

Ingest Data in a Designer Pipeline



Ifedayo Bamikole

CLOUD SOLUTION ARCHITECT

@dayobam7 www.dayobam.com



Overview



Ingestion Options in Azure Machine Learning Studio

Understanding the Dataset Types

Connectivity to Data in Azure ML Studio



Working with Data in Azure ML Studio



Register the Dataset

How to Connect

- Locally (Desktop)
- Web Files e.g. GitHub
- Azure Datastores e.g. Azure SQL Database



Azure Data Stores

Connect to Built-in Data Stores

- Azure Blob Storage
- Azure File Share
- Open Datasets

Connect to Azure Data Stores

- Azure Blob Storage
- Azure Data Lake Storage
- Azure SQL Database
- Azure PostgreSQL Database



Connectivity to Azure Data Stores



Authentication to Azure Data Stores in Azure ML

Storage Type	Authentication
Azure Blob Storage	Account Key or SAS Token
Azure File Share	Account Key or SAS Token
Azure Data Lake Storage	Service Principal
Azure SQL Database	SQL Authentication or Service Principal
Azure PostgreSQL Database	SQL Authentication



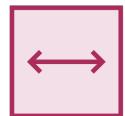
Retrieving Data from Azure Storage in Azure ML



Make sure the users has Storage Blob Data Reader



For Account SAS tokens defaults to no permissions



For Read Access have a minimum of list and read permissions for containers and objects



For Write Permission write and add permissions should be granted as well



Types of Dataset



Types of Dataset in Azure ML Studio

Tabular

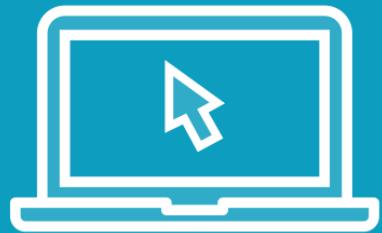
Represents data that is displayed data as a Table

File

References files in a datastore or public URL



Demo



Register a Dataset in Azure ML Studio

- Create a Data Store
- Selecting Data from the Data Store

Explore Open Datasets



Summary



- Reviewed Ingestion Options in Azure Machine Learning Designer
- Understood the Dataset Types
- Registered Dataset in Azure ML Studio



Using the Designer Modules to Define a Pipeline Data Flow



Ifedayo Bamikole

CLOUD SOLUTION ARCHITECT

@dayobam7 www.dayobam.com



Overview

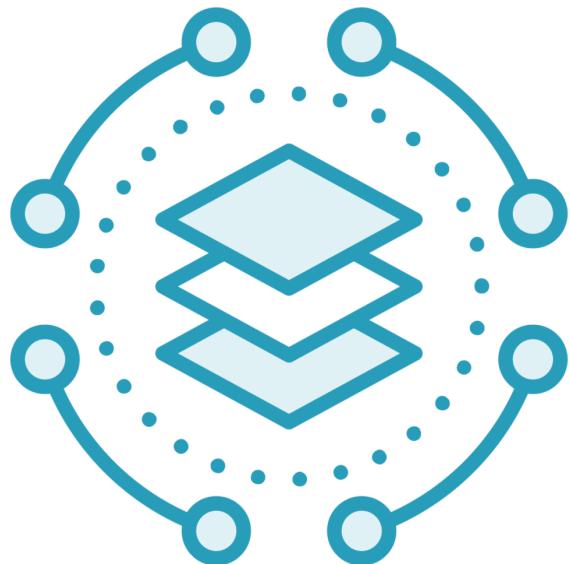


**Defining Azure Machine Learning
Designer Module**

Understand when to use a Module



What Are the Designer Modules?



Modules gives the ability to perform actions on dataset

It can be dragged and dropped onto the Designer Platform

Top Modules

- Select Column in Dataset
- Clean Missing Data
- Split Data
- Train Model
- Execute Custom Code e.g. Python



Scenario

Find out how much a car will cost?



