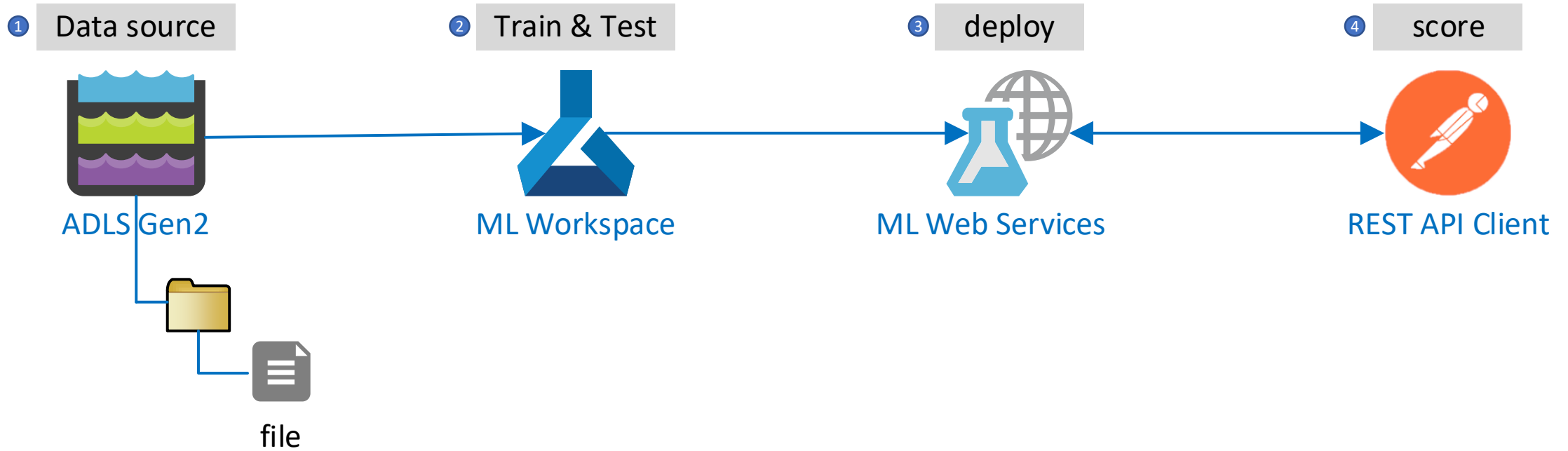


Azure Automated ML Demo

Venky Dahale (DS)
Arpana Brahmabhatt (IE)
Sridhar Kothalanka (CSA)

Mar 2020

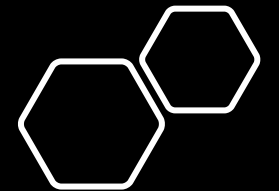




Overview

What is Azure Automated ML?

Azure Automated Machine Learning (AutoML) focuses on making the entire process of machine learning easy, with the goal of bringing efficiency to data scientists as well as enabling non-data scientists to build models



Aircraft Engine Remaining Useful Life (RUL) Dataset

Dataset Description: Simulated aircraft sensor values to predict when an aircraft engine will fail in the future so that maintenance can be planned, in advance.

Binary classification: Predict if an asset will fail within certain timeframe (e.g., 30 days)

