



Azure ML Guide

Part 2 – Automated ML

October 2020

Hubert Duan, Cloud Solution Architect

Architecture

Modeling and management



Azure ML
services

Designer

Automated ML

Big data development



Azure
Databricks

Feature engineering

Spark

Data



Azure Data Lake Gen 2

Getting started

Azure provisioning

This workshop will assume services are organized within a resource group, so go ahead and create a new RG along with these services:

- Azure Data Lake Gen 2
 - Hierarchical namespace enabled
- Azure Databricks
 - Premium tier
- Azure Machine Learning
 - Enterprise edition

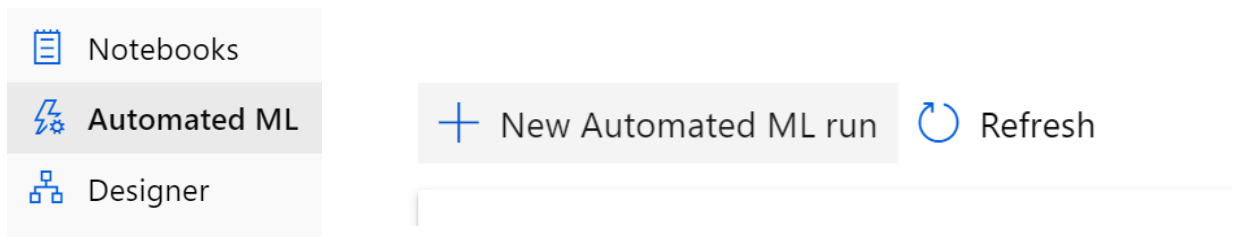
1. AutoML

Azure Machine Learning

Please refer to Part 1 of the Guide series on how to create dataset and compute cluster in Azure ML

- Go to Azure ML studio UI (ml.azure.com)

Automated ML → Computer clusters → Create / New



- You can select a dataset (or even create a new one – Part 1 of the Guide series has the details)

Select dataset

Select a dataset from the list below, or create a new dataset. Au

[+ Create dataset](#)  | ☒ Show supported datasets

Dataset name	Dataset type
<input type="radio"/> dd	Tabular
<input checked="" type="radio"/> sample_flower	Tabular
<input type="radio"/>

1. AutoML

Azure Machine Learning

- Click on create new experiment and give it a name

Dataset

sample_flower ([View dataset](#))

Experiment name *

☐

Select existing

☒

Create new

New experiment name



flower_dataset_experiment

- Select the column of the label and the compute cluster (creating one is detailed in Part 1 of the Guide series)

Target column *

class



Select compute cluster *

example



[Create a new compute](#) [Refresh compute](#)



1. AutoML


Azure Machine Learning


- You can pick the type of machine learning, between Classification, Regression, and Time series forecasting. Select Classification for now


Select task type

Select the machine learning task type for the experiment. Additional settings are available to fine tune the model as needed.

 **Classification**
To predict one of several categories in the target column. yes/no, blue, red, green. 

☐ Enable deep learning 

 **Regression**
To predict continuous numeric values

 **Time series forecasting**
To predict values based on time

- Remember to consider additional configurations and featurization settings

 [View additional configuration settings](#)  [View featurization settings](#)

1. AutoML

Azure Machine Learning

- For configurations, you can set
 - the primary metric,
 - block any algorithms from use (which is a good way as well to see all the considered algorithms), select exit criterion
 - validation type
 - and concurrency.
- Here are some common values

Primary metric ⓘ
Accuracy ▾

☒ Explain best model ⓘ

Blocked algorithms ⓘ
A list of algorithms that Automated ML will not use during training.

Exit criterion

Training job time (hours) ⓘ

Metric score threshold ⓘ

Validation

Validation type ⓘ

Number of cross validations * ⓘ

Concurrency

Max concurrent iterations ⓘ

1. AutoML

Azure Machine Learning

- For featurization, you can set
 - Feature selection
 - Specify feature type
 - And imputation method
- Here are some common values

☒ Enable featurization

Column name	Included	Feature type	Impute with	Data example
sepal_length	<input checked="" type="checkbox"/>	Auto	Auto	
sepal_width	<input checked="" type="checkbox"/>	Auto	Auto	
petal_length	<input checked="" type="checkbox"/>	Auto	Auto	
petal_width	<input checked="" type="checkbox"/>	Auto	Auto	
class (Target column)	<input type="checkbox"/>	Auto	Auto	

- Click on save and submit the automated ML run!
- Once done you can find models here:

New

Home

Author

Notebooks

Automated ML

Designer

azure-ml-hudua > Automated ML > flower_dataset_experiment > Run 1

Run 1 Preparing

Refresh Cancel

Details Data guardrails **Models** Outputs + logs Child runs Snapshot

Deploy Download Explain model