

CLOUD SKILLS CHALLENGE

## Azure Data Scientist

Design and implement a data science solution on Azure. Learn how to build machine learning models, no-code predictive models, and ML solutions, and run data science workloads in the cloud. In under 7 hours, you'll learn how to build and operationalize machine learning models using Microsoft Azure in a secure, scalable, and responsible way.

 Add



**Congratulations, you've completed the challenge!**

Modules you've completed **10 / 10**

Scores updated within a few hours

[See your completed collection](#)

Build more skills on Microsoft Learn



**Time to complete challenge:**

28    22    07    00

DAYS

HOURS

MINUTES

SECONDS

https://ml.azure.com/runs/8a6f0250-c07c-4e69-b17f-6233843a4a2a?wsid=/subscriptions/2396fe70-b8e1-4bdd-bb

Azure AI | Machine Learning Studio

Default Directory > mlw-dp100-labs > Jobs > train-regression-designer > Regression - Automobile Price Prediction (Basic)

All workspaces Refresh Clone Resubmit View profiling Publish Schedule ... Cancel

Home Model catalog PREVIEW

Regression - Automobile Price Prediction (Basic) Running Share Add to compare Job overview

Automating

- Notebooks
- Automated ML
- Designer
- Prompt flow PREVIEW

Assets

- Data
- Jobs
- Components
- Pipelines
- Environments

Linear Regression linear\_regression  
Select Columns in Dataset select\_columns\_in\_dataset  
Clean Missing Data clean\_missing\_data  
Split Data split\_data  
Train Model train\_model  
Score Model score\_model

```
graph TD; A[Automobile price data (Raw)] --> B[Select Columns in Dataset<br>select_columns_in_dataset<br/>Exclude normalized losses which have many missing values]; B --> C[Clean Missing Data<br>clean_missing_data<br/>Remove missing value rows]; C --> D[Split Data<br>split_data<br/>Split the dataset into training set (0.7) and test set (0.3)]; D --> E[Train Model<br>train_model]; E --> F[Score Model<br>score_model];
```

Navigator 71% 1:1

All workspaces

Home

Model catalog PREVIEW

## Authoring

Notebooks

Automated ML

Designer

Prompt flow PREVIEW

## Assets

Data

Jobs

Components

Pipelines

Environments

Default Directory &gt; mlw-dp100-labs &gt; Jobs &gt; train-regression-designer &gt; Regression - Automobile Price Prediction (Basic)

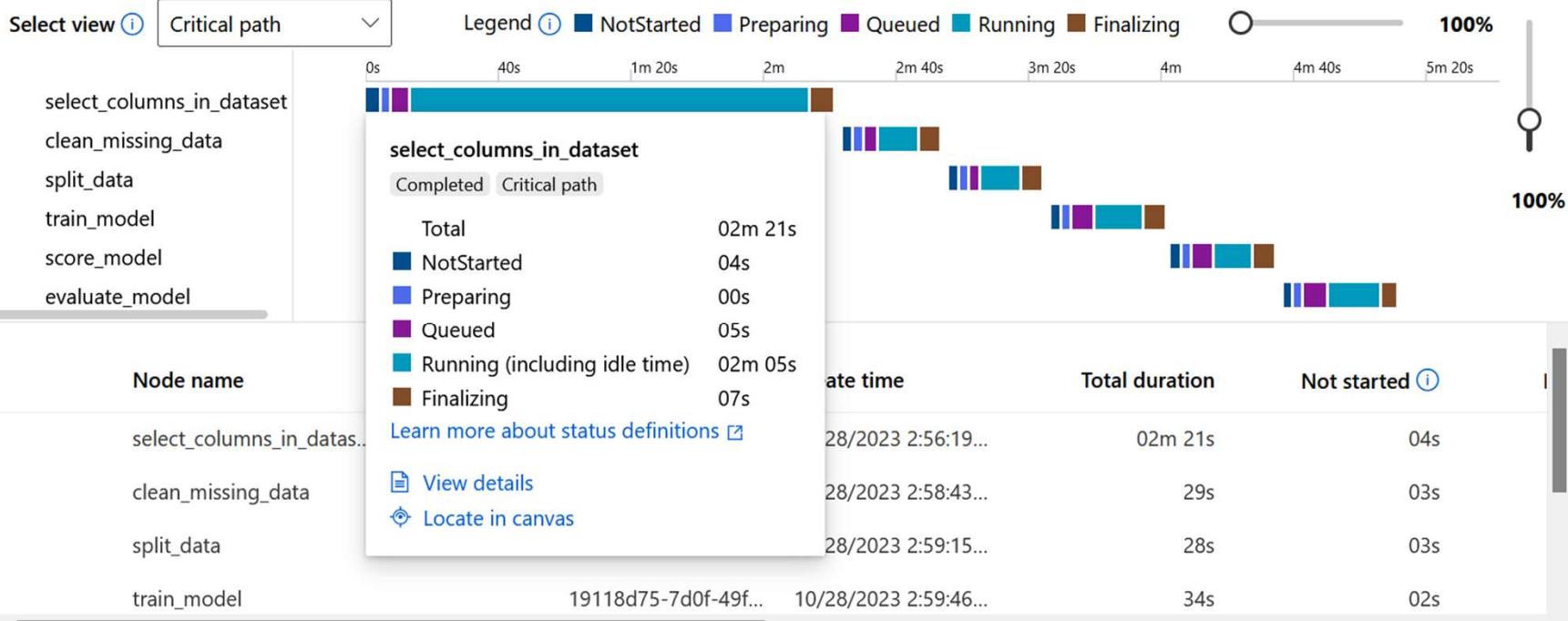
Refresh Clone Resubmit View profiling Publish Schedule ... Show lineage ... Cancel

Regression - ... X

Profiling - Regression - Automobile Price Prediction (Basic) Completed

Search

I → Export CSV



## first-compute

### Resource properties

#### Status

 Running

#### Last operation

Created at Oct 28, 2023 2:49 PM: Succeeded

#### Virtual machine size

Standard\_DS11\_v2 (2 cores, 14 GB RAM, 28 GB disk)

#### Processing unit

CPU - Memory optimized

#### Estimated cost

\$0.19/hr (when running)

#### Additional data storage

--

#### Applications

JupyterLab Jupyter VS Code (Web)  VS Code (Desktop)  Terminal ...

#### Created on

28/10/2023, 2:48:28 pm

#### SSH access

Disabled

#### Private IP address

10.0.0.5 

#### Virtual network/subnet

--

#### Public IP address

## Pipeline job overview

Overview Settings Outputs + logs Metrics Child jobs Images

### Properties

#### Job display Name

Regression - Automobile Price Prediction (Basic)

#### Status

 Completed

#### Created by

Harsh Yadav

#### Total steps

7

#### Job name (Run ID)

8a6f0250-c07c-4e69-b17f-6233843a4a2a

#### Published pipeline

--

#### Experiment

train-regression-designer

#### Registered models

None

#### Created on

Oct 28, 2023 2:56 PM

#### End time

Oct 28, 2023 3:01 PM

**Outline**


[+ Add filter](#)
**Regression - Automobile Price Predi...**

linear\_regression

Automobile price data (Raw)

select\_columns\_in\_dataset

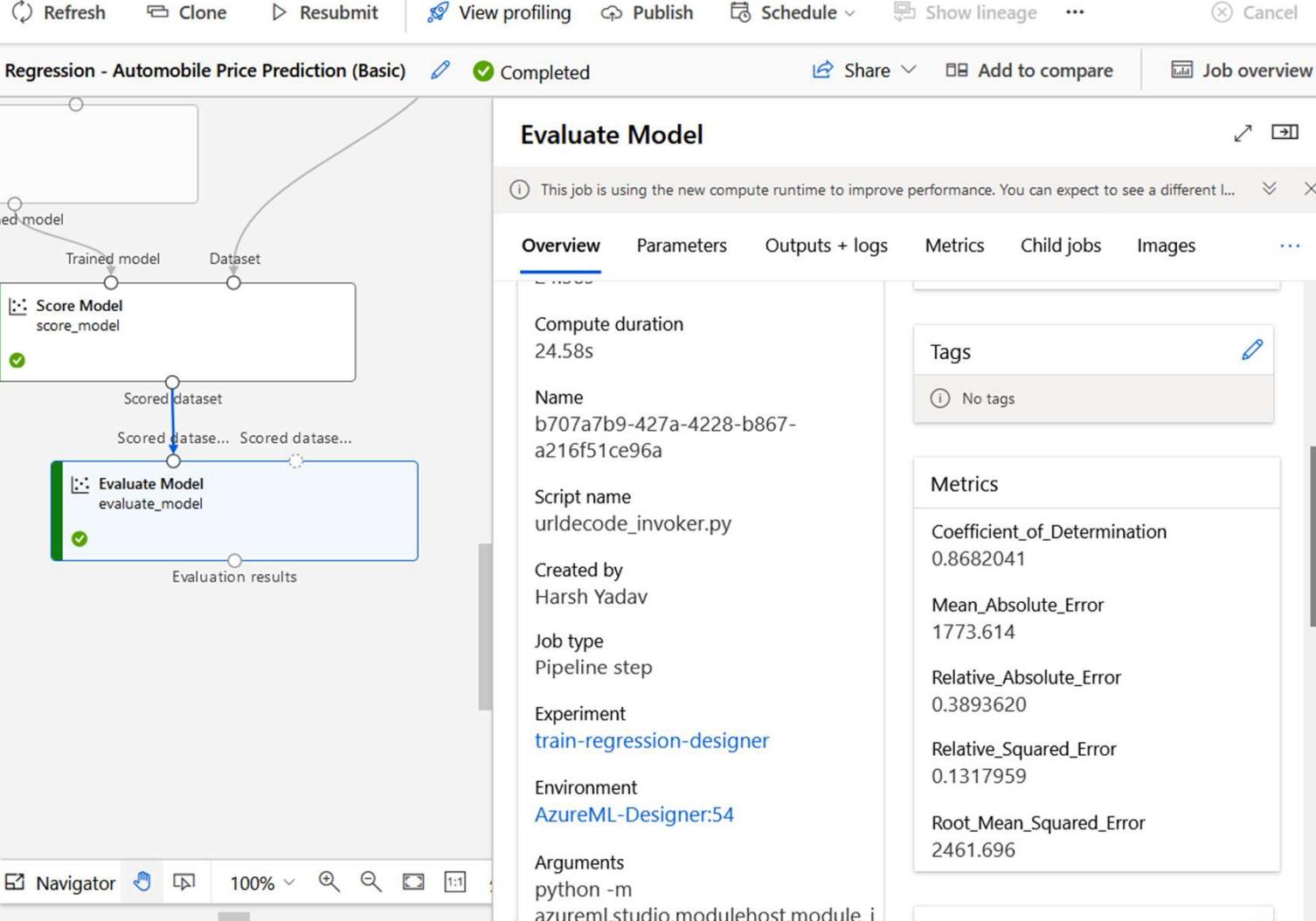
clean\_missing\_data

split\_data

train\_model

score\_model

evaluate\_model



Model catalog PREVIEW

Default Directory &gt; mlw-dp100-l902807c27eee4927a7 &gt; Compute &gt; aml-cluster

aml-cluster ☆

## Authoring

Notebooks

Automated ML

Designer

Prompt flow PREVIEW

## Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

## Manage

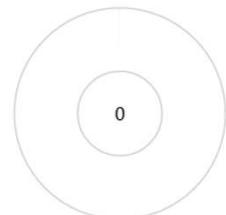
Compute

Monitoring PREVIEW

Data Labeling

Linked Services

## Cluster node status



- Idle
- Leaving
- Preparing
- Running

## Cluster state

## Allocation state

✓ Succeeded (0 nodes)