Sample Dataset

id	cycle	setting1	setting2	setting3	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15	s16	s17	s18	s19	s20	s21	label1
1	1	0.0023	3.00E-04	100	518.67	643.02	1585.29	1398.21	14.62	21.61	553.9	2388.04	9050.17	1.3	47.2	521.72	2388.03	8125.55	8.4052	0.03	392	2388	100	38.86	23.3735	1
1	2	-0.0027	-3.00E-04	100	518.67	641.71	1588.45	1395.42	14.62	21.61	554.85	2388.01	9054.42	1.3	47.5	522.16	2388.06	8139.62	8.3803	0.03	393	2388	100	39.02	23.3916	1
2	1	-9.00E-04	4.00E-04	100	518.67	642.66	1589.3	1407.16	14.62	21.61	553.14	2388.1	9040.2	1.3	47.43	521.62	2388.14	8129.59	8.4283	0.03	392	2388	100	39	23.3923	0
2	2	-0.0011	2.00E-04	100	518.67	642.51	1588.43	1405.47	14.62	21.61	553.53	2388.07	9053.77	1.3	47.45	522.02	2388.08	8120.05	8.4414	0.03	393	2388	100	38.84	23.2902	0
3	95	0.0014	-2.00E-04	100	518.67	642.95	1595.37	1408.57	14.62	21.61	553.01	2388.1	9049.88	1.3	47.61	521	2388.16	8129.84	8.4695	0.03	393	2388	100	38.78	23.1833	0
3	96	-0.0017	-4.00E-04	100	518.67	642.83	1588.76	1406.08	14.62	21.61	552.93	2388.15	9053.68	1.3	47.63	521.34	2388.18	8131.72	8.4491	0.03	394	2388	100	38.77	23.3385	1
4	75	-7.00E-04	5.00E-04	100	518.67	642.85	1587.13	1413.85	14.62	21.61	553.16	2388.07	9054.07	1.3	47.65	521.48	2388.07	8129.5	8.4774	0.03	393	2388	100	38.78	23.2761	0
4	76	-0.0047	-4.00E-04	100	518.67	643.11	1583.44	1408.32	14.62	21.61	552.68	2388.13	9052.06	1.3	47.7	521.49	2388.15	8136.24	8.4456	0.03	393	2388	100	38.92	23.2351	1
5	67	-0.0048	1.00E-04	100	518.67	642.58	1586.59	1410.29	14.62	21.61	553.2	2388.1	9058.17	1.3	47.53	521.75	2388.06	8128.16	8.4509	0.03	394	2388	100	38.89	23.2904	0
5	68	-0.0017	-2.00E-04	100	518.67	642.9	1583.33	1413.17	14.62	21.61	553.86	2388.13	9053.78	1.3	47.6	521.08	2388.11	8131.52	8.432	0.03	394	2388	100	38.89	23.26	1
...																										
100	167	0.0011	1.00E-04	100	518.67	642.66	1590.79	1412.69	14.62	21.61	553.57	2388.04	9108.28	1.3	47.59	521.52	2388.05	8172.7	8.4855	0.03	395	2388	100	38.93	23.3488	0
100	168	-0.001	1.00E-04	100	518.67	642.94	1585.16	1406.95	14.62	21.61	553.81	2388.04	9099.31	1.3	47.47	521.65	2388.07	8181.49	8.452	0.03	393	2388	100	38.97	23.3224	1

Mapping to Bell sampleProcessData.xlsx

Column Name	Description	Bell sampleprocessdata.xlsx (guessing the mapping)
Id	Aircraft engine identifier, range [1,100]	Machine Id
Cycle	Time, in cycles	Date, Timestamp, Load Code
Setting 1, 2, 3	Operational settings (settings1 – settings3)	Setpoint e.g., “Nitrogen Setpoint”
s1 – s21	Sensor measurement (s1 – s21)	Present Value e.g., “Nitrogen PV”
label1	Ground truth - maintenance required in 30 days (1)	<not available>

Create an ADLS-Gen2 resource and initialize it

1

The screenshot shows the Microsoft Azure portal interface for a storage account named 'belldatalake'. The left sidebar contains navigation options like 'Home', 'Dashboard', 'All services', and 'FAVORITES'. The main content area displays the 'Overview' tab for the storage account, showing details such as 'Resource group: Demo-automi-datalake', 'Status: Primary: Available', 'Location: South Central US', and 'Subscription ID: 33124dac-d702-4d7d-851b-0e1159f01295'. Below the overview, there are three cards for 'Containers', 'File shares', and 'Tables'. At the bottom, there are sections for 'Tools and SDKs' (listing Storage Explorer, PowerShell, Azure CLI, .NET, Java, Python, and Node.js) and 'Monitoring' (showing data for the last 1 day).

2

The screenshot shows the 'Containers' page for the 'belldatalake' storage account. It displays a table of containers with columns for 'Name', 'Last modified', 'Public access level', and 'Last state'. A single container named 'bell' is listed, created on 8/18/2020 at 7:42:51 AM with a 'Private' public access level and an 'Available' last state.

