TASK 3 - P

Design and deploy model using

Azure machine learning

Survival prediction on
Titanic dataset

Saurabh Dharmadhikari
Saurabh.dharma01@gmail.com

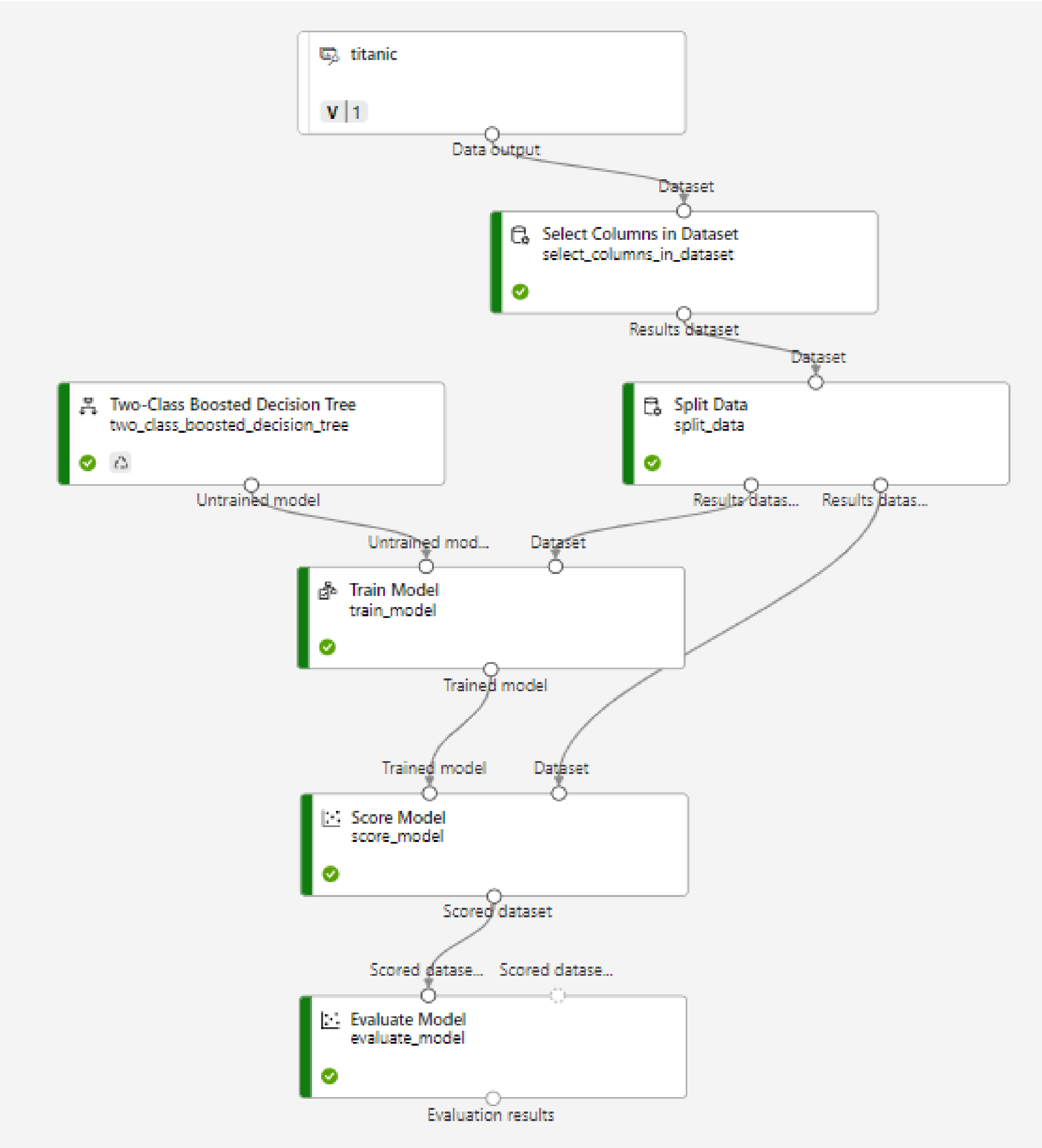
Table of Content:

1. **For this task you need to design and deploy a machine learning model using Azure ML studio (designer). You need to use Microsoft Azure Machine Learning Studio to design and deploy your model. To complete this task, you need to select a dataset and design a decision tree model (classification tree or regression tree) and then deploy the built model and get the API key. To do this task you need to follow the workshop recording and slides and deploy your own model on Azure. You need to provide the screenshots of your designed model, training model, the performance of the built model (e.g., Accuracy, confusion matrix and etc) and deployed model with the API key and test the model. The screenshot of the model should include your Azure account name since the API key is unique to you.**

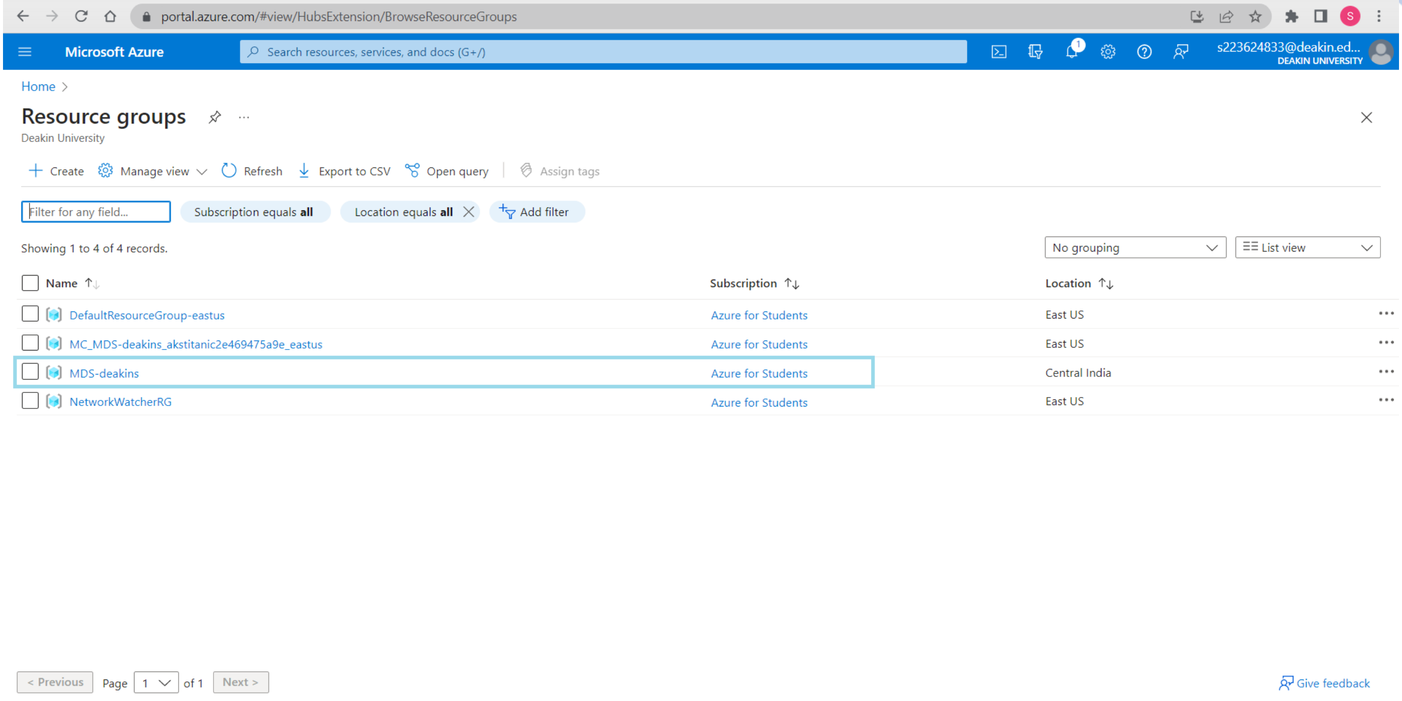
Introduction:

We have used titanic dataset to build a decision tree in Microsoft Azure Machine Learning Studio.

