CODING ANGELS

專屬全台女大專院校生的科技工作坊

線上工作坊：2/28 ~ 3/4, 2022

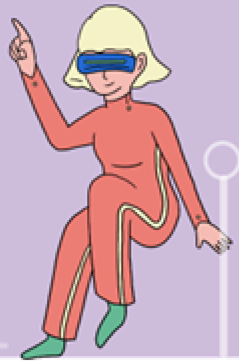
實體工作坊：3/5 & 6, 2022

專屬全台女大專院校生的科技工作坊

講師：Yvonne & Jolin

Agenda

- How Machine Learning works on Azure
- Azure Cognitive Services
- Low Code your AI Apps with Power Apps



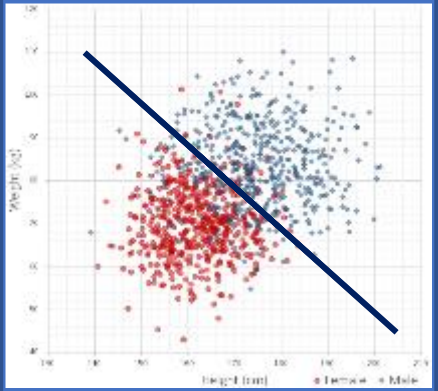
What is Machine Learning ?

Using **known data**, develop a **model** to predict **unknown data**.

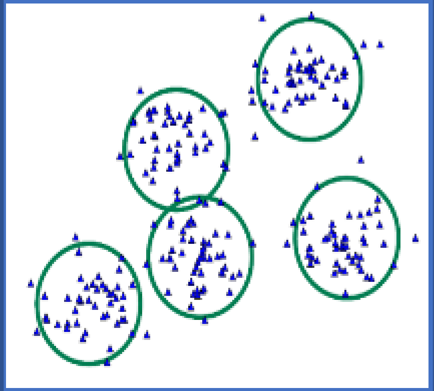
Common Classes of Algorithms

(Supervised|Unsupervised)

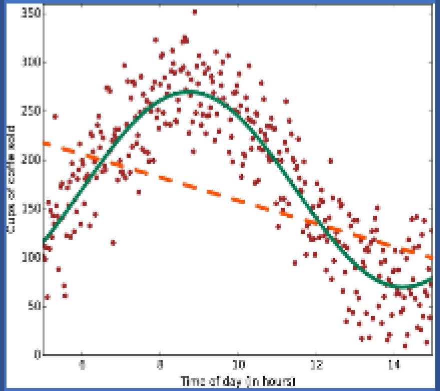
Classification



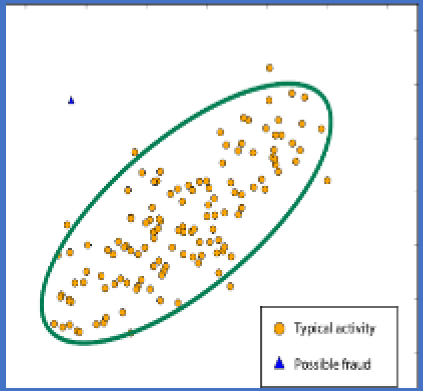
Clustering



Regression



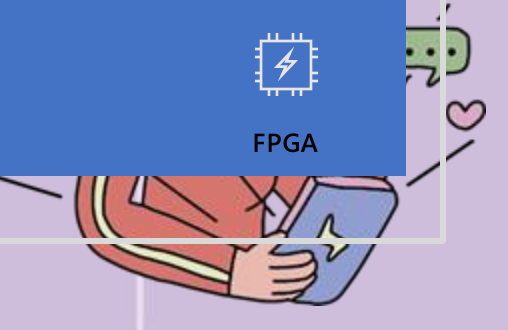
Anomaly Detection



Machine Learning on Azure

Domain specific pretrained models To reduce time to market	Vision	Speech	Language	Search
Familiar Data Science tools To simplify model development	PyCharm	Jupyter	Visual Studio Code	Command line
Popular frameworks To build advanced deep learning solutions	Pytorch	TensorFlow	Scikit-Learn	Onnx
Productive services To empower data science and development teams	Azure Databricks	Azure Machine Learning		Machine Learning VMs
Powerful infrastructure To accelerate deep learning	CPU	GPU		FPGA

