

Raport Lab: Azure ML & DevOps & Mlflow

Mouad GUEDAD

Year : 2023/2024

Major : QFM

Problem:

You work as a data scientist in a small team with three other data scientists for a shipping company based in the port of Turku in Finland. 90% of goods imported into Finland pass through cargo ships in the country's ports. For the transport of goods, weather conditions and logistics can sometimes be difficult at ports. Rainy conditions can distort operations and logistics at ports, which can affect supply chain operations. Predicting rainy conditions in advance helps optimize resources such as human, logistics and transportation for efficient supply chain operations at ports. From a business perspective, forecasting rainy conditions in advance allows ports to reduce operating costs by up to approximately 20% by enabling efficient planning and scheduling of human resources, logistics and shipping resources. transportation for supply chain operations.

Task:

As a data scientist, you are responsible for developing an ML-based solution to predict weather conditions 4 hours in advance at the port of Turku in Finland. This will allow the port to optimize its resources, resulting in savings of up to 20%. To start, you have a set of historical weather data covering a period of 10 years from the port of Turku. Your task is to create an ML solution focused on continuous learning to optimize operations at the Port of Turku

1. Setting up resources and tools

1.1 MLflow

After installing MLflow in my local and I checked I found it is already in my local .

1.2 Azure Machine learning

Microsoft Azure

Upgrade

Search resources, services, and docs (G+)

Mouad.GUEDAD@um6...
UNIVERSITÉ MOHAMMED VI POL...

Create a resource

Subscriptions

SQL databases

Azure DevOps organizations

Resource groups

All resources

Quickstart Center

Virtual machines




App Services

More services

Resources

Recent

Favorite

Name	Type	Last Viewed
 Azure subscription 1	Subscription	2 hours ago
 LearnMLOps3	Resource group	2 hours ago
 MLOps_WS	Azure Machine Learning workspace	2 hours ago

See all

1.3 Azure DevOps

Créer l'organisation

Projets

Actions

Nouveau projet

After I follow the instruction you gave us in the second steep (data preprocessing) I get :

```
C:\Users\ASUS_ZENBOOK>
C:\Users\ASUS_ZENBOOK>git clone https://MouadGUEDAD@dev.azure.com/MouadGUEDAD/Learn_MLOps/_git/Learn_MLOps
Cloning into 'Learn_MLOps'...
remote: Azure Repos
remote: Found 20 objects to send. (14 ms)
Unpacking objects: 100% (20/20), 2.17 MiB | 1.84 MiB/s, done.
C:\Users\ASUS_ZENBOOK>
```

Then after the clonage the the dataset saved on the workspace is :

Azure AI | Machine Learning Studio

Model catalogPREVIEW

Authoring

Notebooks

Automated ML

Designer

Prompt flowPREVIEW

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

...

Université Mohammed VI Polytechnique > MLOps_WS > Data > processed_weather_data_portofTurku

processed_weather_data_portofTurku

Version: 1 (latest) ☆

DetailsConsumeExploreModelsJobs

Profile: This is the quick profile generated by the top 10,000 rows. Please generate profile to view the schema and summary statistics for full data.

RefreshGenerate profile

PreviewProfile

Number of columns: 10Number of rows: First 50

Timestamp	Location	Temperat...	Humidity	Wind_spe...	Wind_bea...	Visibility_...	Pressure_...	Current_w...	Future_w...
2006-04-0...	Port of Tur...	8.756	0.83	11.045	259	15.826	1,016.51	1	1
2006-04-0...	Port of Tur...	9.222	0.85	13.959	258	14.957	1,016.66	1	1
2006-04-0...	Port of Tur...	7.733	0.95	12.365	259	9.982	1,016.72	1	1
2006-04-0...	Port of Tur...	8.772	0.89	14.152	260	9.982	1,016.84	1	1
2006-04-0...	Port of Tur...	10.822	0.82	11.318	259	9.982	1,017.37	1	1
2006-04-0...	Port of Tur...	12.772	0.72	12.526	270	9.982	1,017.22	1	1

3. Pipeline construction

I processed the data on my local computer. For training ML models and implementing the pipeline, I use computing resources provisioned in the cloud (Microsoft Azure). In the following, I will create an Azure compute instance. now my compute instance is ready for cloud computing

Azure AI | Machine Learning Studio

Prompt flowPREVIEW

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Manage

Compute

MonitoringPREVIEW

Data Labeling

Linked Services

Université Mohammed VI Polytechnique > MLOps_WS > Compute

Compute

The "Kubernetes clusters" tab is now where you can access previous versions of "inference clusters" (also known as "AKS clusters") and "attached Kubernetes" compute types along with any previously created compute targets using those types. Learn more about Kubernetes clusters.

Compute instancesCompute clustersKubernetes clustersAttached computes

Choose from a selection of CPU or GPU instances preconfigured with popular tools such as VS Code, JupyterLab, Jupyter, and RStudio, ML packages, deep learning frameworks, and GPU drivers. Learn more about compute instances

+ NewRefreshStartStopRestartSchedule and idle shutdownDeleteView quota

SearchFilterColumns

Name	☆	State	Idle shutdown	Applications
MLOPS		Running	1 hour	JupyterLab Jupyter VS Code (Web)PREVIEW

Once it is provisioned, I selected the JupyterLab option. JupyterLab is an open source web user interface. It comes with features like a text editor, code editor, terminal, and custom components built in an extensible manner. I used it as a programming interface connected to the provisioned compute to train the ML models.

After I cloned the respository to the Azure instance.

The steps can be summarized as follows:

- a. We must connect to the workspace and import the artifacts

We import the required packages, connect to the ML workspace using the `Workspace()` function, and then download the serialized scaler and model to make predictions. Scaler will be used to scale the input data into the same data scale used for training the model. The model file is serialized in ONNX format. Scaler and Model files are imported using the `Model()` function.

- b. We load the artifacts for inference

We open and load the Scaler and Model files into variables that can be used for model inference. The scaler is read and loaded into a variable using `pickle`, and the ONNX runtime is used to load the ONNX file using `InferenceSession()` to make predictions

I followed exactly the instructions in the “Model_evaluation_packaging” folder and after I run the netobook I get :