## Allocation state transition time

28/10/2023, 6:23:00 pm

## Created on

28/10/2023, 6:22:52 pm

## Current node count

0

## Attributes

- Compute name  
aml-cluster
- Resource ID  
--
- Compute type  
Machine Learning compute
- Subscription ID  
2396fe70-b8e1-4bdd-bb3d-3c0aa7bfeb5c
- Resource group  
rg-dp100-l902807c27eee4927a7
- Workspace  
mlw-dp100-l902807c27eee4927a7
- Region  
East US

## Resource properties

- Virtual machine size  
Standard\_DS11\_v2 (2 cores, 14 GB RAM, 28 GB disk)
- Processing unit  
CPU - Memory optimized
- Estimated cost  
\$0.18/hr per node
- OS Type  
Linux
- Virtual machine tier  
Dedicated
- Minimum number of nodes  
0
- Maximum number of nodes  
100

Default Directory > mlw-dp100-l902807c27eee4927a7 > Notebooks

## Notebooks

**Files**      **Samples**

## Samples



1:ci902807c27eee4927a X

\*Work with data.ipynb X



Edit in VS Code

Compute:

ci90 ...



Pyth... Python 3.8 - AzureML

ci902807c27eee4927a7 · Kernel idle CPU 0% RAM 2%

Last saved a few seconds ago

```
'Resource__source_path': None, 'base_path': '/mnt/batch/tasks/shared/LS_root/mounts/clusters/ci902807c27eee4927a7/code/Users/anikaryadav2002/azure-ml-labs/Labs/03', 'creation_context': None, 'serialize': <msrest.serialization.Serializer object at 0x7f5d885cd700>, 'credentials': {'type': 'account_key'}, 'container_name': 'training-data', 'account_name': 'mlwdp100storage74c949317', 'endpoint': 'core.windows.net', 'protocol': 'https'})
```

List the datastores again to verify that a new datastore named `blob_training_data` has been created:



```
1 stores = ml_client.datastores.list()
2 for ds_name in stores:
3     print(ds_name.name)
[6] ✓ <1 sec
...
blob_training_data
workspaceworkingdirectory
workspacefilestore
workspaceblobstore
workspaceartifactstore
```

```
27     returned_job = ml_client.create_or_update(job)
28     aml_url = returned_job.studio_url
29     print("Monitor your job at", aml_url)
```

[16] ✓ 5 sec

```
... https://aka.ms/azuremlexperimental for more information.
Class AutoDeleteConditionSchema: This is an experimental class, and may change at any time. Please see
https://aka.ms/azuremlexperimental for more information.
Class BaseAutoDeleteSettingSchema: This is an experimental class, and may change at any time. Please see
https://aka.ms/azuremlexperimental for more information.
Class IntellectualPropertySchema: This is an experimental class, and may change at any time. Please see
https://aka.ms/azuremlexperimental for more information.
Class ProtectionLevelSchema: This is an experimental class, and may change at any time. Please see https://aka.ms
/azuremlexperimental for more information.
Class BaseIntellectualPropertySchema: This is an experimental class, and may change at any time. Please see
https://aka.ms/azuremlexperimental for more information.
Uploading src (0.0 MBs): 100%|██████████| 1113/1113 [00:00<00:00, 54865.84it/s]
```

Monitor your job at [https://ml.azure.com/runs/happy\\_date\\_6qkjxfvpq7?wsid=/subscriptions/2396fe70-b8e1-4bdd-bb3d-3c0aa7bfeb5c/resourcegroups/rg-dp100-1902807c27eee4927a7/workspaces/mlw-dp100-1902807c27eee4927a7&tid=67619fe8-0e63-4838-8be4-ff1a924699f6](https://ml.azure.com/runs/happy_date_6qkjxfvpq7?wsid=/subscriptions/2396fe70-b8e1-4bdd-bb3d-3c0aa7bfeb5c/resourcegroups/rg-dp100-1902807c27eee4927a7/workspaces/mlw-dp100-1902807c27eee4927a7&tid=67619fe8-0e63-4838-8be4-ff1a924699f6)

## move-diabetes-data 🖊️ ⭐️ ⏴️ Running

Overview Metrics Images Child jobs Outputs + logs Code Explanations (preview) Fairness (preview) Monitoring

**Properties**

Status  
▶️ Running

Created on  
Oct 28, 2023 6:51 PM

Start time  
Oct 28, 2023 6:54 PM

Name  
happy\_date\_6qkjxfvpq7

Command  
`python move-data.py --input_data ${{inputs.local_data}}  
--output_datastore ${outputs.datastore_data}}`

Created by  
Harsh Yadav

**Inputs**  
Input name: local\_data  
Data asset: [diabetes-local:1](#)  
Asset URI:  

**Tags**  
 No tags 

**Metrics**  
 No data 

**Description** 

Microsoft Azure Search resources, services, and docs (G/) 2 ? ? anikaryadav2002@gmail.com DEFAULT DIRECTORY (ANIKARYA...)

Home > Resource groups >

## Resource groups

Default Directory (anikaryadav2002@gmail.onmicrosoft.com)

+ Create Manage view ...

Filter for any field...

Name	Type	Location	...
cloud-shell-storage-centralindia	Action group	Global	...
rg-dp100-l902807c27eee4927a7	Resource group	West US	...

Page 1 of 1

**rg-dp100-l902807c27eee4927a7** Resource group

Search

+ Create Manage view Delete resource group Refresh Export to CSV ...

Filter for any field... Type equals all Add filter More (1)

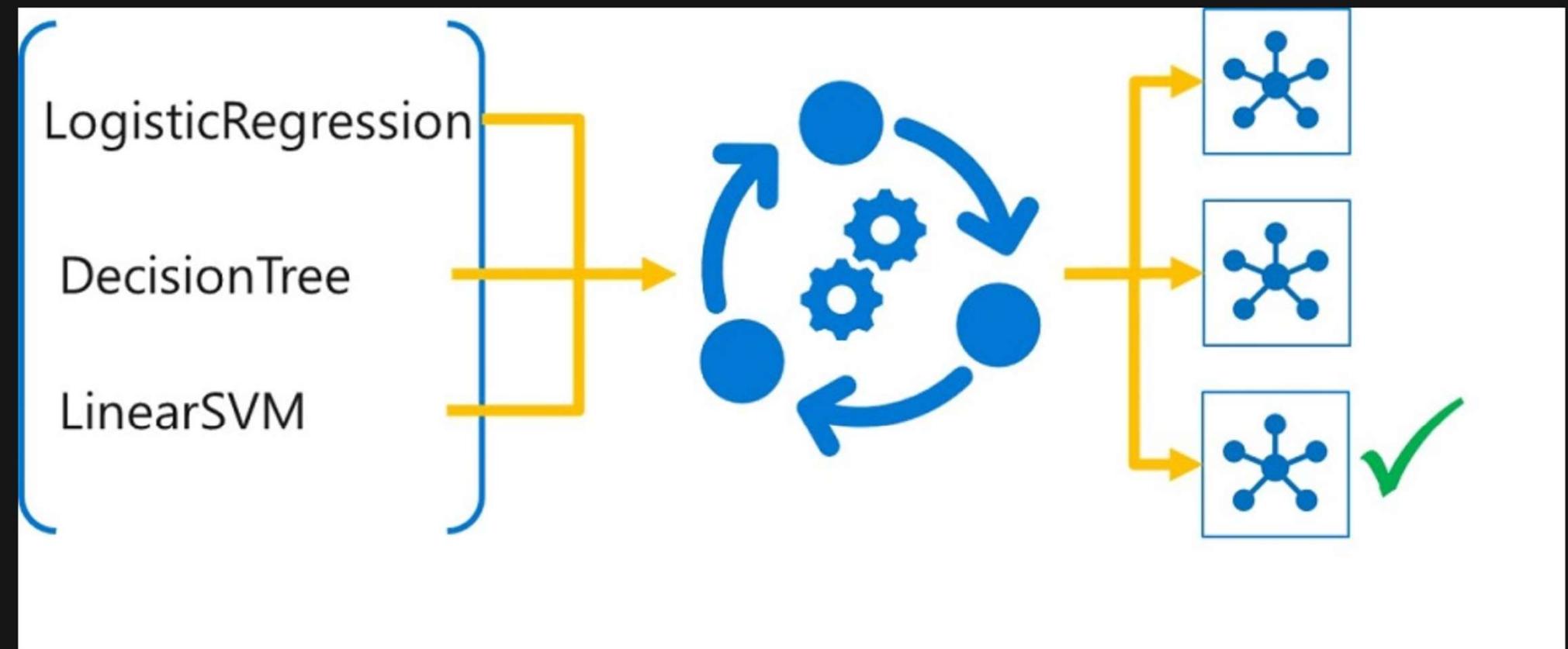
Show 1 to 7 of 7 records.  Show hidden types

No grouping

Name	Type	Location	...
Application Insights Smart Detection	Action group	Global	...
Failure Anomalies - mlwdp100insights3...	Smart detector alert rule	Global	...
mlw-dp100-l902807c27eee4927a7	Azure Machine Learni...	West US	...
mlwdp100insights3f8038d6	Application Insights	West US	...
mlwdp100keyvault896ad9fb	Key vault	West US	...
mlwdp100logalyti67512998	Log Analytics workspace	West US	...