Azure ML Studio - Designer Modules

Dataset

Where is my Data coming from ?
CSV File, Open Dataset, SQL

Clean Missing Data

Remove Empty Spots
Specify what to do with missing Data

Split Data

Specify Percentage Split e.g. 70/30
One for Training, One for Testing

Select Columns in Dataset

What Columns will help my Model
can also be seen as removing
unwanted columns



Normalize Data

View some of your data on a scale of 0-1
Visualize statistics within the data

Train Data

Add Algorithm
e.g. Regression



Azure ML Studio - Designer Modules

Dataset

Where is my Data coming from?

CSV File, Open Dataset, SQL

Clean Missing Data

Remove Empty Spots

Specify what to do with missing Data

Split Data

Specify Percentage Split e.g. 70/30

One for Training, One for Testing

Select Columns in Dataset

What Columns will help my Model

can also be seen as removing unwanted columns



Normalize Data

View some of your data on a scale of 0-1

Visualize statistics within the data

Train Data

Add Algorithm
e.g. Regression



Azure ML Studio - Designer Modules

Dataset

Where is my Data coming from?

CSV File, Open Dataset, SQL

Clean Missing Data

Remove Empty Spots

Specify what to do with missing Data

Split Data

Specify Percentage Split e.g. 70/30

One for Training, One for Testing

Select Columns in Dataset

What Columns will help my Model
can also be seen as removing
unwanted columns



Normalize Data

View some of your data on a scale of 0-1
Visualize statistics within the data

Train Data

Add Algorithm
e.g. Regression



Azure ML Studio - Designer Modules

Dataset

Where is my Data coming from?

CSV File, Open Dataset, SQL

Clean Missing Data

Remove Empty Spots

Specify what to do with missing Data

Split Data

Specify Percentage Split e.g. 70/30

One for Training, One for Testing

Select Columns in Dataset

What Columns will help my Model
can also be seen as removing
unwanted columns



Normalize Data

View some of your data on a scale of 0-1
Visualize statistics within the data

Train Data

Add Algorithm
e.g. Regression



Azure ML Studio - Designer Modules

Dataset

Where is my Data coming from?

CSV File, Open Dataset, SQL

Clean Missing Data

Remove Empty Spots

Specify what to do with missing Data

Split Data

Specify Percentage Split e.g. 70/30

One for Training, One for Testing

Select Columns in Dataset

What Columns will help my Model

can also be seen as removing unwanted columns



Normalize Data

View some of your data on a scale of 0-1

Visualize statistics within the data

Train Data

Add Algorithm
e.g. Regression



Azure ML Studio - Designer Modules

Dataset

Where is my Data coming from?

CSV File, Open Dataset, SQL

Clean Missing Data

Remove Empty Spots

Specify what to do with missing Data

Split Data

Specify Percentage Split e.g. 70/30

One for Training, One for Testing

Select Columns in Dataset

What Columns will help my Model

can also be seen as removing unwanted columns



Normalize Data

View some of your data on a scale of 0-1

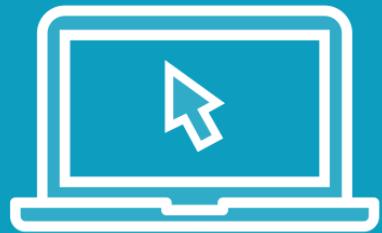
Visualize statistics within the data

Train Data

Add Algorithm
e.g. Regression



Demo



Using Modules on Datasets

- Review common Modules
- Running a Designer Pipeline to view Results
- Execute Custom Code