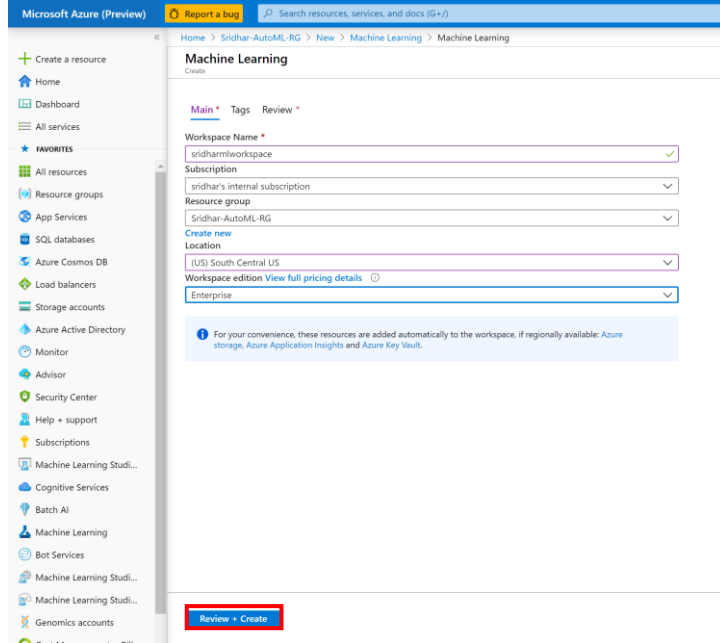
3

The screenshot shows the 'bell' container page. It displays the 'Overview' tab with a table of blobs. A single blob named 'dataset1' is listed, created on 8/18/2020 at 7:42:51 AM with a 'Private' public access level and an 'Available' last state.

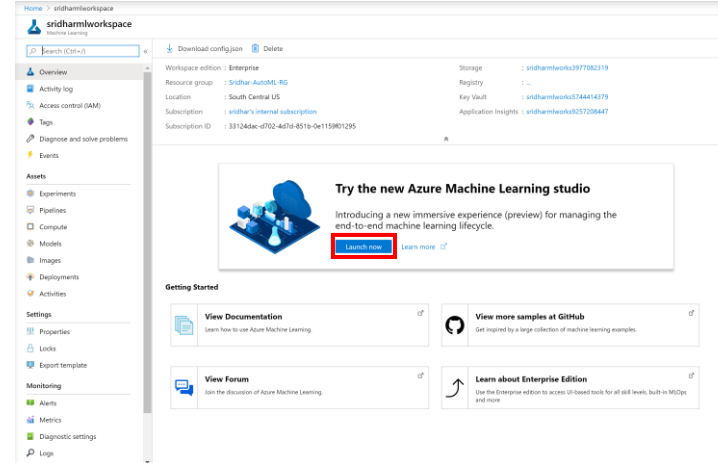
4

The screenshot shows the 'bell' container page with an 'Upload blob' dialog box open. The dialog box has a 'File' section with a 'Select a file' button and an 'Advanced' section with a 'Choose whether if files already exist' dropdown. Below the dialog box, there is a table of blobs. A single blob named 'dataset1' is listed, created on 8/18/2020 at 7:42:51 AM with a 'Private' public access level and an 'Available' last state.

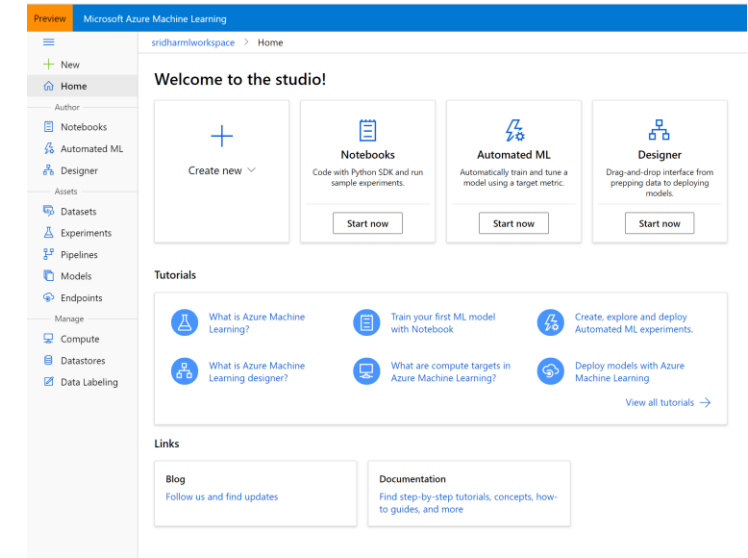
1



2



3



Create a Machine Learning workspace

1

AuthorNotebooksAutomated MLDesignerAssetsDatasetsExperimentsModelsPipelinesData Labeling