< Page 1 > Print

AutoML allows you to try multiple preprocessing transformations and algorithms with your data to find the best machine learning model.



# Configure automated machine learning job

Now, you're ready to configure the automated machine learning experiment.

When you run the code below, it will create an automated machine learning job that:

- Uses the compute cluster named `aml-cluster`
- Sets `Diabetic` as the target column
- Sets `accuracy` as the primary metric
- Times out after `60` minutes of total training time
- Trains a maximum of `5` models
- No model will be trained with the `LogisticRegression` algorithm

```
1  from azure.ai.ml import automl
2
3  # configure the classification job
4  classification_job = automl.classification(
5      compute="aml-cluster",
6      experiment_name="auto-ml-class-dev",
7      training_data=my_training_data_input
```



Edit in VS Code

Compute:

ci08 ...



Pyth... ▾



ci084e3b8c44674341a1 · Kernel idle CPU 0% RAM 2%

Last saved a few seconds ago

Python 3.8 - AzureML

OK, you're ready to go. Let's run the automated machine learning experiment.

**Note:** This may take some time!

```
1 # Submit the AutoML job
2 returned_job = ml_client.jobs.create_or_update(
3     classification_job
4 )
5
6 # submit the job to the backend
7 aml_url = returned_job.studio_url
8 print("Monitor your job at", aml_url)
```

[6]

✓ 2 sec

Monitor your job at [https://ml.azure.com/runs/loyal\\_parsnip\\_42vbf84sn8?wsid=/subscriptions/2396fe70-b8e1-4bdd-bb3d-3c0aa7bfeb5c/resourcegroups/rg-dp100-1084e3b8c44674341a1/workspaces/mlw-dp100-1084e3b8c44674341a1&tid=67619fe8-0e63-4838-8be4-ff1a924699f6](https://ml.azure.com/runs/loyal_parsnip_42vbf84sn8?wsid=/subscriptions/2396fe70-b8e1-4bdd-bb3d-3c0aa7bfeb5c/resourcegroups/rg-dp100-1084e3b8c44674341a1/workspaces/mlw-dp100-1084e3b8c44674341a1&tid=67619fe8-0e63-4838-8be4-ff1a924699f6)

# loyal\_parsnip\_42vbf84sn8



Running

[Overview](#) [Data guardrails](#) [Models](#) [Outputs + logs](#) [Child jobs](#)[Refresh](#)[Edit and submit \(preview\)](#)[Register model](#)[Cancel](#)[Delete](#)

Data guardrails are run by Automated ML when automatic featurization is enabled. This is a sequence of checks over the input data to ensure high quality data is being used to train model.

Type	Status	Description	
Class balancing detection	Passed	Your inputs were analyzed, and all classes are balanced in your training data. <a href="#">Learn more about imbalanced data.</a>	
Missing feature values imputation	Passed	No feature missing values were detected in the training data. <a href="#">Learn more about missing value imputation.</a>	
High cardinality feature detection	Passed	Your inputs were analyzed, and no high cardinality features were detected. <a href="#">Learn more about high cardinality feature detection.</a>	

## loyal\_parsnip\_42vbf84sn8 Completed

Overview Data guardrails **Models** Outputs + logs Child jobs

 Refresh  Deploy  Download  Explain model  View generated code  View options

 Search

 Filter

 Columns

Algorithm name	Explained	Responsible AI	Accuracy ↓	Sampling	Cre
VotingEnsemble	<a href="#">View explanation</a>		0.95300	100.00 %	Oct
StackEnsemble			0.95280	100.00 %	Oct
MaxAbsScaler, LightGBM			0.95180	100.00 %	Oct
MaxAbsScaler, XGBoostClassifier			0.95180	100.00 %	Oct
MaxAbsScaler, ExtremeRandomTrees			0.83090	100.00 %	Oct

6aaaf457

30395894

**DATA STATISTICS**  
Binary classifier  
5000 datapoints  
9 features

**DATASET COHORTS**  
All data  
5000 datapoints  
0 filters

**+ New cohort**

**Edit cohort**

Model performance   Dataset explorer   **Aggregate feature importance**   Individual feature importance

Explore the top-k important features that impact your overall model predictions (a.k.a. global explanation). Use the slider to show descending feature importances. All cohorts' feature importances are shown side by side and can be toggled off by selecting the cohort in the legend. Click on any of the features in the graph to see a density plot below of how values of the selected feature affect prediction.

**Top 4 features by their importance**

Feature	Importance
Pregnancies	1.5432893468177915
Age	1.1
BMI	0.7
SerumInsulin	0.6

Sort by cohort

All data

Chart type

Bar

Box

Class importance weights

Class: 0

# serene\_dolphin\_brqxhy1x



Completed

Overview Model Explanations (preview) Responsible AI (preview) Metrics Data transformation (preview) Test results (preview) Outputs + logs Images Child jobs Code ...

DATA STATISTICS  
Binary classifier  
5000 datapoints  
9 features

DATASET COHORTS  
All data  
5000 datapoints  
0 filters

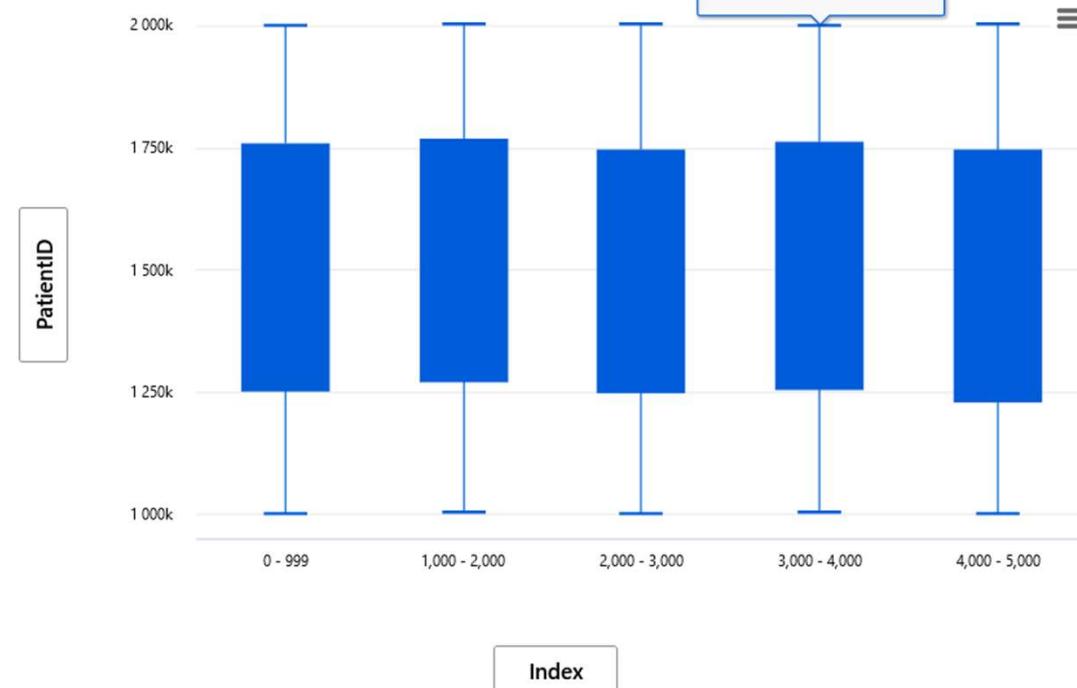
Select a dataset cohort to explore

All data

Test results (preview)  
3,000 - 4,000  
• PatientID  
Upper whisker: 1 998 691  
Upper quartile: 1 759 071  
Median: 1 507 737  
Lower quartile: 1 257 217.5  
Lower whisker: 1 003 003

Chart type

- Aggregate plots
- Individual datapoints



Index

## serene\_dolphin\_brqxhy1x Completed

Overview Model Explanations (preview) Responsible AI (preview) **Metrics** Data transformation (preview) Test results (preview) Outputs + logs Images Child jobs Code

 Refresh  Cancel

 Create custom chart

 View as... 

Current view: Local  Edit view 

 Select metrics

  accuracy <b>0.953</b>	  AUC_macro <b>0.9907651</b>	  AUC_micro <b>0.9917770</b>	  AUC_weighted <b>0.9907651</b>	  average_precision_scor... <b>0.9889315</b>	  average_precision_scor... <b>0.9919709</b>	  average_precision_scor... <b>0.9910882</b>
  balanced_accuracy <b>0.9454798</b>	  f1_score_macro <b>0.9469672</b>	  f1_score_micro <b>0.953</b>	  f1_score_weighted <b>0.9529191</b>	  log_loss <b>0.1155661</b>	  matthews_correlation <b>0.8940360</b>	  norm_macro_recall <b>0.8909597</b>
  precision_score_macro <b>0.9485671</b>	  precision_score_micro <b>0.953</b>	  precision_score_weighted <b>0.9529288</b>	  recall_score_macro <b>0.9454798</b>			

serene\_dolphin\_brqxhy1x ✎ ⚡ ✓ Completed

Overview Model Explanations (preview) Responsible AI (preview) Metrics Data transformation (preview) Test results (preview) Outputs + logs