From the Intelligent Cloud to the Intelligent Edge



Azure AI

AI apps and agents

Azure Cognitive Services

Azure Bot Service

Knowledge mining

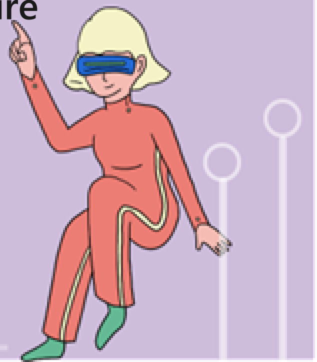
Azure Search

Machine learning

Azure Databricks

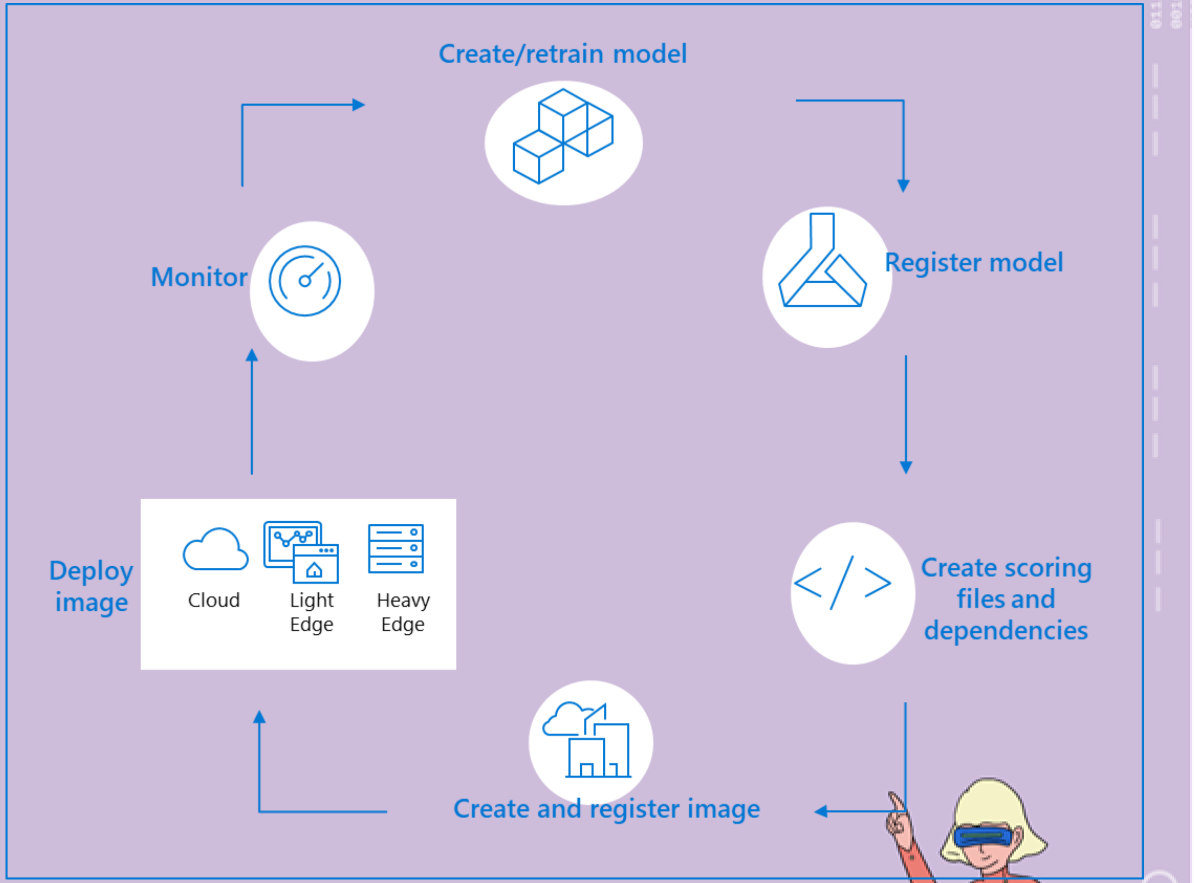
Azure Machine Learning

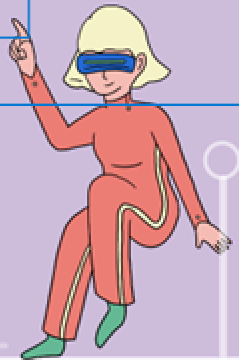
Azure AI Infrastructure



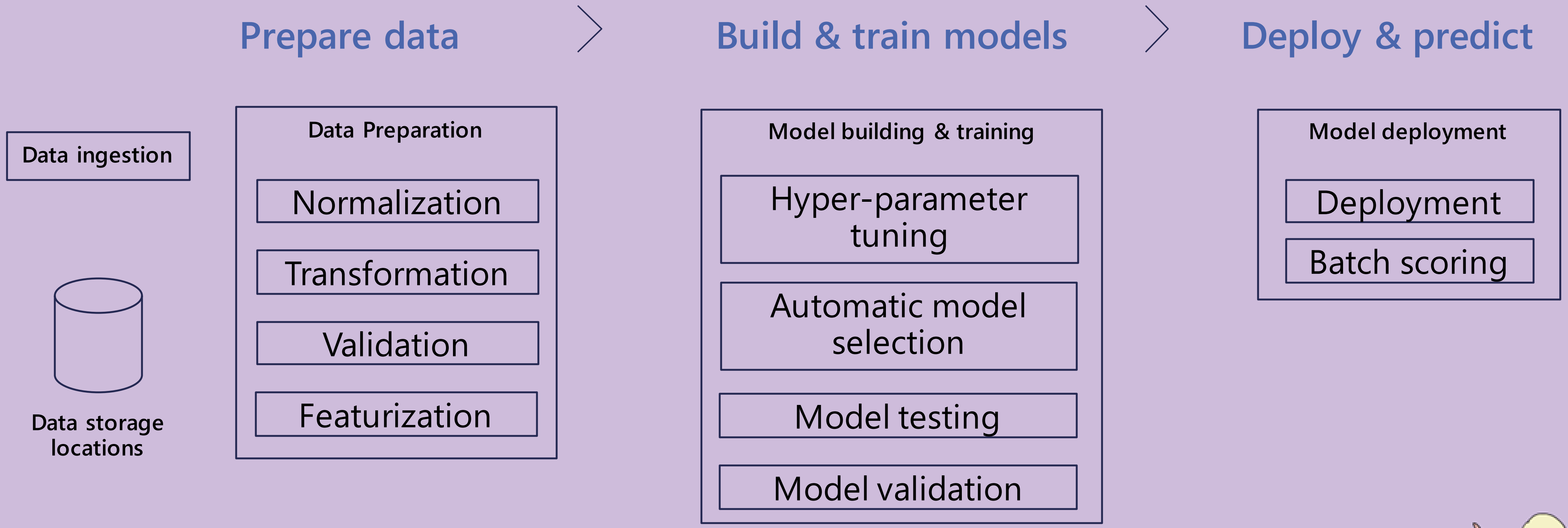
Manage Model Lifecycle

- **Track model versions & metadata** with a centralized **model registry**
- **Package, validate, profile** models to guarantee correct behavior
- **Manage dependencies** for training and inference
- Provide **scalable compute** for training and inference
- Capture **metrics & telemetry** – health, performance, inputs / outputs
- **Encapsulate each step** in the lifecycle to **enable automation** via CI/CD and DevOps





Azure Machine Learning pipelines



Prepare data

Data ingestion

Data storage locations

Data Preparation

- Normalization
- Transformation
- Validation
- Featurization

Build & train models

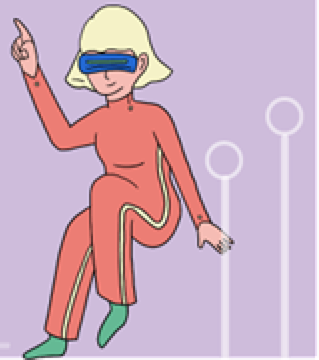
Model building & training

- Hyper-parameter tuning
- Automatic model selection
- Model testing
- Model validation

Deploy & predict

Model deployment

- Deployment
- Batch scoring



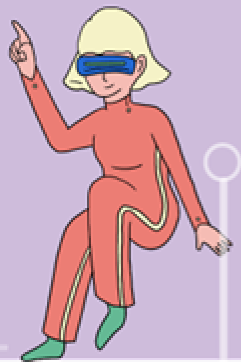
Azure Machine Learning

Make data scientists to be more productive
Enable your organization to manage the ML lifecycle through MLOps

Azure Cloud Services + Python SDK + Cross-Platform CLI

That enables you to

- ✓ Prepare Data
- ✓ Build Models
- ✓ Train Models
- ✓ Manage Models
- ✓ Track Experiments
- ✓ Deploy Models

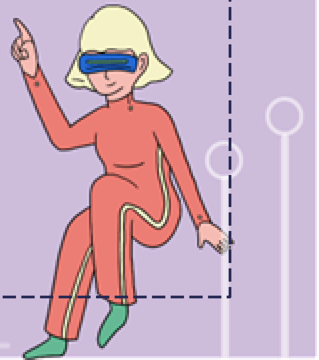


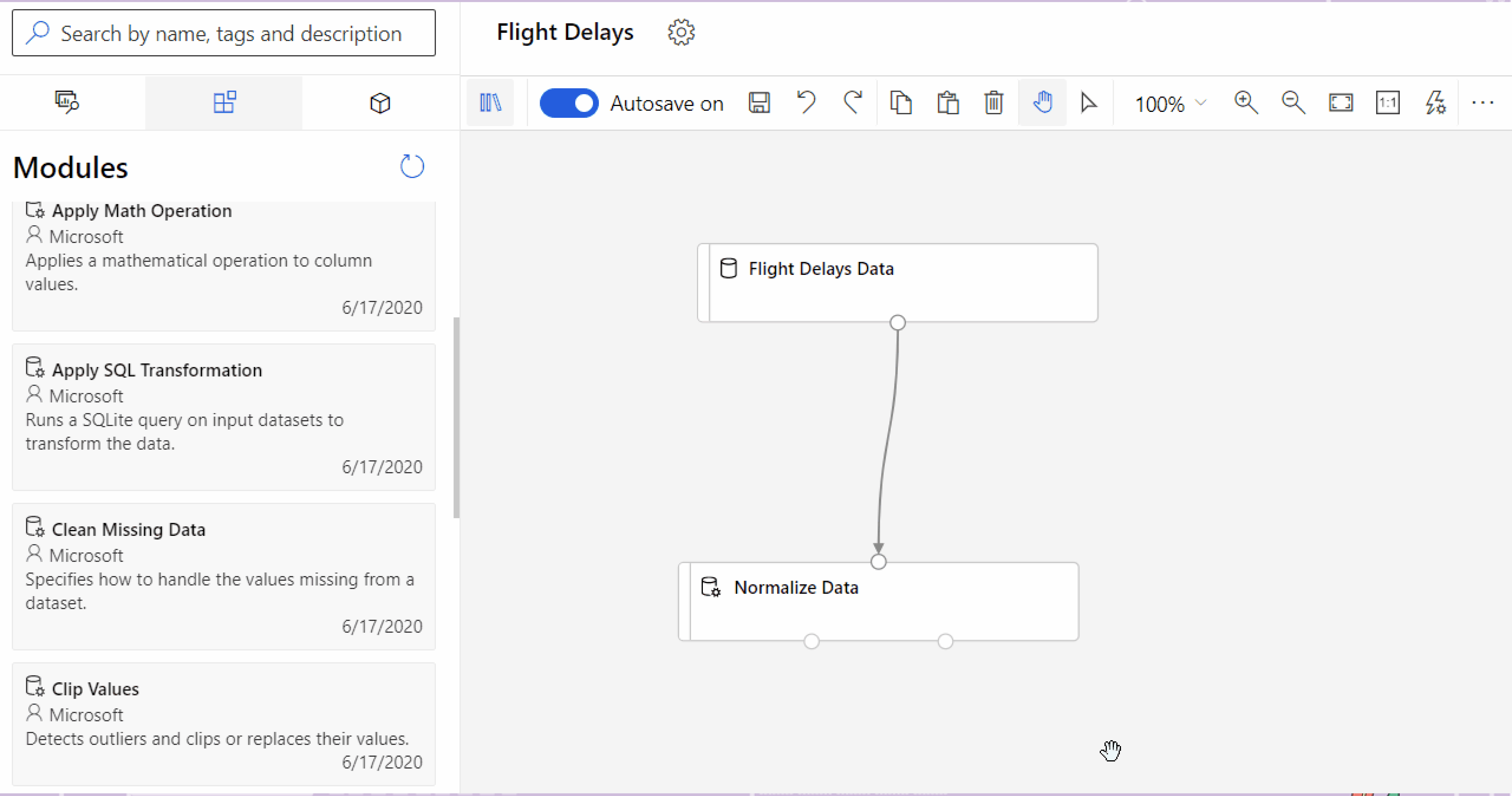
Azure ML service

Key Artifacts

Workspace

- Models
- Experiments
- Pipelines
- Compute Target
- Images
- Deployment
- Data Stores





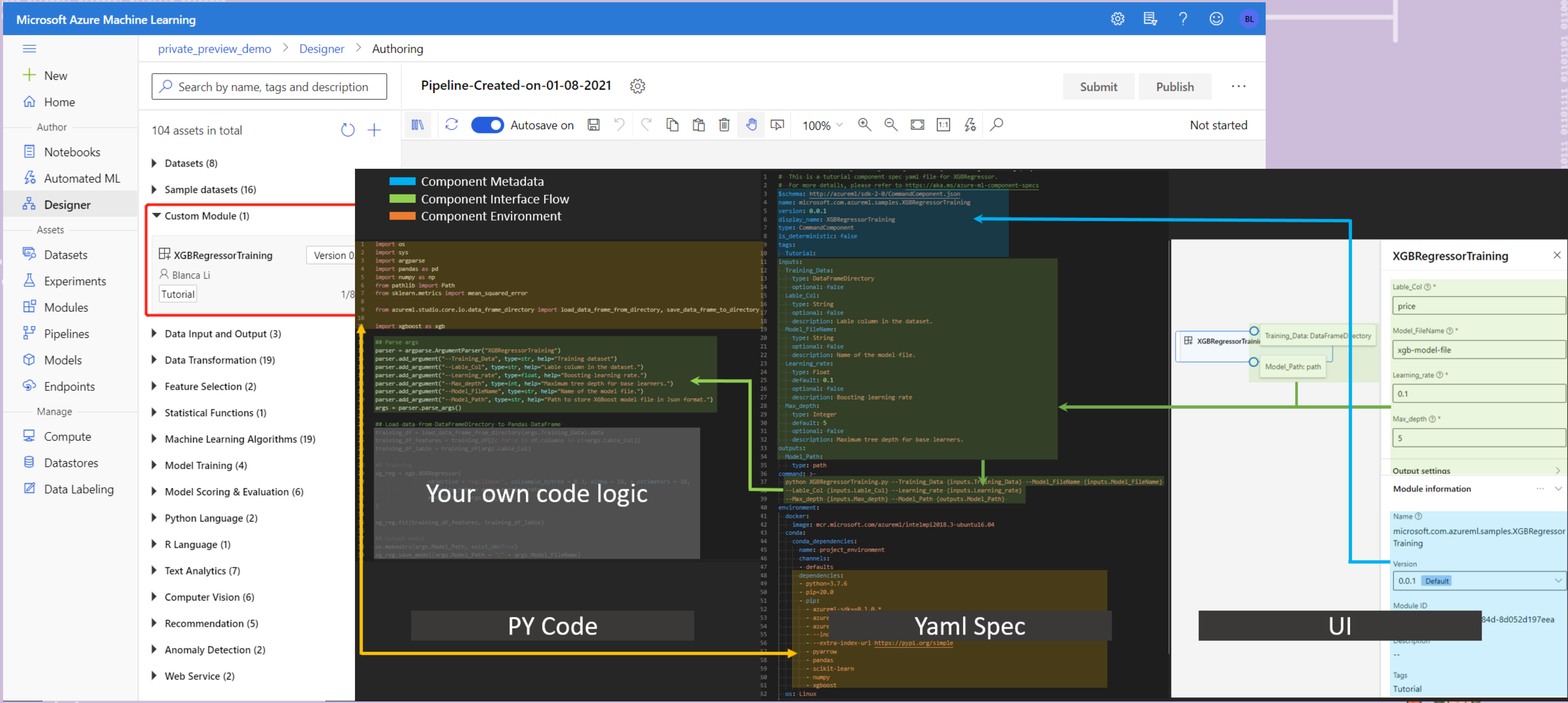
Modules

Apply Math Operation
Microsoft
Applies a mathematical operation to column values.
6/17/2020