Summary



Defined Azure ML Designer Modules

Understood when to use Modules

Demonstrated using Modules in Azure ML Studio



Applying Machine Learning Algorithms



Ifedayo Bamikole

CLOUD SOLUTION ARCHITECT

@dayobam7 www.dayobam.com



Overview

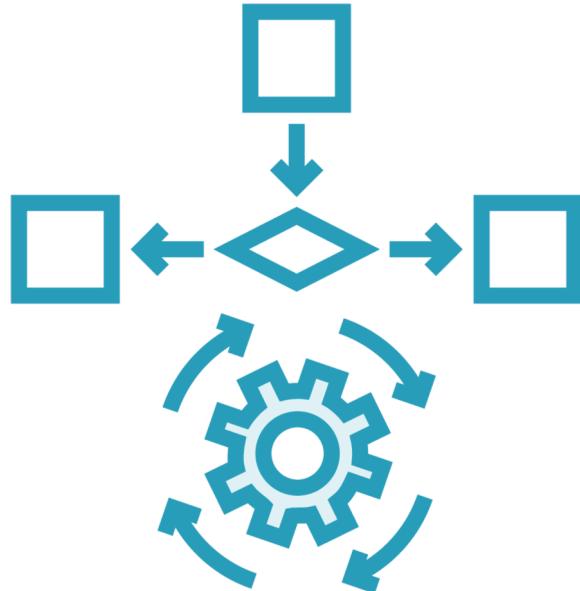


Define ML Algorithms

Implementing ML Algorithms in Azure ML Designer



Machine Learning Algorithms



Using code to help understand complex data set

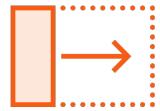
Examples are Regression or Classification

