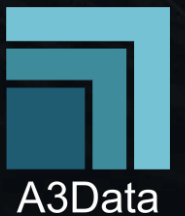


MLOps on the cloud

Prof. Dr. Neylson Crepalde



Neylson Crepalde

Cientista de Dados – A3Data



Neylson Crepalde

Cientista de Dados – A3Data

Coordenador do curso de Ciência de
Dados – Izabela Hendrix



Neylson Crepalde

Cientista de Dados – A3Data

Coordenador do curso de Ciência de
Dados – Izabela Hendrix

Doutor em Sociologia Econômica



Neylson Crepalde

Cientista de Dados – A3Data

Coordenador do curso de Ciência de
Dados – Izabela Hendrix

Doutor em Sociologia Econômica

Cuber!

www.neylsoncrepalde.com



IT WORKS ON MY MACHINE



Machine Learning avança nas organizações ancorado em equipes bem treinadas e *technical skills* necessárias.

A maioria dos *Data Scientists* não tem formação em **engenharia de software**.

Principais problemas



Códigos não
reutilizáveis/reprodutíveis



Trabalho excessivamente
manual



Pouco acompanhamento
da qualidade do código



Baixo monitoramento das
decisões do modelo



Accountability issues



CODING – CODE
DEVELOPMENT AND
REVIEW, SOURCE CODE
MANAGEMENT TOOLS,
CODE MERGING



BUILDING –
CONTINUOUS
INTEGRATION TOOLS,
BUILD STATUS



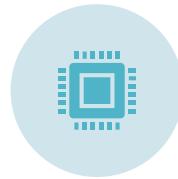
TESTING –
CONTINUOUS TESTING
TOOLS THAT PROVIDE
QUICK AND TIMELY
FEEDBACK ON
BUSINESS RISKS



PACKAGING – ARTIFACT
REPOSITORY,
APPLICATION PRE-
DEPLOYMENT STAGING



RELEASING – CHANGE
MANAGEMENT,
RELEASE APPROVALS,
RELEASE AUTOMATION



CONFIGURING –
INFRASTRUCTURE
CONFIGURATION AND
MANAGEMENT,
INFRASTRUCTURE AS
CODE TOOLS



MONITORING –
APPLICATIONS
PERFORMANCE
MONITORING, END-
USER EXPERIENCE

DevOps

ELITE PERFORMERS

Comparing the elite group against the low performers, we find that elite performers have...



