

MLOps In a Day

M02 Use Case

Learning Units covered in this Module



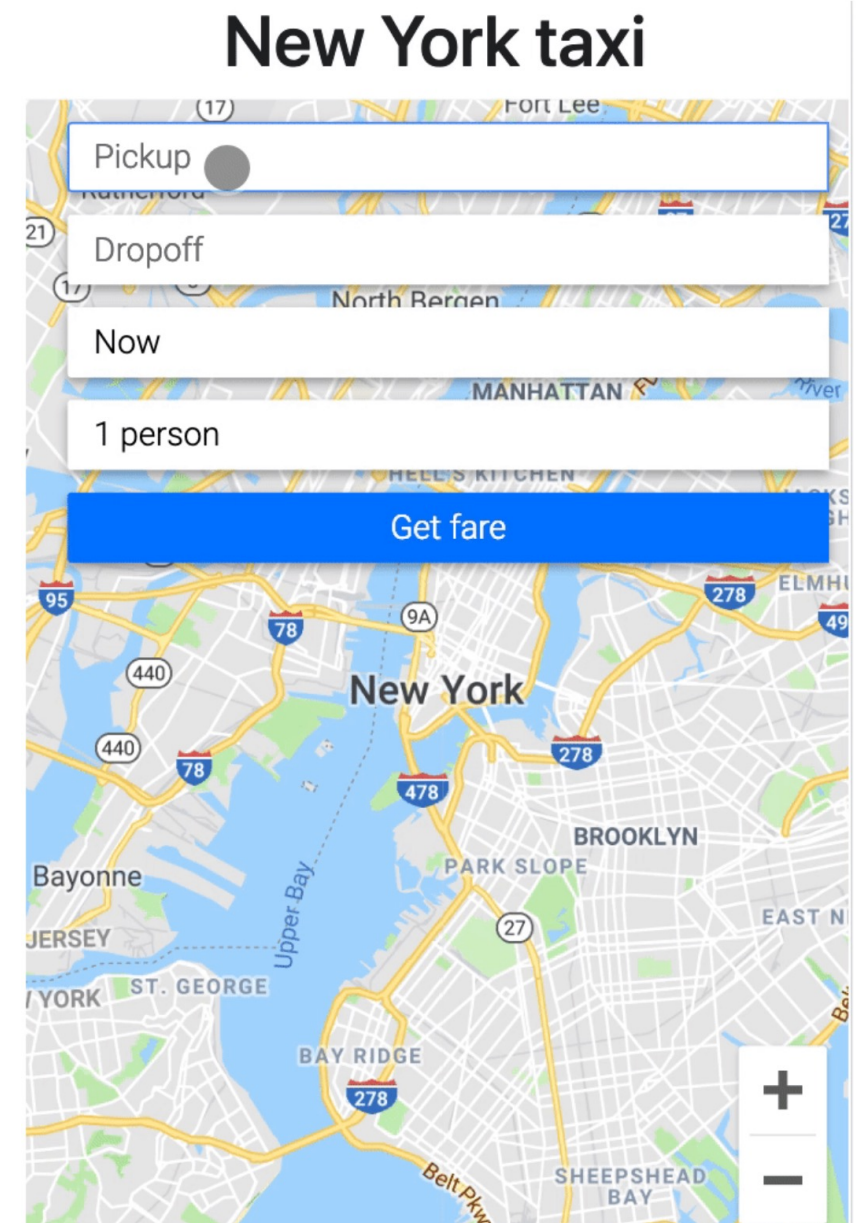
MLOps Use Case



Use Case

Use Case

- Predict taxi fare in New York City
- Tabular data
- Regression model
- Random forest scikit-learn model
- Train model using Azure Machine Learning
- Deploy model using Azure Machine Learning Managed Endpoints
- mlflow for tracking & model management
- Azure Devops for infra setup and CI/CD
- Demo Repository:
 - <https://github.com/Azure/mlops-v2-workshop>