Apply SQL Transformation
Microsoft
Runs a SQLite query on input datasets to transform the data.
6/17/2020

Clean Missing Data
Microsoft
Specifies how to handle the values missing from a dataset.
6/17/2020

Clip Values
Microsoft
Detects outliers and clips or replaces their values.
6/17/2020



🟦 Component Metadata
🟩 Component Interface Flow
🟧 Component Environment

```python
import os
import sys
import argparse
import pandas as pd
import numpy as np
from pathlib import Path
from sklearn.metrics import mean_squared_error

from azureml.studio.core.io.data_frame_directory import load_data_frame_from_directory, save_data_frame_to_directory

import xgboost as xgb

## Parse args
parser = argparse.ArgumentParser("XGBRegressorTraining")
parser.add_argument("--Training_Data", type=str, help="Training dataset")
parser.add_argument("--Lable_Col", type=str, help="Lable column in the dataset.")
parser.add_argument("--Learning_rate", type=float, help="Boosting learning rate.")
parser.add_argument("--Max_depth", type=int, help="Maximum tree depth for base learners.")
parser.add_argument("--Model_FileName", type=str, help="Name of the model file.")
parser.add_argument("--Model_Path", type=str, help="Path to store XGBoost model file in Json format.")
args = parser.parse_args()

## Load data from DataFrameDirectory to Pandas DataFrame
training_df = load_data_frame_from_directory(args.Training_Data).data
training_df_features = training_df[[c for c in df.columns if c!=args.Lable_Col]]
training_df_lable = training_df[args.Lable_Col]

## Training
xg_reg = xgb.XGBRegressor(
                objective ='reg:linear', colsample_bytree = 0.3, alpha = 10, n_estimators = 10,
                learning_rate = args.Learning_rate,
                max_depth = args.Max_depth
)

xg_reg.fit(training_df_features, training_df_lable)

## Output model
os.makedirs(args.Model_Path, exist_ok=True)
xg_reg.save_model(args.Model_Path + "/" + args.Model_FileName)
```

Your own code logic

PY Code

```yaml
#  This is a tutorial component spec yaml file for XGBRegressor.
#  For more details, please refer to https://aka.ms/azure-ml-component-specs
$schema: http://azureml/sdk-2-0/CommandComponent.json
name: microsoft.com.azureml.samples.XGBRegressorTraining
version: 0.0.1
display_name: XGBRegressorTraining
type: CommandComponent
is_deterministic: false
tags:
  Tutorial:
inputs:
  Training_Data:
    type: DataFrameDirectory
    optional: false
  Lable_Col:
    type: String
    optional: false
    description: Lable column in the dataset.
  Model_FileName:
    type: String
    optional: false
    description: Name of the model file.
  Learning_rate:
    type: Float
    default: 0.1
    optional: false
    description: Boosting learning rate
  Max_depth:
    type: Integer
    default: 5
    optional: false
    description: Maximum tree depth for base learners.
outputs:
  Model_Path:
    type: path
command: >-
  python XGBRegressorTraining.py --Training_Data {inputs.Training_Data} --Model_FileName {inputs.Model_FileName}
  --Lable_Col {inputs.Lable_Col} --Learning_rate {inputs.Learning_rate}
  --Max_depth {inputs.Max_depth} --Model_Path {outputs.Model_Path}
environment:
  docker:
    image: mcr.microsoft.com/azureml/intelmpi2018.3-ubuntu16.04
  conda:
    conda_dependencies:
      name: project_environment
      channels:
      - defaults
      dependencies:
      - python=3.7.6
      - pip=20.0
      - pip:
        - azureml-sdk==0.1.0.*
        - azure
        - azure
        - --inc
        - --extra-index-url https://pypi.org/simple
        - pyarrow
        - pandas
        - scikit-learn
        - numpy
        - xgboost
  os: Linux
```

Yaml Spec

XGBRegressorTraining

Lable_Col *
price

Model_FileName *
xgb-model-file

Learning_rate *
0.1

Max_depth *
5

Output settings

Module information

Name
microsoft.com.azureml.samples.XGBRegressorTraining

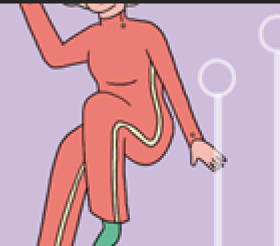
Version
0.0.1 Default

Module ID
84d-8d052d197eea

Description
--

Tags
Tutorial

UI



[AzureMachineLearningGallery/tutorial1-use-existing-components.md at main · Azure/AzureMachineLearningGallery (github.com)](https://github.com)

AutoML

Which features?

Mileage

Condition

Car brand

Year of make

Regulations

...

Which algorithm?

Gradient Boosted

Nearest Neighbors

SVM

Bayesian Regression

LGBM

...

Which parameters?

Parameter 1

Parameter 2

Parameter 3

Parameter 4

Criterion

Loss

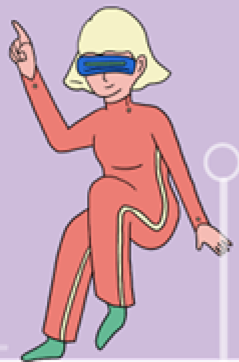
Min Samples Split

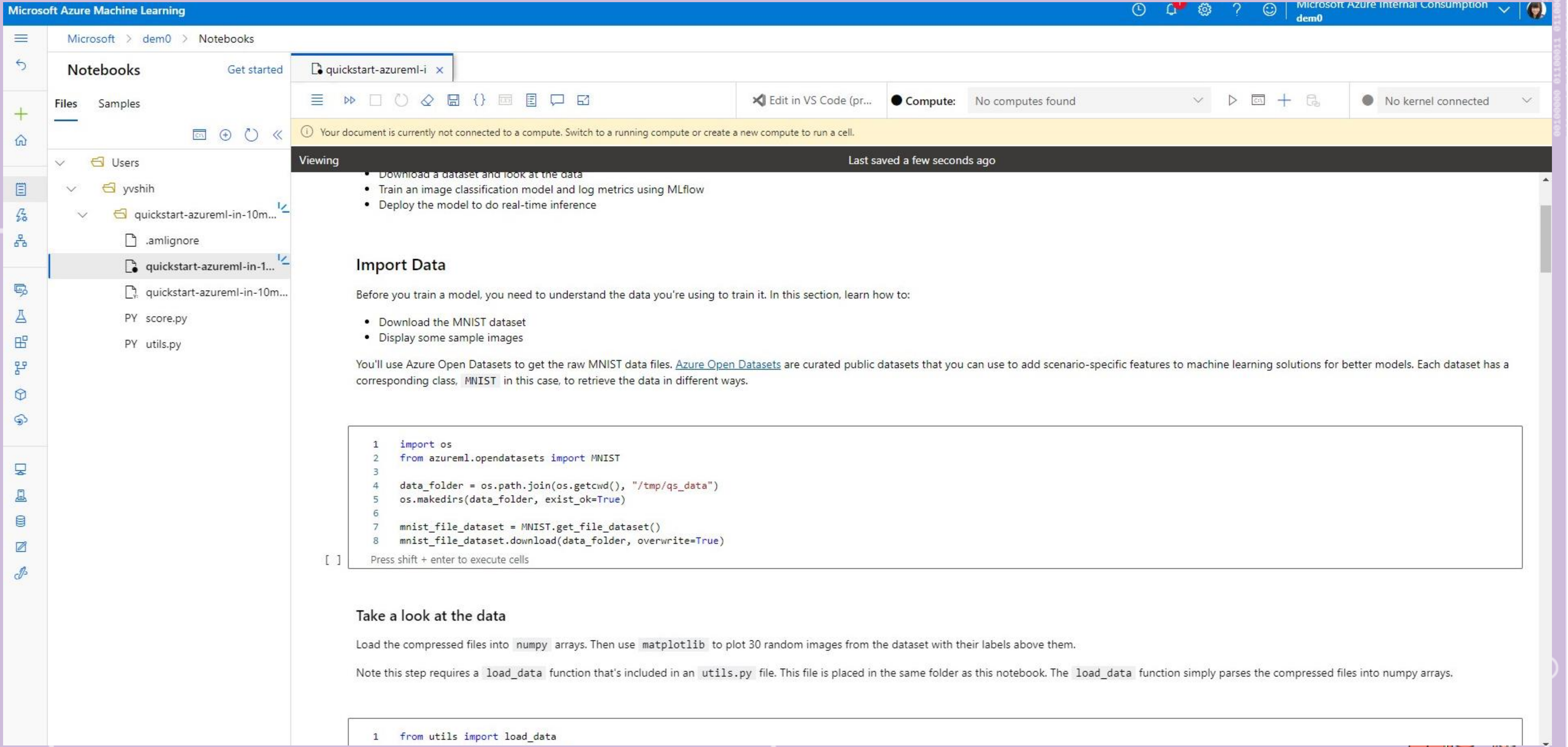
Min Samples Leaf

Others

30%

Model




Microsoft Azure Machine Learning
Microsoft Azure Internal Consumption
dem0
Microsoft > dem0 > Notebooks
Notebooks
Get started
Files
Samples
Users
yvshih
quickstart-azureml-in-10m...
.amlignore
quickstart-azureml-in-1...
quickstart-azureml-in-10m...
PY score.py
PY utils.py
quickstart-azureml-i
Edit in VS Code (pr...
● Compute:
No computes found
No kernel connected
Viewing
Last saved a few seconds ago

Your document is currently not connected to a compute. Switch to a running compute or create a new compute to run a cell.