The below is the image showing the pipeline of building this model.



First of all we create a resource group called MDS-deakins with location as Central India.



portal.azure.com/#view/HubsExtension/BrowseResourceGroups

Microsoft Azure

Search resources, services, and docs (G+/)

s223624833@deakin.ed...
DEAKIN UNIVERSITY

Home >

Resource groups

Deakin University

+ Create · Manage view · Refresh · Export to CSV · Open query · Assign tags

Filter for any field... · Subscription equals **all** · Location equals **all** · Add filter

Showing 1 to 4 of 4 records.

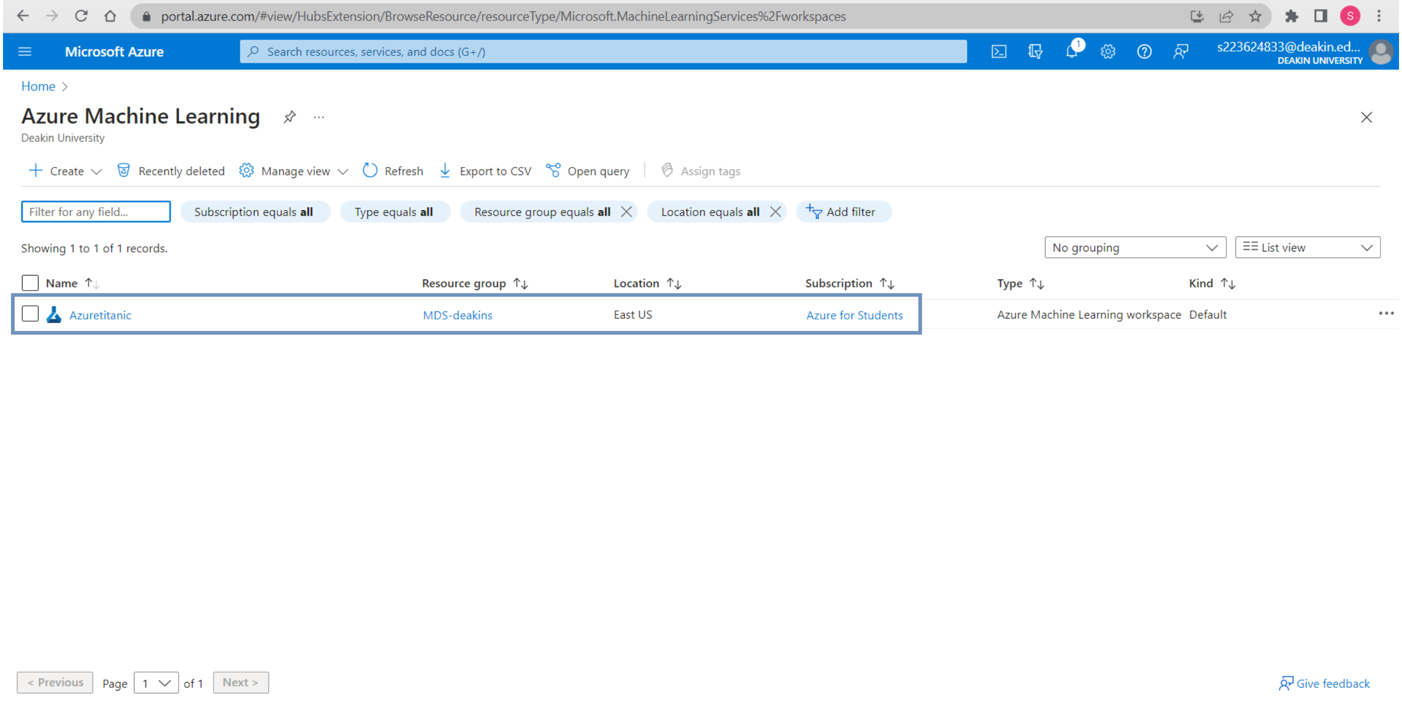
No grouping · List view

Name ↑↓	Subscription ↑↓	Location ↑↓
DefaultResourceGroup-eastus	Azure for Students	East US
MC_MDS-deakins_akstitanic2e469475a9e_eastus	Azure for Students	East US
MDS-deakins	Azure for Students	Central India
NetworkWatcherRG	Azure for Students	East US

< Previous · Page 1 of 1 · Next >

Give feedback

After creating the resource group, we create resource group we create a new Azure machine learning under the resource group we created.



portal.azure.com/#view/HubsExtension/BrowseResource/resourceType/Microsoft.MachineLearningServices%2Fworkspaces

Microsoft Azure

Search resources, services, and docs (G+/)

s223624833@deakin.ed...
DEAKIN UNIVERSITY

Home >

Azure Machine Learning

Deakin University

+ Create · Recently deleted · Manage view · Refresh · Export to CSV · Open query · Assign tags

Filter for any field... · Subscription equals **all** · Type equals **all** · Resource group equals **all** · Location equals **all** · Add filter

Showing 1 to 1 of 1 records.

No grouping · List view

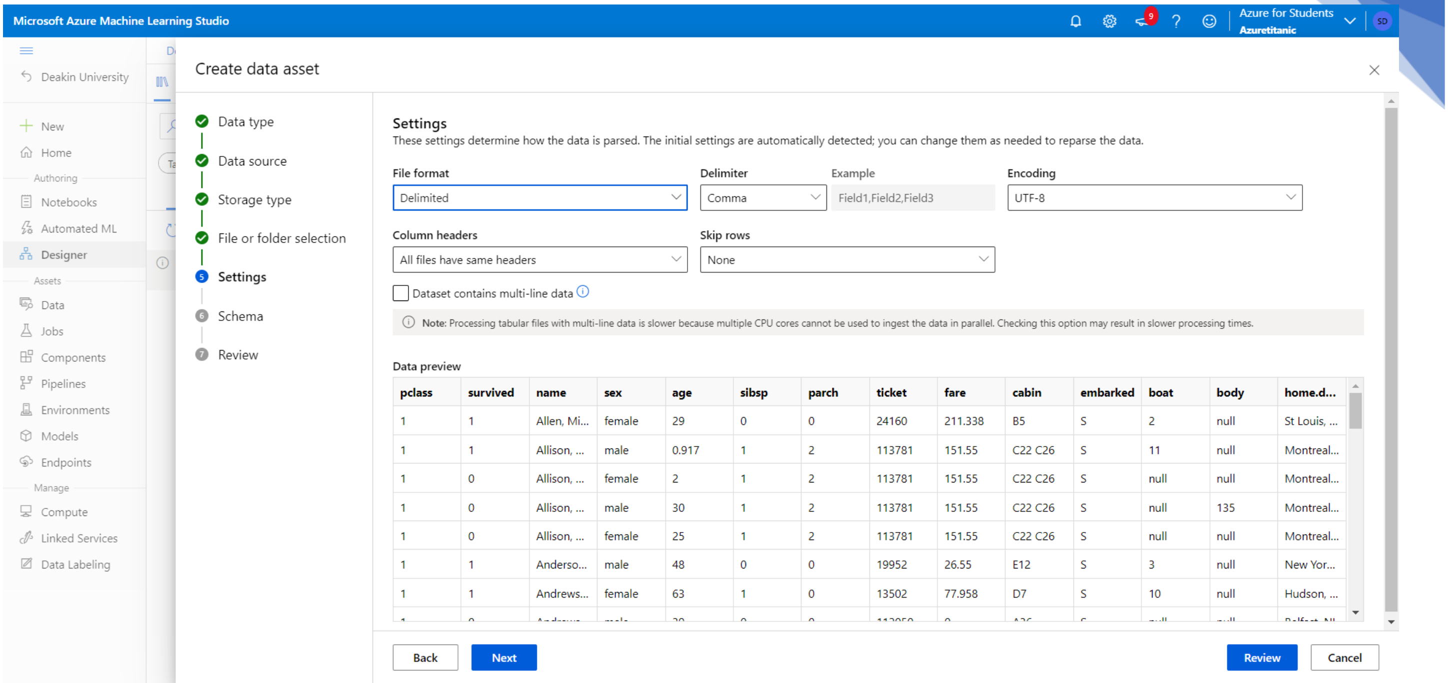
Name ↑↓	Resource group ↑↓	Location ↑↓	Subscription ↑↓	Type ↑↓	Kind ↑↓
Azuretitanic	MDS-deakins	East US	Azure for Students	Azure Machine Learning workspace	Default

< Previous · Page 1 of 1 · Next >

Give feedback

WE have created a work space in Azure machine learning.