Registered datasetsDataset model

+ Create dataset

Refresh

Basic info

Datatore selection

Settings and preview

Schema

Confirm details

Select a datatore *

Previously created datatore

Create new datatore

Customers should not include personal data or other sensitive information in fields marked with * because the content in these fields may be logged and shared across Microsoft systems to facilitate operations and troubleshooting. Learn more

Datatore name *my_data_store

Datatore type *Azure Blob Storage

Account selection method

From Azure subscriptionEnter manually

Subscription ID *GridHar's internal subscription (33124dec-d702-4d7d-851b-0e1159)

Storage account *vsharddatalakegen2 (MyDataLakeGen2-RG)

Blob container *hell

Authentication type *

2

Path selection

Selected path: /

Include files in subfolders

Filter (case-sensitive and prefix-only)...

Name	Created on	Modified on
dataset1/		

< Prev

Next >

3

Selected path: dataset1/SampleProcess Data.xlsx

Include files in subfolders

Filter (case-sensitive and prefix-only)...

Name	Created on	Modified on
CNCData.xlsx	Mar 12, 2020 12:10 PM	Mar 12, 2020 12:10 PM
SampleProcess Data.xlsx	Mar 12, 2020 12:10 PM	Mar 12, 2020 12:10 PM
SampleProcessDataAlarms.xlsx	Mar 12, 2020 12:10 PM	Mar 12, 2020 12:10 PM

< Prev

Next >

Save

Cancel

4

Confirm details

Basic info

Datatore selection

File settings

Profile this dataset after creation

There are no available computes to queue profile generation

Back

Create

5

Column headers

All files have same headers

Skip rows

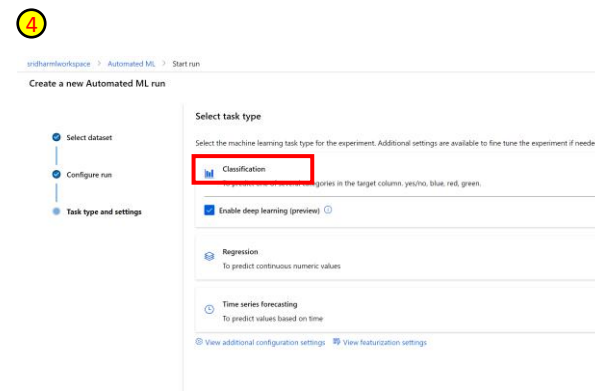
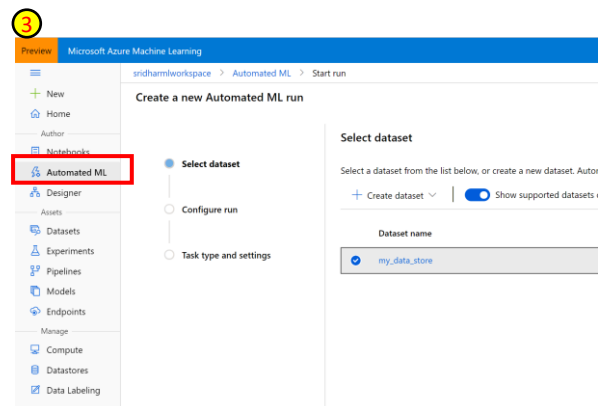
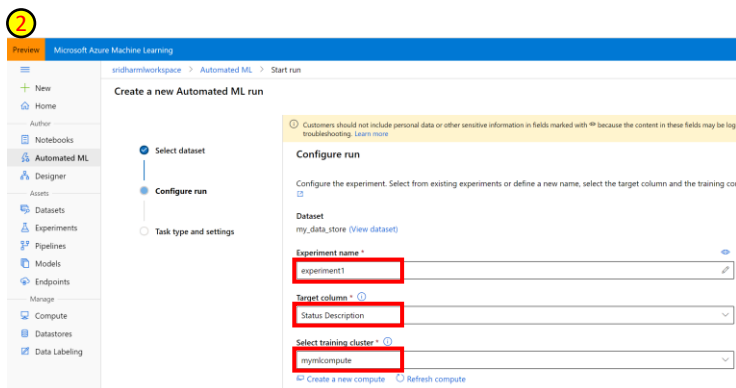
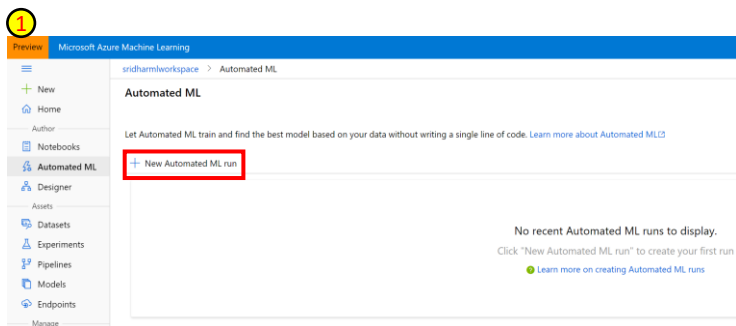
None

