

Data Science Lifecycle & MLOps

Prof. Dr. Jan Kirenz
HdM Stuttgart

ML System = Code + Data

Model &
Hyperparameters

Model-Centric vs Data-Centric

Academic research

ML System = Code + Data
Optimize Fixed

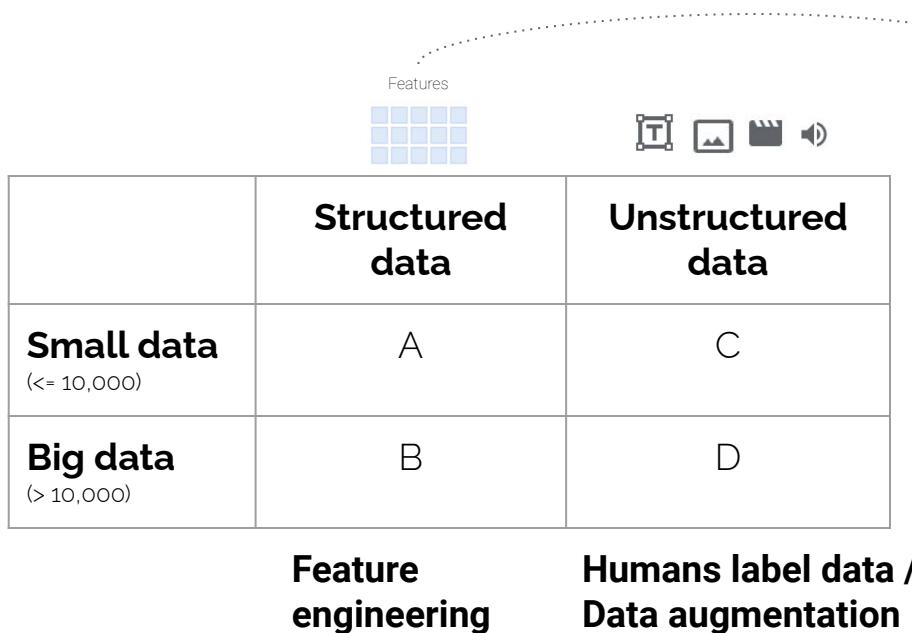
The image is a collage of five screenshots from academic research websites:

- NeurIPS | 2021**: The homepage of the Thirty-fifth Conference on Neural Information Processing Systems. It features a sidebar with links like "Year (2021)", "Help", "Profile", "Contact NeurIPS", "Sponsor Info", "Ethics Guideline", "Publications", "Future Meetings", "Video Archives", "Diversity & Inclusion", "New in ML", "Code of Conduct", "About Us", and "NeurIPS Blog". The main content area shows a "17.861 Impact Factor".
- IEEE Xplore®**: A screenshot of the IEEE Xplore digital library interface, showing a search bar and navigation menus for "Browse", "My Settings", "Help", and "Institutional S...".
- Cornell University**: The Cornell University logo and name.
- arXiv.org > stat > stat.ML**: The arXiv.org page for the statistics section, specifically for machine learning. It includes a "Machine Learning" heading and a "Authors and titles for recent submissions" section.
- IEEE Trans. Pattern Anal. Mach. Intell.**: The IEEE Transactions on Pattern Analysis and Machine Intelligence homepage, featuring a "Machine Learning" section.

	Steel defect detection	Solar panel	Surface inspection
Baseline	76.2%	75.68%	85.05%
Model-centric	+0% (76.2%)	+0.04% (75.72%)	+0.00% (85.05%)
Data-centric	+16.9% (93.1%)	+3.06% (78.74%)	+0.4% (85.45%)

Model optimization vs. data centric approach

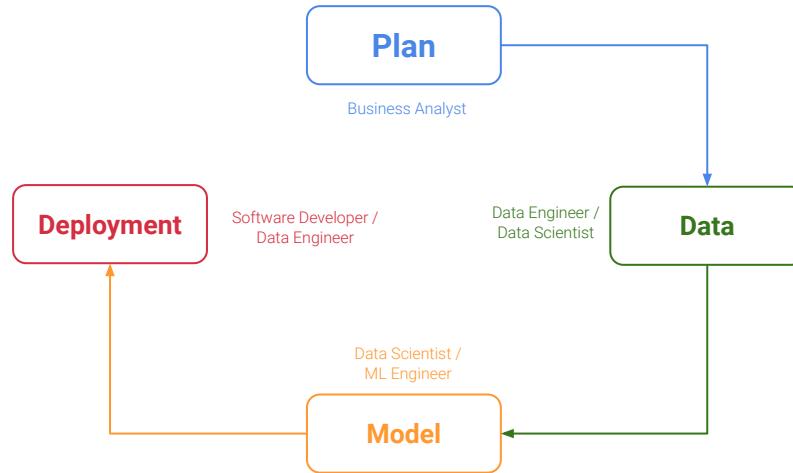
Major types of ML projects



A "feature" refers to the entire column in the dataset

Transaction_id	in_foreign_country	size_compared_to_avg_transaction	fraud?
7485	False	0.8x	False
46854	True	21.2x	True
3521	True	1.1x	False

A "feature value" refers to a single value of a feature column



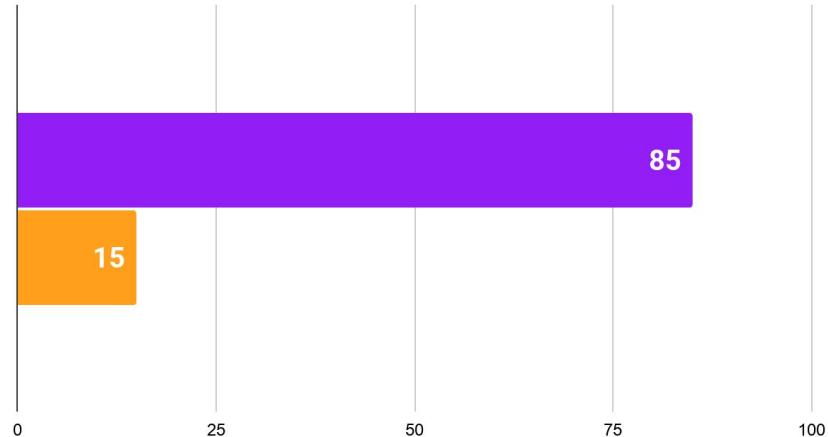
01

Proof of Concept Factory

- **Siloed operating model typically IT-led**
- **Unable to extract value from their data**
- **Struggle to scale as unrealistic expectations on time required**
- **Significant under investment, yielding low returns**

80-85%

PoC Factory Scaling AI



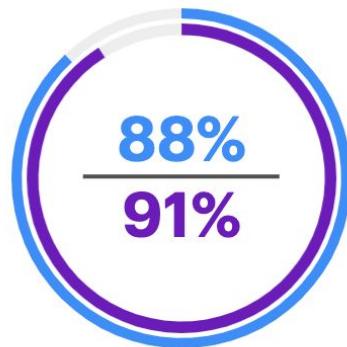
How crucial is Scaling to Your Business?

UNITED STATES

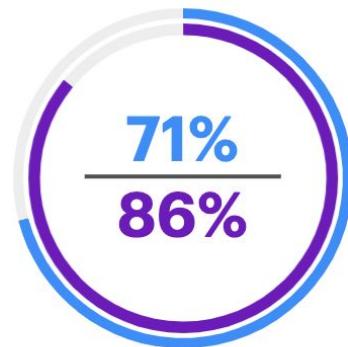


vs.

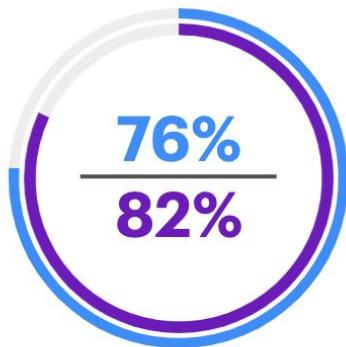
GERMANY



of executives say they won't achieve their growth objectives without scaling AI.