We then launch studio and we create dataset in this environment using data option in designer.



Create data asset

Settings

These settings determine how the data is parsed. The initial settings are automatically detected; you can change them as needed to reparse the data.

File format: Delimited

Delimiter: Comma

Example: Field1,Field2,Field3

Encoding: UTF-8

Column headers: All files have same headers

Skip rows: None

☐ Dataset contains multi-line data

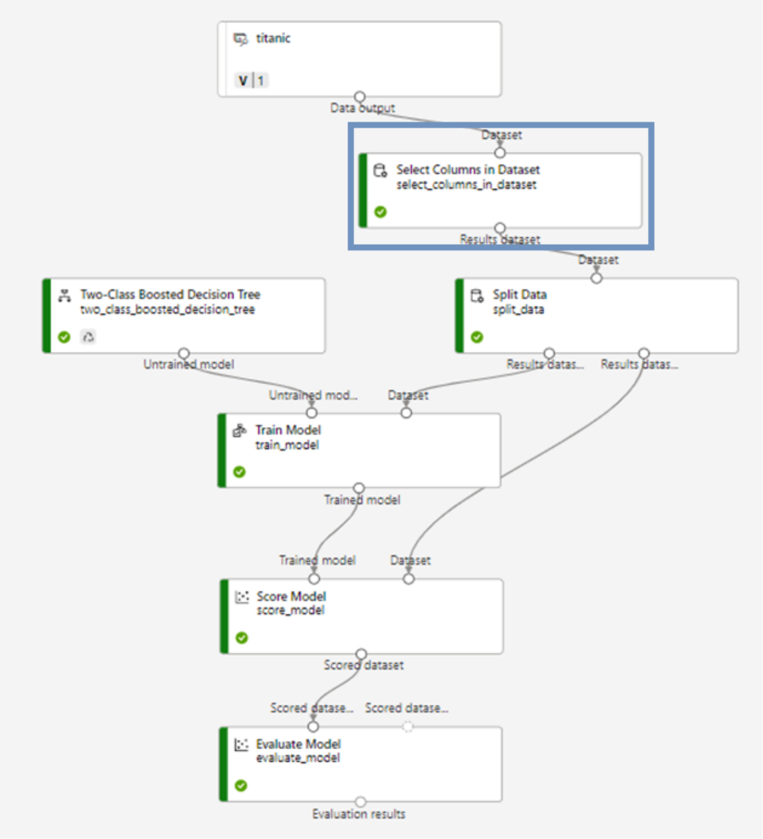
Note: Processing tabular files with multi-line data is slower because multiple CPU cores cannot be used to ingest the data in parallel. Checking this option may result in slower processing times.

Data preview

pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body	home.d...
1	1	Allen, Mi...	female	29	0	0	24160	211.338	B5	S	2	null	St Louis, ...
1	1	Allison, ...	male	0.917	1	2	113781	151.55	C22 C26	S	11	null	Montreal...
1	0	Allison, ...	female	2	1	2	113781	151.55	C22 C26	S	null	null	Montreal...
1	0	Allison, ...	male	30	1	2	113781	151.55	C22 C26	S	null	135	Montreal...
1	0	Allison, ...	female	25	1	2	113781	151.55	C22 C26	S	null	null	Montreal...
1	1	Anderso...	male	48	0	0	19952	26.55	E12	S	3	null	New Yor...
1	1	Andrews...	female	63	1	0	13502	77.958	D7	S	10	null	Hudson, ...

We have taken this data set from our local computer and in the above image we can see a preview of the dataset.

We can now drag and drop the dataset created in our work space to build a pipeline.



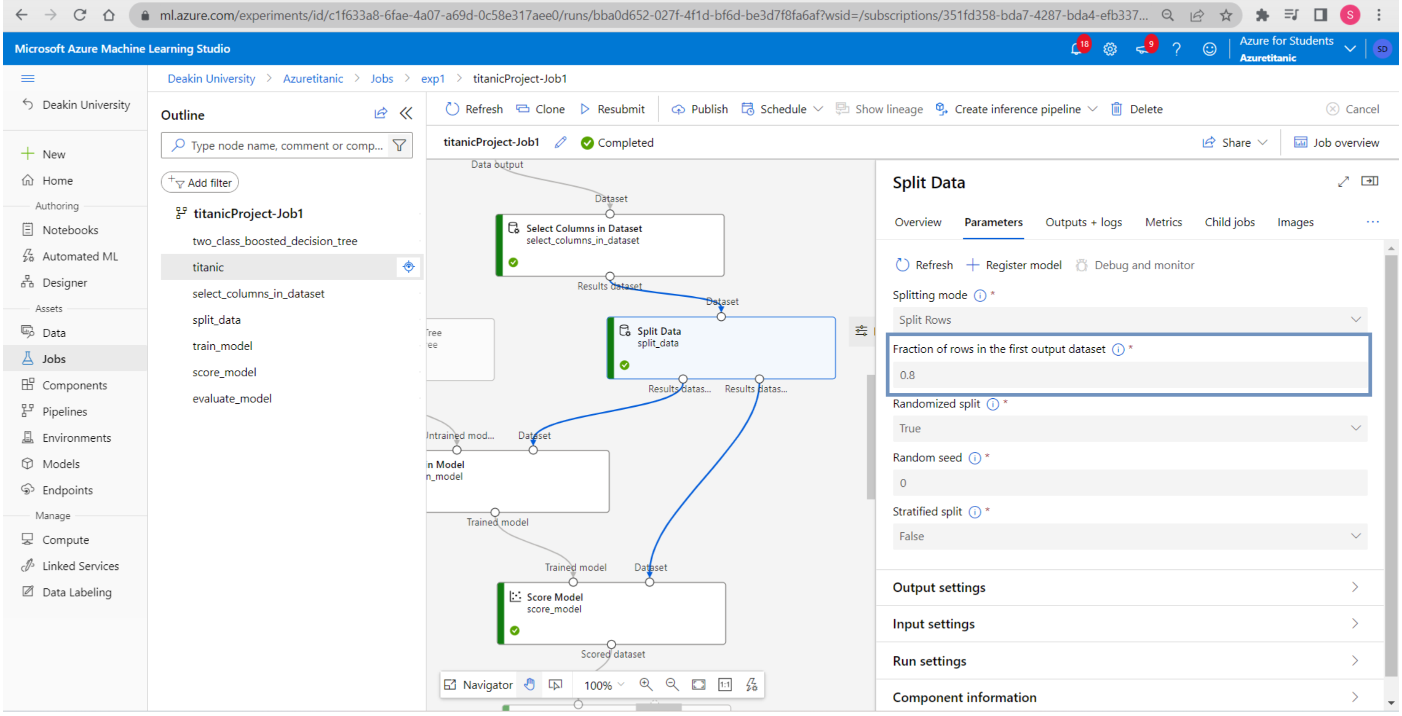
We then drag and drop Select Columns in Dataset from the component option in designer and connect it to dataset.

We select a total of 8 features in the Select Column in Dataset, namely:

1. **pclass**
2. **survived**
3. **sex**
4. **age**
5. **sibsp**
6. **parch**
7. **fare**
8. **embarked**

Including our target variable i.e., survived

We now drag and drop split data option from component and split the data into 80:20 ratio. Where 80 will be our train data and 20 will be our test data.