⟳ Refresh ⌂ Cancel

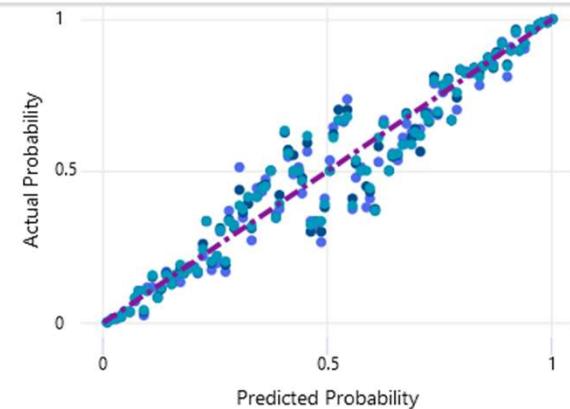
\_CREATE CUSTOM CHART

View as... ▾

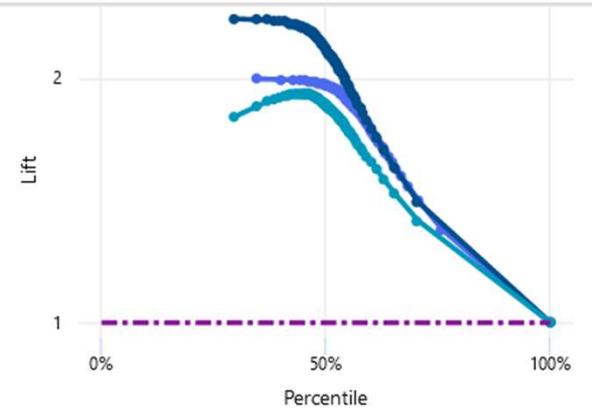
Current view: Local ▾

Edit view ▾

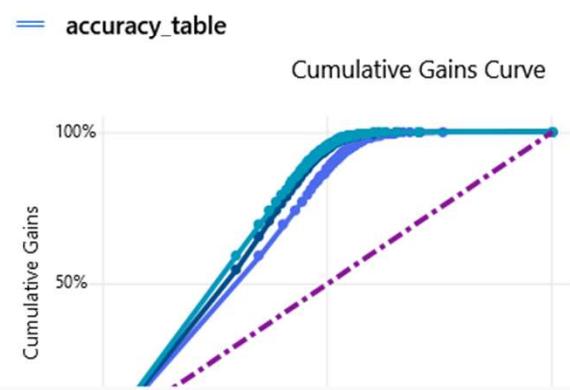
Select metrics



- Weighted Average
- Macro Average
- Micro Average
- Ideal
- 0
- 1



- Weighted Average
- Macro Average
- Micro Average
- Random
- 0
- 1



- Weighted Average
- Macro Average
- Micro Average
- Random
- 0
- 1



Raw

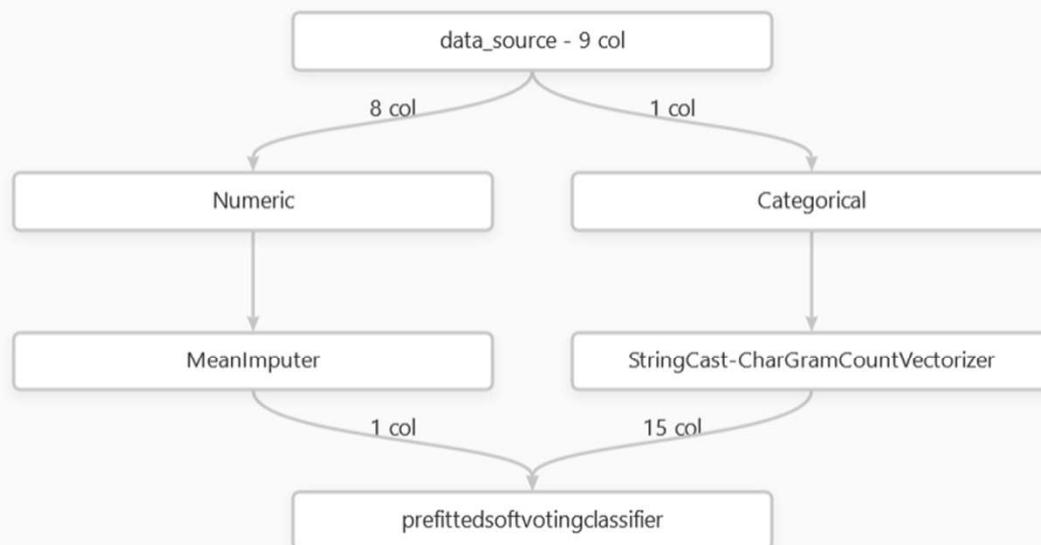
Confusion Matrix

serene\_dolphin\_brqxhy1x 🎒 ⭐️ ✅ Completed

Overview Model Explanations (preview) Responsible AI (preview) Metrics Data transformation (preview) Test results (preview) Outputs + logs Images Child jobs Code

⟳ Refresh ➡ Deploy ⏪ Download ⚡ Explain model # View generated code ✓ Test model (preview) + Register model ✖ Cancel 🗑 Delete

The following diagram illustrates the data preprocessing, feature engineering, scaling techniques and the machine learning algorithm that Automated ML applied to generate this particular model.



(i) This job is using the new compute runtime to improve performance. You can expect to see a different log structure along with the new runtime.

## diabetes-train-script

Overview Metrics Images Child jobs Outputs + logs Code Explanations (preview) Fairness (preview) Monitoring

 Refresh  Debug and monitor  Edit and submit  Perform sweep  Register model  Cancel  Delete

  std\_log.txt 

>  system\_logs

▽  user\_logs

  std\_log.txt

1  
2  
3  
4 \*\*\*\*\*  
5 Reading data...  
6 Splitting data...  
7 Training model...  
8 Accuracy: 0.774  
9 AUC: 0.8483203144435048  
10 \*\*\*\*\*  
11  
12  
13  
14

## diabetes-train-mlflow



Completed

Overview Metrics **Images** Child jobs Outputs + logs Code Explanations (preview) Fairness (preview) Monitoring

⟳ Refresh

🔍 Debug and monitor

✍ Edit and submit

▷ Perform sweep

✚ Register model

✖ Cancel

trash Delete

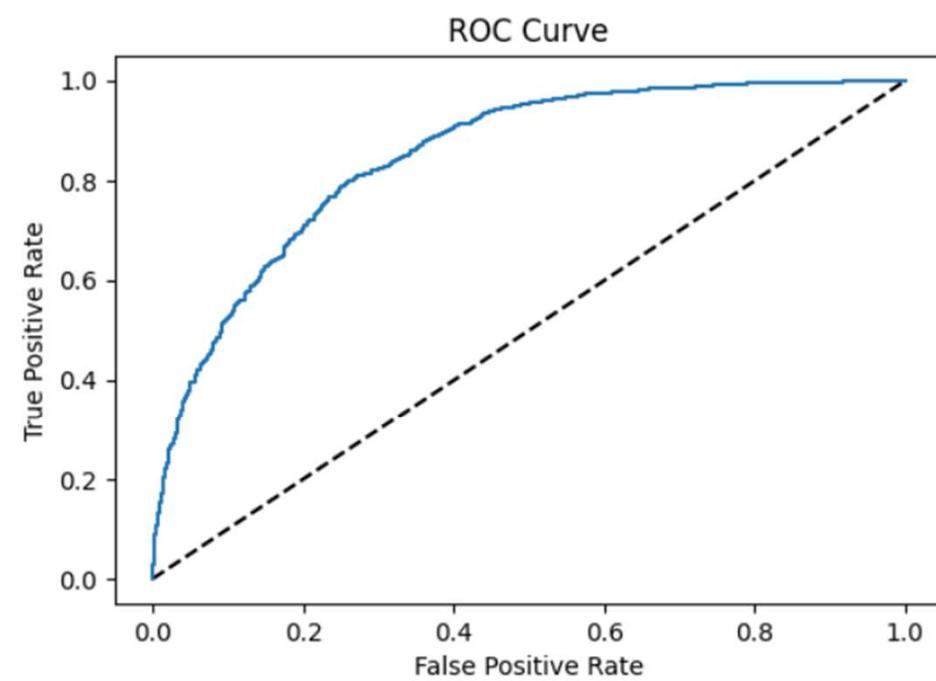


ROC-Curve.png



ROC-Curve.png

...



## diabetes-train-mlflow ✎ ⚡ ✓ Completed

Overview Metrics Images Child jobs Outputs + logs Code Explanations (preview) Fairness (preview) Monitoring

⟳ Refresh

⌚ Debug and monitor

✎ Edit and submit

▷ Perform sweep

✚ Register model

✖ Cancel

trash Delete

⚖ Compare (preview) ▾

### Properties

Status  
✓ Completed

Created on  
Oct 29, 2023 1:10 AM

Start time  
Oct 29, 2023 1:13 AM

Duration  
1m 27.89s

Compute duration  
1m 27.89s

Name  
calm\_ice\_4dxrj7xbtj

Command  
python train-model-mlflow.py  
--training\_data diabetes.csv

Created by  
Harsh Yadav

Job type  
Command

Experiment  
[diabetes-training](#)

Environment  
[AzureML-sklearn-0.24-ubuntu18.04-py37-cpu:49](#)

Registered models  
None

[See all properties](#)

Raw JSON

[See YAML job definition](#)

Job YAML

### Tags

model\_type : LogisticRegression

### Params

Regularization rate : 0.01

### Metrics

AUC	0.8483203	Accuracy	0.774
-----	-----------	----------	-------

### Description

ⓘ Click edit icon to add a description

## diabetes-train-autolog ✎ ⭐ ✓ Completed

Overview Metrics Images Child jobs Outputs + logs Code Explanations (preview) Fairness (preview) Monitoring

### Properties

Status  
Completed

Created on  
Oct 29, 2023 1:10 AM

Start time  
Oct 29, 2023 1:13 AM

Duration  
1m 47.31s

Compute duration  
1m 47.31s

Name  
jovial\_leaf\_bkh3r0801q

Command  
python train-model-autolog.py  
--training\_data diabetes.csv

Created by  
Harsh Yadav

Job type  
Command

Experiment  
diabetes-training

Environment  
AzureML-sklearn-0.24-ubuntu18.04-py37-cpu:49

Registered models  
None

See all properties

Raw JSON

See YAML job definition

Job YAML

### Outputs

Output name: mlflow\_log\_model\_156801412  
Model: [azureml\\_jovial\\_leaf\\_bkh3r0801q\\_output\\_mlflow\\_log\\_model\\_156801412:1](#)  
Asset URI: [azureml:azureml\\_jovial\\_leaf\\_bkh3r0801q\\_output\\_mlflow\\_log\\_model\\_156801412:1](#)

### Tags

estimator\_class : sklearn.linear\_model.logistic.LogisticRegression  
estimator\_name : LogisticRegression

### Params

C : 100.0 class\_weight : None dual : False fit\_intercept : True intercept\_scaling : 1  
l1\_ratio : None max\_iter : 100 multi\_class : auto n\_jobs : None penalty : l2  
random\_state : None regularization\_rate : 0.01 solver : liblinear tol : 0.0001  
verbose : 0 warm\_start : False