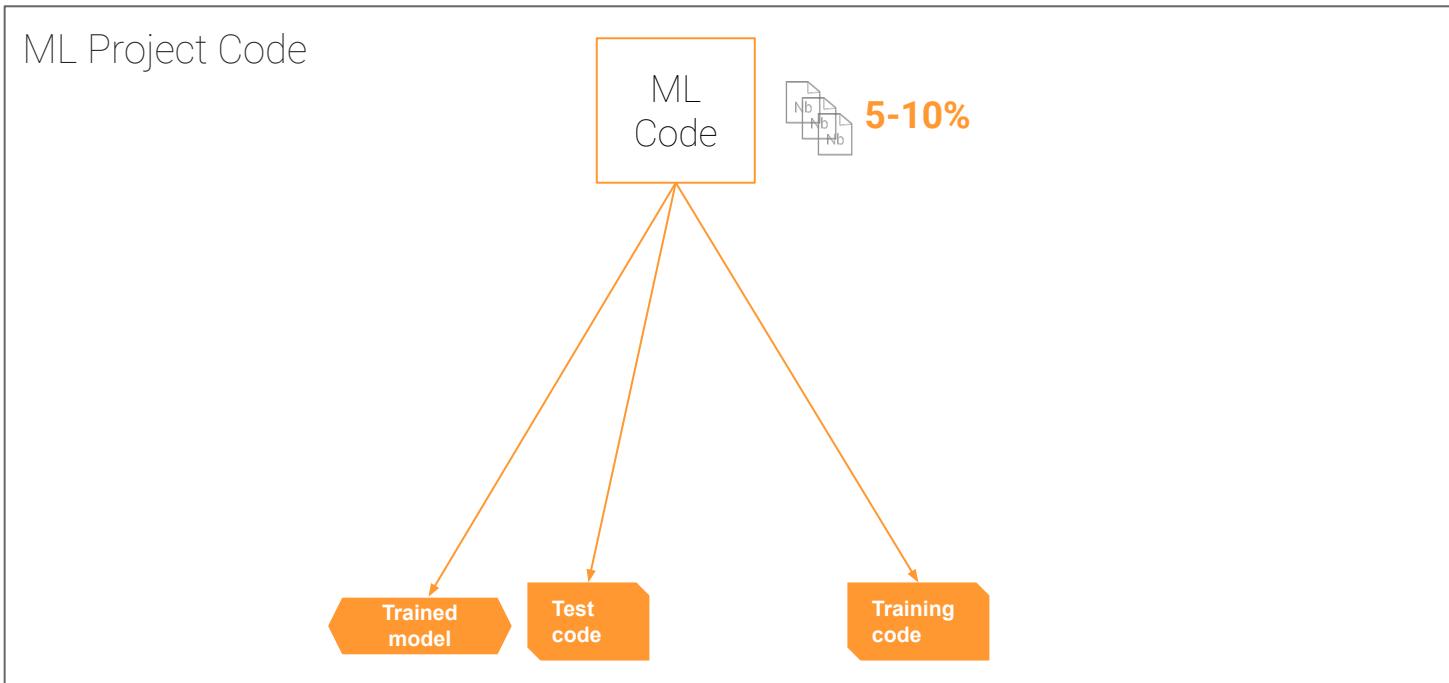


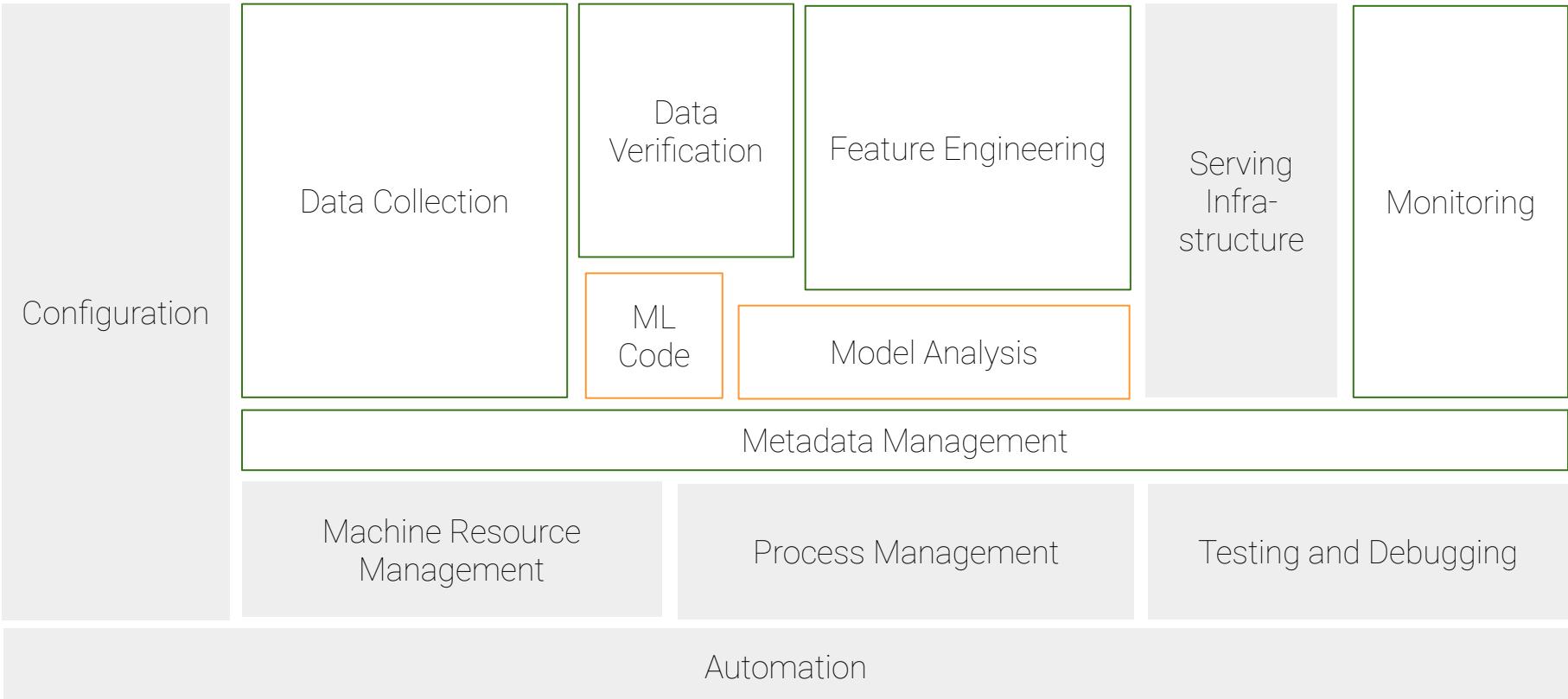
of executives believe they risk going out of business in 5 years if they don't scale AI.



of executives acknowledge they know how to pilot, but struggle to scale AI across the business.

Proof-of-Concept (PoC) to Production Gap



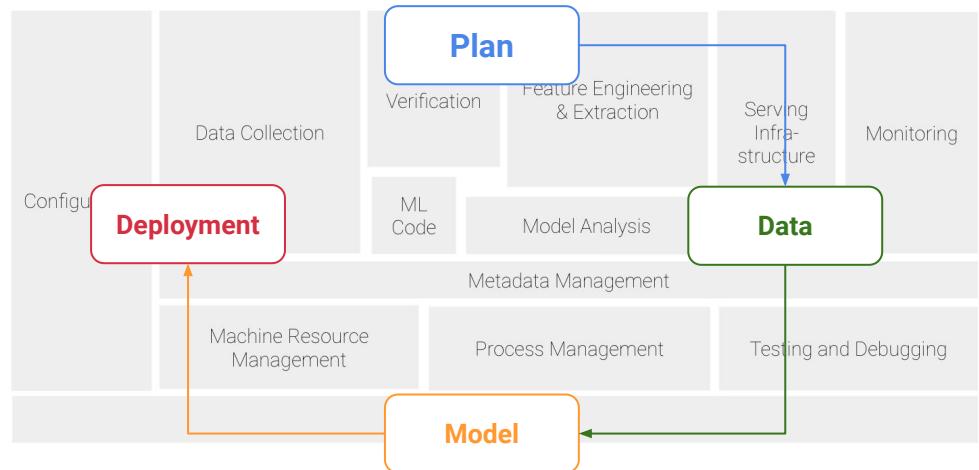


Hidden technical debt in machine learning systems

An **ML Platform** is a horizontal technology (i.e. not specific to a vertical use case) that provides all of the capabilities to cover the full lifecycle of ML applications

Machine learning operations (MLOps)

- Emerging discipline
- Tools and principles to support progress through the ML project lifecycle
- Enables standardization and automation



Machine Learning Operations (MLOps)

Gartner Top 10 Data and Analytics Trends, 2021



gartner.com/SmarterWithGartner

Source: Gartner
© 2021 Gartner, Inc. All rights reserved. CTMKT_1164473

Gartner

Machine Learning Operations (MLOps)

- <**MLOps**> is a ML engineering culture
and practice that aims at
unifying

<ML system development> (Dev) and
<ML system operations> (Ops)

Development and Operations (DevOps)

- <**DevOps**> is a software engineering culture and practice that aims at **unifying** <software development> (Dev) and <software operations> (Ops).

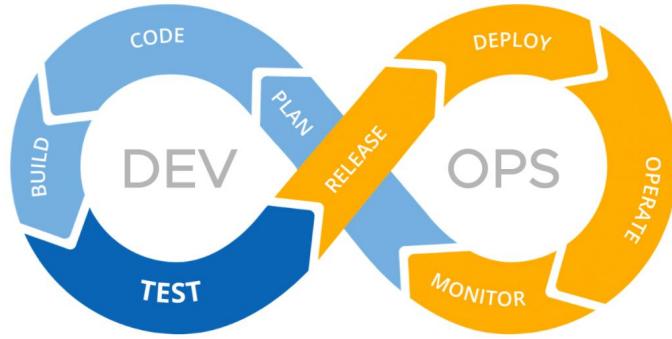


Image credit:: VMWare

Development and Operations (DevOps)

- <DevOps> is to advocate
automation and ***monitoring***
at all steps of <**software**>
construction,
including integration, testing, releasing,
deployment and infrastructure
management.

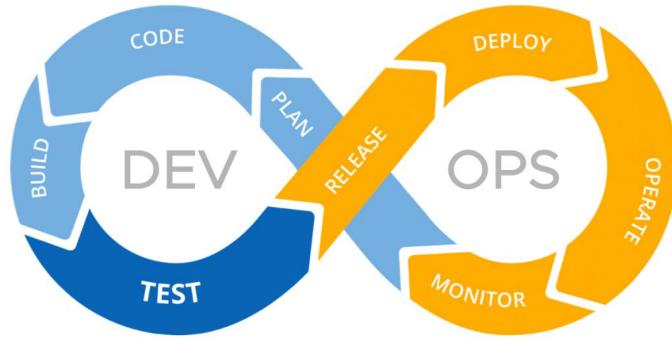


Image credit:: VMWare

Development and Operations (DevOps)

- <MLOps> is to advocate
automation and ***monitoring***
at all steps of <***ML platform***>
construction,
including integration, testing, releasing,
deployment and infrastructure
management.