Split Data

Splitting mode *: Split Rows

Fraction of rows in the first output dataset *: 0.8

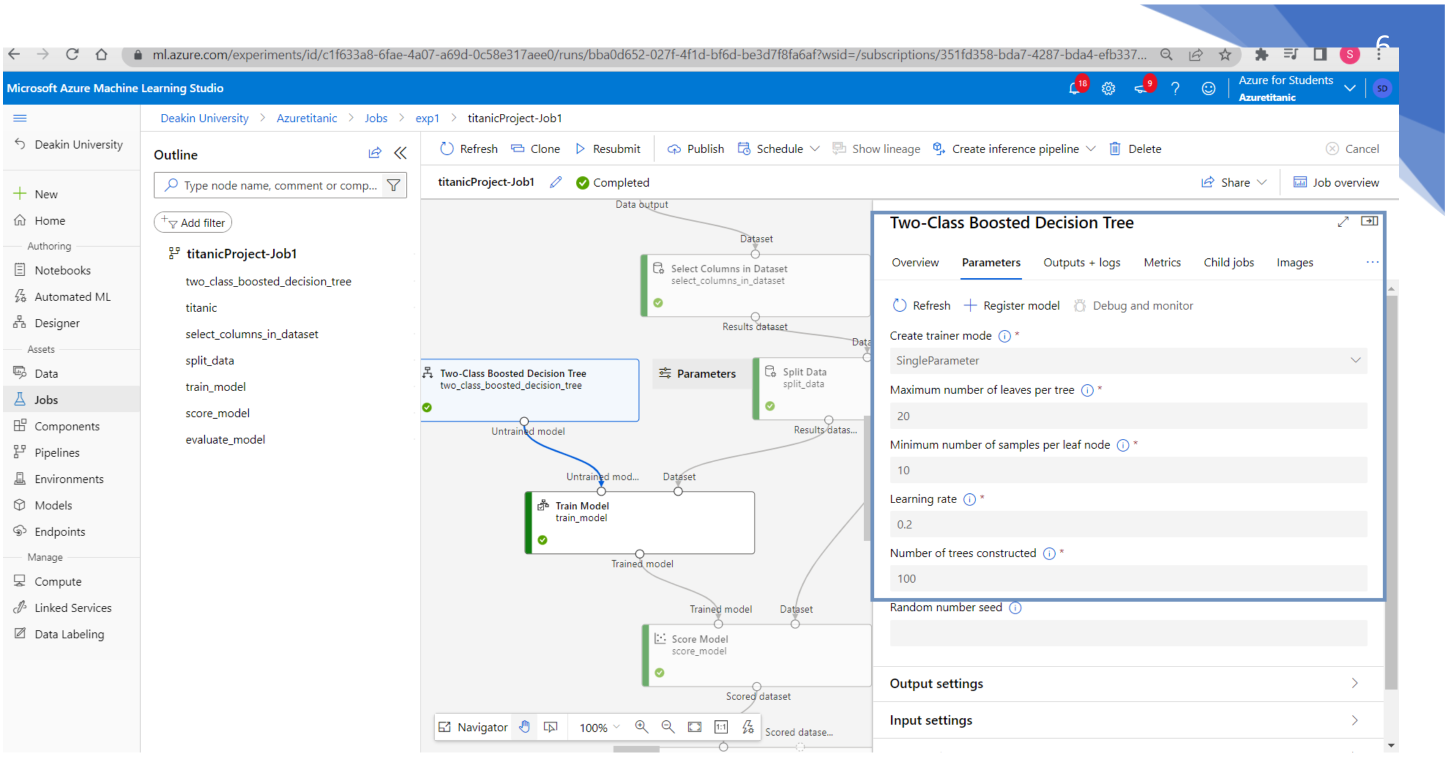
Randomized split *: True

Random seed *: 0

Stratified split *: False

After splitting the data, we drag and drop Two-Class Boosted Decision Tree here under parameters we have specified that:

1.	Max number of leaves pre tree:	20
2.	Min number of samples per leaf node:	10
3.	Learning rate:	0.2
4.	Number of trees constructed:	100



Create trainer mode *
SingleParameter

Maximum number of leaves per tree *
20

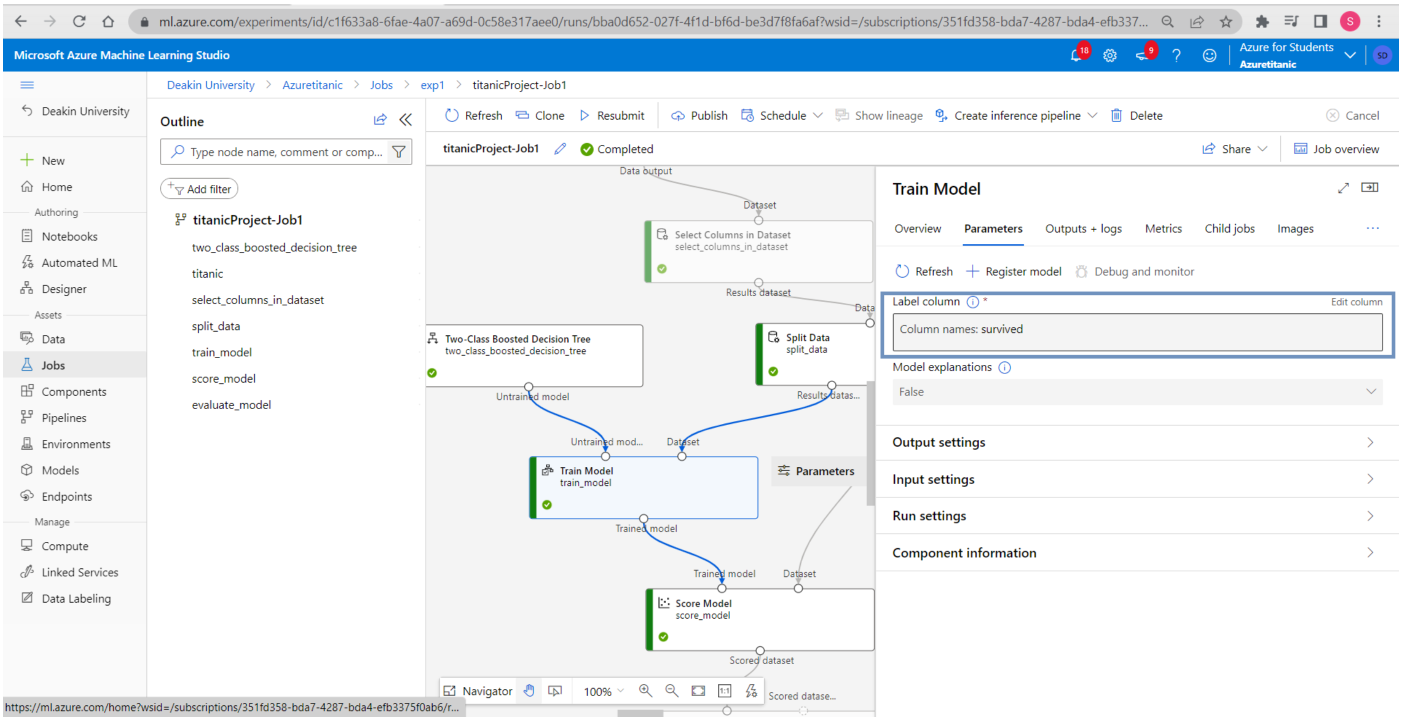
Minimum number of samples per leaf node *
10

Learning rate *
0.2

Number of trees constructed *
100

Random number seed

Hereafter we drag and drop train model and specify the label column as survived. We connect it to Two-Class Boost Decision Tree from where it takes the model for prediction and Split data from where it takes the train set.