Typically fall into these categories

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning



What Can You Do with ML Algorithms



Predict Values



Select/Choose between two options



Group/Classify Items together



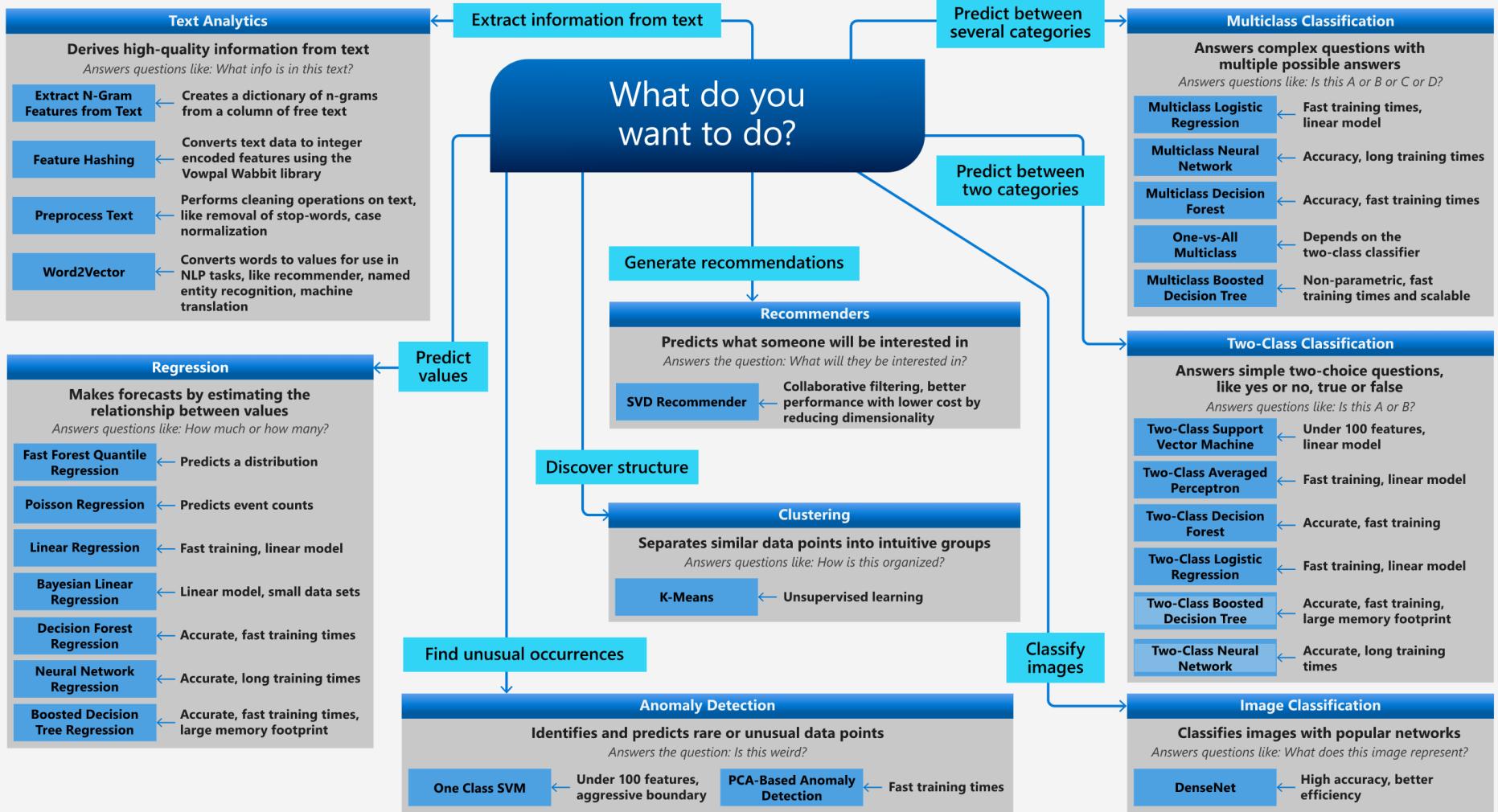
Azure Machine Learning Cheat Sheet





Microsoft Azure Machine Learning Algorithm Cheat Sheet

This cheat sheet helps you choose the best machine learning algorithm for your predictive analytics solution. Your decision is driven by both the nature of your data and the goal you want to achieve with your data.



