"Houston, we have a model..."

Introduction to MLOps

Rui Quintino

AI & Analytics Strategy Advisor
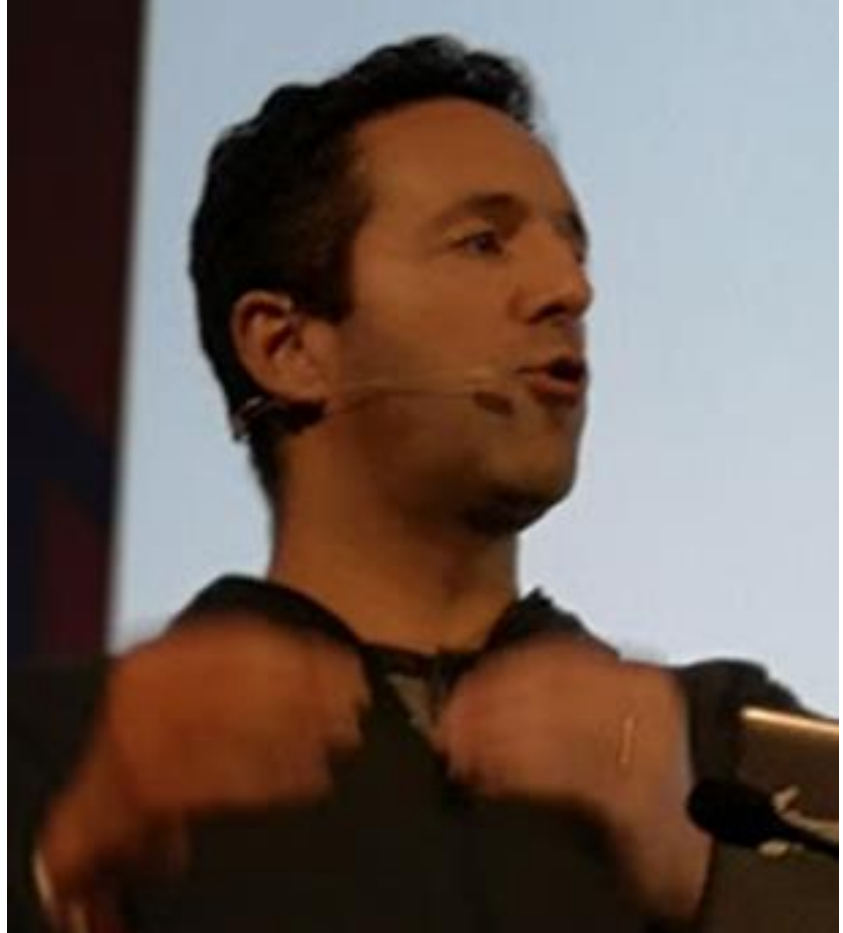**DevScope**

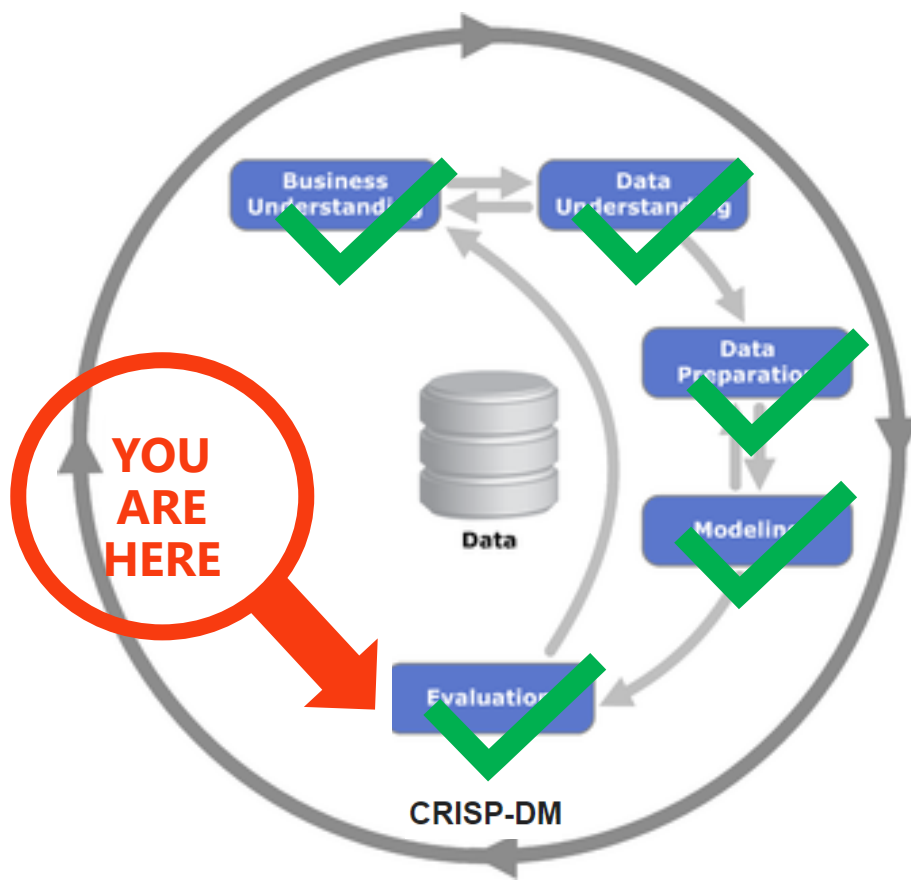**20210219**

dev>scope

AI & Analytics

Strategy Advisor

rui.quintino@devscope.net
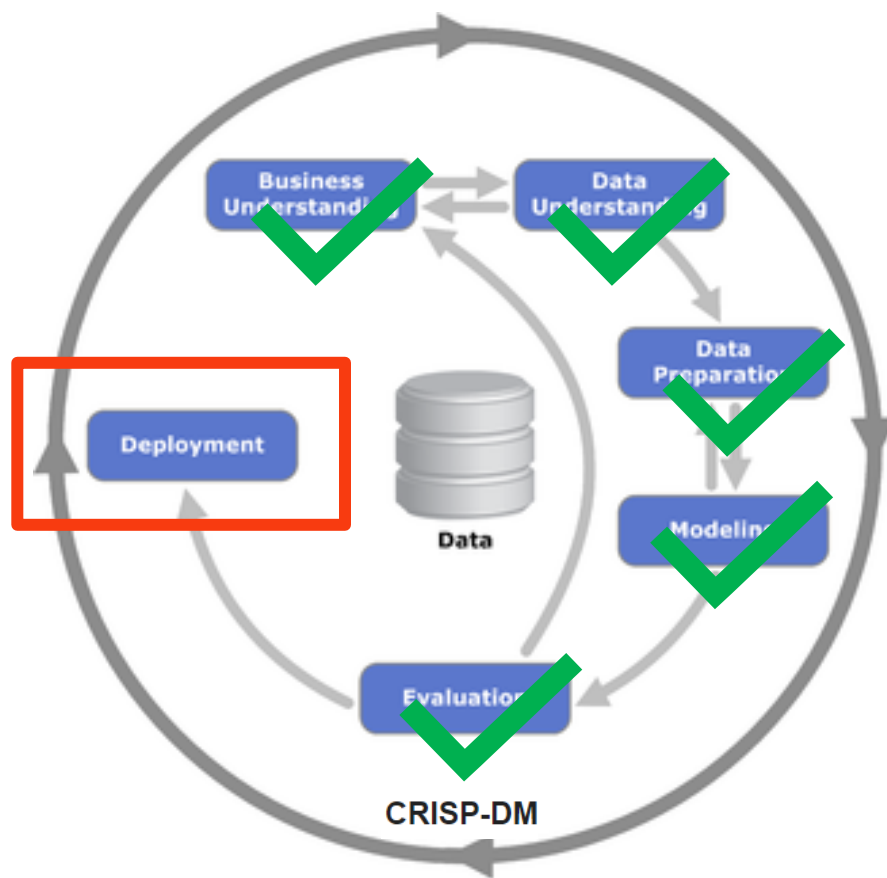
Linkedin.com/ in/rquintino
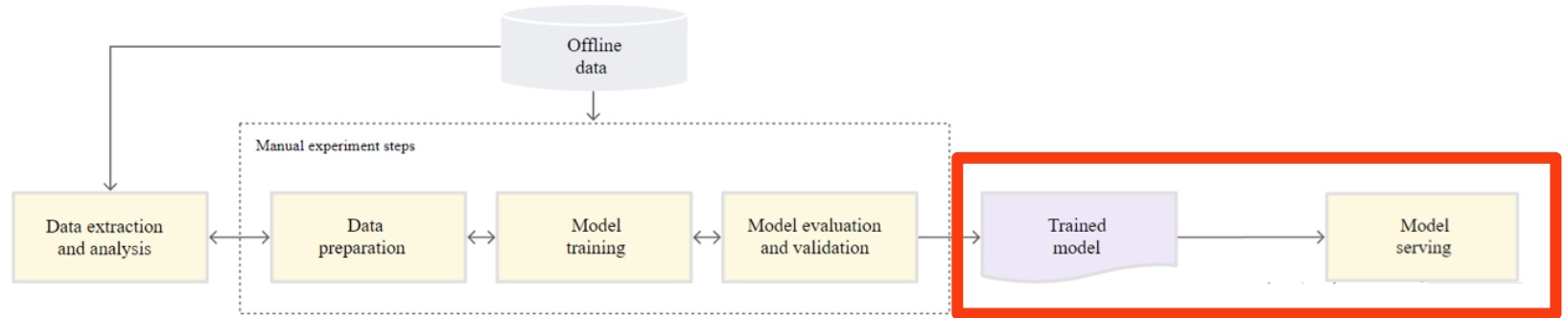
CRISP-DM