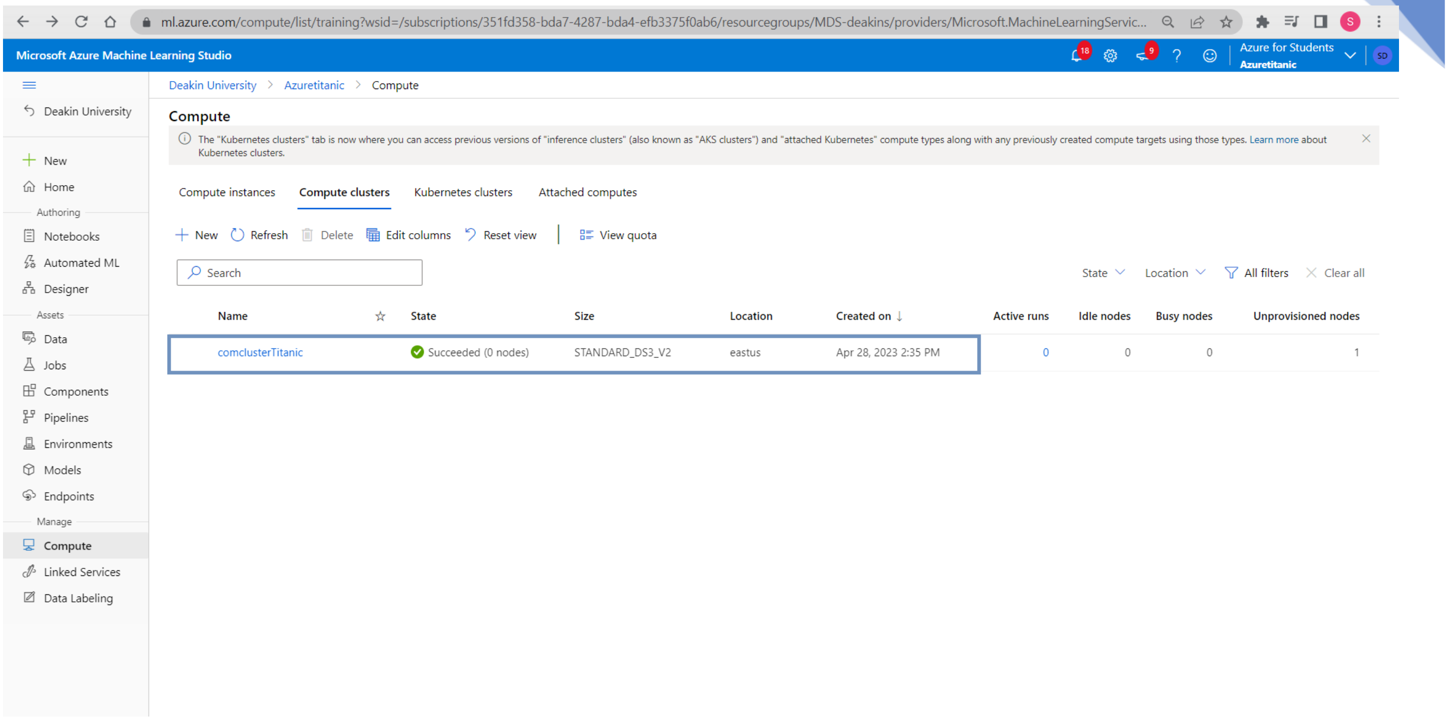
Label column *
Column names: survived

Model explanations
False

Also, we drag and drop score model and connect it with train model and split data to feed it both the trained model and test model. Also, we drag and drop evaluate model and connect it to score model.

Now our pipeline is complete. We save this pipeline.

We now go to the compute option from the side bar and create a compute cluster named comclusterTitanic with location east US and 1 node.



ml.azure.com/compute/list/training?wsid=/subscriptions/351fd358-bda7-4287-bda4-efb3375f0ab6/resourcegroups/MDS-deakins/providers/Microsoft.MachineLearningServic...

Microsoft Azure Machine Learning Studio

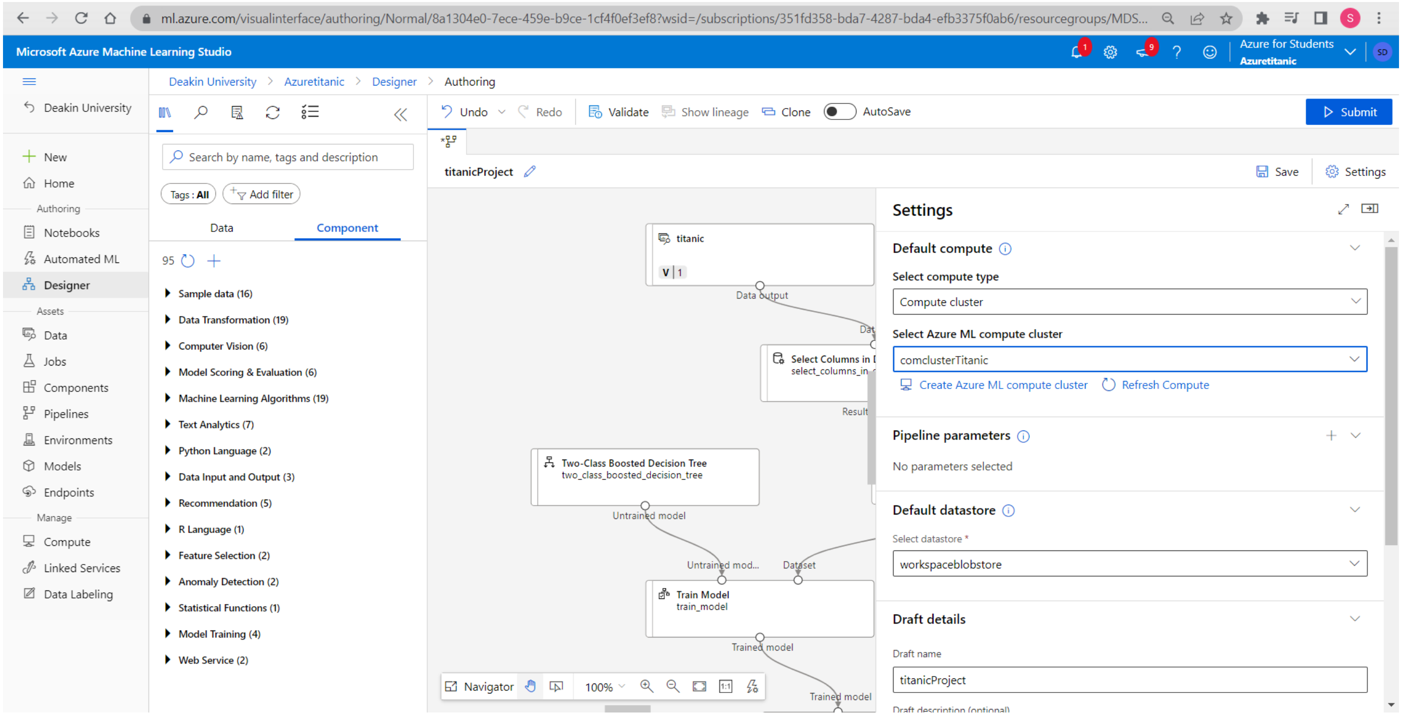
Deakin University > Azuretitanic > Compute

Compute

The "Kubernetes clusters" tab is now where you can access previous versions of "inference clusters" (also known as "AKS clusters") and "attached Kubernetes" compute types along with any previously created compute targets using those types. Learn more about Kubernetes clusters.

Name	State	Size	Location	Created on ↓	Active runs	Idle nodes	Busy nodes	Unprovisioned nodes
comclusterTitanic	Succeeded (0 nodes)	STANDARD_DS3_V2	eastus	Apr 28, 2023 2:35 PM	0	0	0	1

We go back to designer in our pipeline that we saved earlier and click settings on right side of the screen. Here we specify the cluster we created in the previous step and click submit.



ml.azure.com/visualinterface/authoring/Normal/8a1304e0-7ece-459e-b9ce-1cf4f0ef3ef8?wsid=/subscriptions/351fd358-bda7-4287-bda4-efb3375f0ab6/resourcegroups/MDS...