- Download a dataset and look at the data
- Train an image classification model and log metrics using MLflow
- Deploy the model to do real-time inference

Import Data

Before you train a model, you need to understand the data you're using to train it. In this section, learn how to:

- Download the MNIST dataset
- Display some sample images

You'll use Azure Open Datasets to get the raw MNIST data files. Azure Open Datasets are curated public datasets that you can use to add scenario-specific features to machine learning solutions for better models. Each dataset has a corresponding class, `MNIST` in this case, to retrieve the data in different ways.

```
import os
from azureml.opendatasets import MNIST

data_folder = os.path.join(os.getcwd(), "/tmp/qs_data")
os.makedirs(data_folder, exist_ok=True)

mnist_file_dataset = MNIST.get_file_dataset()
mnist_file_dataset.download(data_folder, overwrite=True)
```

[ ] Press shift + enter to execute cells

Take a look at the data

Load the compressed files into `numpy` arrays. Then use `matplotlib` to plot 30 random images from the dataset with their labels above them.

Note this step requires a `load_data` function that's included in an `utils.py` file. This file is placed in the same folder as this notebook. The `load_data` function simply parses the compressed files into numpy arrays.

```
from utils import load_data
```

Datasets in Azure Machine Learning

Manage data

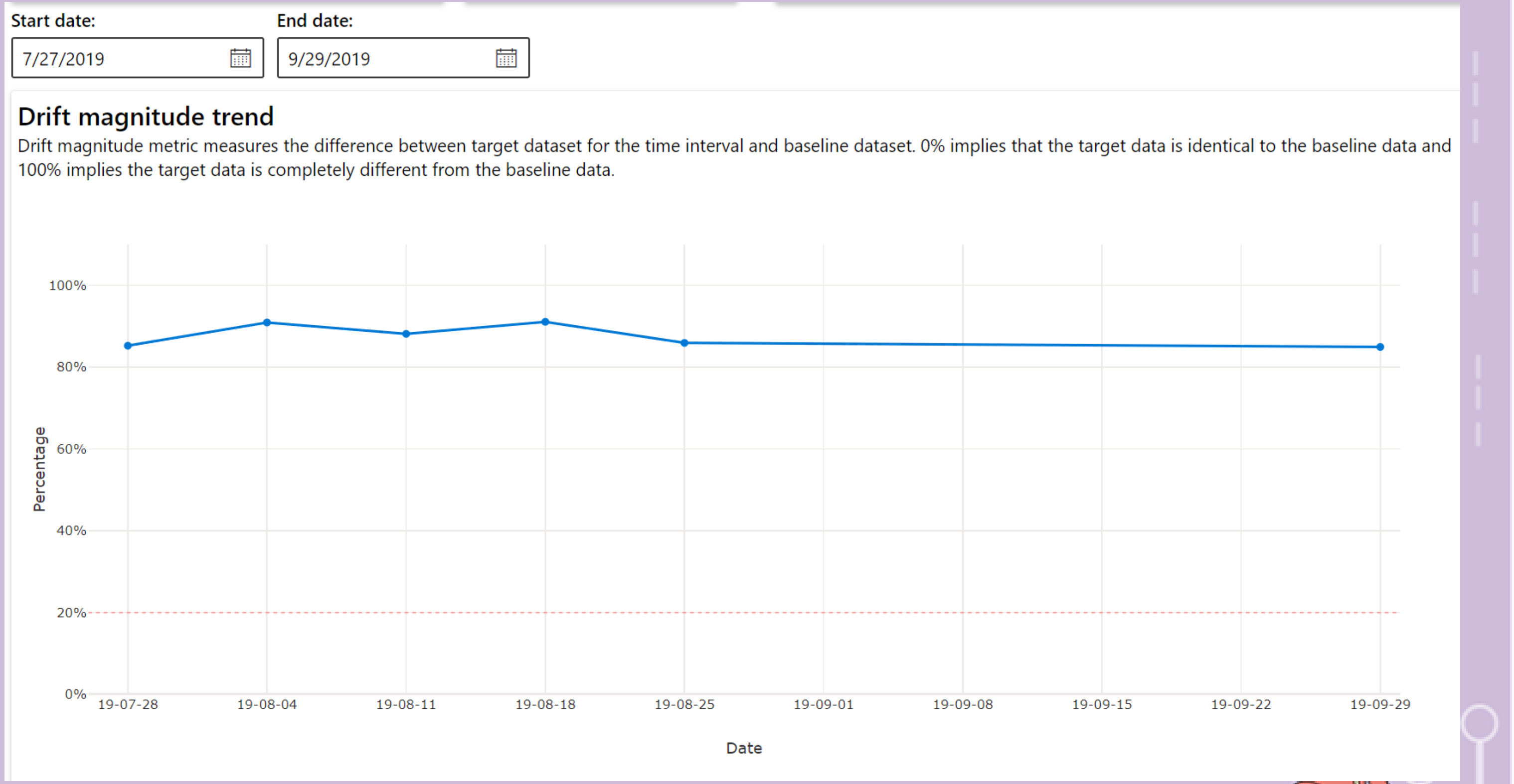
Decrease friction by having consistent data artifacts throughout the ML workflow

Explore and transform at scale

Use one code artifact (which is lazily-evaluated for scalability) locally and on different runtimes

Reproduce and collaborate

Enable teammates to reference and work on shared data artifacts



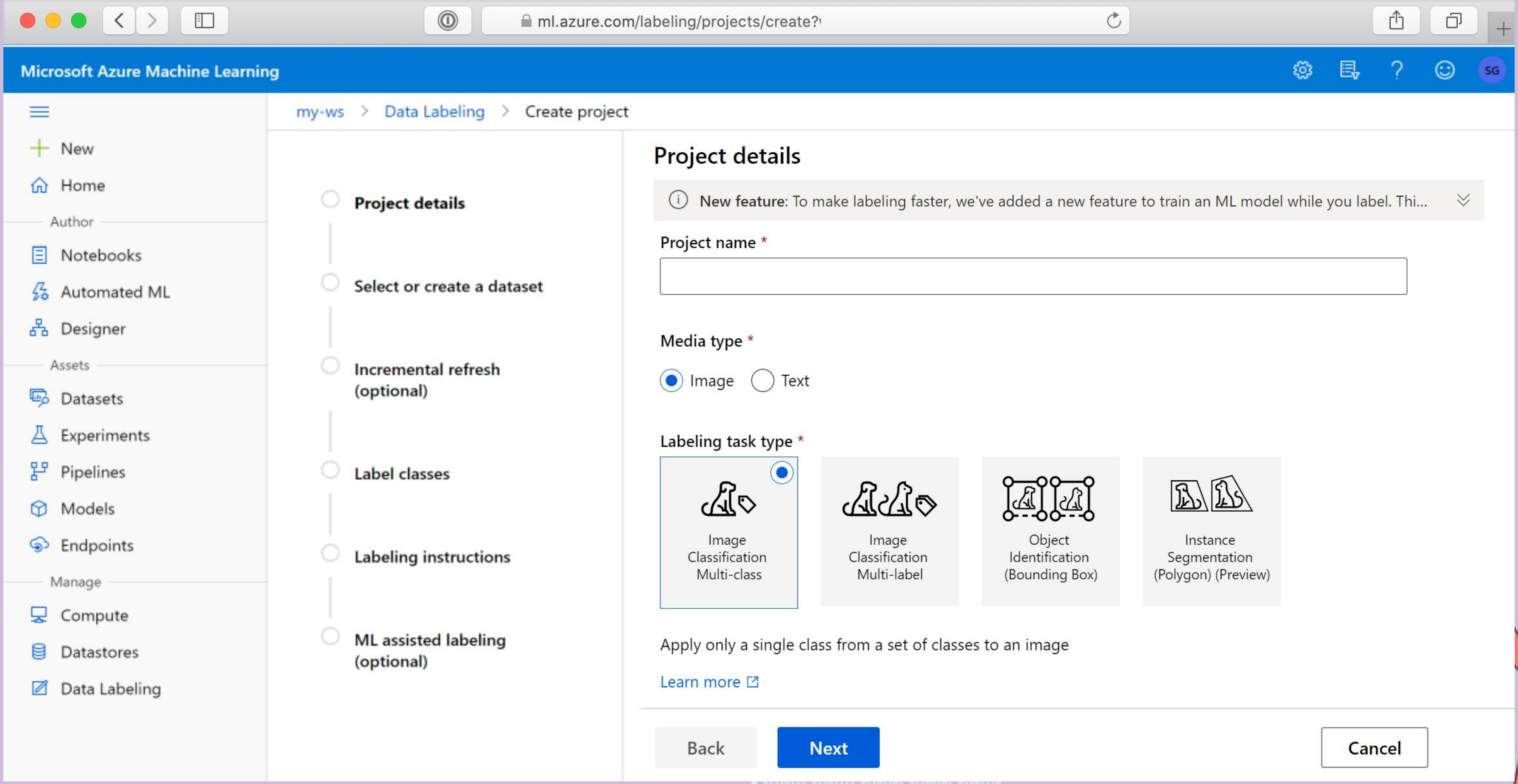
Start date: 7/27/2019

End date: 9/29/2019

Drift magnitude trend

Drift magnitude metric measures the difference between target dataset for the time interval and baseline dataset. 0% implies that the target data is identical to the baseline data and 100% implies the target data is completely different from the baseline data.