We have a model... It's great!

Now what?

CRISP-DM

We just deploy it!

```
model_file_name = "model.pkl"
joblib.dump(svm_model, os.path.join('outputs', model_file_name))
```


Adapted from: https://cloud.google.com/solutions/machine-learning/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

Easy right?

File   Edit   View   Insert   Cell   Kernel   Navigate   Widgets   Help

Code   ▾   Notify: 5 ▾

```python
def stack_images(row):
    row_name=fill_nas(row,"file_name","")
    row_time=fill_nas(row,"diff_time",0).round(1)

    img1=read_img(row_name["file_name"],row.get("rotate"),row.get("crop"))
    img2=read_img(row_name["file_name_prev2"],row.get("rotate"),row.get("crop"))
    img3=read_img(row_name["file_name_prev4"],row.get("rotate"),row.get("crop"))

    array=(np.moveaxis(np.stack([img1,img2,img3]),0,-1))

    # Inject two small squares to encode ocean and relative time info
    # Use max time of ~20 days

    max_time=20*3600*24
    time_marker=np.min([255,150+((row["relative_time"]/max_time)*255).astype("uint8")])
    array[121:140,121:140,2]=time_marker
    array[141:160,141:160,1]=row["ocean"]*50

    return pil_image.fromarray(array)
```

executed in 344ms, finished 14:20:49 2021-02-18
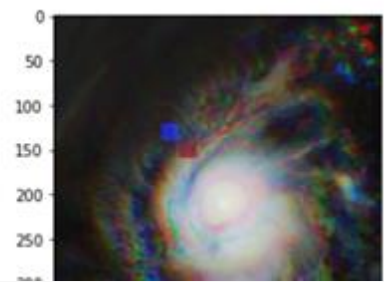
What about the data Pipeline/preproc?

Drift overview

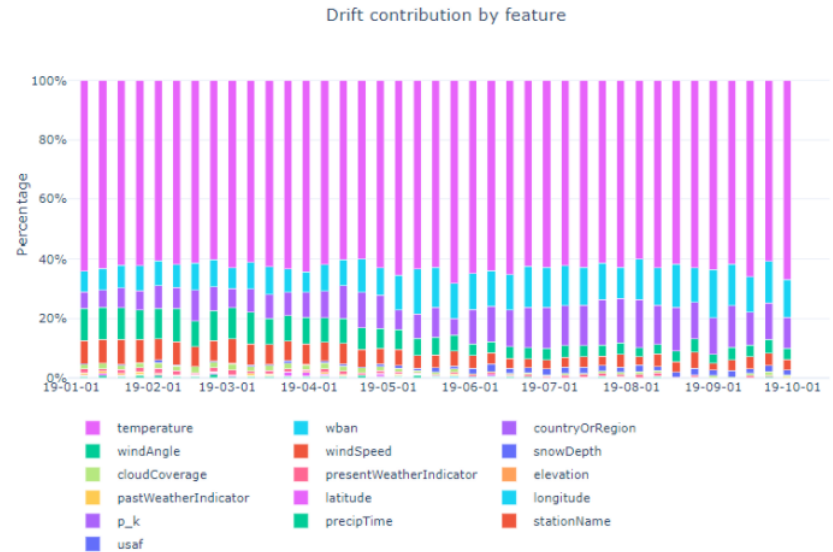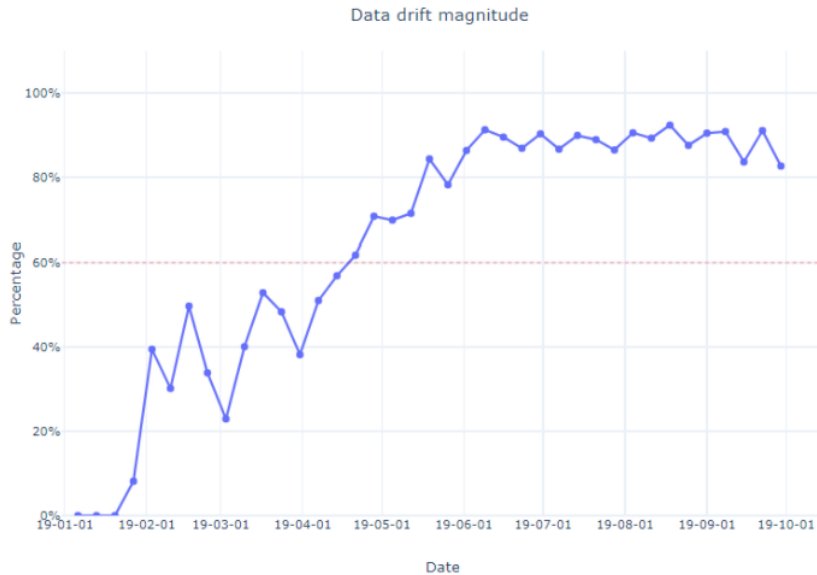https://github.com/Azure/MachineLearningNotebooks/blob/master/how-to-use-azureml/work-with-data/datadrift-tutorial/datadrift-tutorial.ipynb

Or Data Drift -> causing Model Performance degradation?

Plus...

- How/who can we retrain if needed (ex: data drift)?
- What are the software requirements?
- How can we later update lib/os versions? (ex: security)
- How can we monitor model performance in production?
- How do we ensure data checks in production?
- How was the model trained? What was the performance?
- How to scale? And Reproduce?
- ...