Questions to Ask

**What do you want to do with
the Data**

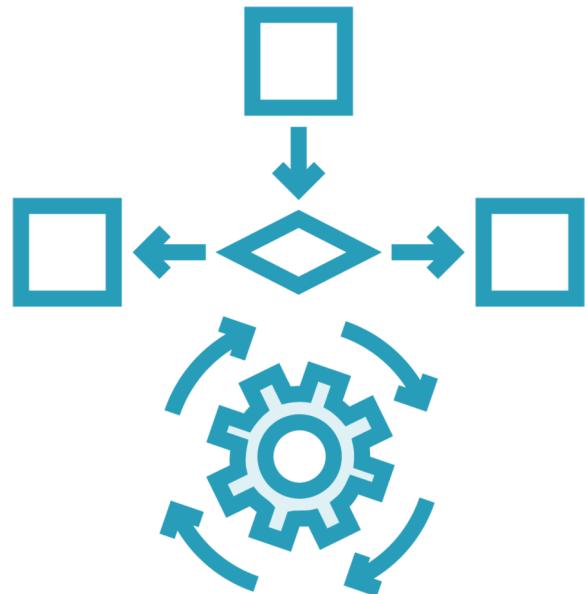
**Review the Algorithm Cheat
Sheet**

**Which Algorithm can Answer
the Question**