Data Labeling



ml.azure.com/labeling/projects/create?

Microsoft Azure Machine Learning

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

my-ws > Data Labeling > Create project

○ Project details

○ Select or create a dataset

○ Incremental refresh (optional)

○ Label classes

○ Labeling instructions

○ ML assisted labeling (optional)

Project details

ⓘ **New feature**: To make labeling faster, we've added a new feature to train an ML model while you label. Thi...

Project name *

Media type *

● Image ○ Text

Labeling task type *

● Image Classification Multi-class

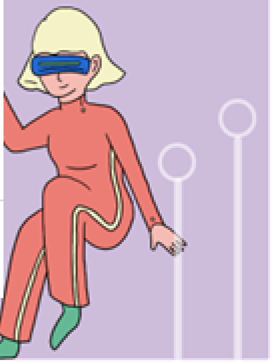
Image Classification Multi-label

Object Identification (Bounding Box)

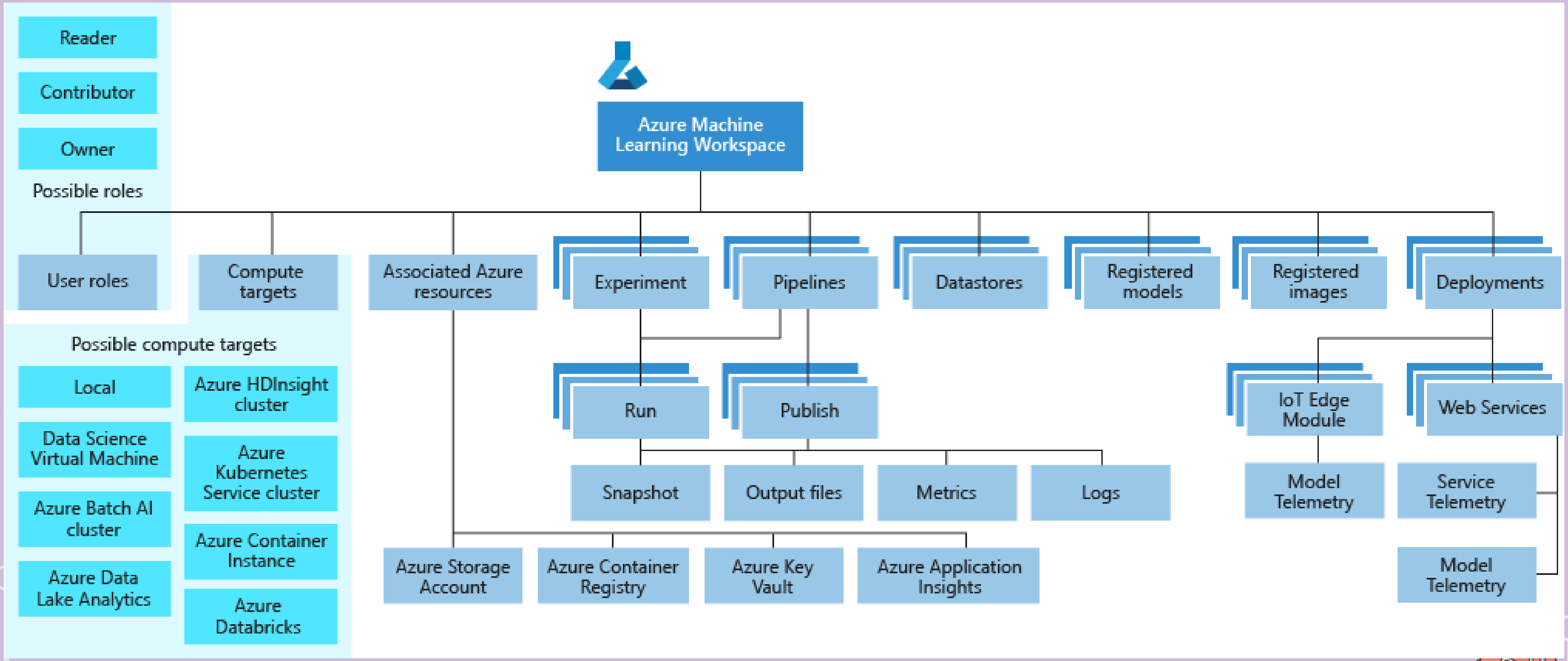
Instance Segmentation (Polygon) (Preview)

Apply only a single class from a set of classes to an image

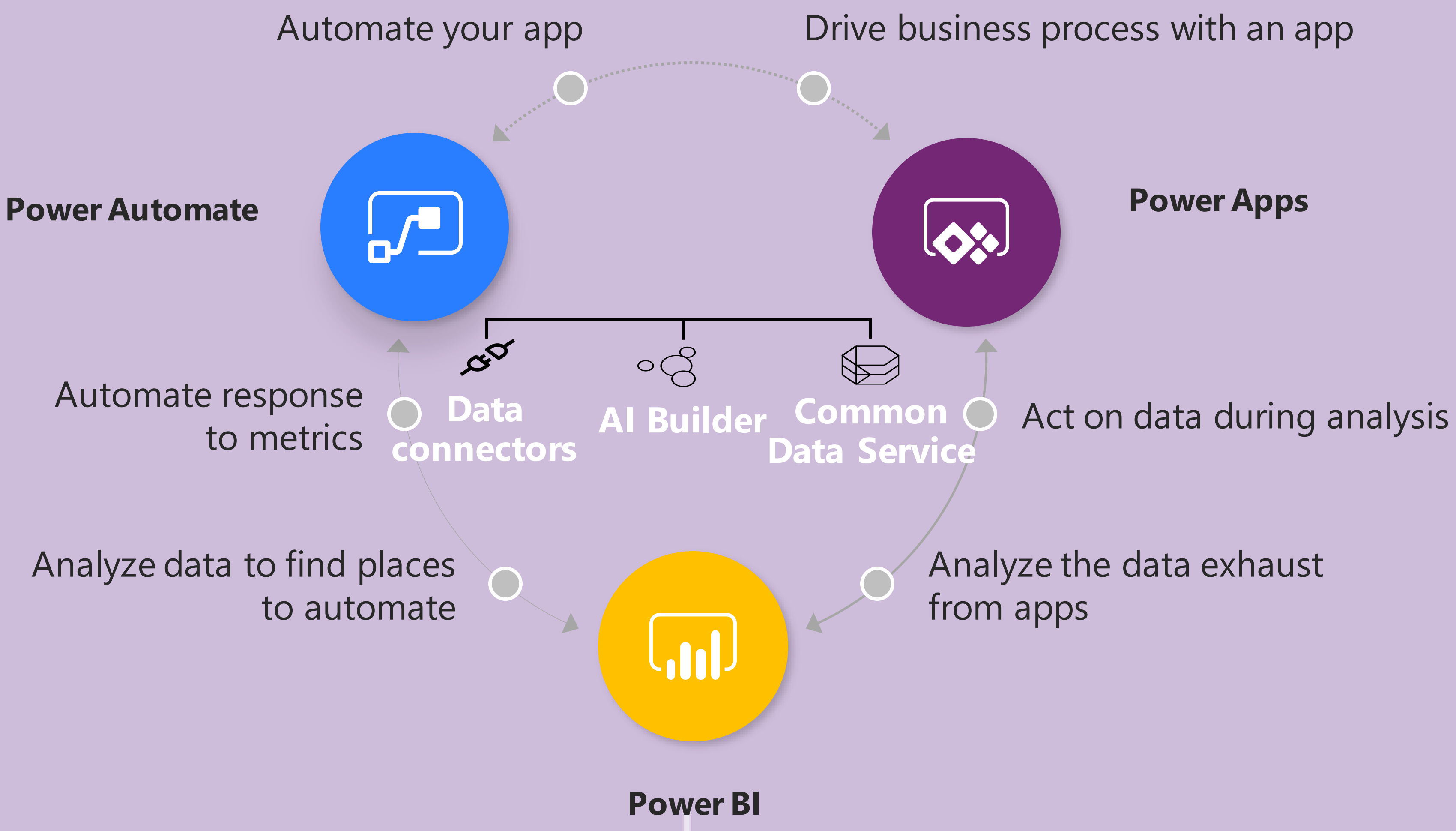
Learn more



Azure ML service Workspace Taxonomy



Integrated low-code platform



Automate your app

Drive business process with an app

Power Automate

Power Apps

Automate response
to metrics

Data connectors

AI Builder

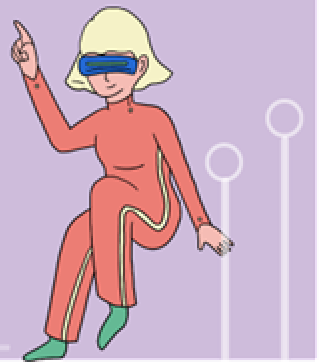
Common Data Service

Act on data during analysis

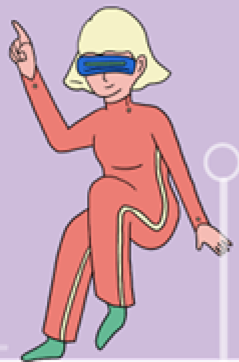
Analyze data to find places
to automate

Analyze the data exhaust
from apps

Power BI



DEMO



Bring AI to every application

Redefine existing applications with AI

Vision

Speech

Language

Knowledge

Pre-built AI services

+

Create new conversational AI experiences

kik·

Conversational AI platform

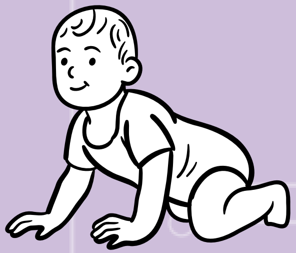
Cognitive Services

Introduction of speaker

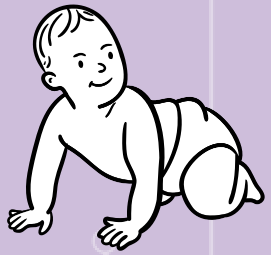
Jolin Cai

- Face API Software Engineer
- Azure Support Engineer

What is Cognitive Services?