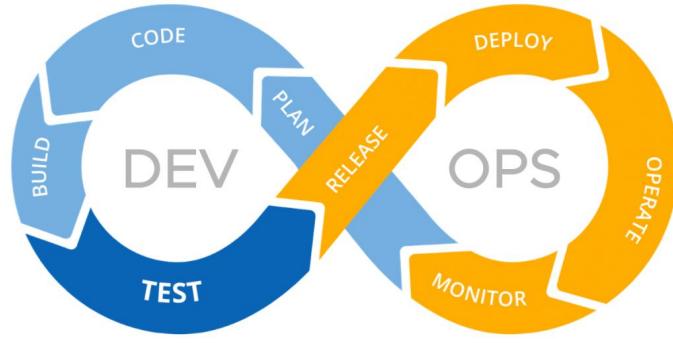


Image credit:: VMWare

DevOps vs MLOps

- **DevOps** = Code
- **MLOps** = Code (Model) + **Data (Features)**

The importance of data (features)

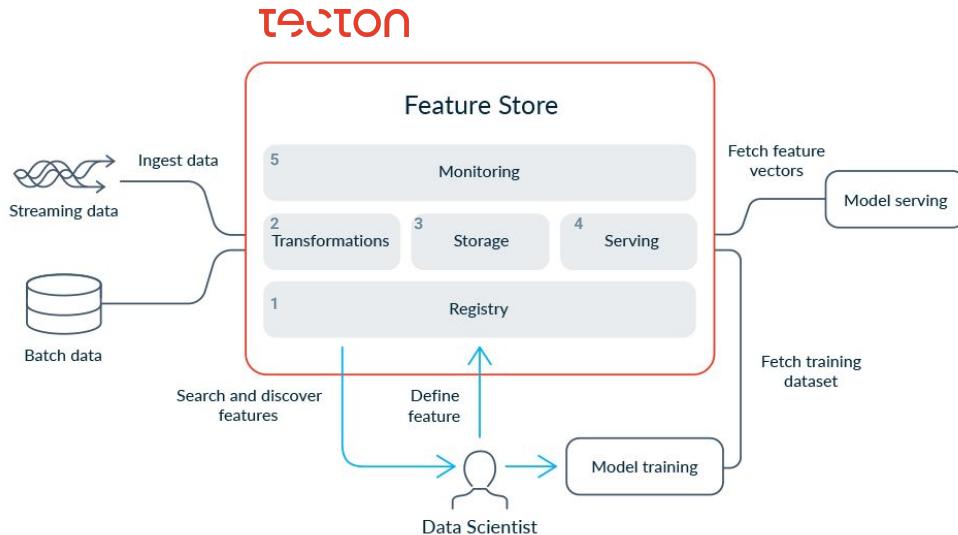
- ML platforms are mainly concerned with data management
- Selecting and transforming features at training time is time consuming
- Pipelines to deliver features to production models are crucial
- Broken data is a common cause of problems in production ML systems



Feature Store

Data management and orchestration system:

1. Register and search features
2. Pipelines for transforming raw data into features or labels
3. Store for storing historical feature and label data
4. Serves feature data consistently for training and inference purposes
5. Monitors production features to detect breakages and drift





Open source solution

Collaborate with **tecton**

Feast (Feature Store)

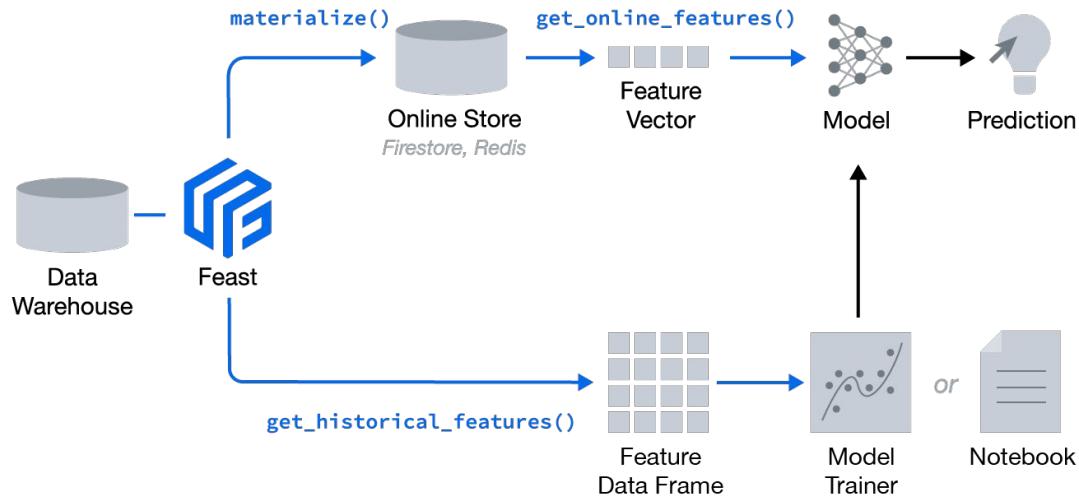
Feast (Feature Store)

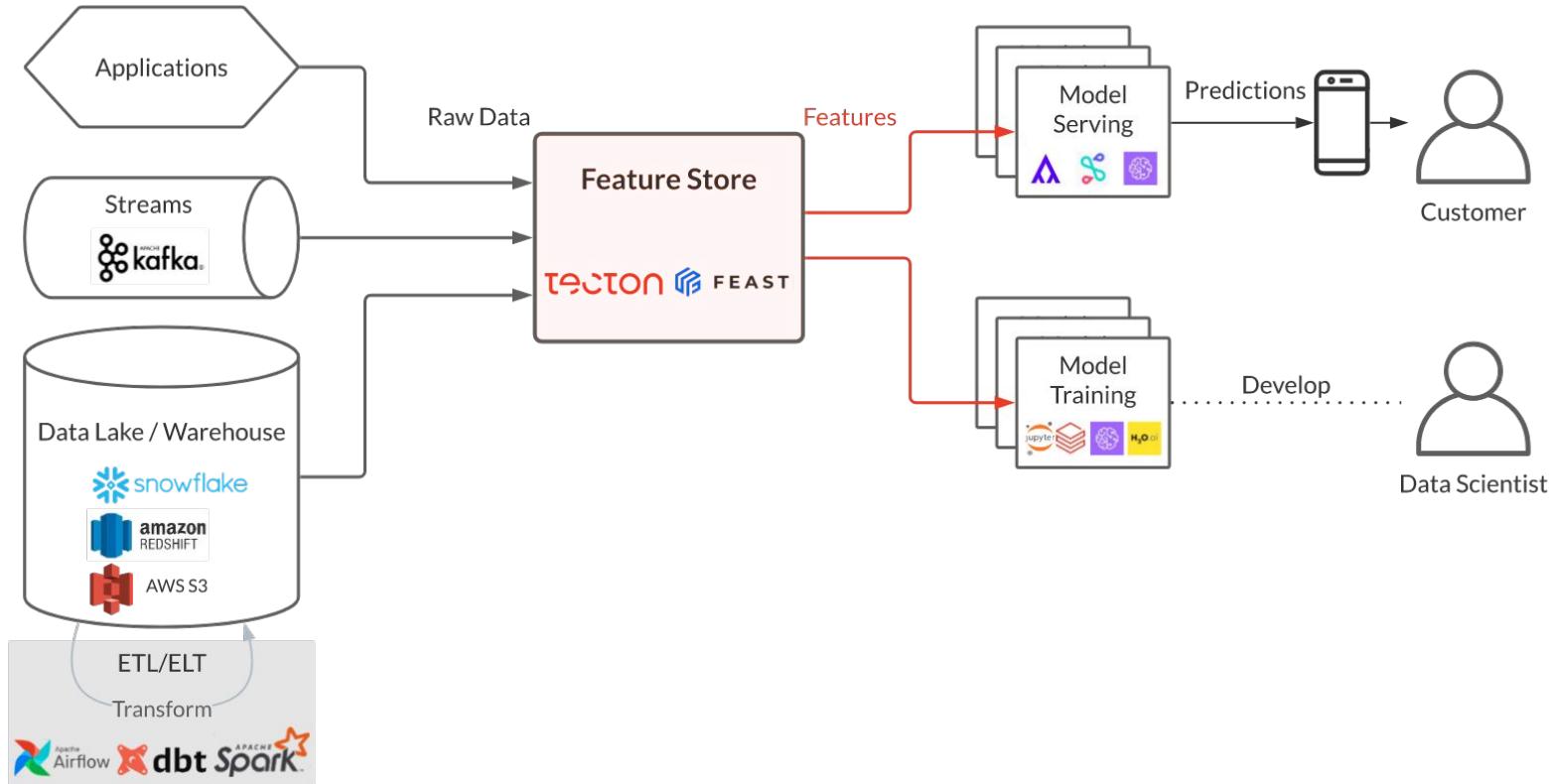
Problems Feast **solves**

- Operational data system for managing and serving machine learning features to models in production

What Feast is **not**

- A general purpose data transformation or pipelining system





Architecture (Tecton/FEAST)

ML Platform Solutions

Internal ML Platforms (Open Source)

NETFLIX



Google



ML Platform solutions (Open Source)

Feature Stores



Internal ML Platforms

Uber



facebook

FBLearner Flow

ML Platform solutions (Enterprise)

Feature Stores



mlflow™

2M+
monthly downloads

200+
code contributors

100+
contributing organizations

1M runs
created per week on Databricks

100k models
registered per week on Databricks

4x YoY
usage growth



TensorFlow



PyTorch



Keras



Apache Spark



sklearn

RAPIDS



H2O.ai



python™



R



Java



mleap



ONNX



GLUON

XGBoost

LightGBM

spaCy

fast.ai



statsmodels



CatBoost



pCARET



ALGORITHMIA



OPTUNA



RAY



CONDA



kubernetes



docker



Amazon SageMaker



Azure Machine Learning



Google Cloud



databricks

mlflow from Databricks



MLflow

MLflow Tracking

Record and query experiments: code, data, config, and results

[Read more](#)

MLflow Projects

Package data science code in a format to reproduce runs on any platform

[Read more](#)

MLflow Models

Deploy machine learning models in diverse serving environments

[Read more](#)