208
TIMES MORE
frequent code deployments

106
TIMES FASTER
lead time from
commit to deploy



2,604
TIMES FASTER
time to recover from incidents

7
TIMES LOWER
change failure rate
(changes are $\frac{1}{7}$ as likely to fail)



Throughput Stability

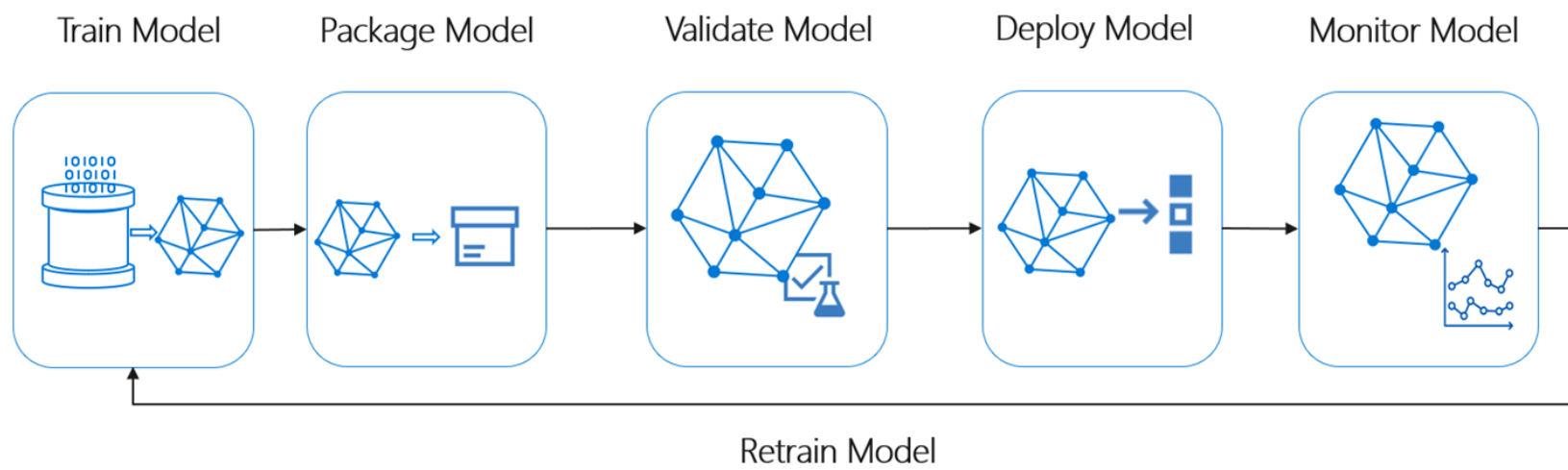
MLOps

MLOps (a compound of “machine learning” and “operationalization”) is the practice of operationalizing and managing the lifecycle of ML in production.




MLOps establishes a culture and environment where ML technologies can generate business benefits by optimizing the ML lifecycle to automate and scale ML initiatives and optimized business return of ML in production.

MLOps enables collaboration across diverse users (such as Data Scientists, Data Engineers, Business Analysts and ITOps) on ML operations and enables a data driven continuous optimization of ML operations' impact or ROI (Return on Investment) to business applications.





azure-pipelines.yml

 Edit[Contents](#) [History](#) [Compare](#) [Blame](#) You updated  azure_tests Yesterday[Create a pull request](#)

```
-
6 trigger:
7   - master
8
9 pool:
10  | vmImage: 'ubuntu-latest'
11 strategy:
12  | matrix:
13    | Python36:
14    |   python.version: '3.6'
15    | Python37:
16    |   python.version: '3.7'
17
18 steps:
19 - task: UsePythonVersion@0
20   inputs:
21   | versionSpec: '$(python.version)'
22   | displayName: 'Use Python $(python.version)'
23
24 - script: |
25   | python -m pip install --upgrade pip
26   | pip install -r requirements.txt
27   | pip install dask[dataframe] --upgrade
28   | displayName: 'Install dependencies'
29
30 - script: |
31   | pip install pytest pytest-cov pytest-azurepipelines
32   | cd src
33   | pytest --junitxml=test-results.xml --cov=. --cov-report=xml
34
```

Summary

4 Run(s) Completed (4 Passed, 0 Failed) [5 unique failing tests in the last 14 days](#)

148

Total tests
+144



148 ● Passed
0 ● Failed
0 ● Others

100%

Pass percentage
↑ 100%

1m 31s

Run duration ⓘ
↑ +1m 28s

0

Tests not reported

Bug Link

Test run Column Options

Filter by test or run name

Tags Test file Owner Aborted (+1)



Hooray! There are no test failures.

Change the test outcome filter to view tests relevant to you.