Azure Cognitive Services enable your applications to see, hear, speak, understand, and even make decisions.



Categories of Cognitive Services

Vision

Language

Speech

Decision

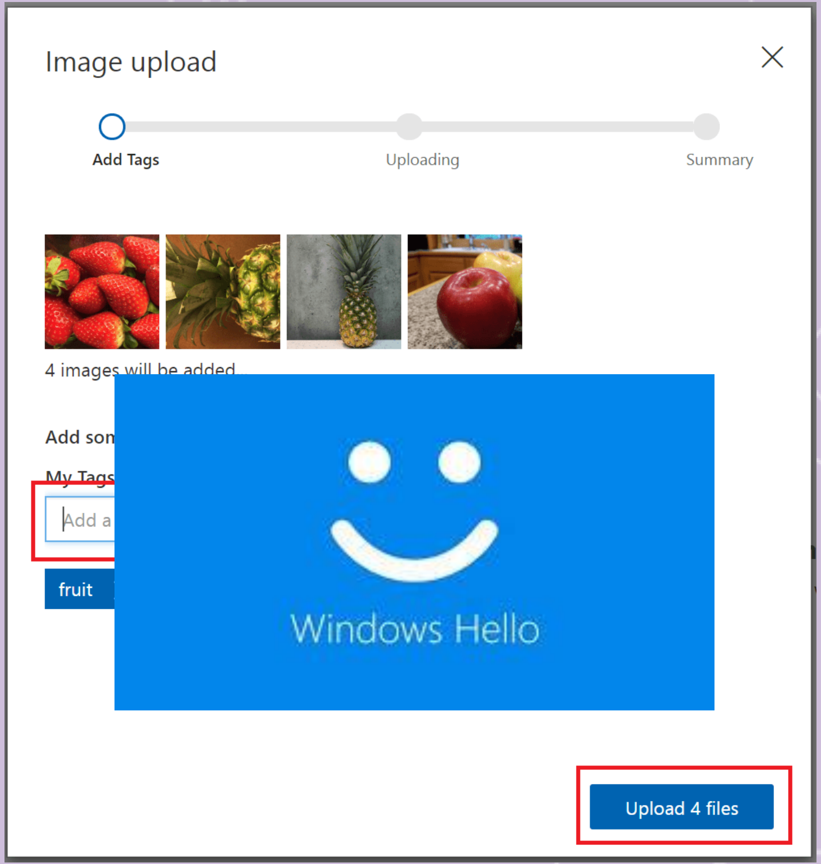
CODING ANGELS
EMPOWER WOMEN IN TECH

Vision

Custom Vision

Computer Vision

Face

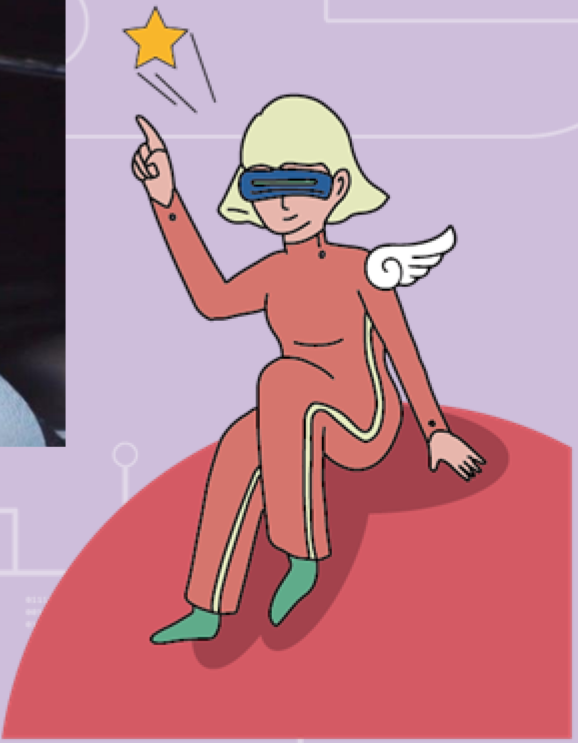


4 images will be added

CODING ANGELS
EMPOWER WOMEN IN TECH

Face – Scenario(Uber)

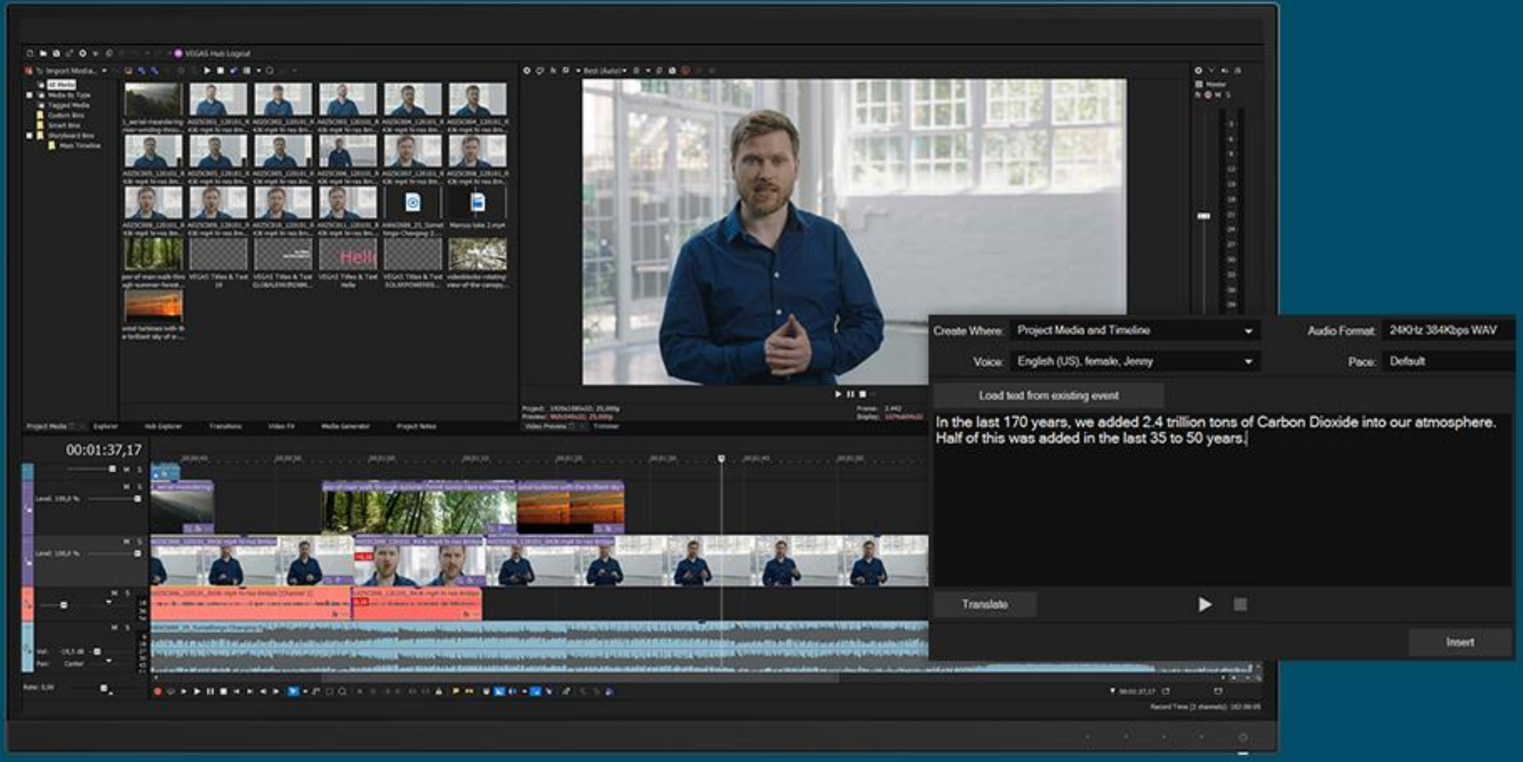
Uber boosts platform security with the Face API, part of Microsoft Cognitive Services



Speech

Speech to Text
Text to Speech

Speech – Scenario(VEGAS)