dev>scope

ML
code

Model & ML code just tip of the iceberg...
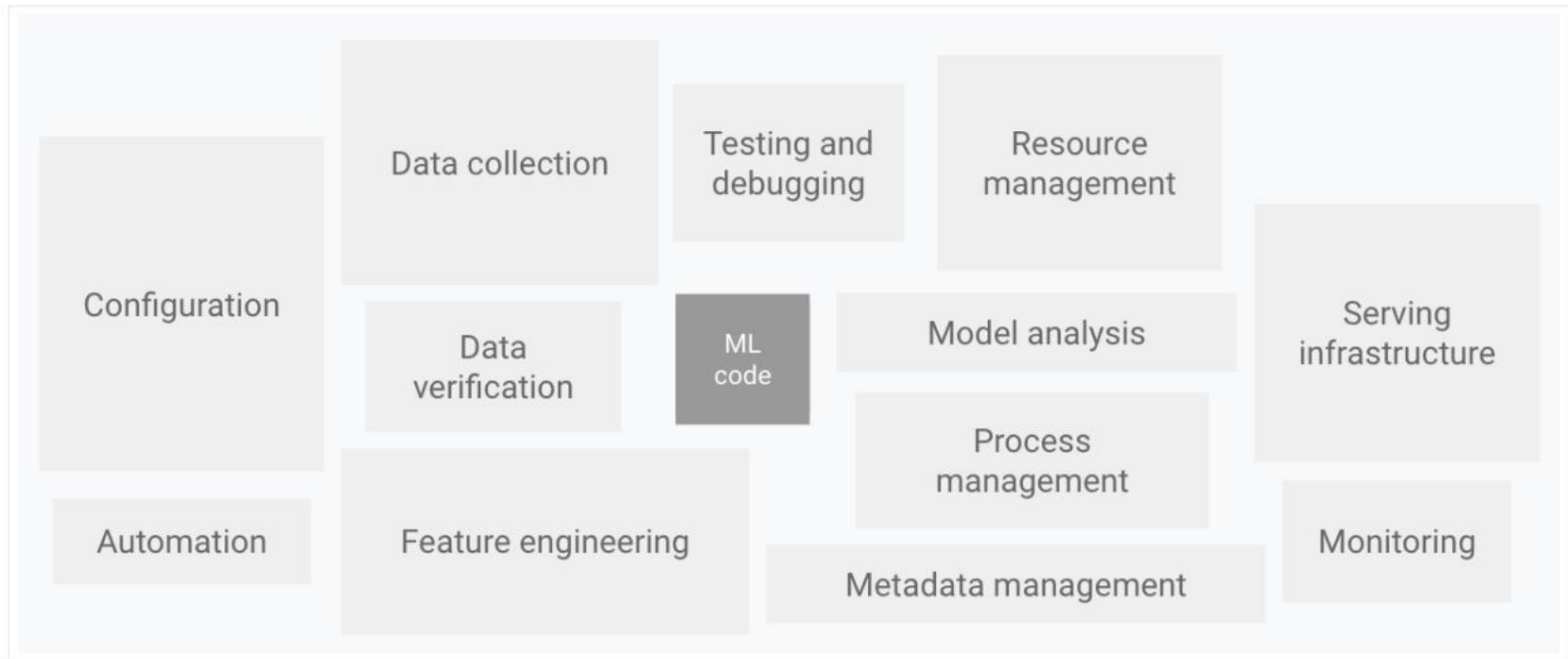
Figure 1. Elements for ML systems. Adapted from Hidden Technical Debt in Machine Learning Systems.

https://cloud.google.com/solutions/machine-learning/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

Model & ML code just tip of the iceberg...

Introducing MLOps – Machine Learning & Operations

DevOps

Code reproducibility

Code testing

App deployment

MLOps

Model reproducibility

Model validation

Model deployment

Model retraining

DevOps

Collaborate → Build app → Test app → Deploy app → Monitor app
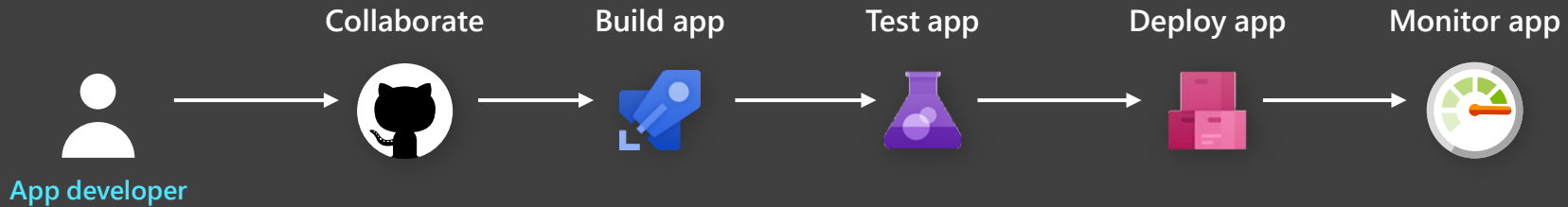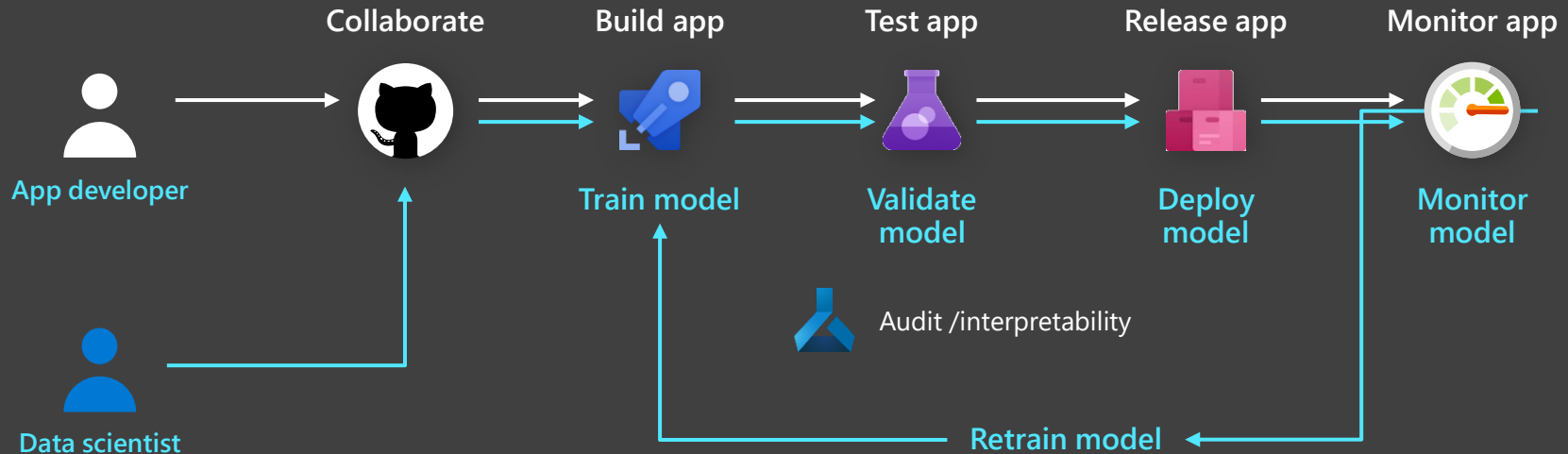
App developer

MLOps with Azure Machine Learning

MLOps Demo

GitHub & Azure ML

dev>scope

rquintino / mlops_azureml   Private

generated from machine-learning-apps/ml-template-azure

Unwatch ▾ | 1   ☆ Star | 0   Fork | 0

<> Code   ⊙ Issues

machine-l.../ml-template-azure   Template

Template for getting started with automated ML Ops on Azure Machine Learning

● Python   ☆ 47   ⑂ 28   Updated on Dec 11, 2020

⚠ Security   Insights   ⚙ Settings

master ▾   1 b

Go to file   Add file ▾   Code ▾

About   ⚙

MLOps-Azure Machine Learning

rquintino Update run_config.yml   ✓ 18b9bce 2 hours ago   ⊙ 6 commits

📖 Readme

⚖ MIT License

.cloud/.azure	Update compute.json	2 hours ago
.github/workflows	Initial commit	2 hours ago
code	Update run_config.yml	2 hours ago
docs/images	Initial commit	2 hours ago
.gitignore	Initial commit	2 hours ago
CODE_OF_CONDUCT.md	Initial commit	2 hours ago
LICENSE	Initial commit	2 hours ago
README.md	Initial commit	2 hours ago
SECURITY.md	Initial commit	2 hours ago

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● Python 100.0%

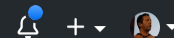README.md   ✏

rquintino / mlops_azureml   Private

generated from machine-learning-apps/ml-template-azure

Unwatch ▾   1   ★ Star   0   Fork   0

<> Code   ⊘ Issues   Pull requests 1   ▶ Actions   Projects   Security   Insights   ⚙ Settings

mlops_azureml / .github / workflows /   train_deploy.yml   Cancel

Start commit ▾

<> Edit file   ⊙ Preview changes   Spaces ▾   2 ▾   No wrap ▾

Marketplace   Documentation

```
 1  # Actions train a model on Azure Machine Learning
 2  name: aml-train-deploy-workflow
 3
 4  on:
 5    push:
 6  #    branches:
 7  #      - master
 8      paths:
 9        - 'code/*'
10
11  jobs:
12    train:
13      runs-on: ubuntu-latest
14      steps:
15      # Checks-out your repository under $GITHUB_WORKSPACE, so your job can access it
16      - name: Check Out Repository
17        id: checkout_repository
18        uses: actions/checkout@v2
19
20      # Connect or Create the Azure Machine Learning Workspace
21      - name: Connect/Create Azure Machine Learning Workspace
```

machine learning

Marketplace / Search results ●

Azure Machine Learning Workspace   ★ 15
By Azure ✓
Connect to or create an Azure
Machine Learning Workspace with this
GitHub Action