**What Additional
Requirements are need**



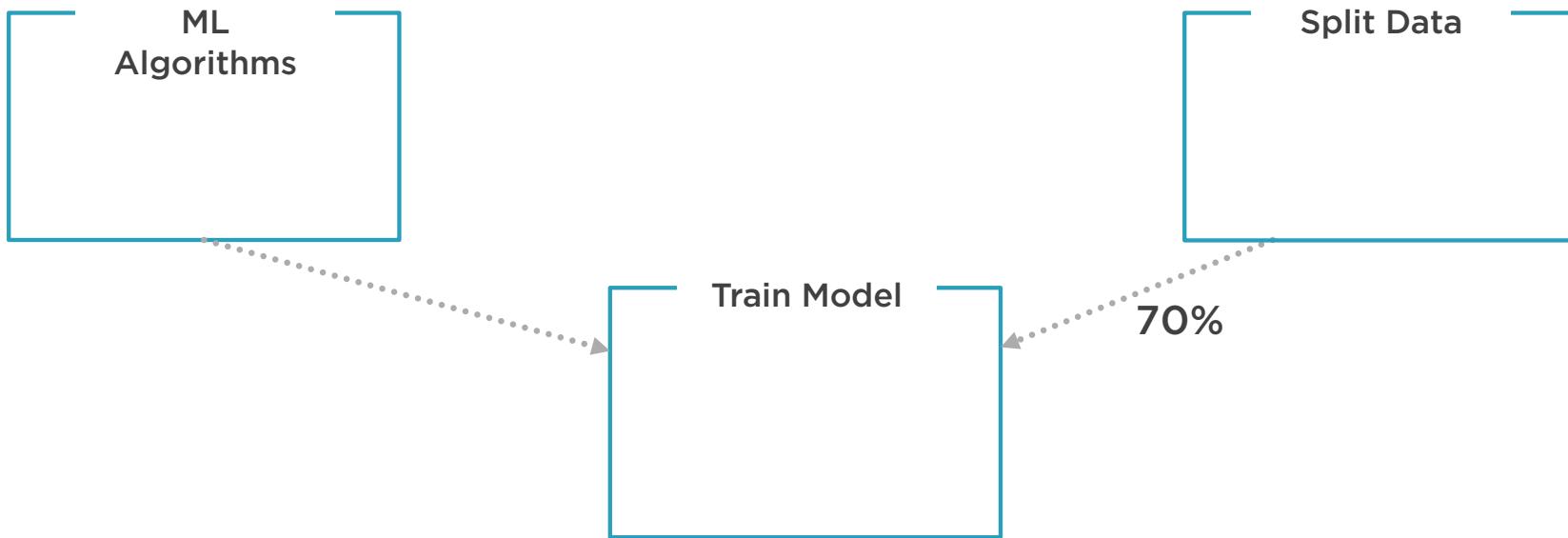
Connecting Algorithms in Azure ML Designer



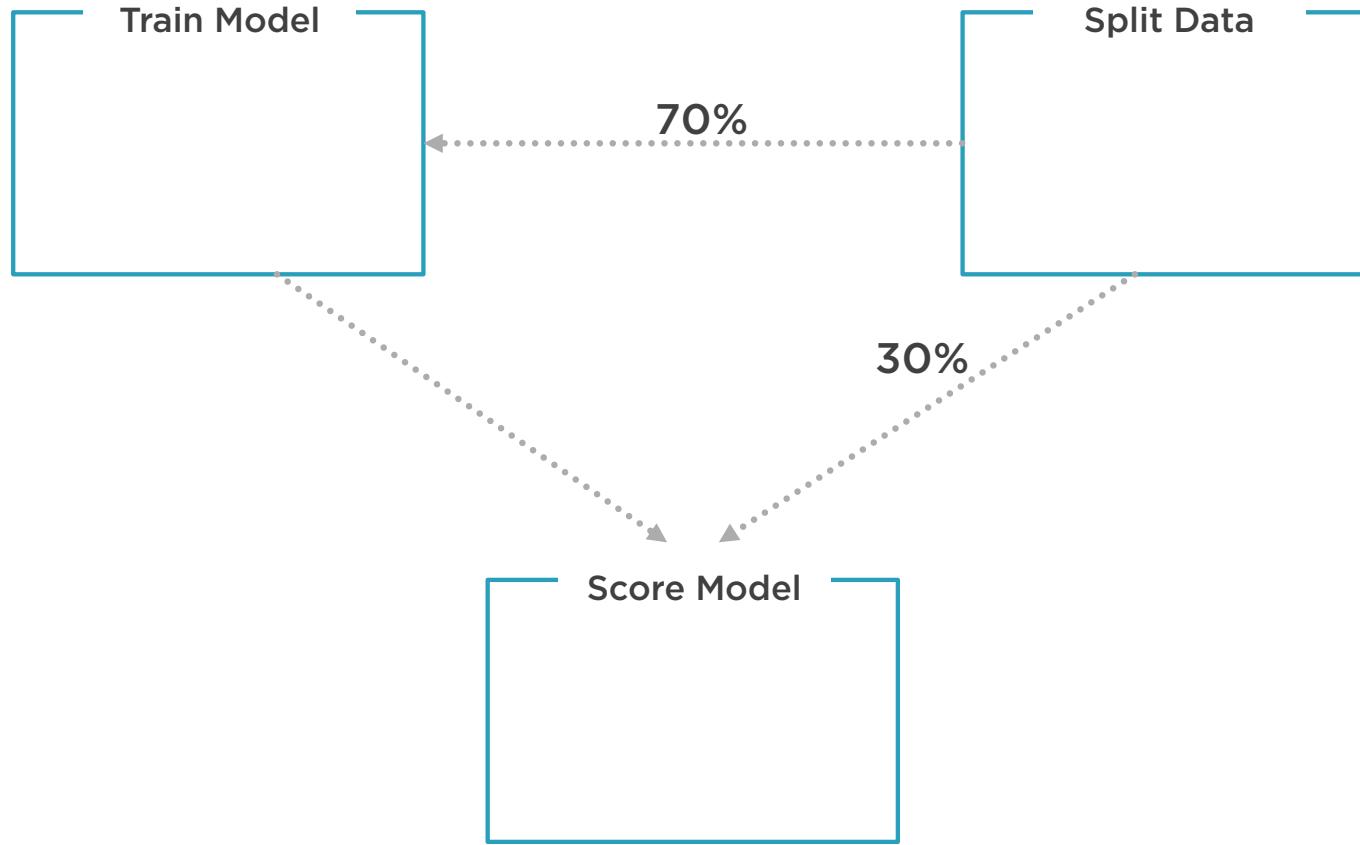
Train the Model
Score the Model
Evaluate the Model