Create Where: Project Media and Timeline

Audio Format: 24KHz 384Kbps WAV

Voice: English (US), female, Jenny

Pace: Default

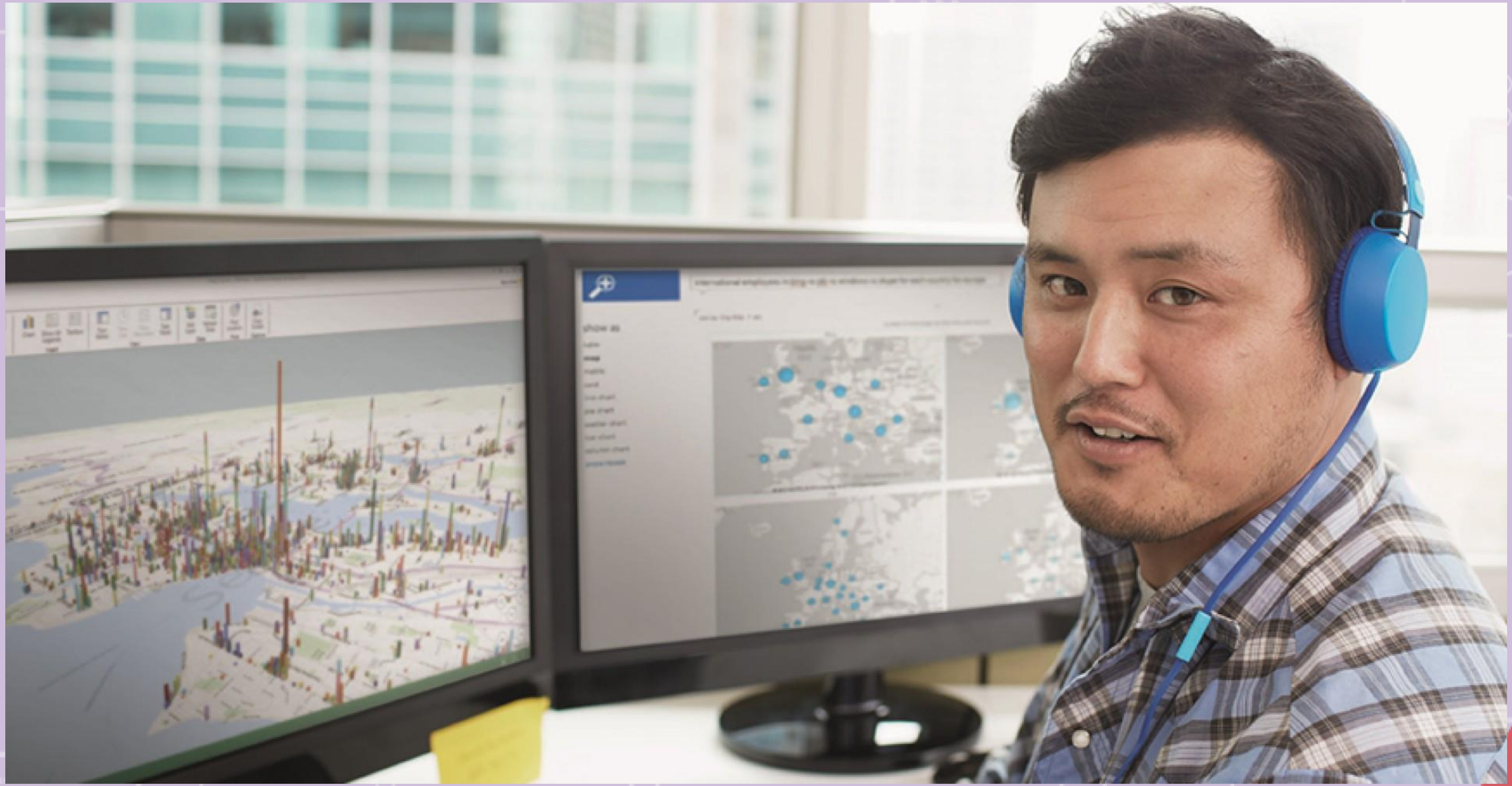
In the last 170 years, we added 2.4 trillion tons of Carbon Dioxide into our atmosphere.
Half of this was added in the last 35 to 50 years.

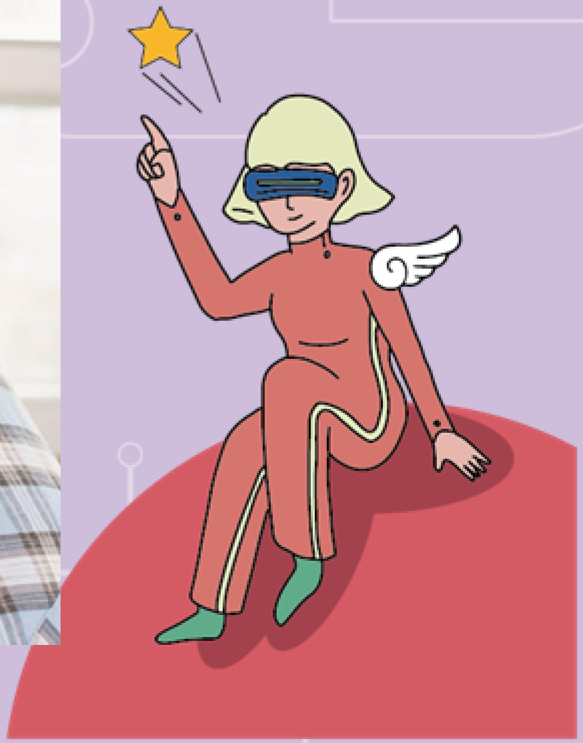
Language

Analyze sentiment and opinions
Conversational language understanding

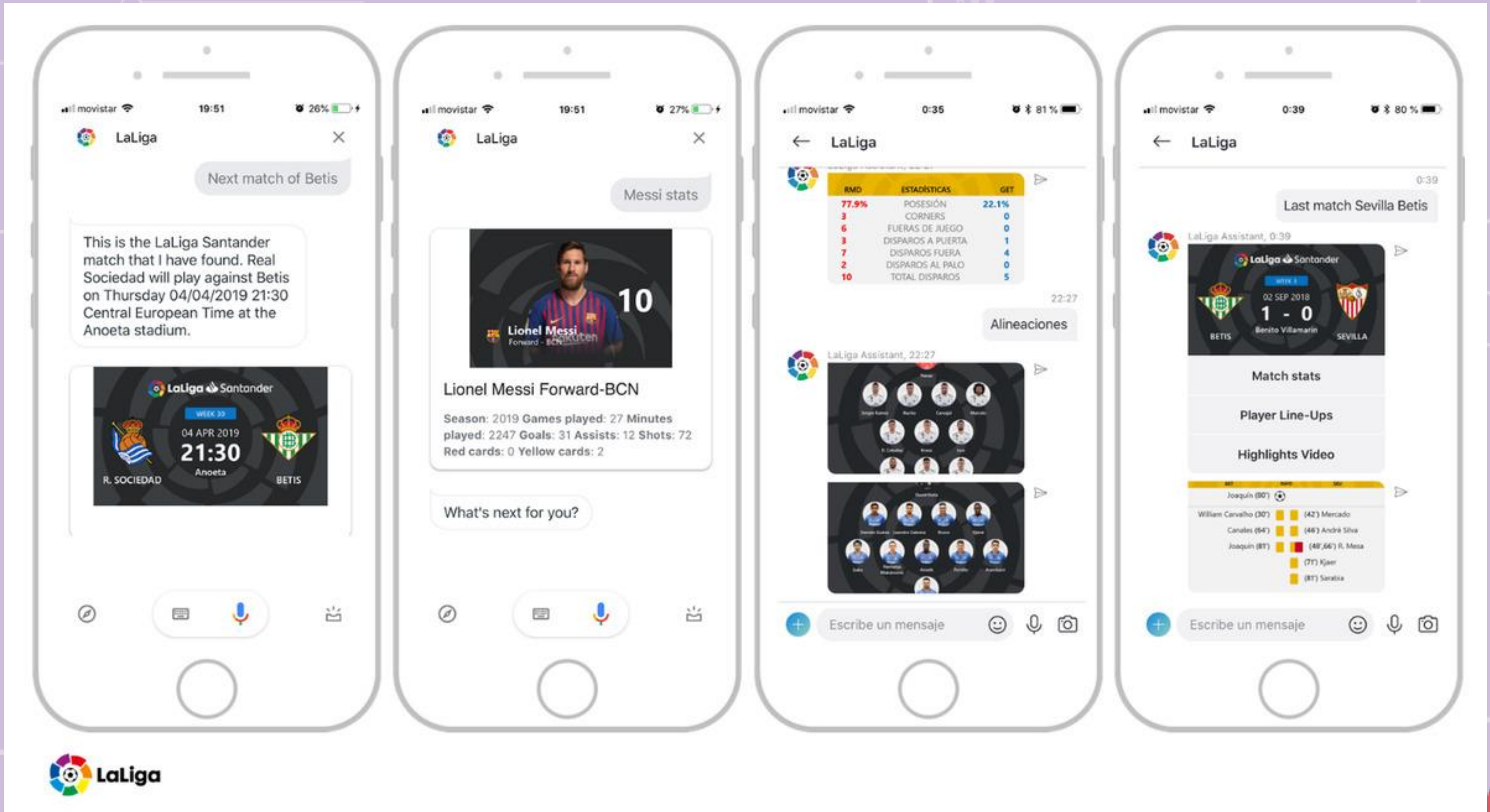
CODING ANGELS
EMPOWER WOMEN IN TECH

Language – Scenario(Microsoft)





Language – Scenario(LaLiga)



Next match of Betis

This is the LaLiga Santander match that I have found. Real Sociedad will play against Betis on Thursday 04/04/2019 21:30 Central European Time at the Anoeta stadium.

Messi stats

Lionel Messi Forward-BCN

Season: 2019 Games played: 27 Minutes played: 2247 Goals: 31 Assists: 12 Shots: 72 Red cards: 0 Yellow cards: 2

What's next for you?