Azure Machine Learning Compute   ★ 16
Action
By Azure ✓
Connect to or create a Compute Target
to your Azure Machine Learning
Workspace with this GitHub Action

Azure Machine Learning Run Action   ★ 13
By Azure ✓
Submit a run to an Azure Machine

Use Control + Space to trigger autocomplete in most situations.

```yaml
      # Connect or Create the Azure Machine Learning Workspace
      - name: Connect/Create Azure Machine Learning Workspace
        id: aml_workspace
        uses: Azure/aml-workspace@v1
        with:
          azure_credentials: ${{ secrets.AZURE_CREDENTIALS }}

      # Connect or Create a Compute Target in Azure Machine Learning
      - name: Connect/Create Azure Machine Learning Compute Target
        id: aml_compute_training
        uses: Azure/aml-compute@v1
        with:
          azure_credentials: ${{ secrets.AZURE_CREDENTIALS }}

      # Submit a training run to the Azure Machine Learning
      - name: Submit training run
        id: aml_run
        uses: Azure/aml-run@v1
        with:
          azure_credentials: ${{ secrets.AZURE_CREDENTIALS }}

      # Register model in Azure Machine Learning model registry
      - name: Register model
```

master

mlops_azureml / .cloud / .azure / compute.json

Go to file

rquintino Update compute.json

Latest commit 8324d31 2 hours ago   History

1 contributor

8 lines (8 sloc) | 173 Bytes
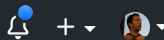
Raw   Blame

```json
1   {
2       "name": "mlopscluster",
3       "compute_type": "amlcluster",
4       "vm_size": "Standard_DS3_v2",
5       "vm_priority": "lowpriority",
6       "min_nodes": 0,
7       "max_nodes": 1
8   }
```

rquintino / mlops_azureml   Private

generated from machine-learning-apps/ml-template-azure

Unwatch ▾ 1   ☆ Star 0   Fork 0

<> Code   ⊙ Issues   ⑂ Pull requests 1   ▶ Actions   ⊞ Projects   🛡 Security   📈 Insights   ⚙ Settings

⑂ master ▾

mlops_azureml / code / train / train.py / ‹ › Jump to ▾

Go to file   ⋯

rquintino Initial commit   Latest commit 1c05e79 3 hours ago   🕐 History

👥 1 contributor

130 lines (103 sloc) | 4.59 KB   Raw   Blame   🖥   ✎   🗑

```python
1    import os
2    import argparse
3    import itertools
4    import numpy as np
5    import pandas as pd
6    import joblib
7    import matplotlib.pyplot as plt
8
9    from sklearn import datasets
10   from sklearn.svm import SVC
11   from sklearn.metrics import confusion_matrix, precision_score, recall_score, f1_score
12   from sklearn.model_selection import train_test_split
13
14   from azureml.core import Dataset, Run
```

```python
x, y = datasets.load_iris(return_X_y=True)

#dividing X,y into train and test data
x_train, x_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=223)
data = {'train': {'X': x_train, 'y': y_train},
        'test': {'X': x_test, 'y': y_test}}

# train a SVM classifier
svm_model = SVC(kernel=args.kernel, C=args.penalty, gamma='auto').fit(data['train']['X'], data['train']['y'])
svm_predictions = svm_model.predict(data['test']['X'])

# accuracy for X_test
accuracy = svm_model.score(data['test']['X'], data['test']['y'])
print('Accuracy of SVM classifier on test set: {:.2f}'.format(accuracy))
run.log('Accuracy', np.float(accuracy))

# precision for X_test
precision = precision_score(svm_predictions, data["test"]["y"], average='weighted')
print('Precision of SVM classifier on test set: {:.2f}'.format(precision))
run.log('precision', precision)

# recall for X_test
recall = recall_score(svm_predictions, data["test"]["y"], average='weighted')
print('Recall of SVM classifier on test set: {:.2f}'.format(recall))
run.log('recall', recall)

# f1-score for X_test
```

Lets try a new hyper param configuration...

```
127
128    if __name__ == '__main__':
129        args = parse_args()
130        main(args=args)
131
```

Commit changes

change gamma to auto

Add an optional extended description...

○ ⊶ Commit directly to the `master` branch.

● ⑆ Create a **new branch** for this commit and start a pull request. Learn more about pull requests.

⑆ rquintino-gamma-auto

Propose changes   Cancel

Open a pull request

The change you just made was written to a new branch named `rquintino-gamma-auto`. Create a pull request below to propose these changes.

base: master ← compare: rquintino-gamma-auto

✓ **Able to merge.** These branches can be automatically merged.

change gamma to auto

Write | Preview

H B I ☰ <> 🔗 ☰ ☰ ☑ @ ↗ ↩

Leave a comment

Attach files by dragging & dropping, selecting or pasting them.

Create pull request

Remember, contributions to this repository should follow its code of conduct.

Reviewers ⚙
No reviews

Assignees ⚙
No one—assign yourself

Labels ⚙
None yet

Projects ⚙
None yet

Milestone ⚙
No milestone

Linked issues ⓘ

Add more commits by pushing to the `rquintino-gamma-auto` branch on **rquintino/mlops_azureml**.

○ **Some checks haven't completed yet**   Hide all checks

1 in progress check

● 🐙 **aml-train-deploy-workflow / train (push)**   *In progress — This check has...*   Details

✓ **This branch has no conflicts with the base branch**

Merging can be performed automatically.

Merge pull request ▾   You can also open this in GitHub Desktop

Write | Preview   H B I ≔ ⟨⟩ 🔗 ☰

Leave a comment

Attach files by dragging & dropping, selecting or pasting them

```
1   # Actions train a model on Azure Machine Learning
2   name: aml-train-deploy-workflow
3
4   on:
5     push:
6   #     branches:
7   #       - master
8       paths:
9         - 'code/*'
10
```

Projects ⚙

None yet

Milestone ⚙

No milestone

Linked issues ⚙

Successfully merging this pull request may close these issues.

None yet

🔒 Lock conversation

Update run_config.yml aml-train-deploy-workflow #6

Cancel workflow

Summary

Jobs

train

---

train
Started 33m 3s ago

Search logs

> ✓ Set up job   7s

> ✓ Build Azure/aml-workspace@v1   24s

> ✓ Build Azure/aml-compute@v1   3s

> ✓ Build Azure/aml-run@v1   2s

> ✓ Build Azure/aml-registermodel@v1   21s

> ✓ Build Azure/aml-deploy@v1   3s

> ✓ Check Out Repository   2s

> ✓ Connect/Create Azure Machine Learning Workspace   4s

> ✓ Connect/Create Azure Machine Learning Compute Target   7s

> ✓ Submit training run   31m 10s

> ✓ Register model   9s

  ⊙ Deploy model   31s

  ○ Post Check Out Repository

rq-azureml > Experiments

Experiments

All experiments   All runs

↻ Refresh   ▭ Archive experiment | ⬤ View archived experiments

🔍 Search

Experiment	Latest run	Last submitted ↓	Created	Cre
mlops_azureml-master	3	Feb 18, 2021 9:49 PM	Feb 18, 2021 9:24 PM	Se
BankMarketing	11	Jan 21, 2021 4:07 PM	Jan 21, 2021 3:00 PM	Ru

Compute, Experiment, Model, Deployment in Azure ML

rq-azureml > Compute

Compute

Compute instances **Compute clusters** Inference clusters Attached compute

+ New   ↻ Refresh   🗑 Delete   ▤ View quota

▽ Search to filter items...