MLOps Journey

Develop

Package

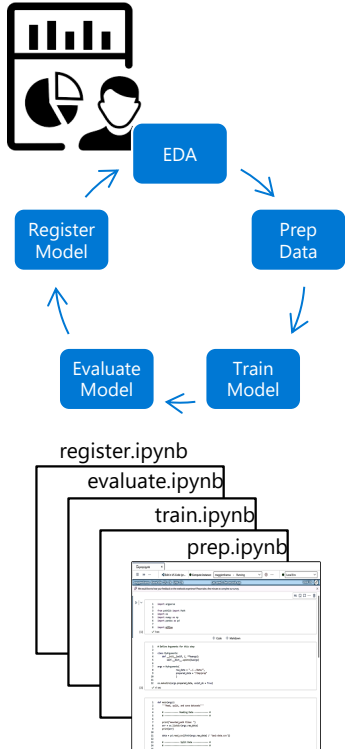
Orchestrate

Deploy & Serve


Monitor


Automate

Experimentation




Azure Machine Learning

 Compute

 Notebooks

MLOps Journey



Experimentation

Training Code

✓ components

> evaluate

> prep

> register

> train

evaluate.yml

prep.yml

register.yml

train.yml

MLOps Journey

Develop

Package

Orchestrate

Deploy & Serve

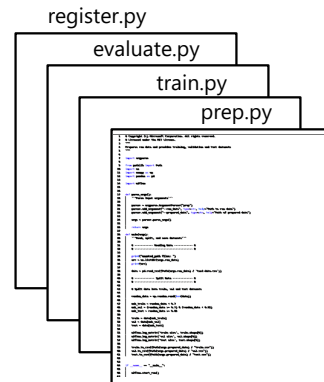
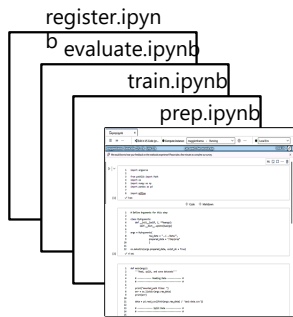
Monitor

Automate

Training Code

Training Code

1. Refactor code & package into scripts
 - input & output arguments
2. Prepare required Environment



train-conda.yml

```
1 channels:
2   - defaults
3   - anaconda
4   - conda-forge
5 dependencies:
6   - python=3.7.5
7   - pip
8   - pip:
9     - azureml-mlflow==1.38.0
10    - azureml-sdk==1.38.0
11    - scikit-learn==0.24.1
12    - pandas==1.2.1
13    - joblib==1.0.0
14    - matplotlib==3.3.3
```



Azure Machine Learning

















Compute



Notebooks

MLOps Journey



- >  ci-cd
- ✓  components
 - >  evaluate
 - >  prep
 - >  register
 - >  train
 -  evaluate.yml
 -  prep.yml
 -  register.yml
 -  train.yml
- >  data
- ✓  environment
 -  train-conda.yml
 -  train-requirements.txt

MLOps Journey

Develop

Package

Orchestrate

Deploy & Serve

Monitor

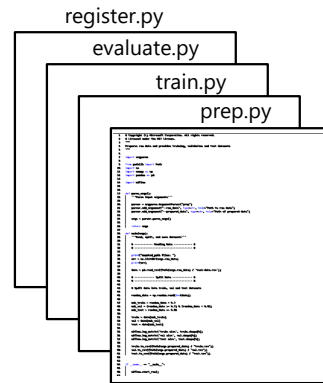
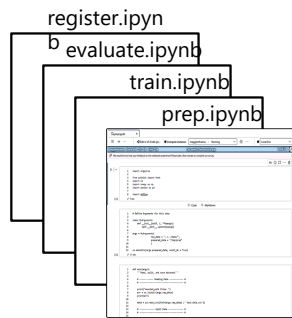
Automate