## diabetes-train-autolog

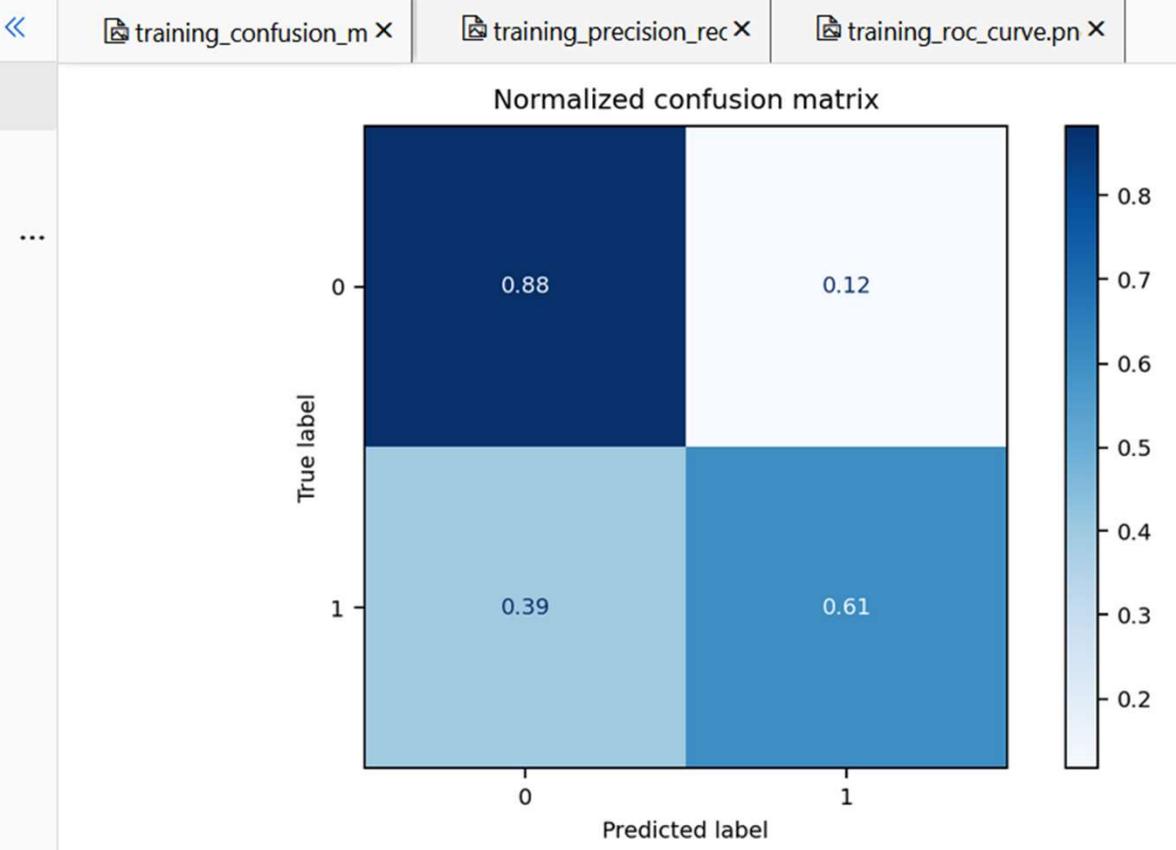
Completed

Overview Metrics Images Child jobs Outputs + logs Code Explanations (preview) Fairness (preview) Monitoring

⟳ Refresh ⚙️ Debug and monitor ✎ Edit and submit ▶ Perform sweep + Register model ✖ Cancel 🗑 Delete

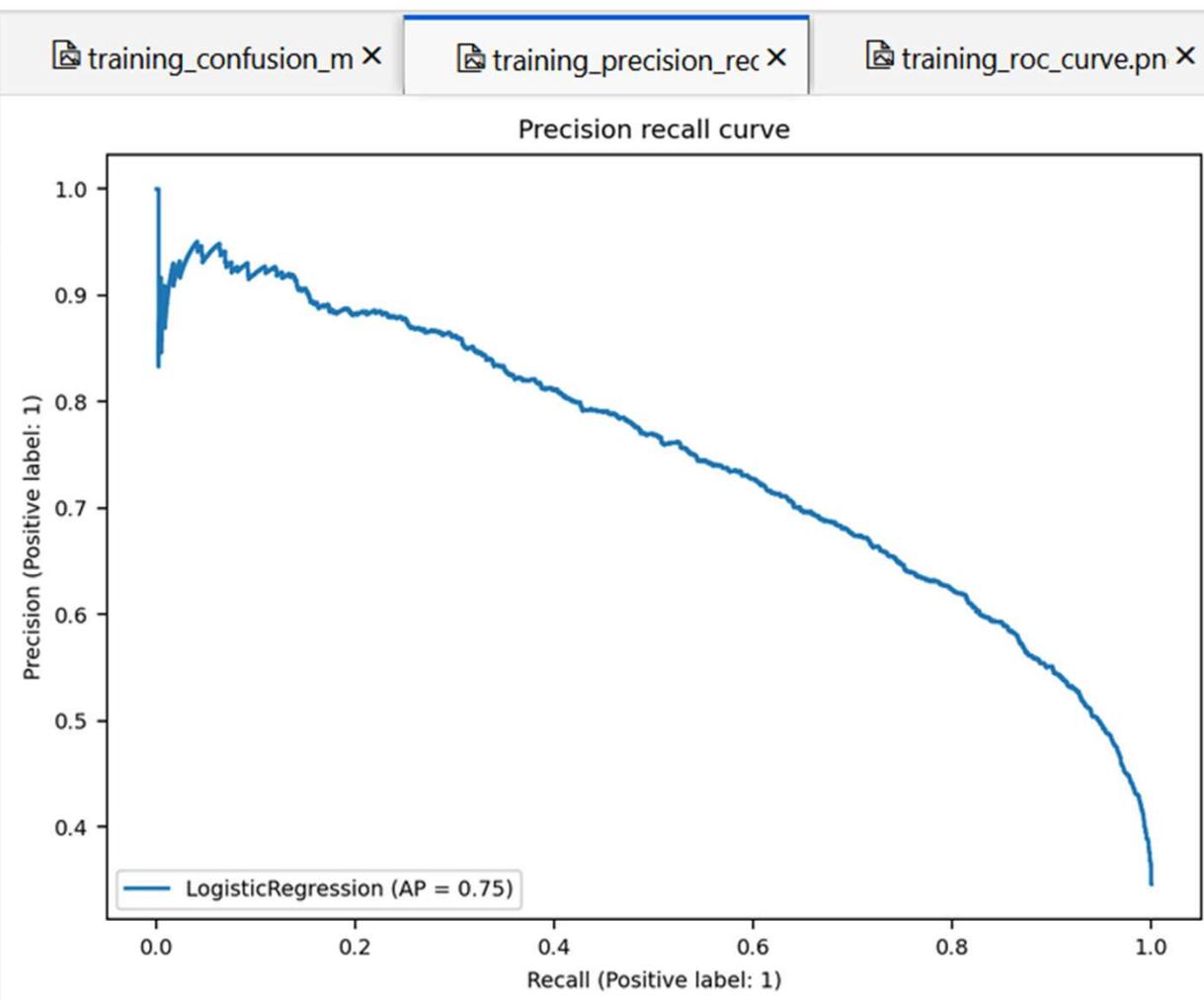


- training\_confusion\_matrix.png
- training\_precision\_recall\_curve.png
- training\_roc\_curve.png
- ...



[Refresh](#)[Debug and monitor](#)[Edit and submit](#)[Perform sweep](#)[Register model](#)[Cancel](#)[Delete](#)

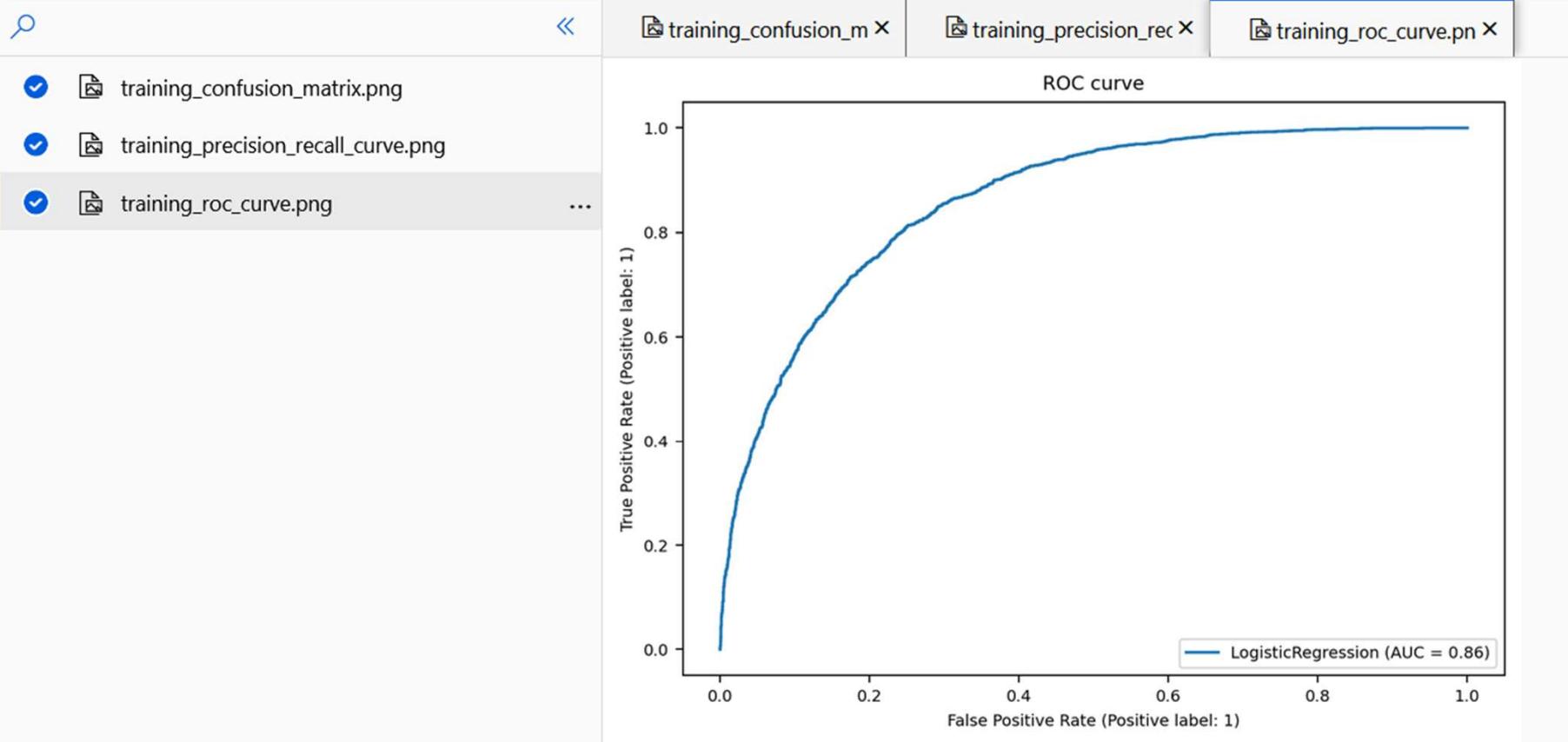
	<a href="#">training_confusion_m</a>	<a href="#">training_precision_rec</a>	<a href="#">training_roc_curve.pn</a>
<input checked="" type="checkbox"/>			
<input checked="" type="checkbox"/>			
<input checked="" type="checkbox"/>			...

