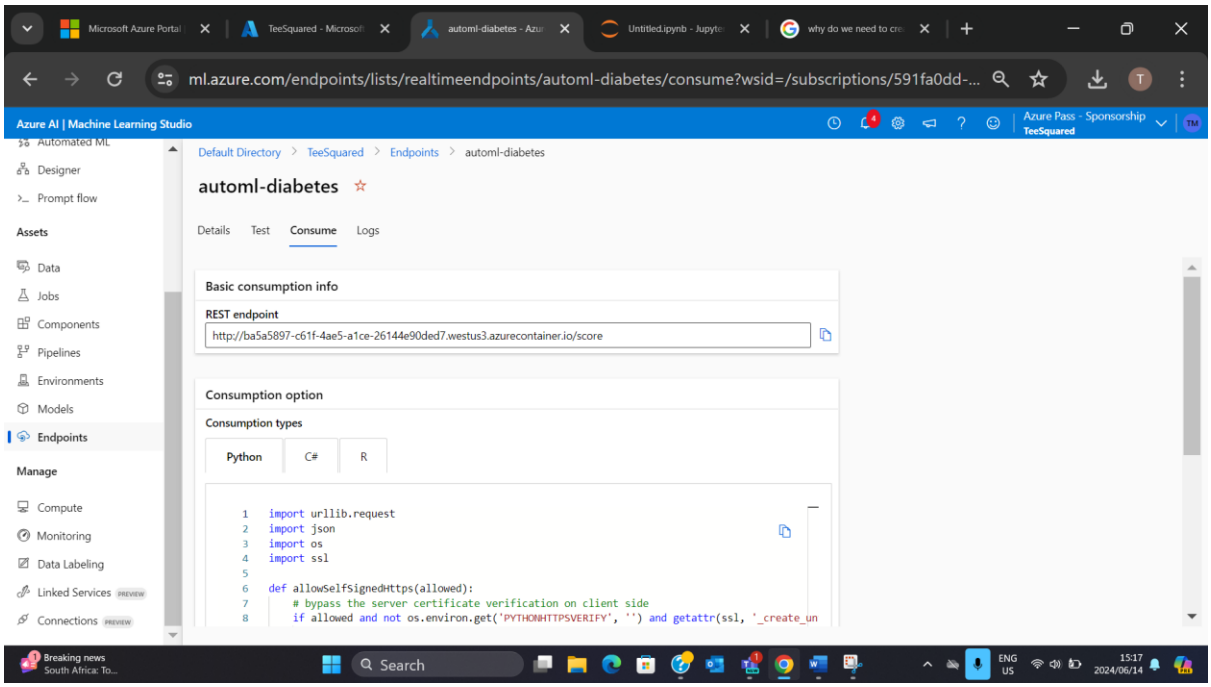
http://ba5a5897-c61f-4ae5-a1ce-26144e90ded7.westus3.azurecontainer.io/score

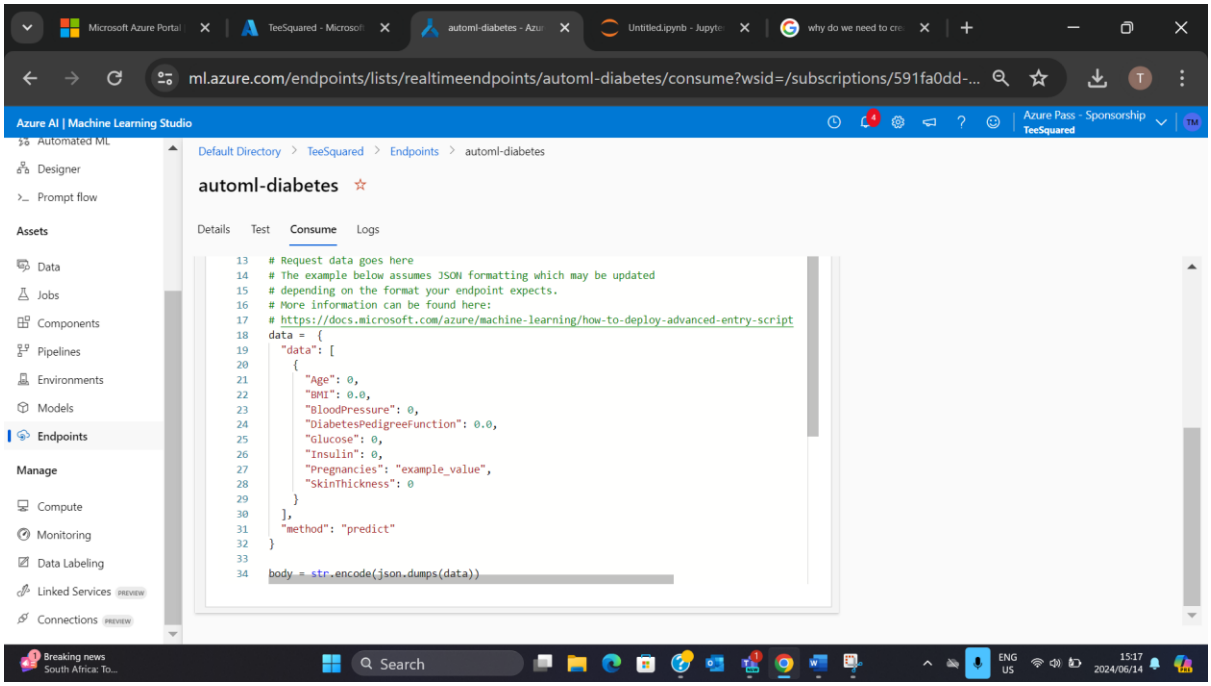
## Testing and Consuming

The model can be tested and




The model can be tested using different codes in the endpoint>consume tab as shown in the snippet below.




Microsoft Azure Portal
TeeSquared - Microsoft
automl-diabetes - Azure
Untitled.ipynb - Jupyter
why do we need to cre
ml.azure.com/endpoints/lists/realtimeendpoints/automl-diabetes/consume?wsid=/subscriptions/591fa0dd-...
Azure AI | Machine Learning Studio
Azure Pass - Sponsorship
TeeSquared
Automated ML
Designer
Prompt flow
Assets
Data
Jobs
Components
Pipelines
Environments
Models
Endpoints
Manage
Compute
Monitoring
Data Labeling
Linked Services
Connections
Default Directory > TeeSquared > Endpoints > automl-diabetes
automl-diabetes
Details
Test
Consume
Logs
# Request data goes here
# The example below assumes JSON formatting which may be updated
# depending on the format your endpoint expects.
# More information can be found here:
# https://docs.microsoft.com/azure/machine-learning/how-to-deploy-advanced-entry-script
data = {
"data": [
{
"Age": 0,
"BMI": 0.0,
"BloodPressure": 0,
"DiabetesPedigreeFunction": 0.0,
"Glucose": 0,
"Insulin": 0,
"Pregnancies": "example_value",
"SkinThickness": 0
}
],
"method": "predict"
}
body = str.encode(json.dumps(data))
Breaking news
South Africa: To...
Search
ENG
US
15:17
2024/06/14