Date	Timestamp	Load Code	Carbon OUTPUT	Carbon PV	Carbon Setpoint	Nitrogen Setpoint	Nitrogen PV	Methanol Setp
202...	1899-12-31...	12.00	100.00	0.00	1.20	45.00	44.96	0.05
202...	1899-12-31...	12.00	91.22	0.00	1.20	45.00	45.02	0.05
202...	1899-12-31...	12.00	42.31	0.00	1.20	45.00	45.00	0.27
202...	1899-12-31...	12.00	0.00	1.27	1.20	45.00	45.06	0.27
202...	1899-12-31...	12.00	0.26	1.30	1.20	45.00	43.50	0.27
202...	1899-12-31...	12.00	2.47	1.20	1.20	45.00	45.00	0.27
202...	1899-12-31...	12.00	2.39	1.20	1.20	45.00	44.99	0.27

6

<input checked="" type="checkbox"/>	Column4	Not applicable to selecte...	Decimal	[object Object], 10
<input checked="" type="checkbox"/>	Column5	Not applicable to selecte...	Decimal	[object Object], 1.0
<input checked="" type="checkbox"/>	Column6	Not applicable to selecte...	Decimal	[object Object], 1.2
<input checked="" type="checkbox"/>	Column7	Not applicable to selecte...	Decimal	[object Object], 45
<input checked="" type="checkbox"/>	Column8	Not applicable to selecte...	Decimal	[object Object], 44
<input checked="" type="checkbox"/>	Column9	Not applicable to selecte...	Decimal	[object Object], 0.0
<input checked="" type="checkbox"/>	Column10	Not applicable to selecte...	Decimal	[object Object], -0
<input checked="" type="checkbox"/>	Column11	Not applicable to selecte...	Decimal	[object Object], 3.1
<input checked="" type="checkbox"/>	Column12	Not applicable to selecte...	Decimal	[object Object], -0
<input checked="" type="checkbox"/>	Column13	Not applicable to selecte...	Decimal	[object Object], 18

Create a Dataset



Create an Auto ML run

1

sridharmworkspace > Automated ML > Run Detail

Run 1 Completed

[Refresh](#) [Cancel](#)

[Details](#) [Data guardrails](#) [Models](#) [Logs](#) [Outputs](#)

Algorithm name	Accuracy ↓
MaxAbsScaler, XGBoostClassifier	0.9947289156626506
MaxAbsScaler, LightGBM	0.9866967871485943
MinMaxScaler, RandomForest	0.9583333333333334
StandardScalerWrapper, ExtremeRandomTrees	0.9535642570281124
StandardScalerWrapper, SGD	0.9515562248995983
MinMaxScaler, SGD	0.947289156626506
RobustScaler, LogisticRegression	0.9455321285140562
StandardScalerWrapper, SGD	0.9440261044176707
SparseNormalizer, LogisticRegression	0.9425200803212851
SparseNormalizer, LogisticRegression	0.9397590361445783

2

sridharmworkspace > Automated ML > Run Detail

Run 1 Completed

[Refresh](#) [Cancel](#)

[Details](#) [Data guardrails](#) [Models](#) [Logs](#) [Outputs](#)

Model summary

Algorithm name
MaxAbsScaler, XGBoostClassifier

Accuracy
0.9947289156626506

Registered models
[AutoMLb041092db1:1](#)

Deploy status
[experiment1](#) Running

[Deploy best model](#) [View model details](#) [Download best model](#)

3

Deploy a model

Name *

Description

Compute type *

Models: AutoMLb041092db1

Enable authentication
☒

☐ Use custom deployment assets

[Advanced](#)

4

sridharmworkspace > Endpoints > experiment1

experiment1

[Details](#) [Consume](#)