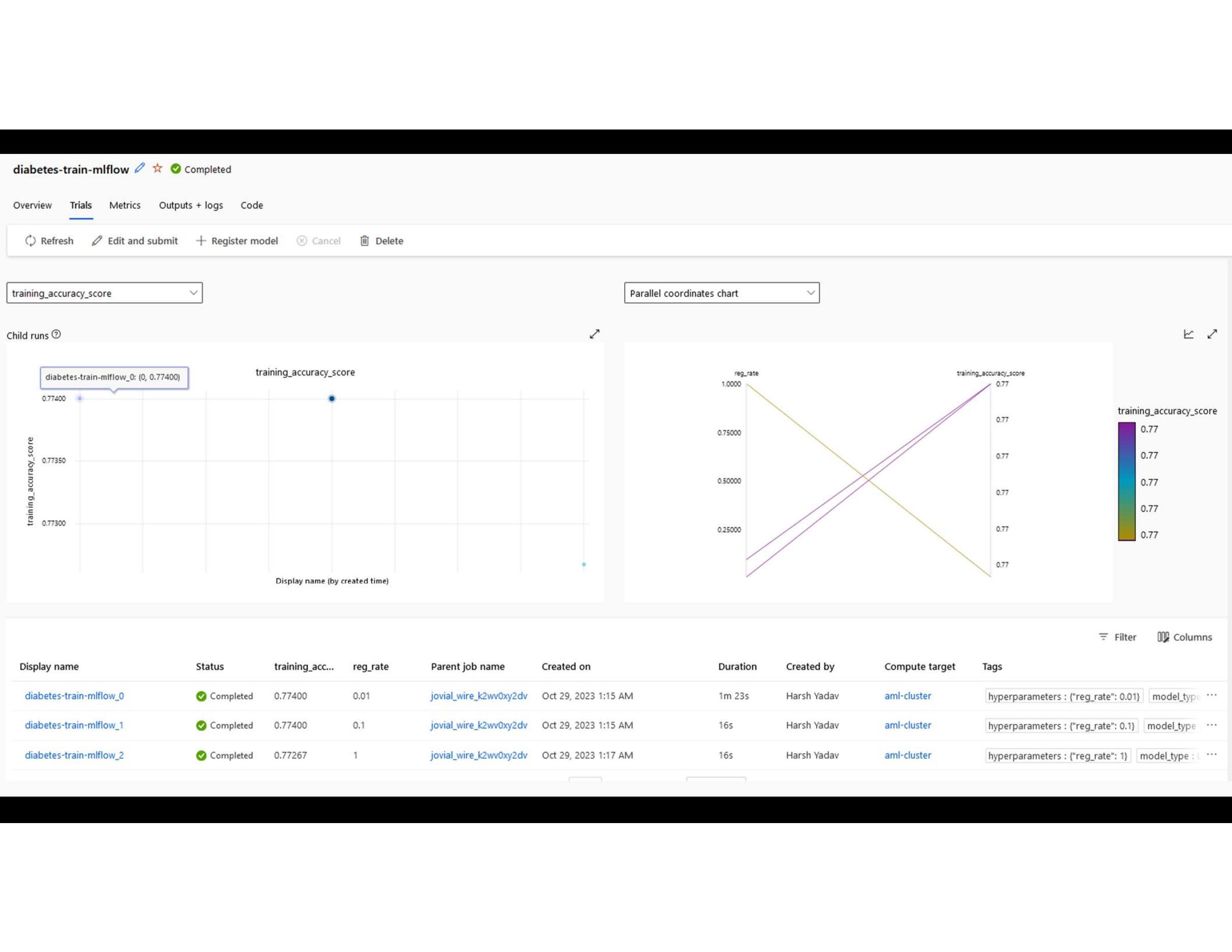


## diabetes-train-autolog ✎ ⭐️ Completed

Overview Metrics **Images** Child jobs Outputs + logs Code Explanations (preview) Fairness (preview) Monitoring

⟳ Refresh ⚡ Debug and monitor 🖊 Edit and submit ➤ Perform sweep + Register model ⏺ Cancel 🗑 Delete





## diabetes-train-mlflow Completed

Overview Trials Metrics Outputs + logs Code

### Properties

Status	Created by
 Completed	Harsh Yadav
Created on	Job type
Oct 29, 2023 1:15 AM	Sweep
Start time	Experiment
Oct 29, 2023 1:15 AM	<a href="#">sweep-diabetes</a>
Duration	Registered models
4m 1.261s	None
Compute duration	See all properties
4m 1.261s	 <a href="#">Raw JSON</a>
Compute target	See YAML job definition
<a href="#">aml-cluster</a>	 <a href="#">Job YAML</a>
Name	
joval_wire_k2wv0xy2dv	

### Tags

model\_type : LogisticRegression 

### Parameter sampling

Sampling policy name  
GRID  
Parameter space  
`{"reg_rate":["choice",[[0.01,0.1,1]]]}`

### Early termination policy

Early termination policy  
Default  
Properties  
`{}`

### Primary metric

Primary metric name  
training\_accuracy\_score  
Primary metric goal  
maximize  
Best trial  
[joval\\_wire\\_k2wv0xy2dv\\_0](#)

Outline



Refresh



Clone



Resubmit



View profiling



Publish



Schedule

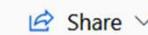


...