Model Registry

Store, annotate, discover, and manage models in a central repository

[Read more](#)

MLflow Tracking

MLflow Tracking

Record and query experiments: code, data, config, and results

[Read more](#)

```
import mlflow.xgboost  
mlflow.xgboost.autolog()
```

```
import mlflow  
mlflow.log_param("learning_rate", 0.001)
```



Parameters



Metrics



Artifacts



Metadata



Models

MLflow Tracking

MLflow Tracking

Record and query experiments: code, data, config, and results

[Read more](#)



Parameters



Metrics



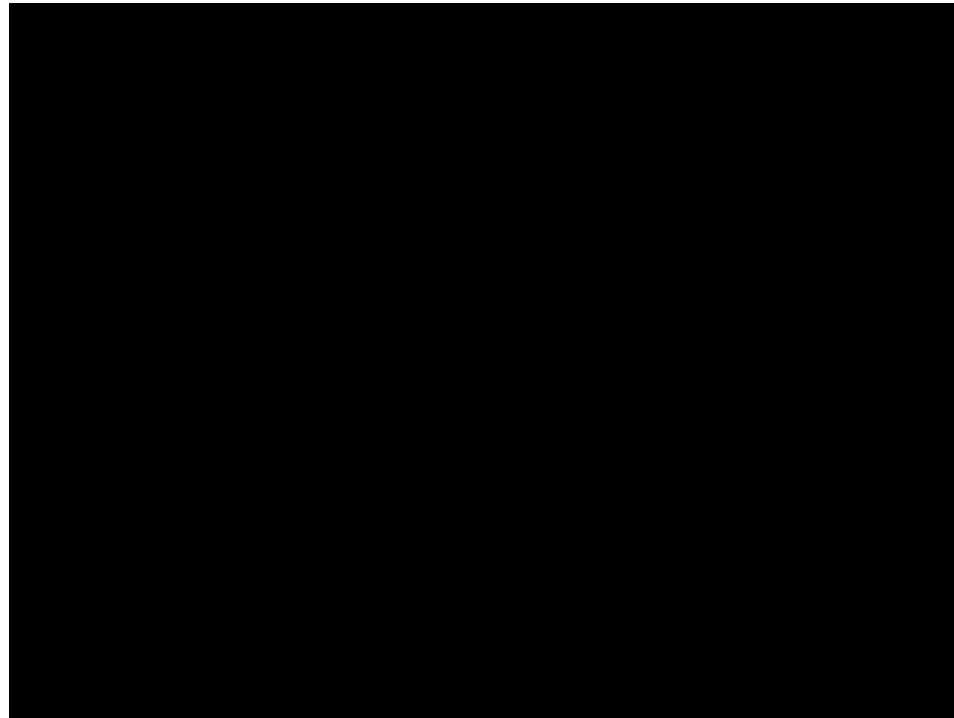
Artifacts



Metadata



Models



MLflow Projects

MLflow Projects

Package data science code
in a format to reproduce runs
on any platform

[Read more](#)

```
mlflow run https://github.com/
```

optional:

```
name: My Project

conda_env: my_env.yaml
# Can have a docker_env instead of a conda_env, e.g.
# docker_env:
#   image: mlflow-docker-example

entry_points:
  main:
    parameters:
      data_file: path
      regularization: {type: float, default: 0.1}
      command: "python train.py -r {regularization} {data_file}"
validate:
  parameters:
    data_file: path
    command: "python validate.py {data_file}"
```

MLflow Models

MLflow Models

Deploy machine learning models in diverse serving environments

[Read more](#)

```
import mlflow.sklearn  
mlflow.sklearn.log_model(model, "myModel")
```

The screenshot shows the Databricks workspace interface. On the left, a sidebar menu includes Home, Workspace, Recents, Data, Clusters, Jobs, Models, and Search. The main area displays two code cells in a notebook titled 'VADER (Python)'. The first cell contains the code for logging a scikit-learn model:1 mlflow.pyfunc.log_model()

```
2 artifact_path="vader",  
3 python_model=model,  
4 conda_env=conda_env,  
5 registered_model_name='text2sentiment',  
6 signature=get_model_signature(model),  
7 )
```

Output from Cmd 12 shows the model registered successfully:Successfully registered model 'text2sentiment'.
Created version '1' of model 'text2sentiment'.
Command took 1.72 seconds -- by sueann@databricks.com at 6/7/2020, 11:57:33 PM on Spark3

The second cell contains the code for getting the model's signature:1 get_model_signature(model)

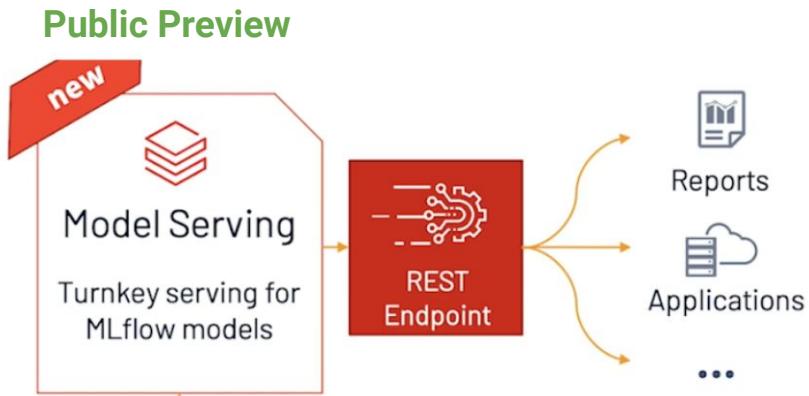
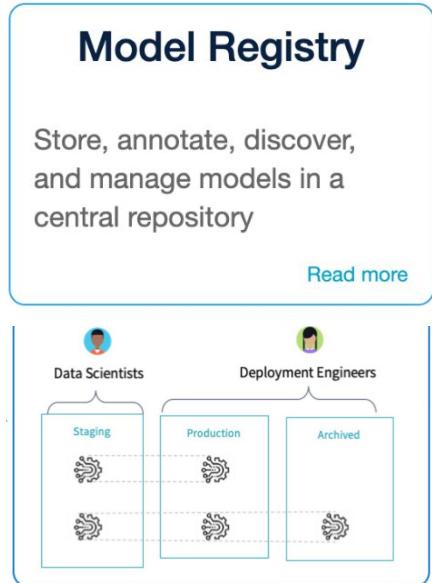
Output from Cmd 13 shows the inputs and outputs defined for the model:Out[51]: inputs:
['text': string]
outputs:
[string]
Command took 0.02 seconds -- by sueann@databricks.com at 6/7/2020, 11:57:16 PM on Spark3

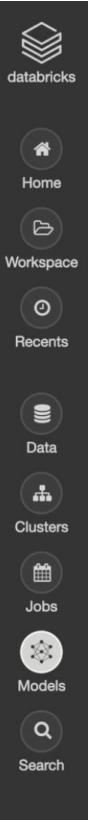
On the right side of the interface, there is a 'Runs' panel showing a single run details:

Run ID	Date	Details
2020-06-07 23:56:23 PDT	Date	alg: vaderSentiment, ... accuracy: 0.758

At the bottom, a note says: 'Showing 1 runs, for more information go to Experiment UI'.

MLflow Registry & Serving





Registered Models > text2sentiment > Version 2 ▾

Registered At: 2020-06-08 00:06:56

Last Modified: 2020-06-08 00:07:06

Creator: [REDACTED] n

Source Run: Run 6fb7c64e7765450988ab37a7735265c

Stage: **Staging** ↕

Request transition to → None

Request transition to → Production

Request transition to → Archived

Transition to → None

Transition to → Production

Transition to → Archived

▼ Description

TorchMoji - a PyTorch implementation of DeepMoji, a BiLSTM-based text to emoji translation model.

(DeepMoji)

Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann.

Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion at Conference on Empirical Methods in Natural Language Processing (EMNLP), 2017.

▼ Pending Requests

Request	Request by	Actions
No pending request.		

▼ Activities

s.com applied a stage transition **None** → **Staging** 28 minutes ago



databricks



Registered Models > text2sentiment ▾

Details

Serving

Enable realtime model serving behind a REST API interface. This will launch a single-node cluster.

[Enable Serving](#)

Workspace

Recents

Data

Clusters

Jobs

Models

Search

xhr: 2, ws: 121, bytes: 67830, xhr-open: 0, ws-open: 1

[OPS] Debug Metrics Requests



databricks



Registered Models > text2sentiment ▾

Details Serving

Status: ● Ready - Stop

Cluster: mlflow-model-text2sentiment ?

Model Versions

Model Events

Model Versions

Version 1	● Ready
	Production
Version 2	● Ready
	Staging
Version 3	● Ready
	None

Model URL: ?

<https://pr-58078.dev.databricks.com/model/text2sentiment/1/invocations><https://pr-58078.dev.databricks.com/model/text2sentiment/Production/invocations>

Call the model

Request ?

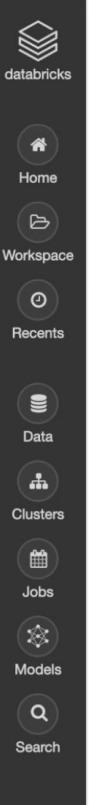
Response ?

Send Request

Show Example

xhr: 3, ws: 534, bytes: 242952, xhr-open: 0, ws-open: 1

[OPS] Debug Metrics Requests



The sidebar contains the Databricks logo at the top, followed by a vertical list of icons with labels: Home, Workspace, Recents, Data, Clusters, Jobs, Models, and Search.