Name	Provisioning state	Virtual machine size	Created on ↓	Idle nodes	Busy nodes
mlopscluster	◯ Resizing (0 -> 1 node)	STANDARD_DS3_V2	Feb 18, 2021 9:24 PM	0	0

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

Compute

Compute instances | **Compute clusters** | Inference clusters | Attached compute

+ New   ⟳ Refresh   🗑 Delete   ▦ View quota   ▽ Search to filter items...

Name	Provisioning state	Virtual machine size	Created on ↓	Idle nodes	Busy nodes
mlopscluster	✓ Succeeded (1 node)	STANDARD_DS3_V2	Feb 18, 2021 9:24 PM	0	1

Navigation (sidebar)

+ New

🏠 Home

Author

📓 Notebooks

⚡ Automated ML

🎨 Designer

Assets

📊 Datasets

🧪 Experiments

🔀 Pipelines

📦 Models

☁ Endpoints

Manage

🖥 Compute

🗄 Datastores

🖊 Data Labeling

Run 3 🔄 Running

🔄 Refresh   ▶ Resubmit   ⊗ Cancel

Details   Metrics   Images   Child runs   Outputs + logs   Snapshot   Explanations (preview)   Fairness (preview)

Properties

Status
🔄 Running

Created
Feb 18, 2021 9:49 PM

Started
Feb 18, 2021 10:16 PM

Compute target
mlopscluster

Run ID
mlops_azureml-master_1613684959_56405295

Script name
train.py

Created by
Service Principal

Input datasets
None

Tags

GITHUB_ACTOR : rquintino   GITHUB_REF : refs/heads/master

GITHUB_REPOSITORY : rquintino/mlops_azureml

GITHUB_SHA : 18b9bce9a9969e24d745430ed8ae238573fd3074

Metrics

ⓘ No data

Description

ⓘ Click edit icon to add a description

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Experiments > mlops_azureml-master > Run 3

Run 3 🏃 Running

Refresh | Resubmit | Cancel | ⬤ Enable log streaming | ◯ Word wrap

Details | Metrics | Images | Child runs | **Outputs + logs** | Snapshot | Explanations (preview) | Fairness (preview)

🔍

> 📁 azureml-logs
> 📁 logs
> 📁 outputs
  📄 confusion_matrix
  ✅ 📄 confusion_matrix_normalized_1613686734.png ...
  ✅ 📄 confusion_m   confusion_matrix_normalized_1613686734.png
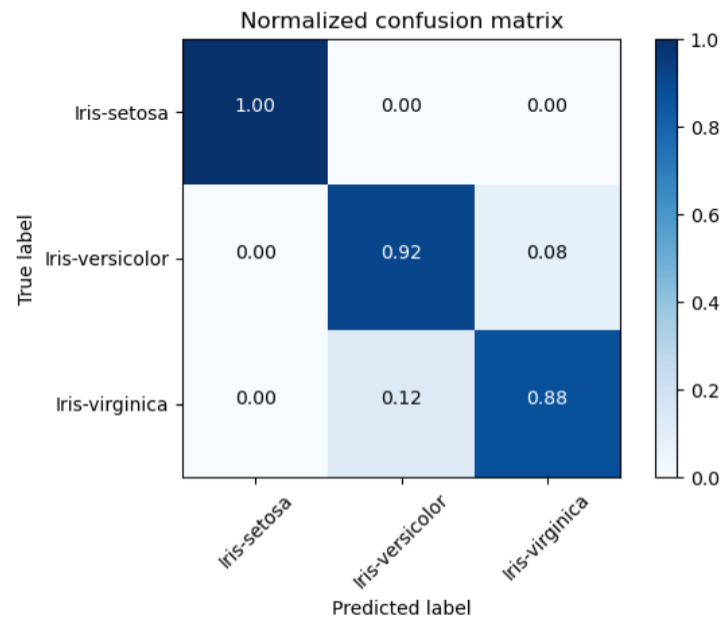
📄 confusion_matrix_nc ✕   📄 confusion_matrix_ur ✕

Normalized confusion matrix

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Experiments > mlops_azureml-master > Run 3

Run 3 ✅ Completed

🔄 Refresh   ▶ Resubmit   ✕ Cancel

Details   Metrics   Images   Child runs   Outputs + logs   Snapshot   Explanations (preview)   Fairness (preview)

Properties

Status
✅ Completed

Created
Feb 18, 2021 9:49 PM

Started
Feb 18, 2021 10:16 PM

Duration
2m 42.44s

Compute target
mlopscluster

Run ID
mlops_azureml-master_1613684959_56405295

Script name
train.py

Created by
Service Principal

Tags

GITHUB_ACTOR : rquintino   GITHUB_REF : refs/heads/master

GITHUB_REPOSITORY : rquintino/mlops_azureml

GITHUB_SHA : 18b9bce9a9969e24d745430ed8ae238573fd3074

Metrics

confusion_matrix_normalized
Type: Image

confusion_matrix_unnormalized
Type: Image

confusion_matrix
azureml.v2.confusion_matrix

f1-score
0.9333333333333333

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

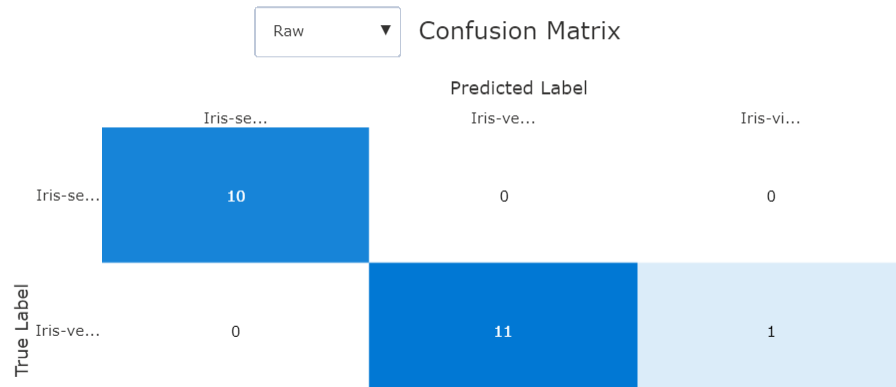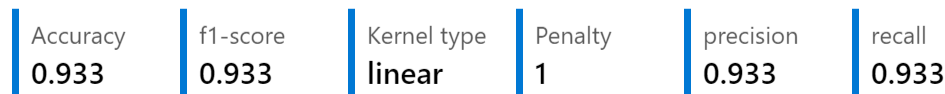Data Labeling

Run 3 ✓ Completed

Refresh   Resubmit   Cancel

Details   **Metrics**   Images   Child runs   Outputs + logs   Snapshot   Explanations (preview)   Fairness (preview)

Select a metric to see a visualization or table of the data.

🔍 Search

☑ Accuracy

☑ confusion_matrix

☑ f1-score

☑ Kernel type

☑ Penalty

☑ precision

☑ recall

View as: ● Chart   ○ Table

Accuracy	f1-score	Kernel type	Penalty	precision	recall
0.933	0.933	linear	1	0.933	0.933

Raw ▼   Confusion Matrix

Predicted Label

	Iris-se...	Iris-ve...	Iris-vi...
Iris-se...	10	0	0
Iris-ve...	0	11	1

True Label

Raw ▼   Confusion Matrix

Predicted Label

	Iris-se...	Iris-ve...	Iris-vi...
Iris-se...	10	0	0
Iris-ve...	0	11	1
Iris-vi...	0	1	7