Training Code

Training Code

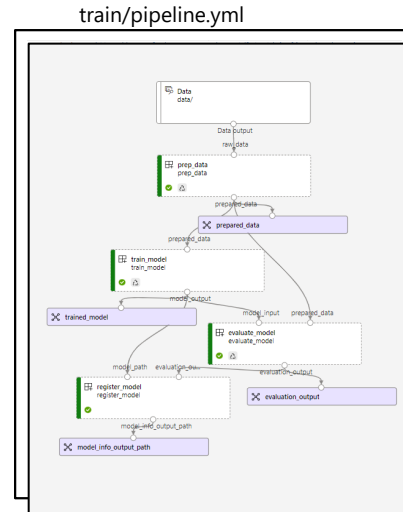
Training Code

```
az ml compute create -f compute.yml
az ml environment create -f environment.yml
az ml job create -f pipeline.yml
```

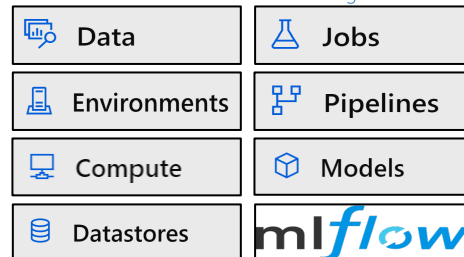


train-conda.yml

```
1 channels:
2   - defaults
3   - anaconda
4   - conda-forge
5 dependencies:
6   - python=3.7.5
7   - pip
8   - pip:
9     - azureml-mlflow==1.38.0
10    - azureml-sdk==1.38.0
11    - scikit-learn==0.24.1
12    - pandas==1.2.1
13    - joblib==1.0.0
14    - matplotlib==3.3.3
```



Azure Machine Learning



MLOps Journey



- > ci-cd
- ✓ components
 - > evaluate
 - > prep
 - > register
 - > train
 - evaluate.yml
 - prep.yml
 - register.yml
 - train.yml
- > data
- ✓ environment
 - train-conda.yml
 - train-requirements.txt

MLOps Journey

Develop

Package

Orchestrate

Deploy & Serve

Monitor

Automate

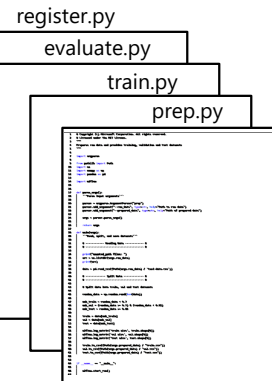
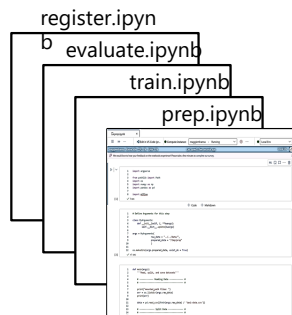
Training Code

Training Code

Training Code

Model

```
# create online endpoint & deployment
# create batch endpoint & deployment
az ml batch-endpoint create -f batch-endpoint.yml
az ml batch-deployment create -f batch-deployment.yml
```



train-conda.yml

```
1 channels:
2   - defaults
3   - anaconda
4   - conda-forge
5 dependencies:
6   - python=3.7.5
7   - pip
8   - pip:
9     - azureml-mlflow==1.38.0
10    - azureml-sdk==1.38.0
11    - scikit-learn==0.24.1
12    - pandas==1.2.1
13    - joblib==1.0.0
14    - matplotlib==3.3.3
```