Microsoft Azure Machine Learning Studio

Deakin University > Azuretitanic > Designer > Authoring

titanicProject

Settings

Default compute

Select compute type: Compute cluster

Select Azure ML compute cluster: comclusterTitanic

Pipeline parameters

No parameters selected

Default datastore

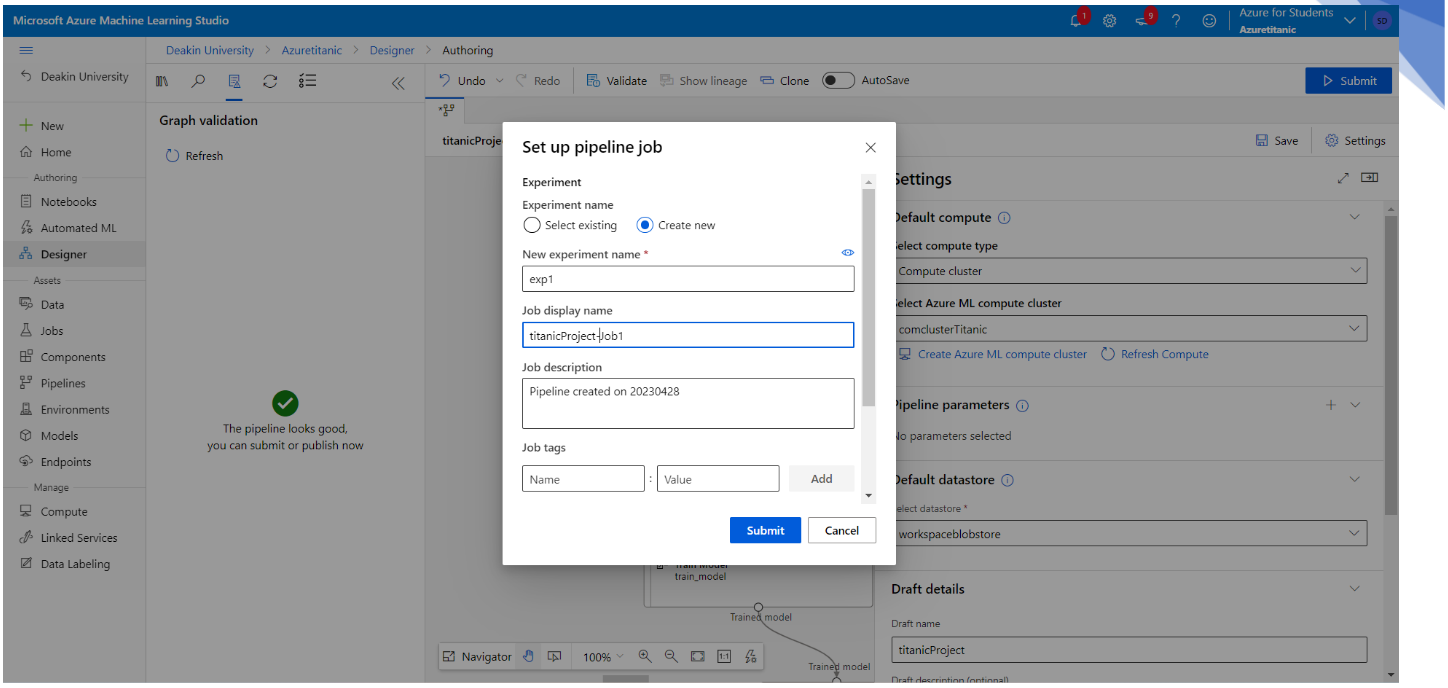
Select datastore *: workspaceblobstore

Draft details

Draft name: titanicProject

Draft description (optional)

We name the experiment as exp1 and also give it a display name and click on submit.



Set up pipeline job

Experiment

Experiment name

○ Select existing ● Create new

New experiment name * exp1

Job display name titanicProject-Job1

Job description Pipeline created on 20230428

Job tags Name : Value Add

Submit Cancel

Settings

Default compute

Select compute type: Compute cluster

Select Azure ML compute cluster: comclusterTitanic

Pipeline parameters

No parameters selected

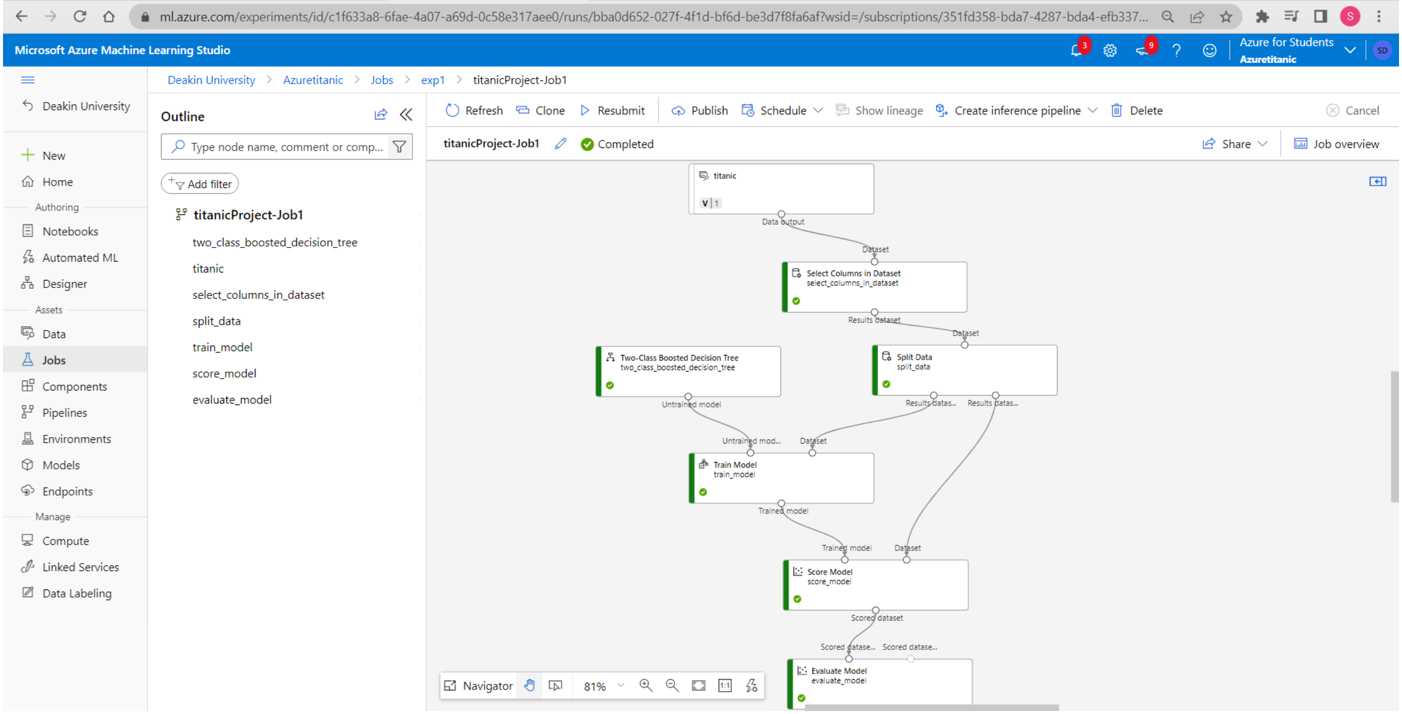
Default datastore

Select datastore *: workspaceblobstore

Draft details

Draft name: titanicProject

exp1 is created in jobs now so we go to jobs and click on exp1, we can see that the model is trained now.