True Label

10
8
6
4
2
0

🔄 Re-run jobs ▾ ⋯
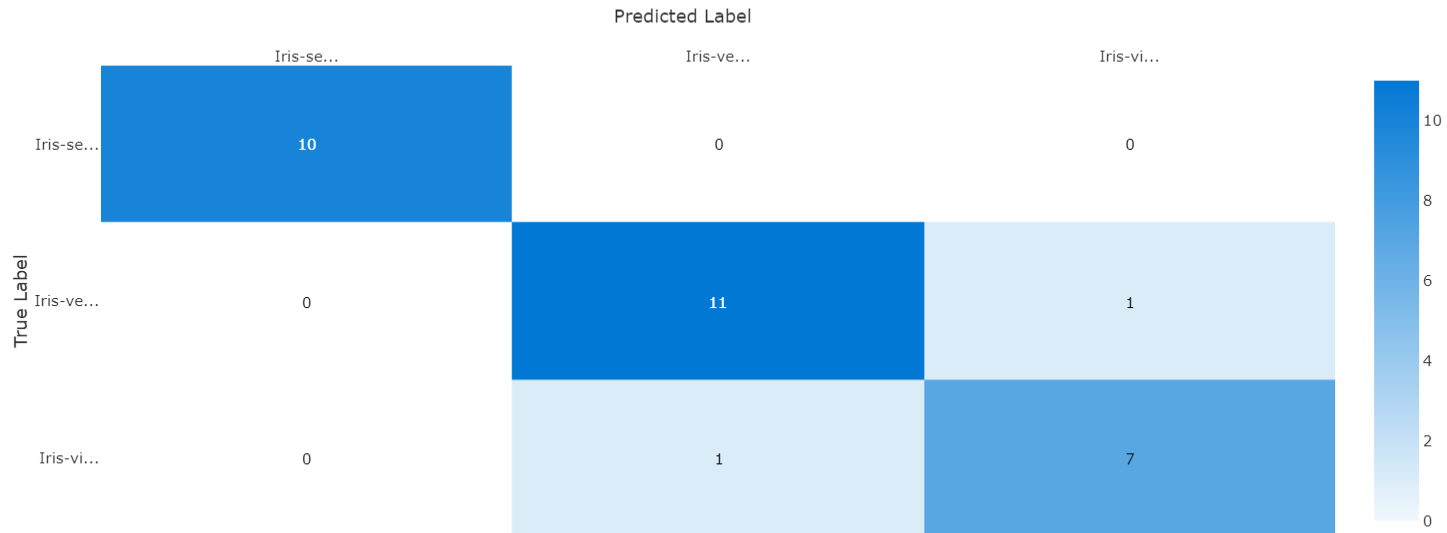
Summary

Jobs

✅ train

train
succeeded 5 minutes ago in 43m 6s

🔍 Search logs ⚙️

> ✅ Set up job   7s

> ✅ Build Azure/aml-workspace@v1   24s

> ✅ Build Azure/aml-compute@v1   3s

> ✅ Build Azure/aml-run@v1   2s

> ✅ Build Azure/aml-registermodel@v1   21s

> ✅ Build Azure/aml-deploy@v1   3s

> ✅ Check Out Repository   2s

> ✅ Connect/Create Azure Machine Learning Workspace   4s

> ✅ Connect/Create Azure Machine Learning Compute Target   7s

> ✅ Submit training run   31m 10s

> ✅ Register model   9s

> ✅ Deploy model   10m 34s

> ✅ Post Check Out Repository   0s

> ✅ Complete job   0s

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Models

Model List

+ Register model   🗑 Delete   ▷ Deploy   ↻ Refresh

+ Add filter

Name	Version	Experiment	Run ID	Created on	Tags
mlops_azureml-master	1	mlops_azureml-master	mlops_azureml-master_161368...	Feb 18, 2021 10:20 PM	

https://ml.azure.com/model/list?wsid=/subscriptions/69df3b8f-2020-4de0-b...

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

mlops_azureml-master:1

Details | Artifacts | Endpoints | Explanations (preview) | Fairness (preview) | Datasets

⟳ Refresh   ▶ Deploy   ⬇ Download

Attributes

Version
1

ID
mlops_azureml-master:1

Date registered
2/18/2021, 10:20:33 PM

Framework
Custom

Framework version
--

Experiment name
mlops_azureml-master

Run ID
mlops_azureml-master_1613684959_56405295

Created by
6b3d82ad-c37e-4817-9f27-941025e1df47

Tags

✎

ⓘ No tags

Properties

ⓘ No properties

https://ml.azure.com/model/list?wsid=/subscriptions/69df3b8f-2020-4de0-b47c-29decde8a654/resourcegroups/rq-azureml/workspaces/rq-azureml&tid=3a364c67-349a-484d-8dcb-873c88970c00

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Models > mlops_azureml-master:1

mlops_azureml-master:1

Details **Artifacts** Endpoints Explanations (preview) Fairness (preview) Datasets

Refresh | Deploy | Download

model.pkl

model.pkl ✕

ⓘ The file is either binary or uses an unsupported text encoding

Preview not available

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Endpoints

Endpoints

Real-time endpoints   Pipeline endpoints

🗑 Delete   ▦ Edit columns   ↻ Refresh   ↰ Reset to default view

⊕ Add filter

Showing 1-1 endpoints

Page size: 25 ⌄

Name	Description	Created on	Created by	Updated on ↓	Compute type	C
mlops-azureml-master		Feb 18, 2021 10:20 PM	Service Principal	Feb 18, 2021 10:20 PM	ACI	

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Endpoints > mlops-azureml-master

mlops-azureml-master

Details   Test   Consume   Deployment logs

Attributes

Service ID
mlops-azureml-master

Description
--

Deployment state
Transitioning ⓘ

Compute type
ACI

Created by
6b3d82ad-c37e-4817-9f27-941025e1df47

Model ID
mlops_azureml-master:1

Created on
2/18/2021 10:20:44 PM

Last updated on
2/18/2021 10:20:44 PM

Image ID

Tags

ⓘ No data

Properties

ⓘ No data

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Endpoints > mlops-azureml-master

mlops-azureml-master

Details   Test   Consume   Deployment logs

Attributes

Service ID
mlops-azureml-master

Description
--

Deployment state
Healthy ⓘ

Compute type
ACI

Created by
6b3d82ad-c37e-4817-9f27-941025e1df47

Model ID
mlops_azureml-master:1

Created on
2/18/2021 10:20:44 PM

Last updated on
2/18/2021 10:20:44 PM

Image ID

Tags

ⓘ No data

Properties

hasInferenceSchema
True

hasHttps
False

https://ml.azure.com/model/mlops_azureml-master:1/details?wsid=/subscri...

New

Home

Author

Notebooks

Automated ML

Designer

Assets

Datasets

Experiments

Pipelines

Models

Endpoints

Manage

Compute

Datastores

Data Labeling

rq-azureml > Compute

Compute

Compute instances | **Compute clusters** | Inference clusters | Attached compute

+ New | ↻ Refresh | 🗑 Delete | ▥ View quota

▽ Search to filter items...

Name	Provisioning state	Virtual machine size	Created on ↓	Idle nodes	Busy nodes
mlopscluster	✓ Succeeded (0 nodes)	STANDARD_DS3_V2	Feb 18, 2021 9:24 PM	0	0

mlflow

search_models

Experiment ID: 72   Artifact Location: file:///code/data/mlruns/72

▼ Notes ✎

None

Wind-dependent Variables: Predict Wind Speeds of Tropical Storms

HOSTED BY RADIANT EARTH FOUNDATION

Search Runs: `metrics.rmse < 1 and params.model = "tree" and tags.mlflow.source.type = "LOCAL"`   State: Active ▼   Search   Clear

Showing 29 matching runs   Compare   Delete   Download CSV 📥   ☰ ⊞   ⚙ Columns

	Date	Run Name	User	Source	Version	FOLD	NOTEBC	SEED	avg_trai	avg_val	avg_val	AILAB_[	AILAB_E	AILAB_E
☐	2021-01-2	MODEL=nasnetalarge BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	1.77...	8.51...	9.00...	202...	doc...	sear...
☐	✓ 2021-0	MODEL=resnet101d BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	4.10...	9.15...	9.66...	202...	doc...	sear...
☐	✓ 2021-0	MODEL=efficientnet_b3a BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	2.13...	8.04...	8.45...	202...	doc...	sear...
☐	✓ 2021-0	MODEL=tresnet_xl_448 BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	-	-	-	202...	doc...	sear...
☐	✓ 2021-0	MODEL=xception71 BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	1.94...	7.92...	8.35...	202...	doc...	sear...
☐	✓ 2021-0	MODEL=xception65 BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	2.01...	8.57...	9.75...	202...	doc...	sear...
☐	✓ 2021-0	MODEL=xception BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	1.96...	7.92...	8.38...	202...	doc...	sear...
☐	✓ 2021-0	MODEL=hrnet_w48 BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	2.46...	8.12...	8.57...	202...	doc...	sear...
☐	2021-0	MODEL=tf_efficientnet_b8_ap BATCH_SIZE=8	rqui...	🖥 ipyke	-	1	1-Tr...	1	1.76...	8.41...	8.92...	202...	doc...	sear...