train/pipeline.yml

```
1 schema: https://azuremlschemas.azureedge.net/latest/pipeline.schema.json
2 type: pipeline
3 experiment_name: taxi-fare-training
4 description: training pipeline to train a model that predicts taxi fare price
5
6 # inputs_and_outputs
7 inputs:
8   input: using local data, will create an anonymous data asset
9   type: uri_folder
10  path: ../././data/
11
12 # outputs
13 prepared_data:
14   training_model:
15     evaluation_output:
16       model_info_output_path:
17 # </inputs_and_outputs>
18
19 # jobs
20 settings:
21   default_datastore: azureml_workspaceblobstore
22   default_compute: azureml-rgs-cluster
23   continue_on_step_failure: false
24
25 jobs:
26   prep_data:
27     name: prep_data
28     display_name: prep_data
29     code: ../././data-science/src/prepare
30     command: >
31       python prep.py
32       --raw_data $(inputs.raw_data)
33       --prepared_data $(outputs.prepared_data)
34     environment: azureml-taxi-train-env@latest
35     inputs:
36       raw_data: $(parent.inputs.input)
37     outputs:
38       prepared_data: $(parent.outputs.prepared_data)
39
40 train_model:
41   name: train_model
42   display_name: train_model
43   code: ../././data-science/src/train
44   command: >
45     python train.py
46     --prepared_data $(inputs.prepared_data)
```

train-env.yml

```
1 schema: https://azuremlschemas.azureedge.net/
2 name: taxi-train-env
3 image: mcr.microsoft.com/azureml/openmpi3.1.2-1
4 conda_file: ../././data-science/environment/t
5 description: Environment created from a Docker
```

deploy/online/endpoint.yml

```
1 schema: https://azuremlschemas.azureedge.net/latest/
2 name: blue
3 endpoint_name: taxi-fare-online
4 model: azureml-taxi-model@latest
5 instance_type: Standard_F2s_v2
```

deploy/batch/endpoint.yml

```
1 schema: https://azuremlschemas.azureedge.
2 name: batch-gp
3 endpoint_name: taxi-fare-batch
4 model: azureml-taxi-model@latest
5 computer: azureml-batch-cluster
6 resources:
7   instance_count: 1
8   max_concurrency_per_instance: 2
9   min_batch_size: 10
10  output_action: append_row
11  output_file_name: predictions.csv
12  retry_settings:
13    max_retries: 3
14    timeout: 30
15  error_threshold: -1
16  logging_level: info
```
















Azure Machine Learning



MLOps Journey



- ✓  ml-pipelines
 - ✓  deploy
 - ✓  batch
 -  batch-deployment.yml
 -  batch-endpoint.yml
 - ✓  online
 -  online-deployment.yml
 -  online-endpoint.yml
 - >  train
 -  azureml-cliv2.ipynb
 -  azureml-sdkv2.ipynb
 -  azureml.sh
 -  train-sdkv2.ipynb

MLOps Journey

Develop

Package

Orchestrate

Deploy & Serve

Monitor

Automate

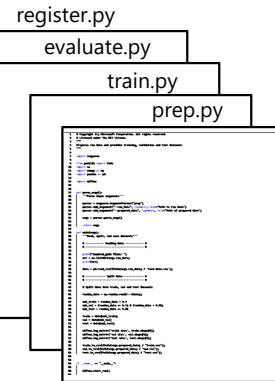
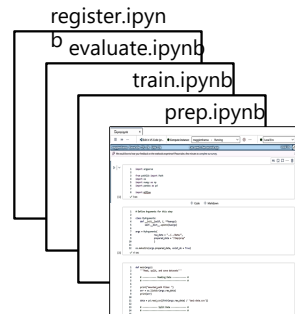
Training Code

Training Code

Training Code

Model

Training & Deployment



train-conda.yml

```
1 channels:
2   - defaults
3   - anaconda
4   - conda-forge
5 dependencies:
6   - python=3.7.5
7   - pip
8   - pip:
9     - azureml-mlflow==1.38.0
10    - azureml-sdk==1.38.0
11    - scikit-learn==0.24.1
12    - pandas==1.2.1
13    - joblib==1.0.0
14    - matplotlib==3.3.3
```