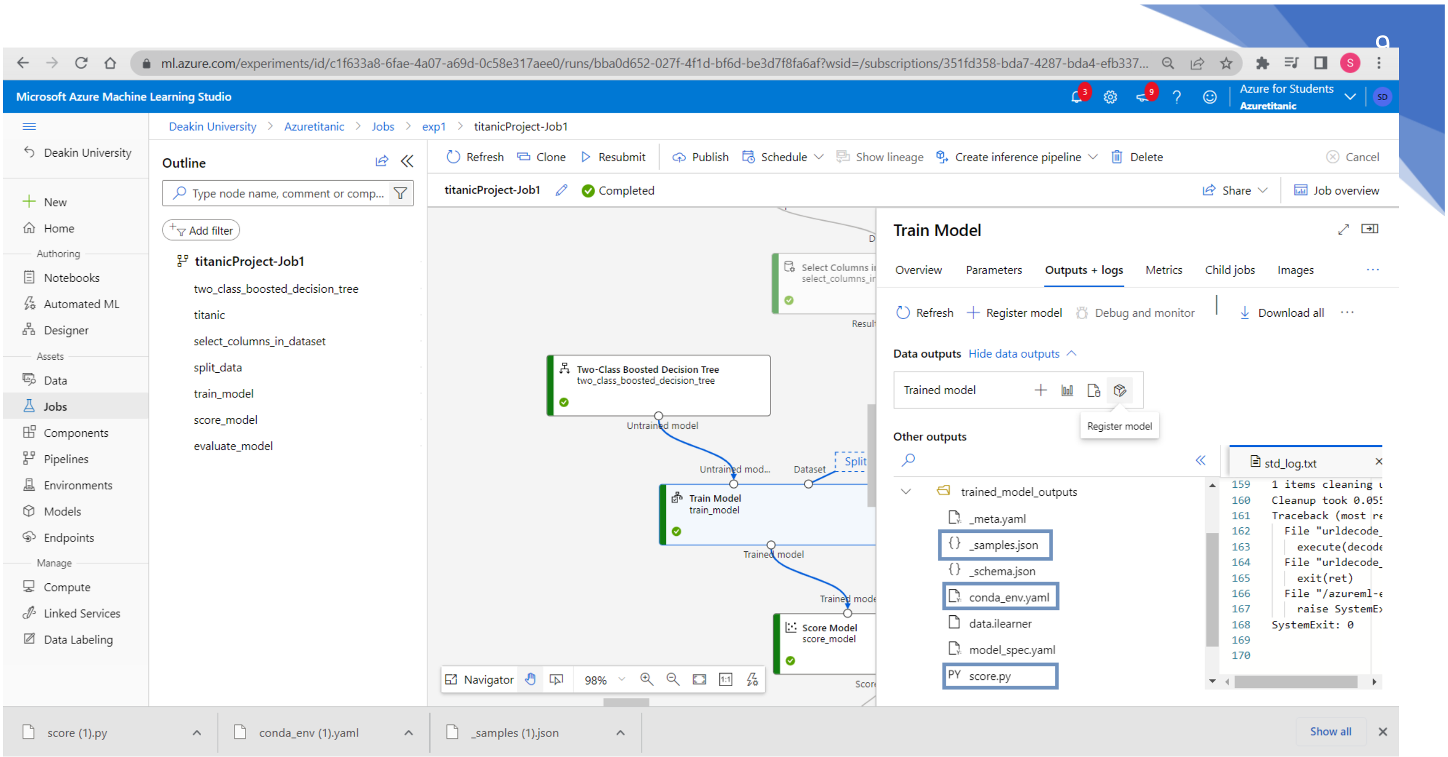
titanicProject-Job1 ✓ Completed

Outline

titanicProject-Job1
- two_class_boosted_decision_tree
- titanic
- select_columns_in_dataset
- split_data
- train_model
- score_model
- evaluate_model

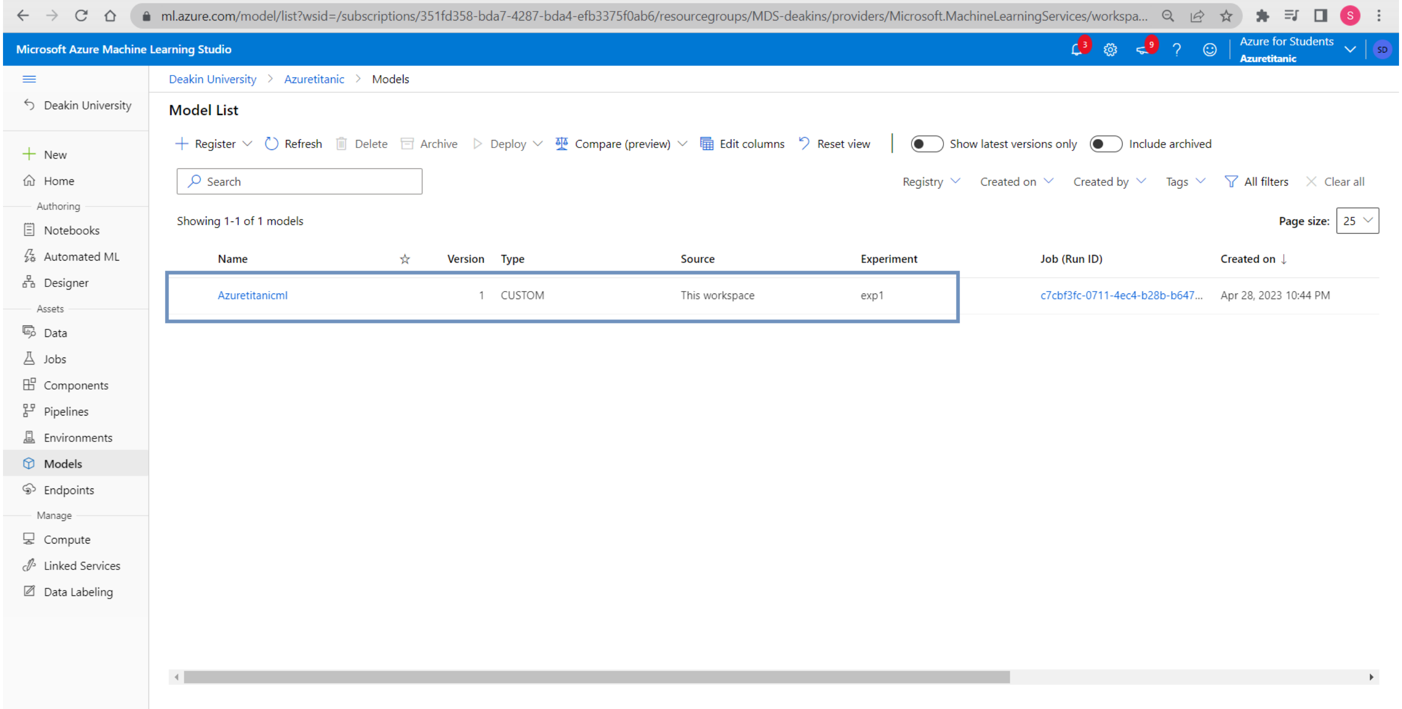
ml.azure.com/experiments/id/c1f633a8-6fae-4a07-a69d-0c58e317aee0/runs/bba0d652-027f-4f1d-bf6d-be3d7f8fa6af?wsid=/subscriptions/351fd358-bda7-4287-bda4-efb337...
Microsoft Azure Machine Learning Studio
Deakin University > Azuretitanic > Jobs > exp1 > titanicProject-Job1
Refresh
Clone
Resubmit
Publish
Schedule
Show lineage
Create inference pipeline
Delete
Cancel
Share
Job overview
Add filter
titanic
Data output
Select Columns in Dataset
Two-Class Boosted Decision Tree
Split Data
Untrained model
Dataset
Train Model
Trained model
Score Model
Scored dataset
Evaluate Model
Navigator
81%

We then double click on train model and download samplejson, conda dependencies and score.py.



```
1 items cleaning
Cleanup took 0.05
Traceback (most re
  File "urldecode_
    execute(decode
  File "urldecode_
    exit(ret)
  File "/azureml-e
    raise SystemEx
SystemExit: 0
```