diabetes\_classification



Completed



Share



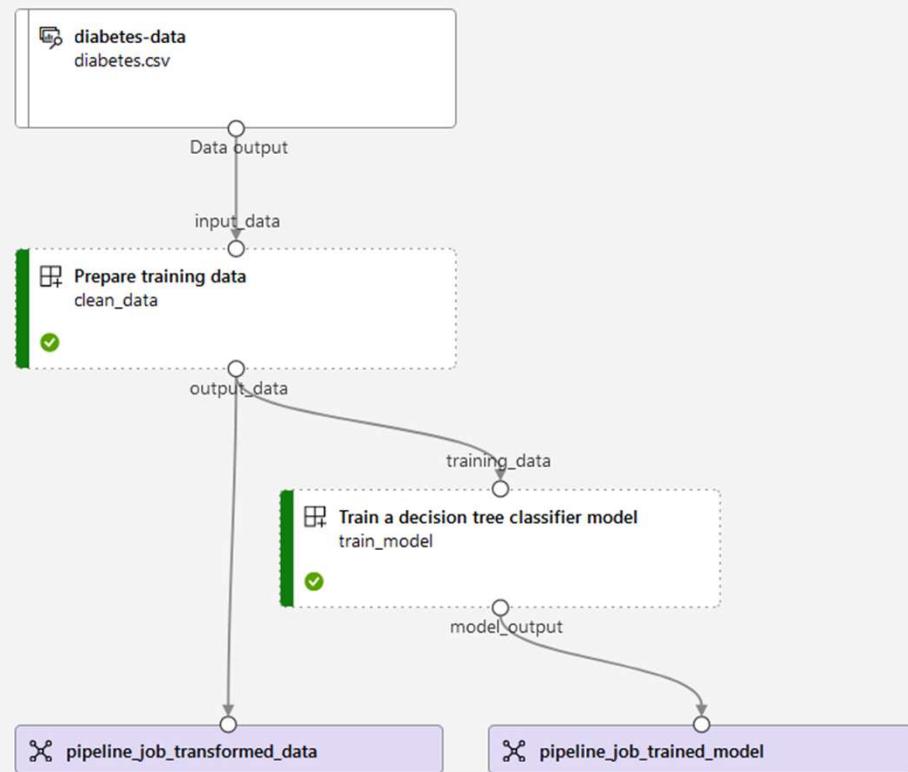
+ Add filter

### diabetes\_classification

diabetes-data

clean\_data

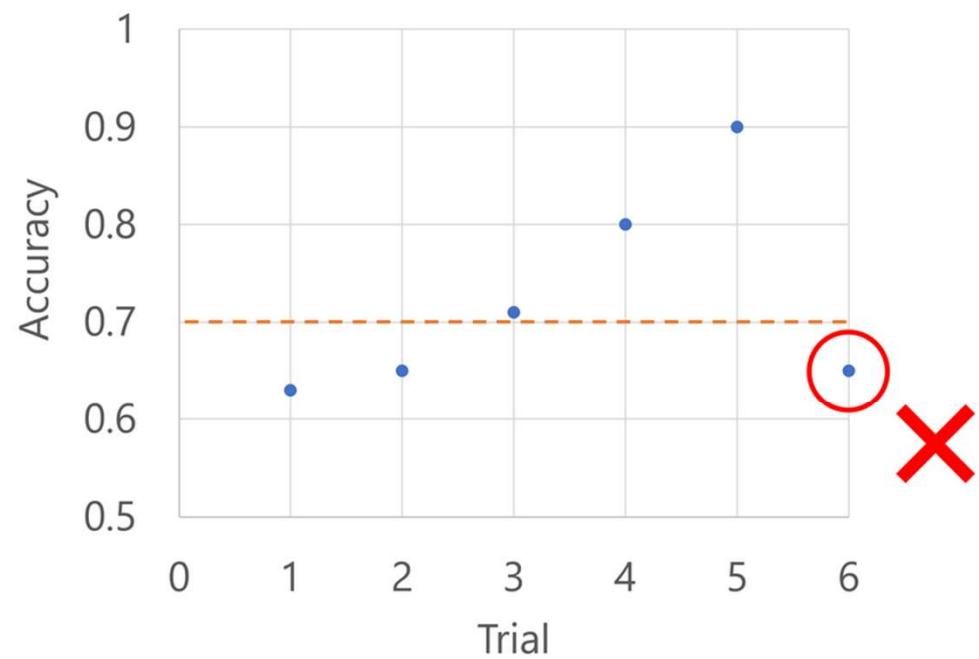
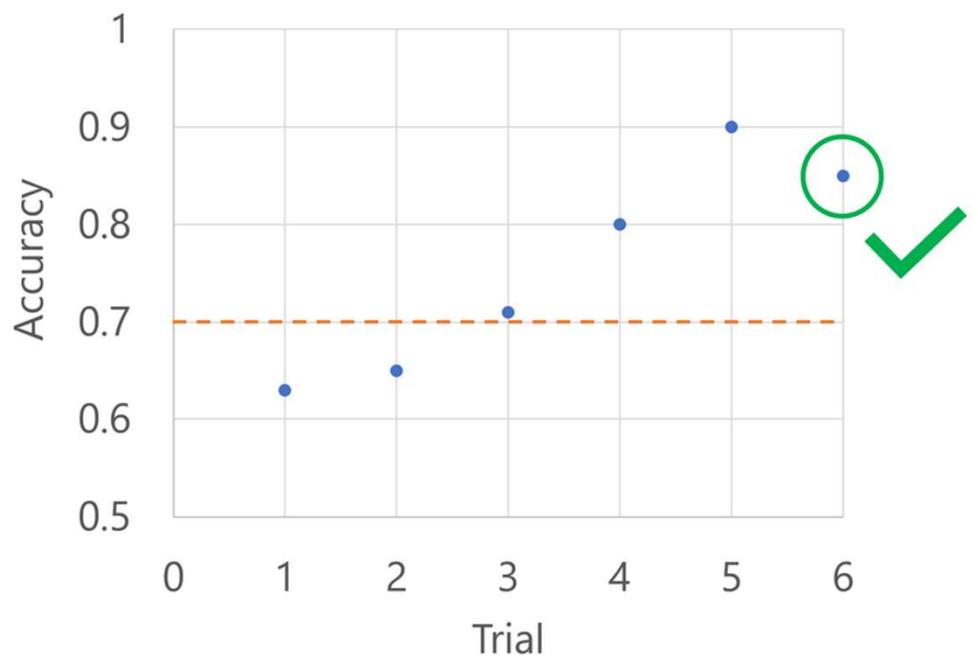
train\_model

 Navigator

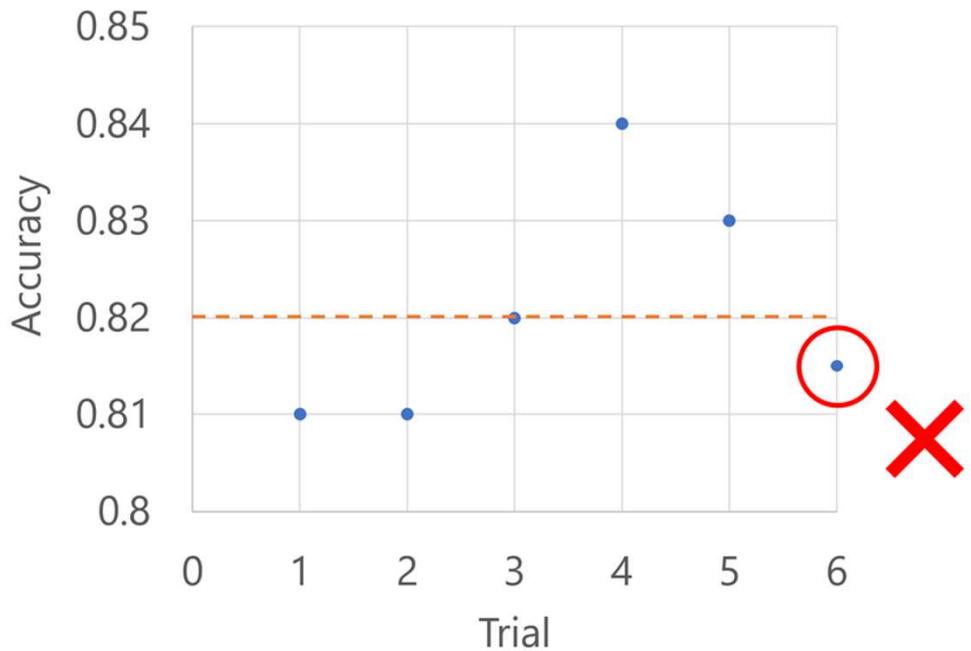
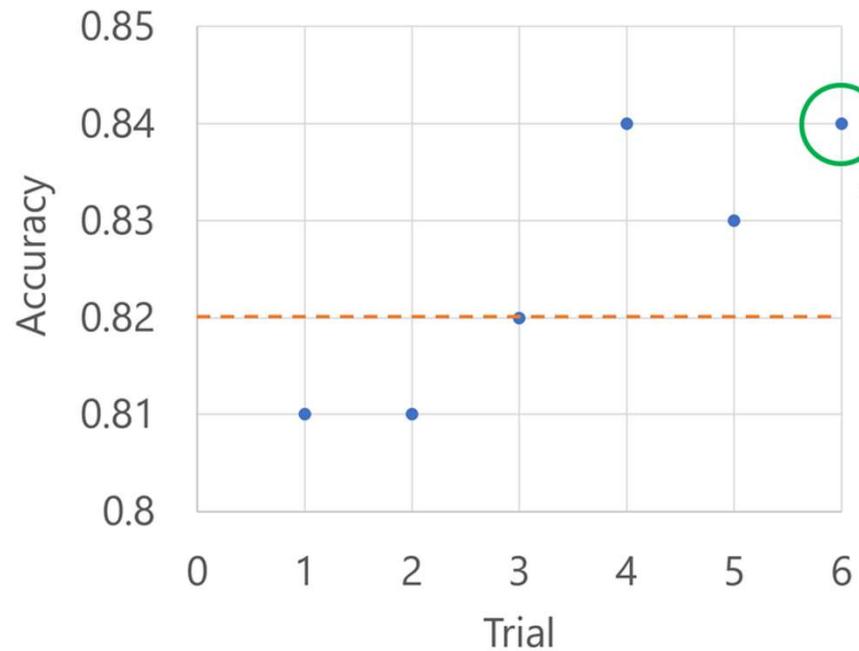
100% ▾



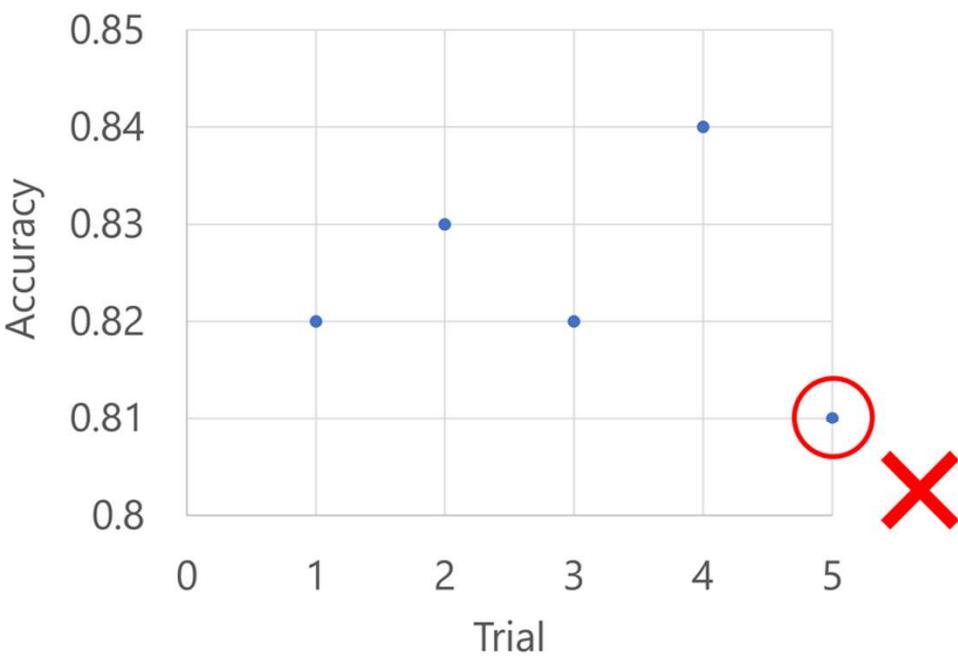
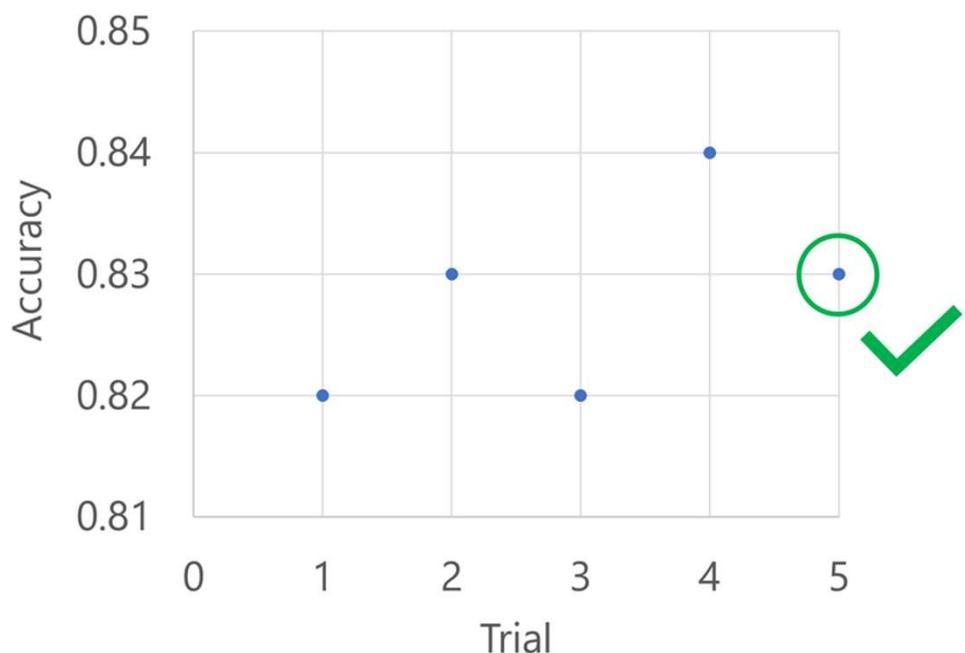
Imagine the primary metric is the accuracy of the model. When after the first five trials, the best performing model has an accuracy of 0.9, any new model needs to perform better than  $(0.9 - 0.2)$  or 0.7. If the new model's accuracy is higher than 0.7, the sweep job will continue. If the new model has an accuracy score lower than 0.7, the policy will terminate the sweep job.