Parameters   Metrics ＞   Tags ＞

Tracking critical for MLOps, but also for Experimenting

Experimentation
With GitHub & CML

- Automation for Experimentation

- Ex: Smoke Test Locally (Fast)

- Queue Reproducible in Github

- Using notebooks

dev>scope

generated from machine-learning-apps/ml-template-azure

Unwatch 1   Star 0   Fork 0

Code   Issues   Pull requests 1   Actions   Projects   Security   Insights   Settings

mlops_azureml / .github / workflows / experiment-train.yml

View runs   Go to file

rquintino Update experiment-train.yml   Latest commit b19c525 5 hours ago   History

1 contributor

21 lines (18 sloc)   526 Bytes

Raw   Blame

```
1   name: experiment-train
2   on: [push]
3
4   jobs:
5     run:
6       runs-on: [ubuntu-latest]
7       container: docker://dvcorg/cml-py3:latest
8       steps:
9         - uses: actions/checkout@v2
10        - name: cml_run
11          env:
12            repo_token: ${{ secrets.GITHUB_TOKEN }}
13          run: |
14
15            # Your ML workflow goes here
16            pip install -r requirements.txt
17            papermill ./code/experiment/train.ipynb train_output.ipynb
18
19            # Write your CML report
20            cat "Hello" >> report.md
21            cml-send-comment report.md
```

Check Continuous ML (CML) framework in Github

localhost:8888/notebooks/code/experiment/train.ipynb

train Last Checkpoint: 3 hours ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Navigate Widgets Help Snippets

Not Trusted

Python 3.7.3 64-bit ('base': conda)

Code

Notify: Disabled

Snippets

Contents

1 Params
2 Read data
3 Train
4 Report

executed in 94ms, finished 11:36:38 2021-02-19

3 Train

```
In [5]: model=RandomForestClassifier()
```
executed in 11ms, finished 11:36:38 2021-02-19

```
In [6]: cv_scores=cross_val_score(model,X,y,cv=5,scoring="roc_auc")
        cv_scores
```
executed in 9.18s, finished 11:36:48 2021-02-19

```
Out[6]: array([0.98643411, 0.98013566, 0.99865862, 0.99027498, 0.99865862])
```
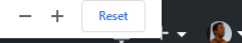
4 Report

```
In [7]: mean_scores,std_scores=np.mean(cv_scores),np.std(cv_scores)
        mean_scores,std_scores
```
executed in 18ms, finished 11:36:48 2021-02-19

```
Out[7]: (0.9908323974856893, 0.007163453627193791)
```

```
In [8]: !echo cv_core_mean: {mean_scores} >> report.md
        !echo cv_core_std: {std_scores} >> report.md
```
executed in 212ms, finished 11:36:48 2021-02-19

```
In [ ]:
```

File   Edit   View   Insert   Cell   Kernel   Navigate   Widgets   Help   Snippets

Not Trusted   Python 3.7.3 64-bit ('base': conda)

Markdown

Contents

```
TEST=1
```
executed in 6ms, finished 11:36:38 2021-02-19

In [3]:
```
print(TEST)
```
executed in 49ms, finished 11:36:38 2021-02-19

```
1
```

2 Read data

In [4]:
```
X,y=load_breast_cancer(return_X_y=True)
```
executed in 94ms, finished 11:36:38 2021-02-19

3 Train

In [5]:
```
model=LogisticRegression()
```
executed in 11ms, finished 11:36:38 2021-02-19

In [6]:
```
cv_scores=cross_val_score(model,X,y,cv=5,scoring="roc_auc")
cv_scores
```
executed in 9.18s, finished 11:36:48 2021-02-19

Out[6]:
```
array([0.98643411, 0.98013566, 0.99865862, 0.99027498, 0.99865862])
```

4 Report

```
mean_scores,std_scores
```

Branch for new experiment, change code, queue in GitHub

Commit & Push branch to queue workflow in GitHub

Notebook execution using papermill on Github agent

```
15        "from sklearn.model_selection import train_test_split\n",
16        "from sklearn.datasets import load_breast_cancer\n",
17        "from sklearn.ensemble import RandomForestClassifier\n",

18        "import numpy as np"
19      ]
20    },
```

```
15        "from sklearn.model_selection import train_test_split\n",
16        "from sklearn.datasets import load_breast_cancer\n",
17        "from sklearn.ensemble import RandomForestClassifier\n",
18  +     "from sklearn.linear_model import LogisticRegression\n",
19        "import numpy as np"
20      ]
21    },
```

@@ -110,7 +111,7 @@

```
110      },
111      "outputs": [],
112      "source": [
113  -      "model=RandomForestClassifier()"
114      ]
115    },
116    {
```

```
111      },
112      "outputs": [],
113      "source": [
114  +      "model=LogisticRegression()"
115      ]
116    },
```

1 comment on commit `793afde`

github-actions `bot` commented on 793afde 6 minutes ago

Train Report

cv_core_mean: 0.9910938715843727
cv_core_std: 0.006007611645631809

Train Finished

Re: [rquintino/mlops_azureml] test logistic (793afde) ✉ Inbox ✕

github-actions[bot] <notifications@github.com> Unsubscribe
to rquintino/mlops_azureml, Subscribed ▾

Train Report

cv_core_mean: 0.9910938715843727
cv_core_std: 0.006007611645631809

Train Finished

CML watermark

—

You are receiving this because you are subscribed to this thread.
Reply to this email directly, view it on GitHub, or unsubscribe.

Write   Preview   H B I

Leave a comment

Train performance results on commits (also PRs)

MLOps Recap

- From Models to Pipelines

- Execution of Pipelines produce models

- Train for Deployment is fully automated/reproducible

- Pipelines enable auto model retraining

- Automation, Automation, Automation…

- Monitoring for Data Drift, Concept Drift, Model Drift

- Tracking, Metadata & Artifact Stores

dev>scope

Adapted from: https://cloud.google.com/solutions/machine-learning/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning
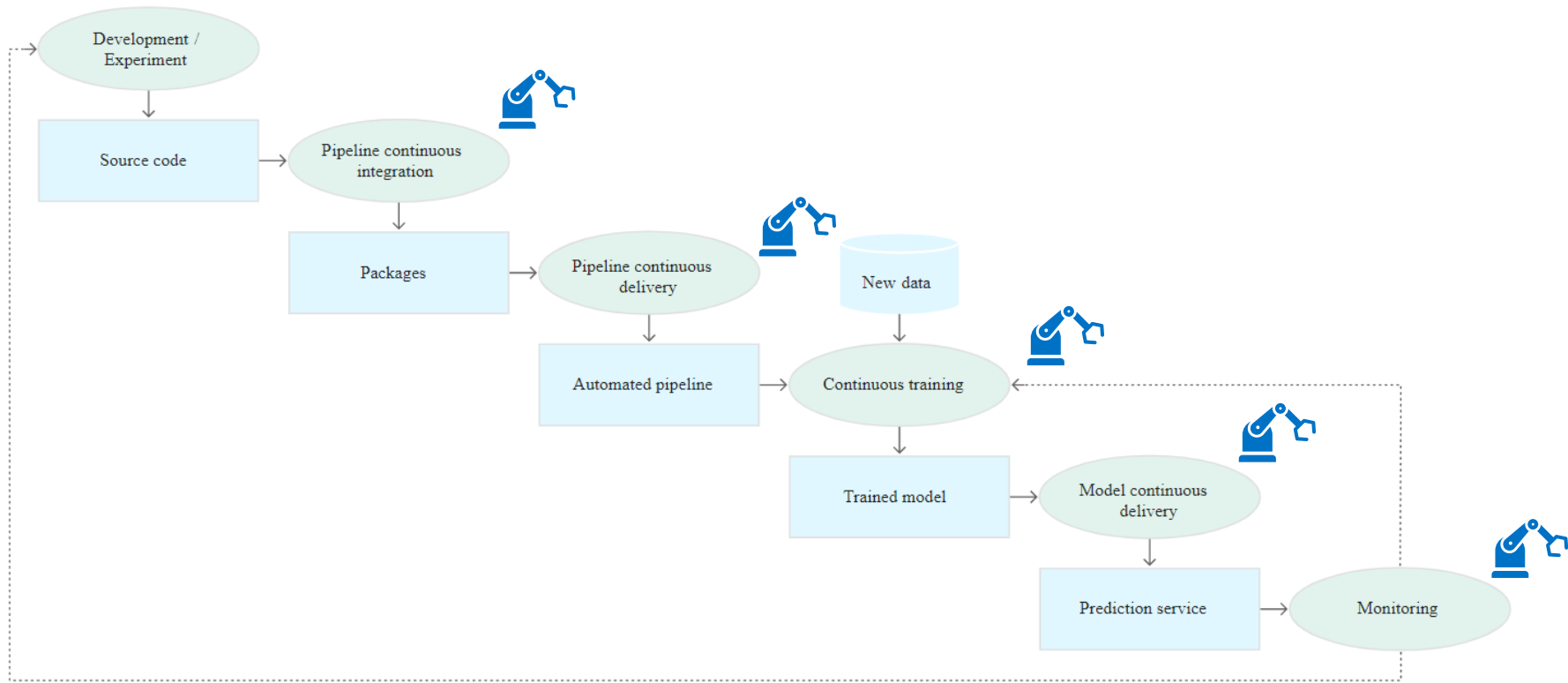
Mature MLOps Example (Level 1)