Model 1

Model 2

Model 3



Model Version:	1	2	3
Run Name:			
Start Time:	2020-06-07 23:56:23	2020-06-07 23:40:54	2020-06-07 23:41:03
Parameters			
alg	"vaderSentiment"	"torchmoji"	"torchmoji++"
num_sentiments	2	64	64
Schema			
<input type="checkbox"/> Inputs		<input type="checkbox"/> Ignore column ordering	<input checked="" type="checkbox"/> Show diff only
inputs [0]	text: string	text: string	text: string
inputs [1]	-	-	full_text: string
inputs [2]	-	-	domain: string
<input type="checkbox"/> Outputs			
outputs [0]	undefined: string	undefined: string	undefined: string
Metrics			
accuracy	0.758	0.873	0.908



databricks



Registered Models > text2sentiment ▾

[Details](#)[Serving](#)Status: Ready - Stop

Cluster: mlflow-model-text2sentiment ?

[Model Versions](#)[Model Events](#)

Model Versions

Version 2

Ready

[Production](#)

Version 3

Ready

[None](#)

Model URL: ?

<https://pr-58078.dev.databricks.com/model/text2sentiment/2/invocations><https://pr-58078.dev.databricks.com/model/text2sentiment/Production/invocations>

Call the model

Request ?

[{"text":"Cute puppy!"}]

Response ?

[{"0":"heart_eyes relaxed"}]

[Send Request](#)[Show Example](#)

Easy deployment of models



Model versioning and collaboration



Seamless application updates





Kubeflow

KubeFlow



TensorFlow
Extended



[Kubeflow on AWS](#)
[Kubeflow on Azure](#)
[Kubeflow on GCP](#)
[Kubeflow on IBM Cloud](#)
[Kubeflow Operator](#)
[Kubeflow on OpenShift](#)

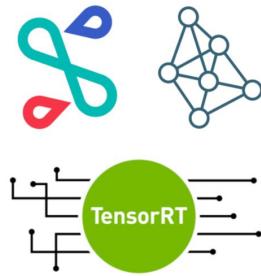
ML toolkit for Kubernetes



Notebooks



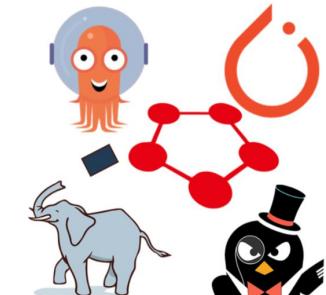
TensorFlow model training



Model serving



Pipelines



Multi-framework



Home

Pipelines

Notebook Servers

Katib

Artifact Store

Manage Contributors

GitHub Privacy • Usage Reporting
build version 0.7.0

Dashboard

Activity

Quick shortcuts

 Upload a pipeline

Pipelines

 View all pipeline runs

Pipelines

 Create a new Notebook server

Notebook Servers

 View Katib Studies

Katib

 View Metadata Artifacts

Artifact Store

Recent Notebooks

No Notebooks in namespace `kubeflow-sarahmaddox`

Recent Pipelines

 [Sample] Basic - Exit Handler

Created 22/12/2019, 06:50:18

 [Sample] Basic - Conditional execution

Created 22/12/2019, 06:50:17

 [Sample] Basic - Parallel execution

Created 22/12/2019, 06:50:16

 [Sample] Basic - Sequential execution

Created 22/12/2019, 06:50:15

 [Sample] ML - XGBoost - Training with ...

Created 22/12/2019, 06:50:14

Documentation

Getting Started with Kubeflow

Get your machine-learning workflow up and running on Kubeflow 

MiniKF

A fast and easy way to deploy Kubeflow locally 

MicrOK8s for Kubeflow

Quickly get Kubeflow running locally on native hypervisors 

Minikube for Kubeflow

Quickly get Kubeflow running locally 

Kubeflow on GCP

Running Kubeflow on Kubernetes Engine and Google Cloud Platform 

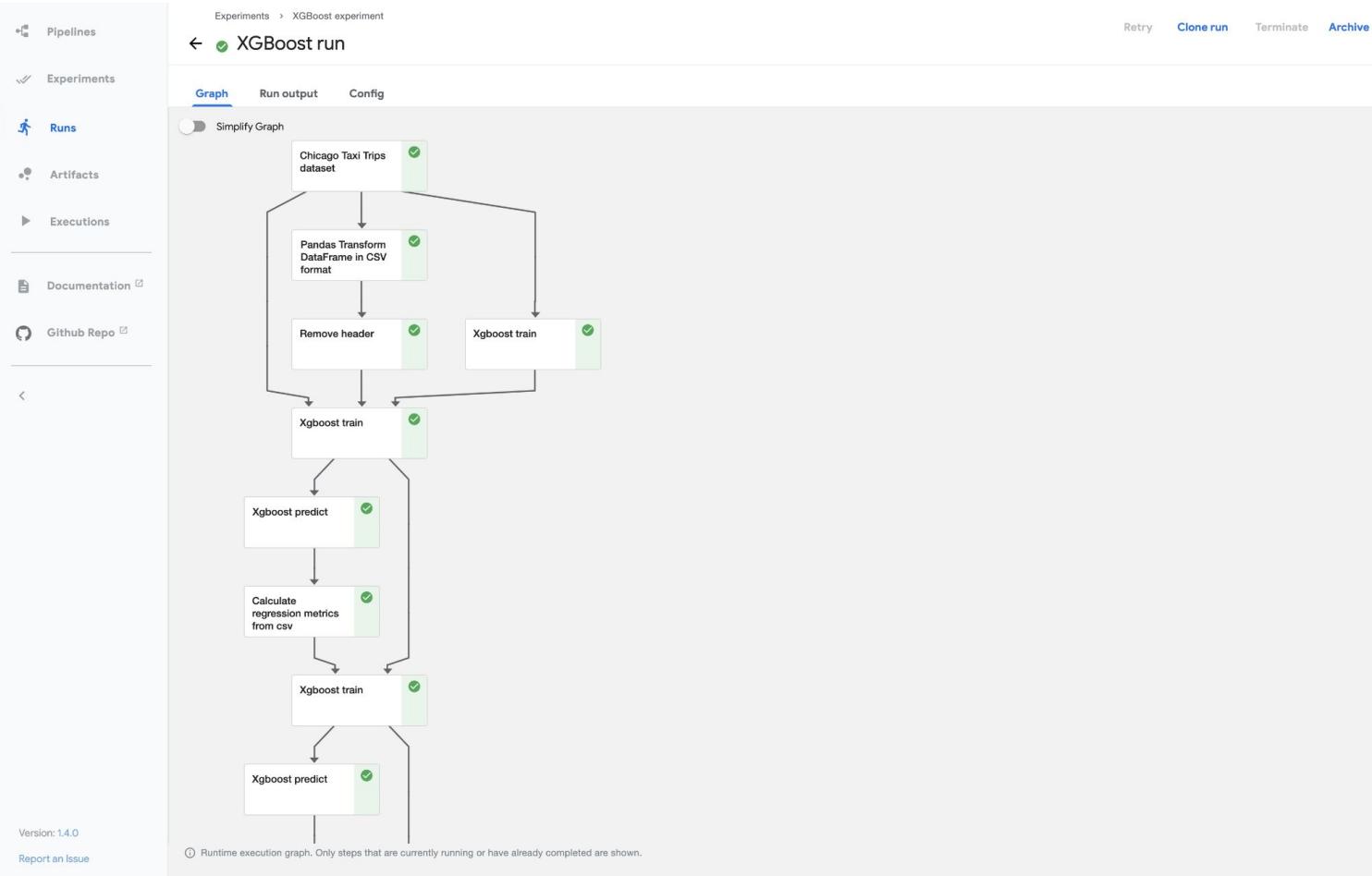
Kubeflow on AWS

Running Kubeflow on Elastic Container Service and Amazon Web Services 

Requirements for Kubeflow



Kubeflow user interface (UI)





The image shows the Kale Deployment Panel integrated with a Jupyter Notebook environment. On the left, the Kale panel has an 'Enable' toggle switch, a 'Pipeline Metadata' section with 'Select experiment' set to 'Titanic', 'Pipeline Name' as 'titanic-ml', and a 'Pipeline Description' of 'Predict which passengers survived the titanic'. It also includes sections for 'Volumes' (with toggles for 'Use this notebook's volumes' and 'Take Rok snapshots before each step') and 'Advanced Settings'. At the bottom is a blue 'COMPILE AND RUN' button. To the right is a Jupyter notebook tab titled 'titanic_dataset_ml.ipynb'. The notebook contains code for importing dependencies and loading data:import numpy as np
import pandas as pd
import seaborn as sns
from matplotlib import pyplot as plt
from matplotlib import style

from sklearn import linear_model
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import RandomForestClassifier
from sklearn.linear_model import Perceptron
from sklearn.linear_model import SGDClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.naive_bayes import GaussianNB

Below this is a step configuration for 'loaddata':

```
step: loaddata  
Cell type: Pipeline Step  
Step name: loaddata  
Depends on:
```

```
[2]: path = "data/"  
  
PREDICTION_LABEL = 'Survived'  
  
test_df = pd.read_csv(path + "test.csv")  
train_df = pd.read_csv(path + "train.csv")
```

The notebook also includes a section titled 'Let's explore the data' with the following text:

These are features of the dataset:

```
survival: Survival  
PassengerId: Unique Id of a passenger.  
pclass: Ticket class
```

Kale (Kubeflow Automated pipeLines Engine)



Katib

☰ Kubeflow anonymous ▾



☰ Katib

Welcome to Katib

Choose type of experiment

Hyperparameter
Tuning

Neural Architecture
Search

For usage instructions, see the [Kubeflow docs](#)

To contribute to Katib, visit [GitHub](#)