Kedro



Kubeflow

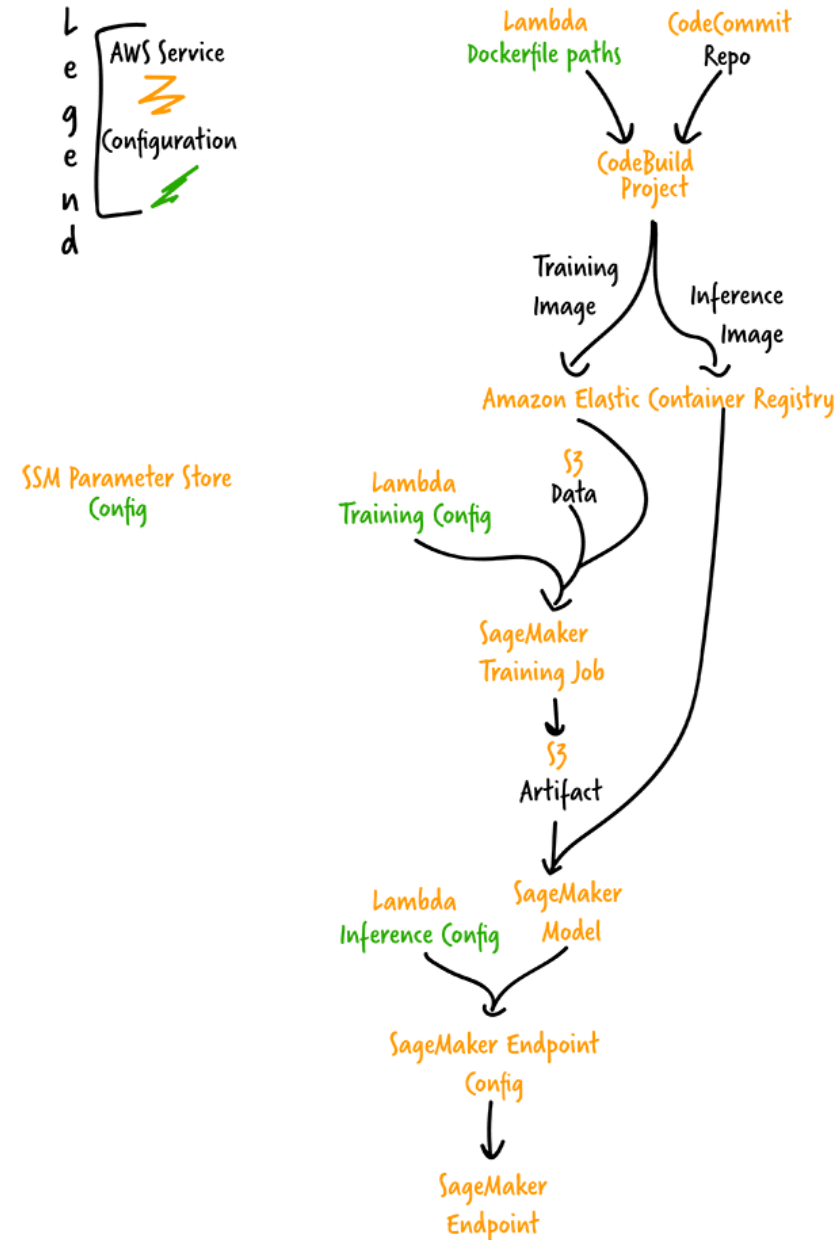


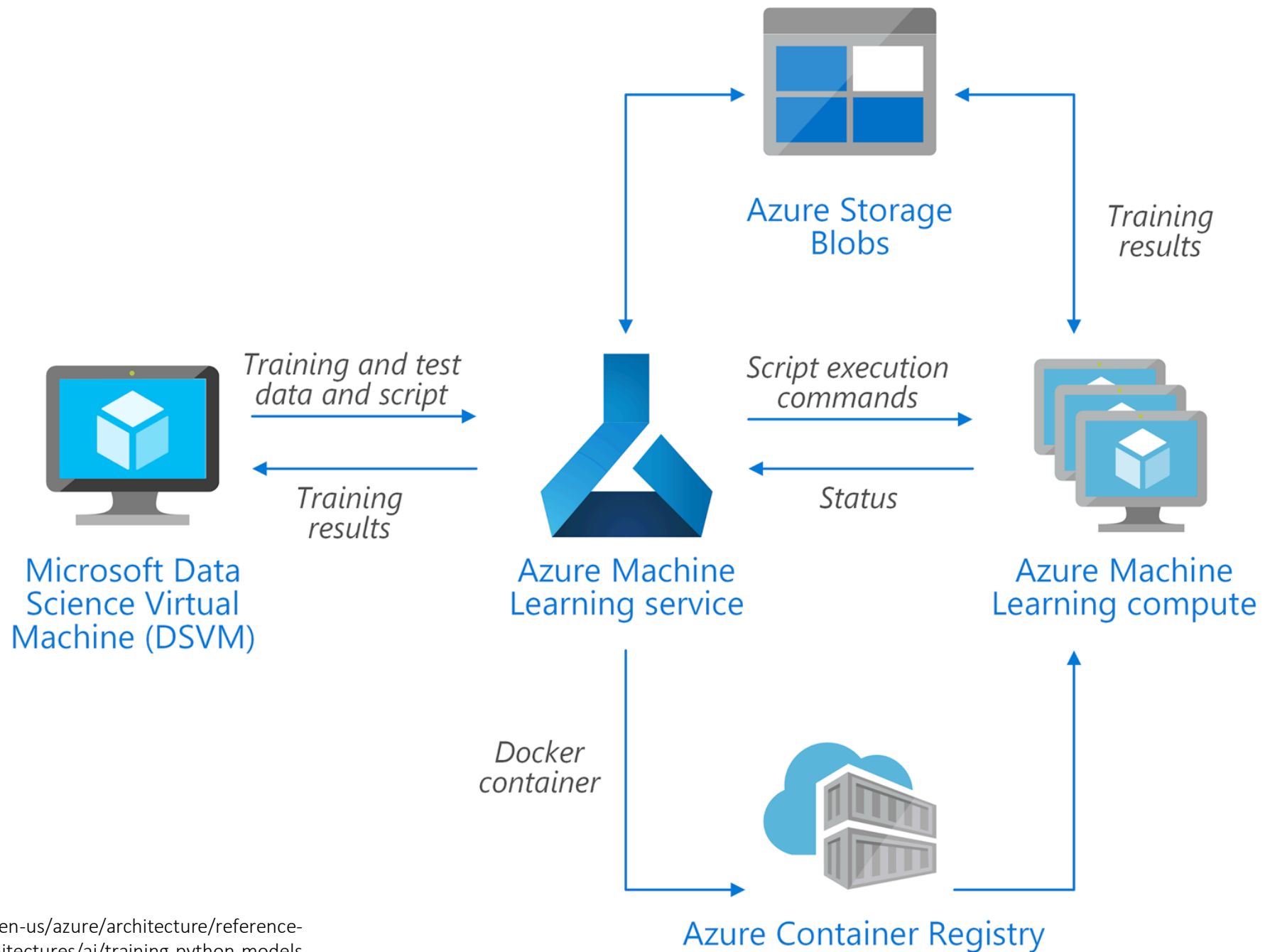
METAFLOW

mlflow



The "sagebuild" architecture





☐ Incluir execuções filhas
 Execuções mais recentes 100

MIN. RMSE

15.16

MIN. MAE

3.482

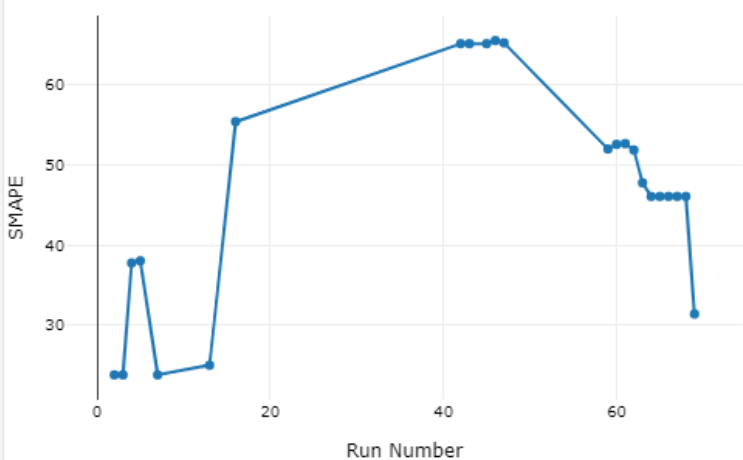
MIN. MAPE

45.54

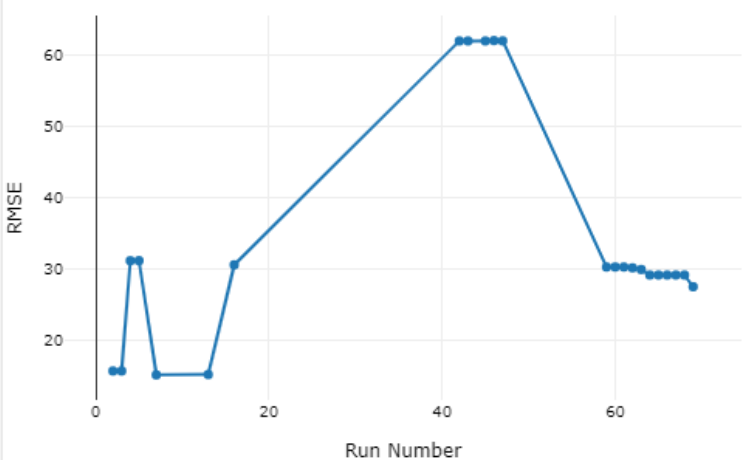
MIN. SMAPE

23.84

SMAPE



RMSE



☐ NÚMERO DA EXECUÇÃO ↑↓

STATUS ↑↓

DURAÇÃO ↑↓

RMSE ↑↓

MAE ↑↓

MAPE ↑↓

SMAPE ↑↓

<input type="checkbox"/> 70	Starting	-				
<input type="checkbox"/> 69	Failed	8h 30m 39.81s	27.49	6.544	110.0	31.44
<input type="checkbox"/> 68	Completed	14m 55.36s	29.13	7.920	219.6	46.11
<input type="checkbox"/> 67	Completed	14m 2.655s	29.13	7.920	219.6	46.11
<input type="checkbox"/> 66	Completed	15m 33.29s	29.13	7.920	219.6	46.11