Run 3 ⊚ Queued

⟳ Refresh   ▶ Resubmit   ⊗ Cancel | 🔘 Enable log streaming   ⚪ Word wrap

Details   Metrics   Images   Child runs   **Outputs + logs**   Snapshot   Explanations (preview)   Fairness (preview)

🔍 [   ]   «

📄 20_image_build_log.   ✕

▾ 🗀 azureml-logs

📄 20_image_build_log.txt

```
670    latest: digest: sha256:7df4cc95e19514ea3b95cb16faf838c26d4a32e786553383ea5c619
671    2021/02/18 22:12:24 Successfully pushed image: 05cf3362d2b44e3a97168a21e22f74e
672    2021/02/18 22:12:24 Step ID: acb_step_0 marked as successful (elapsed time in
673    2021/02/18 22:12:24 Populating digests for step ID: acb_step_0...
674    2021/02/18 22:12:25 Successfully populated digests for step ID: acb_step_0
675    2021/02/18 22:12:25 Step ID: acb_step_1 marked as successful (elapsed time in
676    2021/02/18 22:12:25 The following dependencies were found:
677    2021/02/18 22:12:25
678    - image:
679        registry: 05cf3362d2b44e3a97168a21e22f74e6.azurecr.io
680        repository: azureml/azureml_b40693e99d8d288ddefab00c6595c11c
681        tag: latest
682        digest: sha256:7df4cc95e19514ea3b95cb16faf838c26d4a32e786553383ea5c619d975
683    runtime-dependency:
684        registry: mcr.microsoft.com
685        repository: azureml/base
686        tag: intelmpi2018.3-ubuntu16.04
687        digest: sha256:a1b514f3ba884b9a7695cbba5638933ddaf222e8ce3e8c81e8cdf861679
688    git: {}
689
690    Run ID: cb3 was successful after 22m55s
691
```

Not best fit for interative/model development

Final notes

- MLOps doesn't guarantee ML success

- MLOps should improve feedback and learning
  (not block DataScientists from business/data understanding)

- Dont reinvent the wheel...

- Still a young field
  Dont believe everything you read in docs (and ppt...)

- When/If to introduce MLOps
  (it depends...)

dev>scope

References

MLOps

dev>scope

Machine Learning Ops

A collection of resources on how to facilitate Machine Learning Ops with GitHub.

Learn how to use GitHub for automation, collaboration and reproducibility in your machine learning workflows.

https://mlops.githubapp.com/

DSPT#84 Webinar- The rise of MLOps

494 views • Streamed live on Feb 3, 2021

LICENSE	initial commit	12 months ago
README.md	Contrib/Fix typos (#395) (#396)	5 days ago
bitbucket-pipelines.yml	Bb pr comments (#347)	3 months ago
package-lock.json	Remove new line coerce in private ssh (#408)	18 hours ago
package.json	Remove new line coerce in private ssh (#408)	18 hours ago
prettier.config.js	Pretifier update (#239)	6 months ago
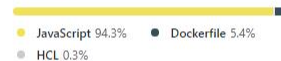
README.md

What is CML? Continuous Machine Learning (CML) is an open-source library for implementing continuous integration & delivery (CI/CD) in machine learning projects. Use it to automate parts of your development workflow, including model training and evaluation, comparing ML experiments across your project history, and monitoring changing datasets.

androvhopf commented on c5eebf3 21 hours ago   Author Member

Style transfer

Workspace vs. Master

Training metrics

Path	Param	Old	New
params.yaml	iter	1	30

GPU info

Also

- https://azure.microsoft.com/en-us/services/machine-learning/mlops/
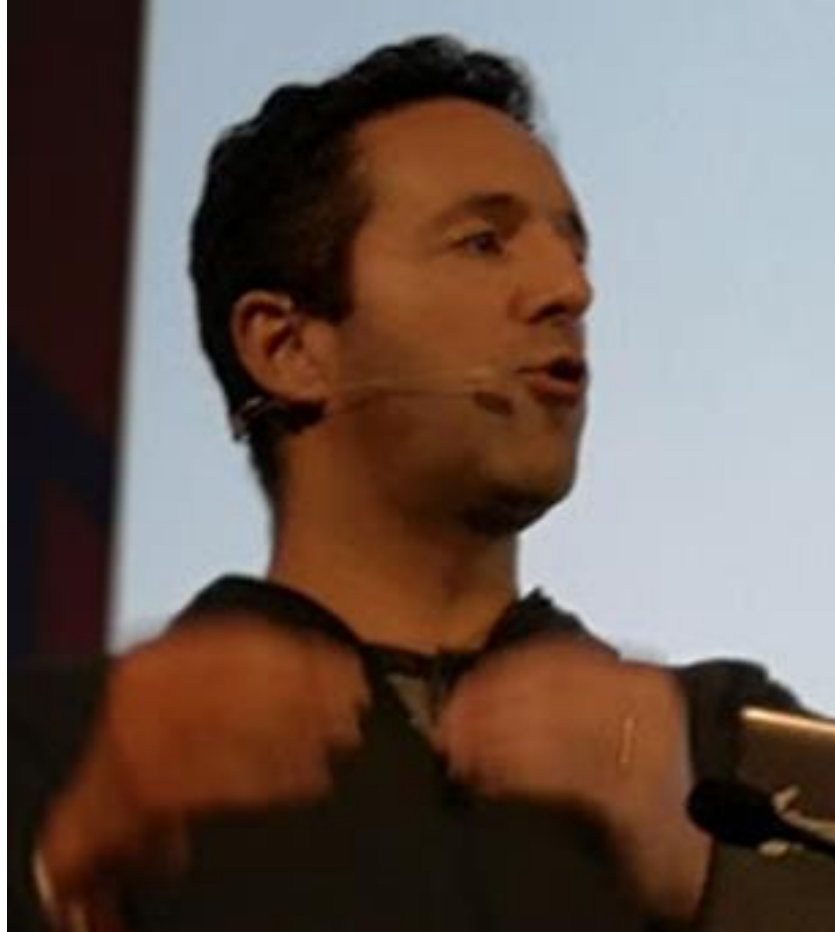
- https://cloud.google.com/solutions/machine-learning/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

- https://github.com/microsoft/MLOps

- https://databricks.com/blog/2020/10/13/using-mlops-with-mlflow-and-azure.html

- https://github.com/Azure/MachineLearningNotebooks

dev>scope

Thank you!

Q&A

rui.quintino@devscope.net

Linkedin.com/ in/rquintino

devscope.net

devscope

🏠 Rua Passos Manuel Nº 223 – 4º Andar
4000-385 Porto

Av. Sidónio Pais, Nº 2 – 3º Andar
1050-214 5 Lisboa

📞 T. +351 223 751 350/51
F. +351 223 751 352

✉️ info@devscope.net
www.devscope.net

Helping customers succeed