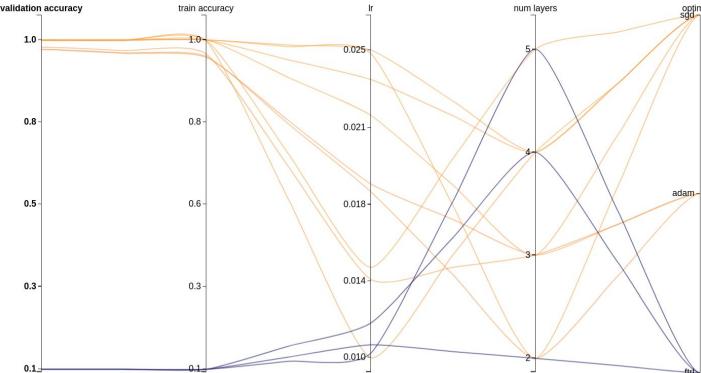
Katib (hyperparameter tuning)



Katib

Experiment details

DELETE



OVERVIEW

TRIALS

DETAILS

YAML

Name bayesianoptimization-example

Status ✓ Experiment has succeeded because max trial count has reached

Best trial bayesianoptimization-example-6935d45b

Best trial's params lr: 1.44659e-2 num-layers: 5 optimizer: sgd

Best trial performance Validation-accuracy: 0.97939 Train-accuracy: 0.99287

User defined goal Validation-accuracy > 0.99

Running trials 0

Failed trials 0

Succeeded trials 12

Experiment Conditions

Status	Type	Last Transition Time	Reason	Message
✓	Created	1 month ago	ExperimentCreated	Experiment is created
▲	Running	1 month ago	ExperimentRunning	Experiment is running
✓	Succeeded	1 month ago	ExperimentMaxTrialsReached	Experiment has succeeded because max trial count has reached

Katib (hyperparameter tuning)



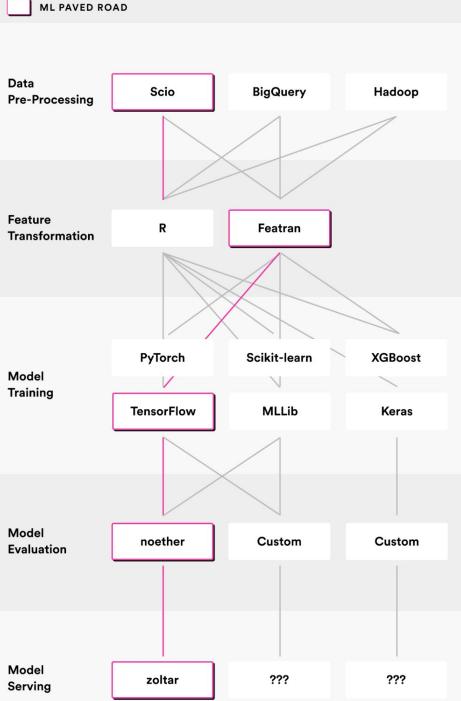
Spotify®

MLops @ Spotify | Powered by:

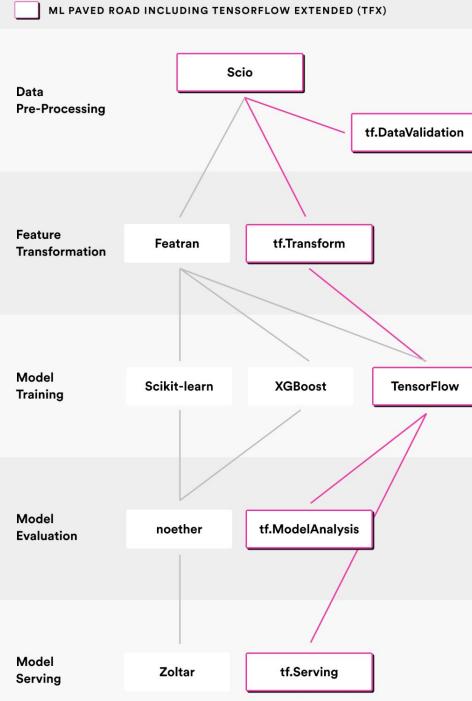




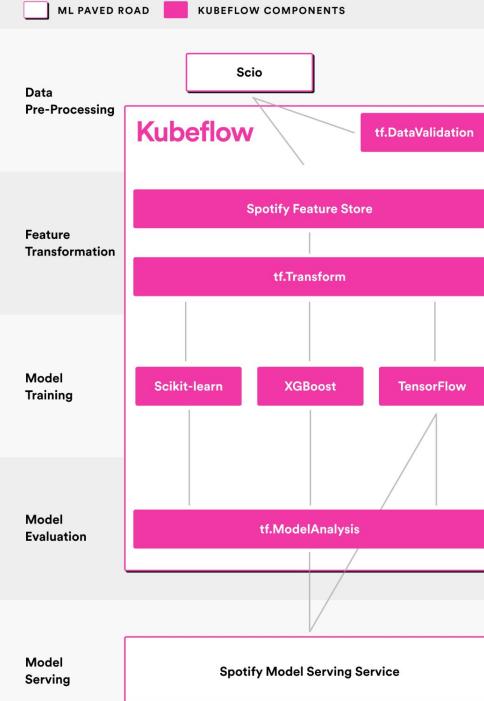
ML Platform 2018



ML Platform 2019



ML Platform 2020



B
beam

Scio

Kubeflow

TensorFlow Extended

learn dmlc XGBoost

More resources about Spotify & Kubeflow

Baer, J. (2019). Kubeflow at Spotify.

https://docs.google.com/presentation/d/12dwhS_x4568G6XQjl9SEUacD-n4hFQczBcRBLdbHNEM/edit#slide=id.g5182096320_0_52

Baer, J. & Samuel Ngahane, S. (2019). The Winding Road to Better Machine Learning Infrastructure Through Tensorflow Extended and Kubeflow.

<https://engineering.at spotify.com/2019/12/13/the-winding-road-to-better-machine-learning-infrastructure-through-tensorflow-extended-and-kubeflow/>

Clough, R. & Dai, K. (2019). Kubeflow @ Spotify. Building & Managing a Centralized Platform.

https://github.com/sbueringer/kubecon-slides/blob/master/slides/2019-kubecon-na/Building%20and%20Managing%20a%20Centralized%20Kubeflow%20Platform%20at%20Spotify%20-%20Keshi%20Dai%20%26%20Ryan%20Clough%2C%20Spotify%20-%20Kubeflow_at_Spotify_Building_and_Managing_a_Centralized_Platform.pdf



Continuous Integration
and Continuous Delivery
pipeline for an ML/AI
project in Microsoft Azure



Azure DevOps



Azure Boards



Azure Repos



Azure Pipelines



Azure Test Plans



Azure Artifacts

Plan, track, and discuss work across teams, deliver value to your users faster.

Unlimited cloud-hosted private Git repos. Collaborative pull requests. Advanced file management, and more.

CI/CD that works with any language, platform, and cloud. Connect to GitHub, GitLab, and any provider and deploy continuously to any cloud.

The test management and exploratory testing tool that lets you ship with confidence.

Create, host, and share packages. Easily add artifacts to CI/CD pipelines.



Timetracker



Docker Integration



Slack Integration



Sentry

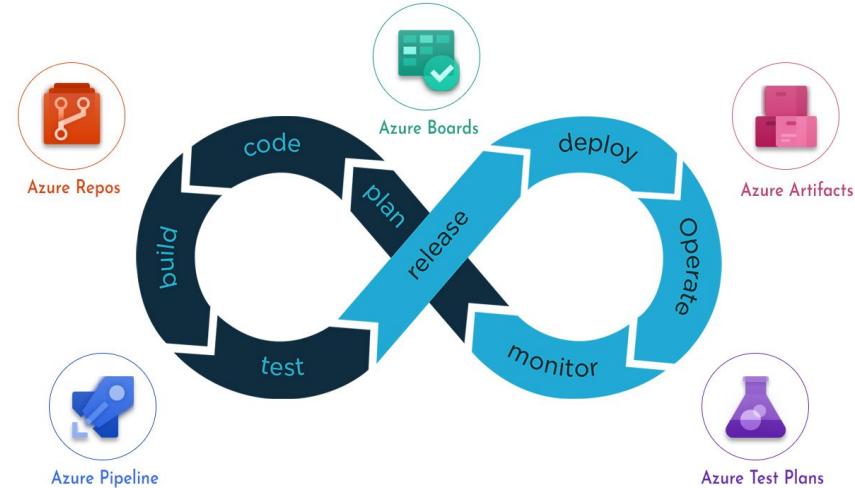


Github Integration

Extensions Marketplace

Access 1,000+ extensions or create your own.

[View all extensions >](#)



CI/CD and source code management



Azure Machine Learning

Microsoft Azure Machine Learning

milprod > Home

Azure Machine Learning studio

+ Create new

Notebooks Start now

Automated ML Start now

Designer Start now

My recent resources

Runs

Run	Run ID	Experiment	Status	Submitted time	Submitted by
Run 1	893f2c6e...	logging-a...	Completed		

Compute

Name: gpu-cluster

Navigation sidebar:

- New
- Home
- Author
- Notebooks
- Automated ML
- Designer
- Assets
- Datasets
- Experiments
- Pipelines
- Models
- Endpoints
- Manage
- Compute
- Datastores
- Data Labeling

Development tools



Languages



Frameworks



mlflow



Machine Learning for all skill levels

Productivity for all skill levels, with Jupyter notebooks, drag-and-drop [designer](#), and [automated machine learning](#).



End-to-end MLOps

Robust [MLOps](#) capabilities that enable creation and deployments of models at scale using automated and reproducible machine learning workflows.