<input type="checkbox"/> 65	Completed	14m 23.87s	29.13	7.920	219.6	46.11
<input type="checkbox"/> 64	Completed	14m 36.77s	29.13	7.921	219.7	46.10
<input type="checkbox"/> 63	Completed	28m 11.36s	29.91	8.463	247.5	47.80

ATTRIBUTES	
Status	Completed
Hora de Criação	Apr 6, 2020 11:00 PM
Duração	14m 55.36s
Destino	cluster
ID da execução	!4781_38e...
Número da Execução	68
Nome do Script	azure_train.py
Criado por	Neylson João Batista Filho Crepalde
Argumentos	--alg ridge --alpha 0.1 --data-folder \$...

TRACKED METRICS	
RMSE	29.13
MAE	7.920
MAPE	219,6
SMAPE	46.11
P: Alpha	0.1000
P: Alg	ridge

NOME	BAIXAR
▼ azureml-logs	
55_azureml-exec...	↓
65_job_prep-tvm...	↓
70_driver_log.txt	↓
75_job_post-tvmp...	↓
process_info.json	↓
process_status.json	↓
► logs	

Visualização

```
1 Starting job release. Current time:2020-04-07T02:15:20.351640
2 Logging experiment finalizing status in history service.
3 Starting the daemon thread to refresh tokens in background for process with pid = 274
4 Job release is complete. Current time:2020-04-07T02:15:21.944005
5
```

NOME	BAIXAR
► azureml-logs	
► logs	
MLmodel	↓
conda.yaml	↓
model_artifacts.pkl	↓
python_model.pkl	↓

Visualização

```
1 channels:
2 - defaults
3 dependencies:
4 - python=3.7.6
5 - pip:
6   - mlflow
7   - cloudpickle==1.3.0
8 name: mlflow-env
9
```

Desafios para MLOps *on the cloud*



INTEGRAÇÃO COM TIME
DE *MACHINE LEARNING*
ENGINEERS



O GRANDE DESAFIO DA
INTERDISCIPLINARIDADE



DOCKER



IAC (*INFRASTRUCTURE*
AS CODE)



CUSTO

Obrigado!