Deployment state
Healthy

Compute type
ACI

Service ID
experiment1

Tags

Created on
3/13/2020 1:19:06 PM

Last updated on
3/13/2020 1:19:06 PM

Compute target
N/A

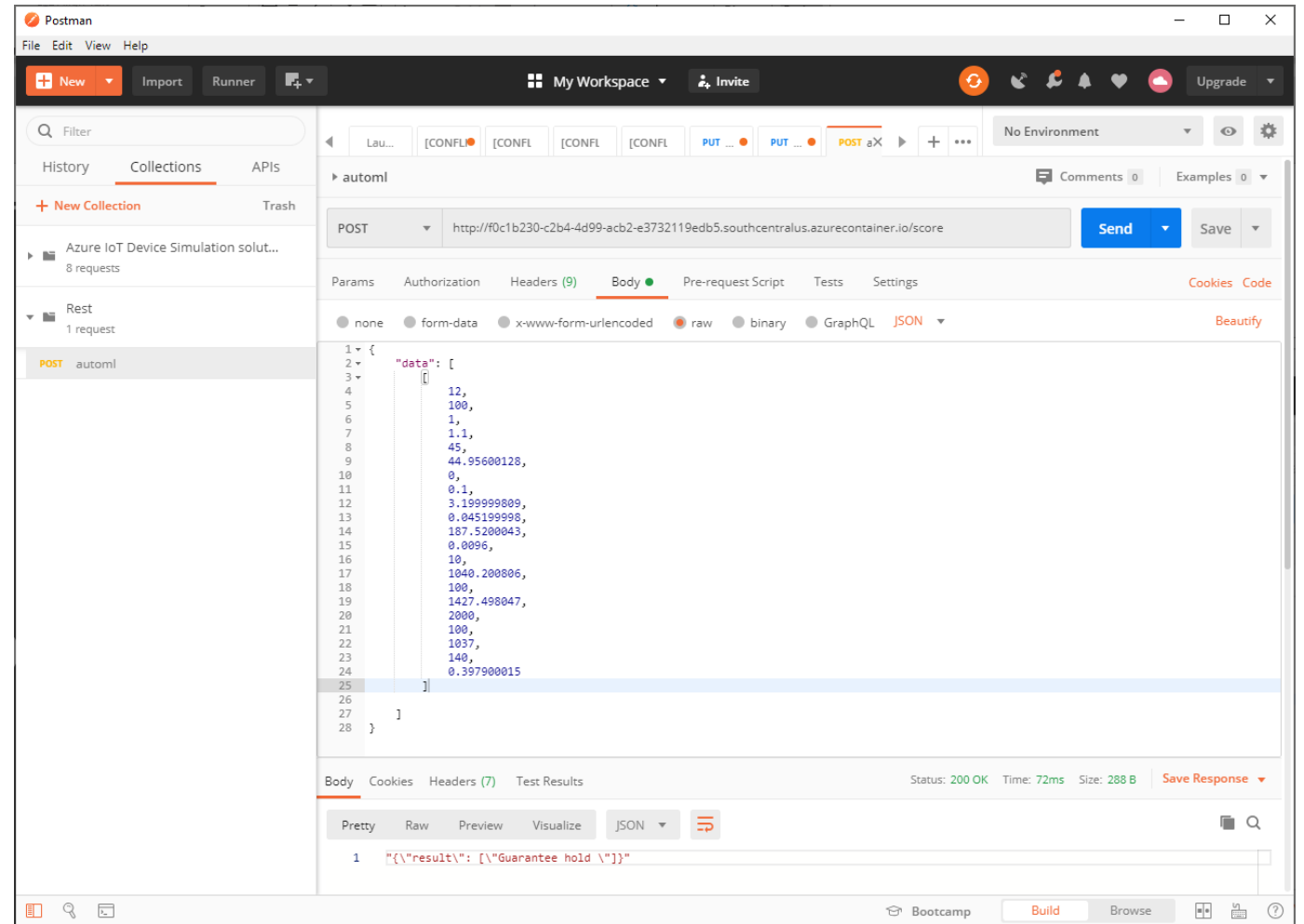
REST endpoint
<http://f0c1b230-c2b4-4d99-acb2-e3732119edb5.southcentralus.azurecontainer.io/score>

Key-based authentication enabled
false

Token-based authentication enabled
false

Deploy the best model

Score



Thank You