Imagine the primary metric is the accuracy of the model. When the accuracy is logged for the sixth trial, the metric needs to be higher than the median of the accuracy scores so far. Suppose the median of the accuracy scores so far is 0.82. If the new model's accuracy is higher than 0.82, the sweep job will continue. If the new model has an accuracy score lower than 0.82, the policy will stop the sweep job, and no new models will be trained.



Imagine the primary metric is the accuracy of the model. When the accuracy is logged for the fifth trial, the metric should **not** be in the worst 20% of the trials so far. In this case, 20% translates to one trial. In other words, if the fifth trial is **not** the worst performing model so far, the sweep job will continue. If the fifth trial has the lowest accuracy score of all trials so far, the sweep job will stop.



# Blue/green deployment

One endpoint can have multiple deployments. One approach is the **blue/green deployment**.

Let's take the example of the restaurant recommender model. After experimentation, you select the best performing model. You use the blue deployment for this first version of the model. As new data is collected, the model can be retrained, and a new version is registered in the Azure Machine Learning workspace. To test the new model, you can use the green deployment for the second version of the model.

Both versions of the model are deployed to the same endpoint, which is integrated with the application. Within the application, a user selects a restaurant, sending a request to the endpoint to get new real-time recommendations of other restaurants the user may like.

When a request is sent to the endpoint, 90% of the traffic can go to the blue deployment\*, and 10% of the traffic can go to the *green deployment*. With two versions of the model deployed on the same endpoint, you can easily test the model.

After testing, you can also seamlessly transition to the new version of the model by redirecting 90% of the traffic to the green deployment. If it turns out that the new version doesn't perform better, you can easily roll back to the first version of the model by redirecting most of the traffic back to the blue deployment.

Blue/green deployment allows for multiple models to be deployed to an endpoint. You can decide how much traffic to forward to each deployed model. This way, you can switch to a new version of the model without interrupting service to the consumer.

---

The screenshot shows the Azure Machine Learning Studio interface. At the top, there's a toolbar with icons for file operations, a 'Deploy to online end...' button, and a 'Compute' dropdown set to 'ci9ea02d2799ea4b02bd - Running'. Below the toolbar, the status bar displays 'ci9ea02d2799ea4b02bd · Kernel idle CPU 0% RAM 10%' on the left, 'Last saved a few seconds ago' in the center, and 'Python 3.8 - AzureML' on the right. A yellow banner at the top of the main area reads: 'IMPORTANT! Wait until the blue deployment is configured before continuing! A green notification should appear in the studio.' The main content area is titled 'Test the deployment' and contains a code snippet for testing a deployed model:

```
1 # test the blue deployment with some sample data
2 response = ml_client.online_endpoints.invoke(
3     endpoint_name=online_endpoint_name,
4     deployment_name="blue",
5     request_file="sample-data.json",
6 )
7
8 if response[1]=='1':
9     print("Diabetic")
10 else:
11     print ("Not diabetic")
```

The code output shows a green checkmark and the text 'Diabetic'.

# endpoint-10282034947851

Details Test Consume Monitoring Logs

## Endpoint attributes

Service ID  
endpoint-10282034947851

Description  
Online endpoint for MLflow diabetes model

Provisioning state  
Succeeded

Error details

--

Compute type  
Managed

Created by  
Harsh Yadav

Created on  
Oct 29, 2023 2:04 AM

Last updated on  
Oct 29, 2023 2:04 AM

Authentication type  
Key

Public network access  
Enabled ⓘ

Swagger URI  
<https://endpoint-10282034947851.northeurope.inference.ml.azure.com/swagger.json>

REST endpoint  
[https://endpoint-10282034947851.northeurope.inference.ml.azure.com\(score](https://endpoint-10282034947851.northeurope.inference.ml.azure.com(score)

Metrics

## Deployment summary

Live traffic allocation

✓ blue (0%)

Mirrored traffic allocation

--

## Deployment blue



Name  
blue

Live traffic  
0%

Scoring script  
Auto-generated

Provisioning state  
✓ Succeeded

Error details

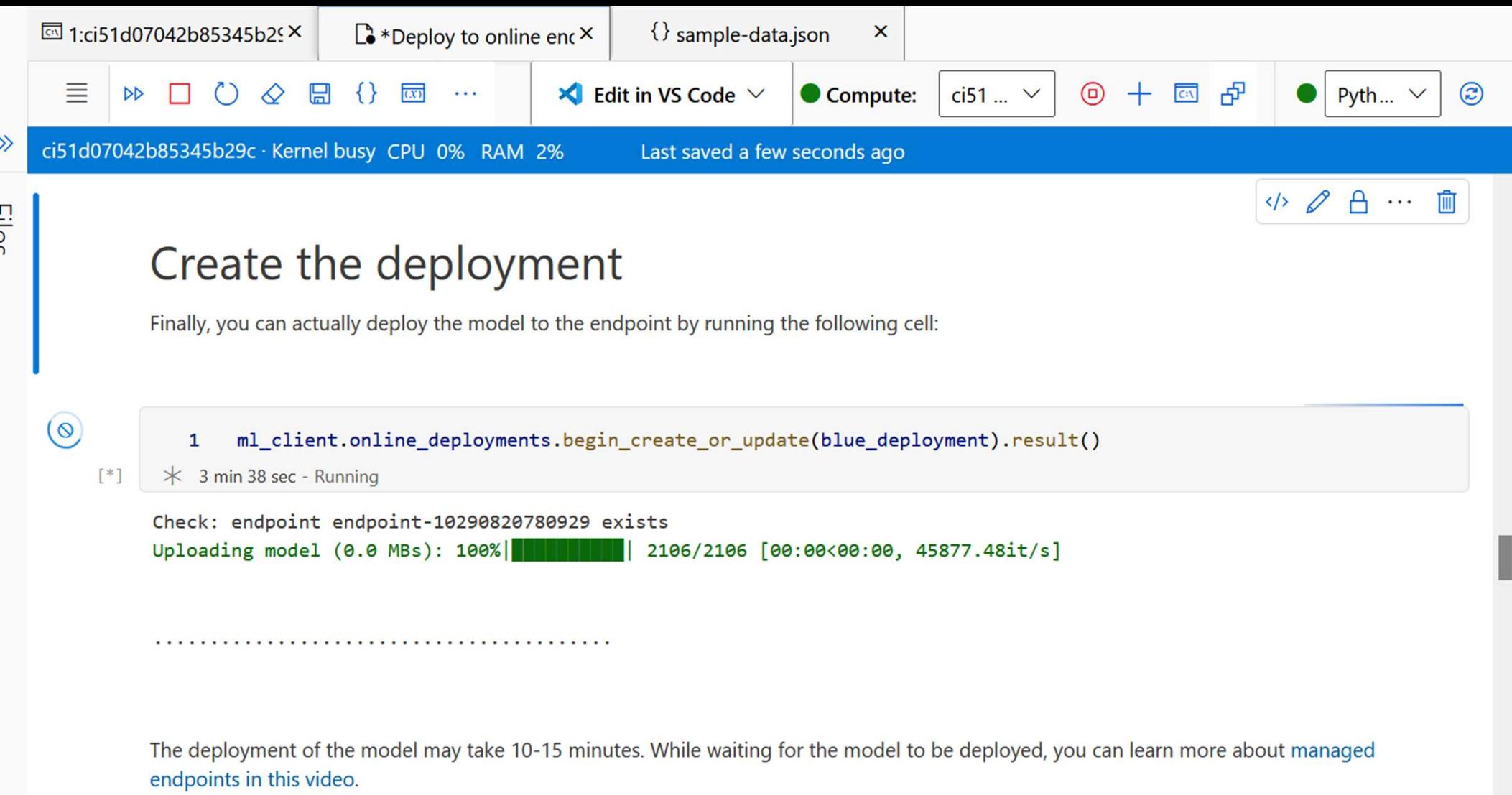
--

SKU  
Standard\_F4s\_v2

Quota type  
Dedicated

Egress public network access  
Enabled ⓘ

Instance count  
1



The screenshot shows a Jupyter Notebook interface with the following details:

- Top Bar:** Contains tabs for "1:ci51d07042b85345b29c", "\*Deploy to online enc...", and "{} sample-data.json".
- Toolbar:** Includes icons for file operations (New, Open, Save, etc.), "Edit in VS Code", "Compute" (set to "ci51..."), and Python environment selection.
- Header:** Shows the kernel name "ci51d07042b85345b29c · Kernel busy CPU 0% RAM 2%", the last save time "Last saved a few seconds ago", and a toolbar with code editor, lock, and delete icons.
- Section Header:** A large h2 "Create the deployment" is centered on the page.
- Text:** A descriptive paragraph: "Finally, you can actually deploy the model to the endpoint by running the following cell:"
- Code Cell:** Displays the Python code `ml\_client.online\_deployments.begin\_create\_or\_update(blue\_deployment).result()`. The cell status is marked with a blue asterisk and "3 min 38 sec - Running". Below the cell, the output shows: "Check: endpoint endpoint-10290820780929 exists" and "Uploading model (0.0 MBs): 100% [██████████] 2106/2106 [00:00<00:00, 45877.48it/s]".
- Footnote:** A note at the bottom states: "The deployment of the model may take 10-15 minutes. While waiting for the model to be deployed, you can learn more about [managed endpoints in this video](#)".