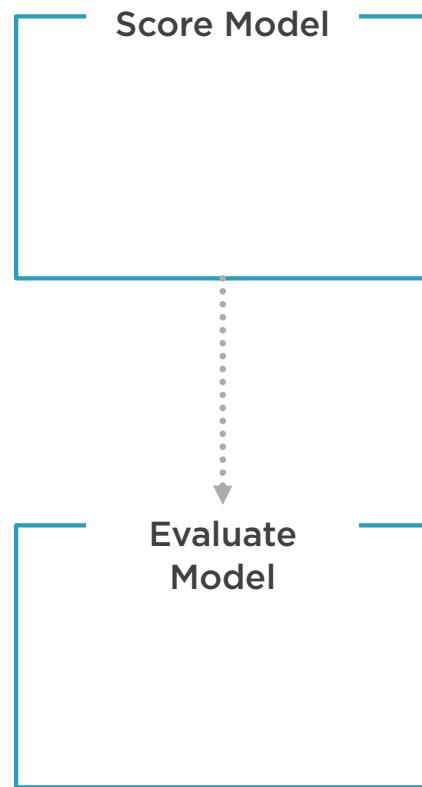
Train the Model



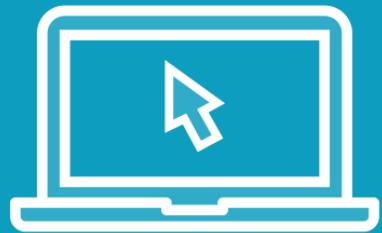
Score the Model



Evaluate the Model



Demo



Selecting an ML Algorithm in Azure ML Designer

**Connect Algorithm Module to be Trained,
Scored and Evaluated**



Summary



Defined ML Algorithms

Implemented ML Algorithm in Azure ML Studio



Publish and Deploy Machine Learning Models in Azure Machine Learning Designer



Ifedayo Bamikole

CLOUD SOLUTION ARCHITECT

@dayobam7 www.dayobam.com



Overview



Publish Azure ML Models

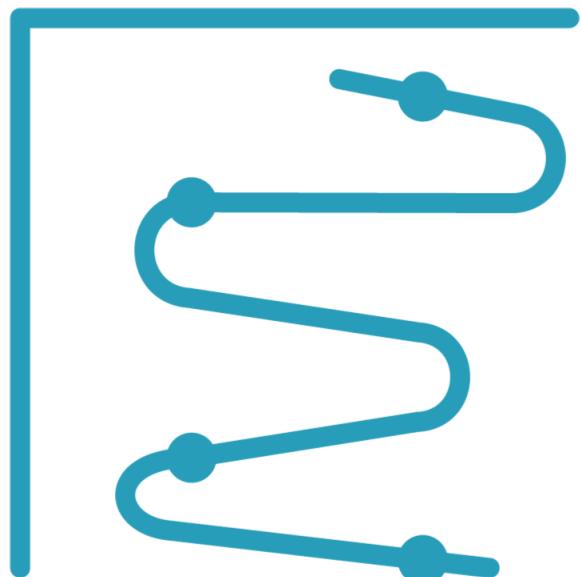
Deploy a Predictive Service

Create a Real Time Pipeline

Create a Batch Pipeline



Release your ML Pipeline



Pipeline must have been submitted

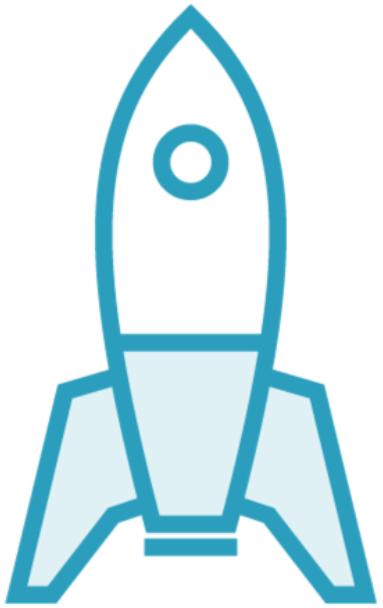
Deploy a Real-time Inference Pipeline

- Requires Azure Kubernetes Cluster

Publish a Batch Inference Pipeline

- Requires Compute Cluster





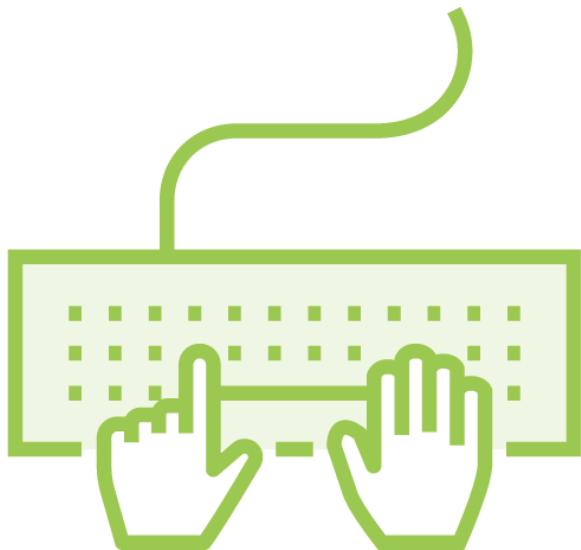
Deploy a Real-time Inference Pipeline

Ability to connect from External App to Scored Model

Predict Results

REST API Key used to call through endpoint

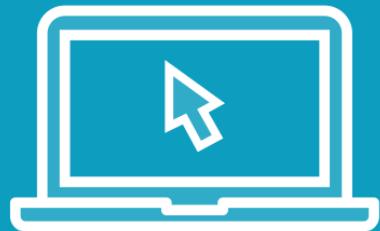




- Publish Batch Inference Pipeline**
- Runs from External Application with REST calls**
- Cannot send or receive data in real-time with endpoint**
- Flexible Pipelines**
 - Publish multiple pipeline to a single pipeline endpoint
 - Specify which pipeline version to run



Demo



Deploy & Publish an Inference Pipeline



Summary



Published Azure ML Modules

Created Real Time and Batch Pipelines

Deployed a Predictive Service