Responsible machine learning innovation

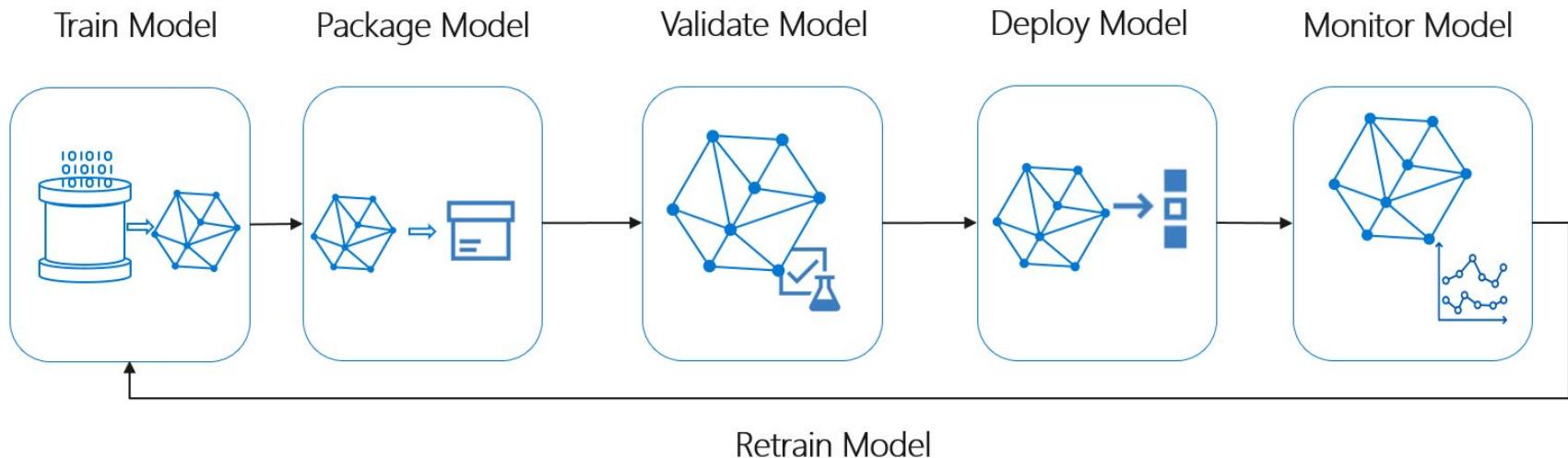
Rich set of built-in [responsible](#) capabilities to understand, protect, and control data, models and processes.

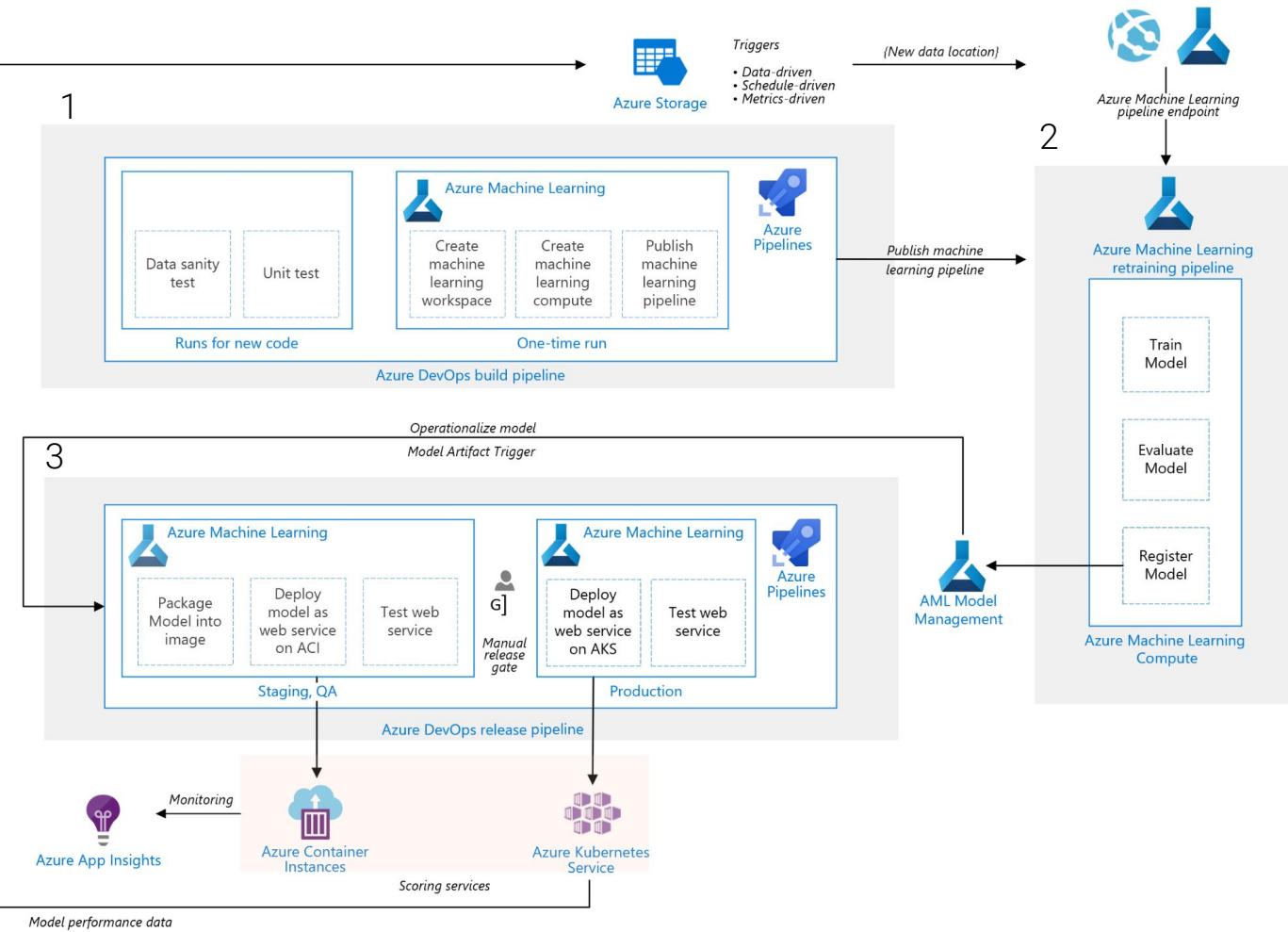


Open and interoperable

Best-in-class support for open-source frameworks and languages including MLflow, Kubeflow, ONNX, PyTorch, TensorFlow, Python, and R.

Cloud based machine learning service





Search or jump to... Watch 55 Unstar 561 Fork 581

[microsoft / MLOpsPython](#) Template

Code Issues 32 Pull requests 3 Actions Projects Wiki Security Insights

master 28 branches 115 tags Go to file Add file Code Use this template

j-so Pin AzureML SDK > 1.18.0 for pyyaml fix (#356) 2892688 on 11 Mar 421 commits

.pipelines Fix CD deployment break (#305) 10 months ago

bootstrap Update docs and pipeline status badge (#303) 10 months ago

charts Canary pipeline fixes (#224) 14 months ago

data Fix Batch Scoring docs (#333) 7 months ago

diabetes_regression Pin AzureML SDK > 1.18.0 for pyyaml fix (#356) 2 months ago

docs Added AKS compute name details (#355) 2 months ago

environment_setup Update docs and pipeline status badge (#303) 10 months ago

experimentation Making changes to experiment notebook based on changes to tutorial... 13 months ago

ml_service Fix default parameters to batch scoring pipeline + doc fixes (#310) 10 months ago

.env.example fix TRAIN_SCRIPT_PATH value in .env.example (#348) 3 months ago

.gitignore Manage environments in conda YAML files (#158) 15 months ago

LICENSE Initial commit 2 years ago

README.md Update docs and pipeline status badge (#303) 10 months ago

README.md

page_type	languages	products	description
sample	python	azure azure-machine-learning-service azure-devops	Code which demonstrates how to set up and operationalize an MLops flow leveraging Azure Machine Learning and Azure DevOps.

MLOps with Azure ML

Ci: Azure Pipelines Succeeded!

MLOpsPython + Pipelines in your region may be impacted by a live site incident, resulting in possible pipeline delays. Check the status here.

Model CI Pipeline

← Jobs in run #20210406.1
kirenz.MLOpsPython (1)

Model CI

- Model CI Pipeline 2m 54s
 - Initialize job 3s
 - Initialize containers 1m 25s
 - Checkout kirenz/MLO... 1s
 - Run lint tests <1s
 - Run unit tests 1s
 - Publish test results 2s
 - Publish coverage report 2s
 - Publish Azure Ma... 1m 14s
 - Post-job: Checkout k... <1s
 - Stop Containers <1s
 - Finalize Job <1s

Train and evaluate model

- > Get Pipeline ID for e... 1m 31s
- > Trigger ML Trainin... 10m 13s
- ✓ Publish artifact if ne... 3m 2s
 - Initialize job 2s
 - Initialize containers 1m 57s
 - Checkout kirenz/MLO... 2s
 - Install AzureML CLI 30s

Project settings <

View raw log

MLOpsPython +

⚠ Pipelines in your region may be impacted by a live site incident, resulting in possible pipeline delays. Check the status here.