RMD	ESTADÍSTICAS	GET
77.9%	POSESIÓN	22.1%
3	CORNERS	0
6	FUERAS DE JUEGO	0
3	DISPAROS A PUERTA	1
7	DISPAROS FUERA	4
2	DISPAROS AL PALO	0
10	TOTAL DISPAROS	5

22:27

Alineaciones

LaLiga Assistant, 22:27

0:39

Last match Sevilla Betis

LaLiga Assistant, 0:39

Match stats

Player Line-Ups

Highlights Video

Joaquín (80')

William Carvalho (30') (42') Mercado

Canales (64') (46') André Silva

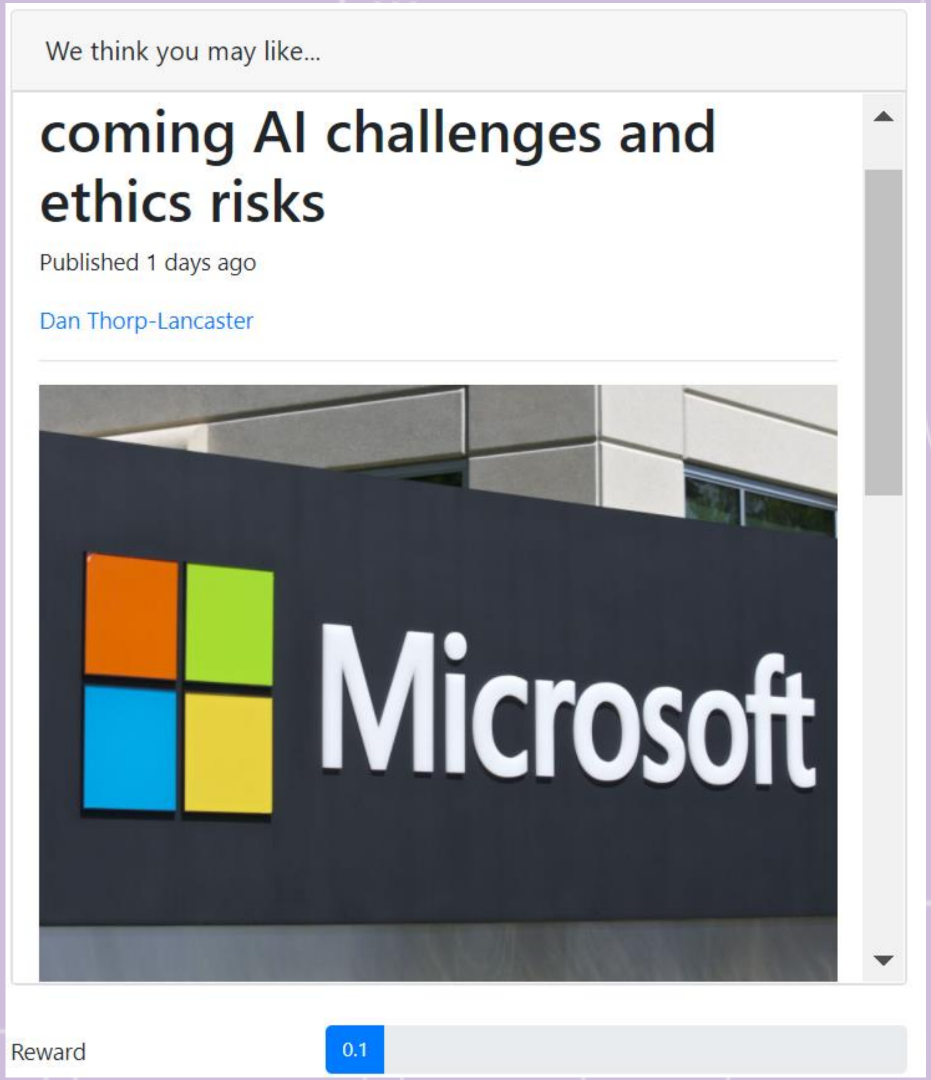
Joaquín (81') (48',66') R. Mesa

(71') Kjaer

(81') Sarabia

Decision APIs

Content Moderator
Personalizer



We think you may like...

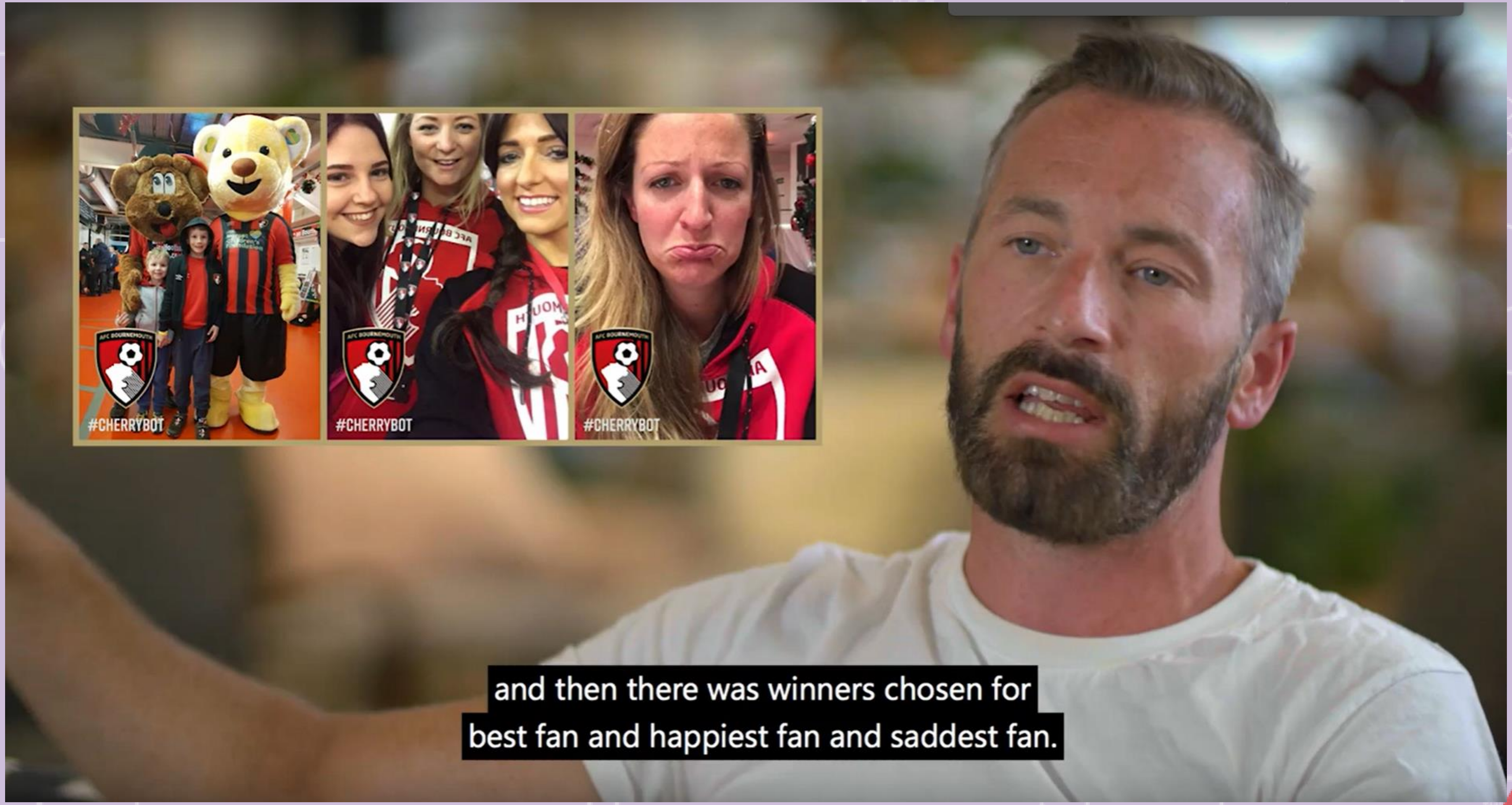
coming AI challenges and ethics risks

Published 1 days ago

Dan Thorp-Lancaster

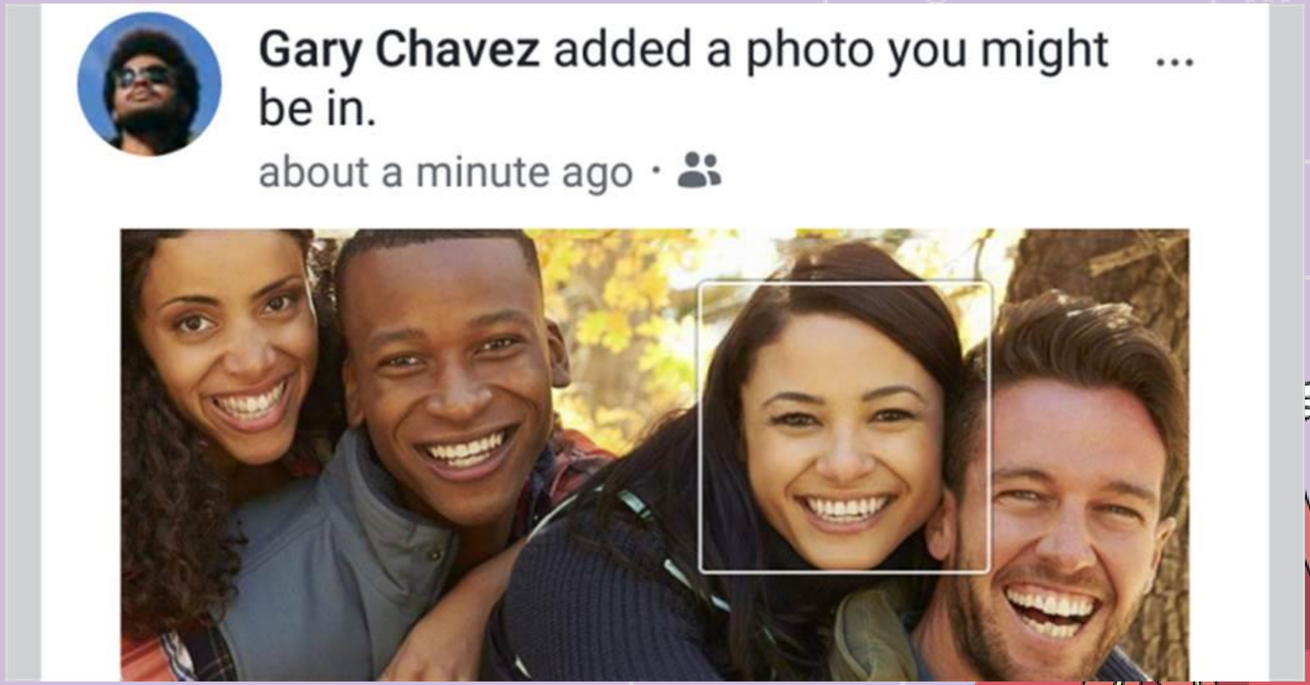
Reward 0.1

Decision APIs – Scenario(Greenwood Campbell)



and then there was winners chosen for
best fan and happiest fan and saddest fan.

Challenges of Privacy



Gary Chavez added a photo you might be in. ...

about a minute ago ·

CODING ANGELS
EMPOWER WOMEN IN TECH

Thanks for your attention!