train/pipeline.yml

```
1 $schema: https://azuremlschemas.azureedge.net/latest/pipeline.schema.json
2 type: pipeline
3 experiment_name: taxi-fare-training
4 description: Training Pipeline to train a model that predicts taxi fare price
5
6 # inputs_and_outputs
7 inputs:
8   # Input using local data, will create an anonymous data asset
9   # type: uri_folder
10  path: ../data/
11
12 outputs:
13   prepared_data:
14     trained_model:
15     evaluation_output:
16     model_info_output_path:
17 # # inputs_and_outputs
18
19 # jobs
20 settings:
21   default_datastore: azureml:workspaceblobstore
22   default_compute: azureml:cpu-cluster
23   continue_on_step_failure: false
24
25 jobs:
26   prep_data:
27     name: prep_data
28     display_name: prep_data
29     code: ../data-science/src/prepare
30     command: >
31       python prep.py
32       --raw_data $(inputs.raw_data)
33       --prepared_data $(outputs.prepared_data)
34     environment: azureml:taxi-train-env@latest
35     inputs:
36       raw_data: $(parent.inputs.raw_data)
37     outputs:
38       prepared_data: $(parent.outputs.prepared_data)
39
40   train_model:
41     name: train_model
42     display_name: train_model
43     code: ../data-science/src/train
44     command: >
45       python train.py
46       --prepared_data $(inputs.prepared_data)
```

train-env.yml

```
1 $schema: https://azuremlschemas.azureedge.net/
2 name: taxi-train-env
3 image: mcr.microsoft.com/azureml/openmpi3.1.2-ubuntu18.04
4 conda_file: ../data-science/environment/t
5 description: Environment created from a Docker
```

```
az ml compute create -f compute.yml
az ml environment create -f environment.yml
az ml job create -f pipeline.yml
```

deploy/online/endpoint.yml

```
$schema: https://azuremlschemas.azureedge.net/latest/
name: taxi-fare-online
endpoint_name: taxi-fare-online
model: azureml:taxi-model@latest
instance_type: Standard_F4s_v2
```

deploy/batch/endpoint.yml

```
$schema: https://azuremlschemas.azureedge.net/latest/
name: taxi-fare-batch
endpoint_name: taxi-fare-batch
model: azureml:taxi-model@latest
compute: azureml:batch-cluster
resources:
  instance_count: 1
  max_concurrency_per_instance: 2
  min_batch_size: 10
  output_action: append_row
  output_file_name: predictions.csv
  retry_settings:
    max_retries: 3
    timeout: 30
  error_threshold: -1
  logging_level: info
```

```
# create online endpoint & deployment
az ml online-endpoint create -f online-endpoint.yml
# create batch endpoint & deployment
az ml batch-endpoint create -f batch-endpoint.yml
az ml batch-deployment create -f batch-deployment.yml
```

unit-tests.yml
deploy-model-training-pipeline.yml

deploy-batch-endpoint-pipeline.yml
deploy-online-endpoint-pipeline.yml

```
# Copyright (c) Microsoft Corporation. All rights reserved.
# Licensed under the MIT License.

variables:
- ${{ if eq(variables['Build.SourceBranchName'], 'main') }}:
  # 'main' branch: PROD environment
  - template: ../config/infra-prod.yml
- ${{ if ne(variables['Build.SourceBranchName'], 'main') }}:
  # 'develop' or feature branches: DEV environment
  - template: ../config/infra-dev.yml

trigger:
- none

pool:
  vmImage: $(ap_vm_image)

resources:
  repositories:
  - repository: mlops-templates
    name: Azure/mlops-templates
    endpoint: github-connection
```



MLOps Journey



- ✓ mlops-v2-training
 - > ci-cd
 - > components
 - > data
 - > environment
 - > imgs
 - ✓ ml-pipelines
 - ✓ deploy
 - ✓ batch
 - batch-deployment.yml
 - batch-endpoint.yml
 - online
 - online-deployment.yml
 - online-endpoint.yml
 - > train
 - azureml-cliv2.ipynb
 - azureml-sdkv2.ipynb
 - azureml.sh
 - train-sdkv2.ipynb