← Jobs in run #20210406.1

> Model CI Pipeline 2m 54s

Train and evaluate model

> Get Pipeline ID for e... 1m 31s

- Initialize job 2s
- Initialize containers 1m 13s
- Checkout kirenz/MLO... 2s
- Get Pipeline ID 11s
- Post-job: Checkout k... <1s
- Stop Containers <1s
- Finalize Job <1s

> Trigger ML Trainin... 10m 13s

- Invoke ML pipeli... 10m 13s

> Publish artifact if ne... 3m 2s

- Initialize job 2s
- Initialize containers 1m 57s
- Checkout kirenz/MLO... 2s
- Install AzureML CLI 30s
- Determine if evaluati... 20s
- PublishPipelineArtifact 7s
- Post-job: Checkout k... <1s
- Stop Containers 1s
- Finalize Job <1s

View raw log

```
1 Pool: Azure_Pipelines
2 Image: ubuntu-latest
3 Queued: Today at 10:34 [manage_parallel_jobs]
4 Agent: Hosted Agent
5 Started: Today at 10:34
6 Duration: 3m 2s
7
8 The agent request is already running or has already completed.
9 ▶ Job preparation parameters
10 □ 1 artifact produced
11 Job live console data:
12 Starting: Publish artifact if new model was registered
13 Finishing: Publish artifact if new model was registered
```

MLOpsPython



 Pipelines in your region may be impacted by a live site incident, resulting in possible pipeline delays. [Check the status here.](#)

Overview

Boards

Repos

Pipelines

Pipelines

Environments

Releases

Library

Task groups

Deployment groups

Test Plans

Artifacts

Pipelines

[New pipeline](#)[Recent](#)[All](#)[Runs](#)[Filter pipelines](#)

Recently run pipelines

Pipeline

Last run

 Model-Train-Register-CI

#20210406.1 • Changed auto trigger

Manually triggered for  master

21m ago

18m 35s

 kirenz.MLOpsPython

#20210406.1 • Changed auto trigger

Manually triggered for  master

31m ago

1m 20s

[New pipeline](#)

- [MLOpsPython](#)
- [Overview](#)
- [Boards](#)
- [Repos](#)
- [Pipelines](#)
- [Pipelines](#)
- [Environments](#)
- [Releases](#)
- [Library](#)
- [Task groups](#)
- [Deployment groups](#)
- [Test Plans](#)
- [Artifacts](#)

 [Project settings](#) <<

⚠ Pipelines in your region may be impacted by a live site incident, resulting in possible pipeline delays. [Check the status here.](#)

#20210406.1 Changed auto trigger

on Model-Train-Register-CI ✘ Retained

[Run new](#)

⋮

[Summary](#) [Tests](#) [Code Coverage](#)

Manually run by Jan Kirenz

[View change](#)

Repository and version

kirenz/MLOpsPython
` master → c9d5976

Time started and elapsed

Today at 10:18
⌚ 18m 35s

Related

0 work items
 2 published

Tests and coverage

100% passed
 47.44% covered

[Warnings 1](#)

Please install dotnet core to enable automatic generation of Html report.
Model CI - Model CI Pipeline - Publish coverage report

[Stages](#) [Jobs](#)

Model CI

1 job completed 2m 57s
 100% tests passed
 1 artifact

Train and evaluate ...

3 jobs completed 15m 17s
 1 artifact
 Get Pipeline ID for exe...
 Trigger ML Training P...
 Publish artifact if new m...

[Rerun stage](#)



Home > Pipelines

Pipelines

[Pipeline runs](#) [Pipeline endpoints](#) [Pipeline drafts](#)[+ New pipeline](#) Refresh[+ Add filter](#)

Run	Run ID	Experiment	Status	Description	Submitted time ↓
Run 1	f8ea2919-0b2b-4864-8550-fcd0b4b37ce9	mllopspython	Completed		Apr 6, 2021 10:23 AM



Home > Modelle

Modelliste

Modell registrieren Löschen Bereitstellen Aktualisieren

Filter hinzufügen

Name	Version	Experiment	Ausführungs-ID	Erstellt am	Tags
diabetes_regression_model.pkl	1	--		6. Apr. 2021 10:33	area : diabetes_regression ...



M MLOpsPython +

- Overview
- Boards
- Repos
- Pipelines
- Pipelines
- Environments
- Releases
- Library
- Task groups
- Deployment groups
- Test Plans
- Artifacts

⚠ Pipelines in your region may be impacted by a live site incident, resulting in possible pipeline delays. [Check the status here.](#)

← Jobs in run #2021...

kirenz.MLOpsPython (1)

✓ Deploy to ACI



[View raw log](#)

Deploy to ACI

✓ [Deploy to ACI] 5m 25s

- ✓ Initialize job 2s
- ✓ Initialize co... 1m 20s
- ✓ Checkout kirenz.MLOpsPython 3s
- ✓ Download Pipeline 6s
- ✓ Parse Json for ... <1s
- ✓ Install AzureM... 30s
- ✓ Deploy to ACI 2m 57s
- ✓ Smoke test 21s
- ✓ Post-job: Check... <1s
- ✓ Stop Containers <1s
- ✓ Finalize Job <1s

1 Pool: [Azure Pipelines](#)

2 Image: Ubuntu16

3 Agent: Hosted Agent

4 Started: Today at 10:45

5 Duration: 5m 25s

6

7 ► Job preparation parameters

8 Job live console data:

9 Finishing: Deploy to ACI

MLOpsPython +

Overview

Boards

Repos

Pipelines

Pipelines

Environments

Releases

Library

Task groups

Deployment groups

Test Plans

Artifacts

Pipelines in your region may be impacted by a live site incident, resulting in possible pipeline delays. [Check the status here.](#)

#20210406.1 Changed auto trigger

on kirenz.MLOpsPython (1) ✘ Retained

[Run new](#)

Summary

Manually run by Jan Kirenz

[View 2 changes](#)

Repository and version

 kirenz/MLOpsPython
master ✘ c9d5976

Time started and elapsed

 Today at 10:45
 5m 32s

Related

 0 work items
 2 consumed

Tests and coverage

Warnings 1

Please use Download Build Artifact task for downloading Build Artifact type artifact. <https://docs.microsoft.com/en-us/azure/devops/pipelines/tasks/utility/download-build-artifacts?view=azure-devops&tabs=classic>
Deploy to ACI · Deploy to ACI · Download Pipeline Artifacts

Stages

Jobs

Deploy to ACI

1 job completed

5m 28s

Deploy to AKS

Skipped

Deploy to Webapp

Skipped

[Rerun stage](#)



End of Demo

Continuous Integration
and Continuous Delivery
pipeline for an ML/AI
project in Microsoft Azure

Backup



An open source
machine learning



Enterprise Data Sciences Platform Resources Pricing Company

Login

CREATE A FREE ACCOUNT



MLflow is an open source platform to manage the ML lifecycle. MLflow currently offers four components:

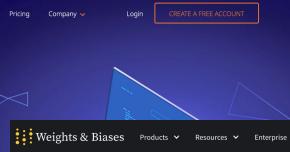
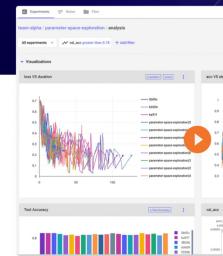
MLflow Tracking

Record and query experiments: code, data, config, and results

[Read more](#)

MLflow Projects

Package data science code in a format to reproduce runs on any platform

[Read more](#)

Build better models faster

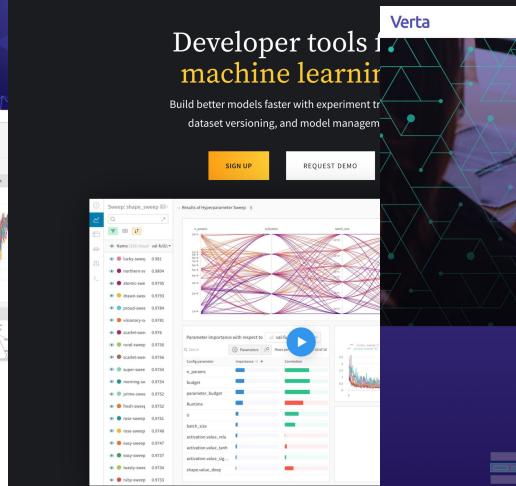
models faster

Comet enables data scientists and teams to track, compare, explain and optimize experiments and models across the model's entire lifecycle. From training to production. With just two lines of code, you can start building better models today.

GET STARTED FOR FREE



Login



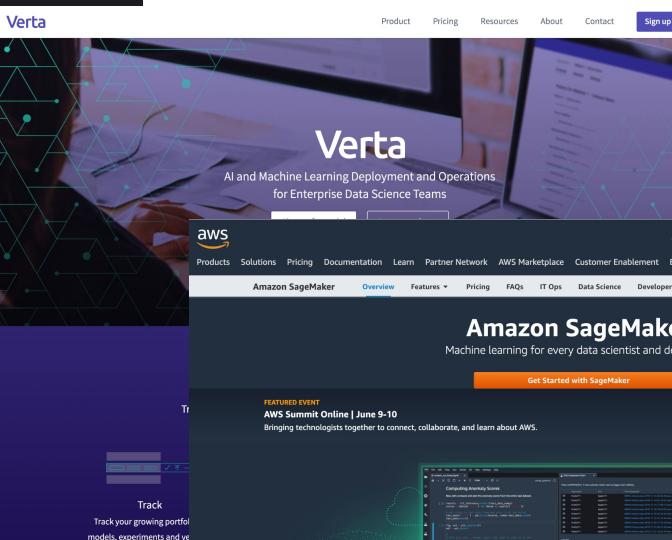
Verta

Developer tools for machine learning

Build better models faster with experiment tracking, dataset versioning, and model management

SIGN UP

REQUEST DEMO



Verta

AI and Machine Learning Deployment and Operations
for Enterprise Data Science Teams



Products Solutions Pricing Documentation Learn Partner Network AWS Marketplace Customer Enablement Events Explore More

Contact Us Support English My Account

Create an AWS Account

Amazon SageMaker

Machine learning for every data scientist and developer

Get Started with SageMaker

FEATURED EVENT

AWS Summit Online | June 9-10

Bringing technologists together to connect, collaborate, and learn about AWS.

Register now >



Comparison of 15 tools (April 2020)

	MLflow	TensorBoard	Guild AI	PolyAxon	TRAINS
Overview					
Focus	Entire Lifecycle	Experiment Management	Experiment Management	Experiment Management	Experiment Management
Price	Free	Free	Free	Enterprise: NA	Free
Free plan limitation					
Open – source	✓	✓	✓	limited	✓
Easy integration	✓	✓	✓		
Lightweight	✓	✓	✓		✓
Experiment Tracking Features					
Data Versioning			limited	✓	
Notebook Versioning					
Model Versioning	limited		limited	✓	limited