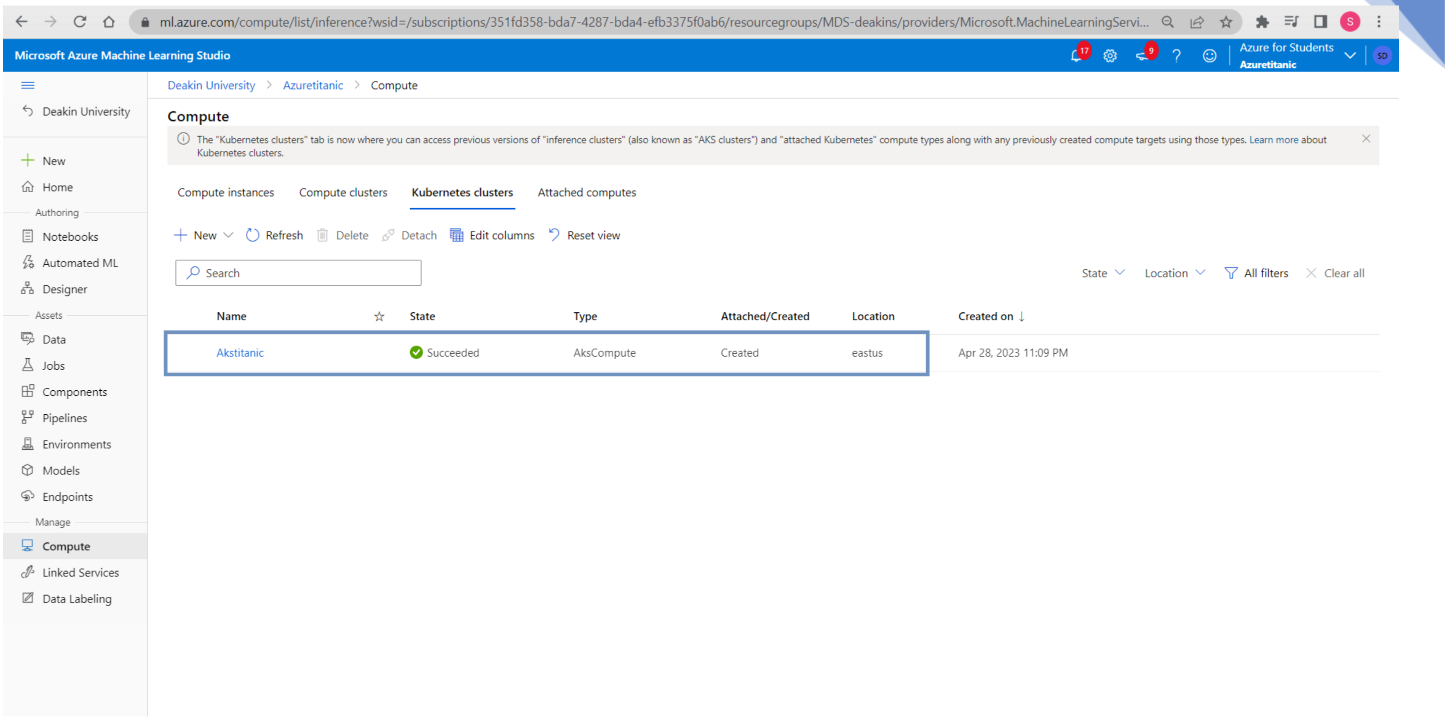
Now we register the model as Azuretitanicml.



Name	☆	Version	Type	Source	Experiment	Job (Run ID)	Created on ↓
Azuretitanicml		1	CUSTOM	This workspace	exp1	c7cbf3fc-0711-4ec4-b28b-b647...	Apr 28, 2023 10:44 PM

We can find the registered model in models option.

We will first create a Kubernetes cluster before deploying the model as shown below. Kubernetes cluster is named Akstitanic that we created for deployment.

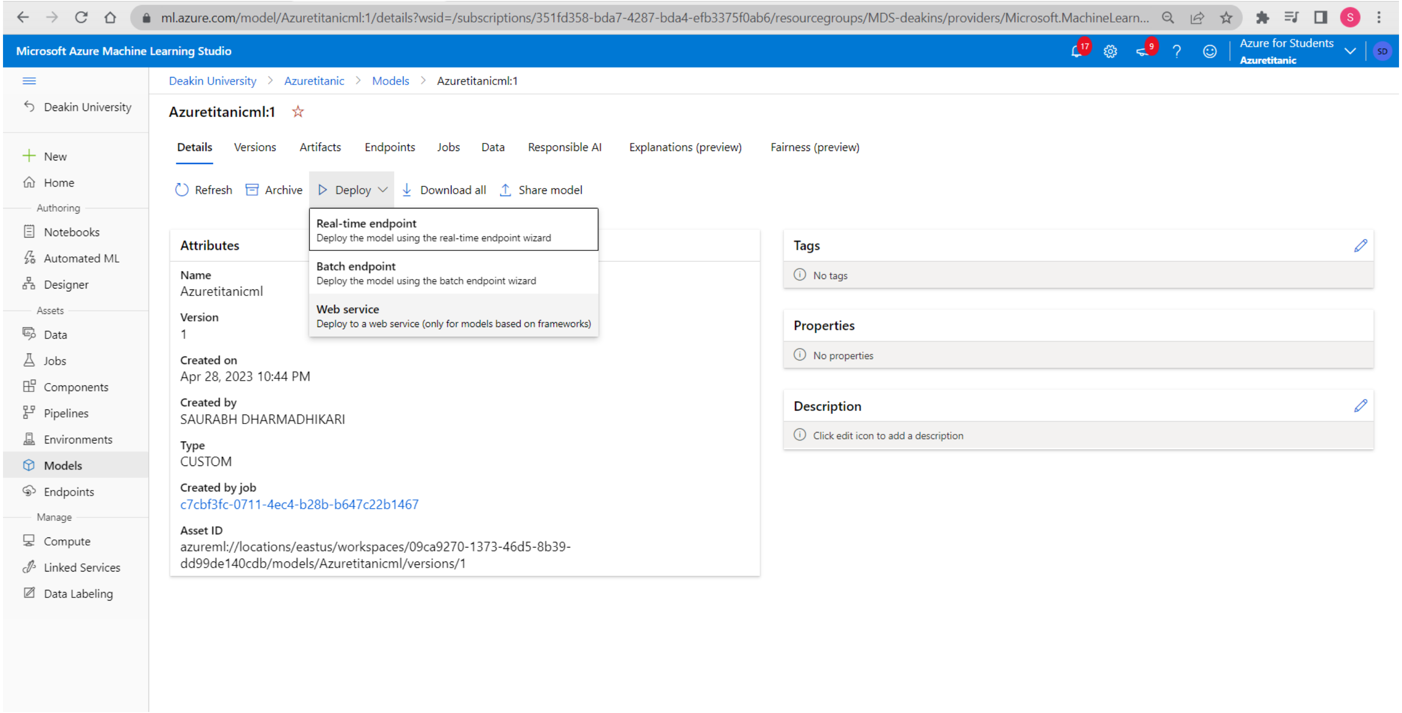


Compute

The "Kubernetes clusters" tab is now where you can access previous versions of "inference clusters" (also known as "AKS clusters") and "attached Kubernetes" compute types along with any previously created compute targets using those types. Learn more about Kubernetes clusters.

Name	State	Type	Attached/Created	Location	Created on ↓
Akstitanic	Succeeded	AksCompute	Created	eastus	Apr 28, 2023 11:09 PM

We find the model created earlier in models after we enter the model, we can deploy it as webservice.



Azuretitanicml:1

Real-time endpoint
Deploy the model using the real-time endpoint wizard

Batch endpoint
Deploy the model using the batch endpoint wizard

Web service
Deploy to a web service (only for models based on frameworks)

Attributes

Name
Azuretitanicml

Version
1

Created on
Apr 28, 2023 10:44 PM

Created by
SAURABH DHARMADHIKARI

Type
CUSTOM

Created by job
c7cbf3fc-0711-4ec4-b28b-b647c22b1467

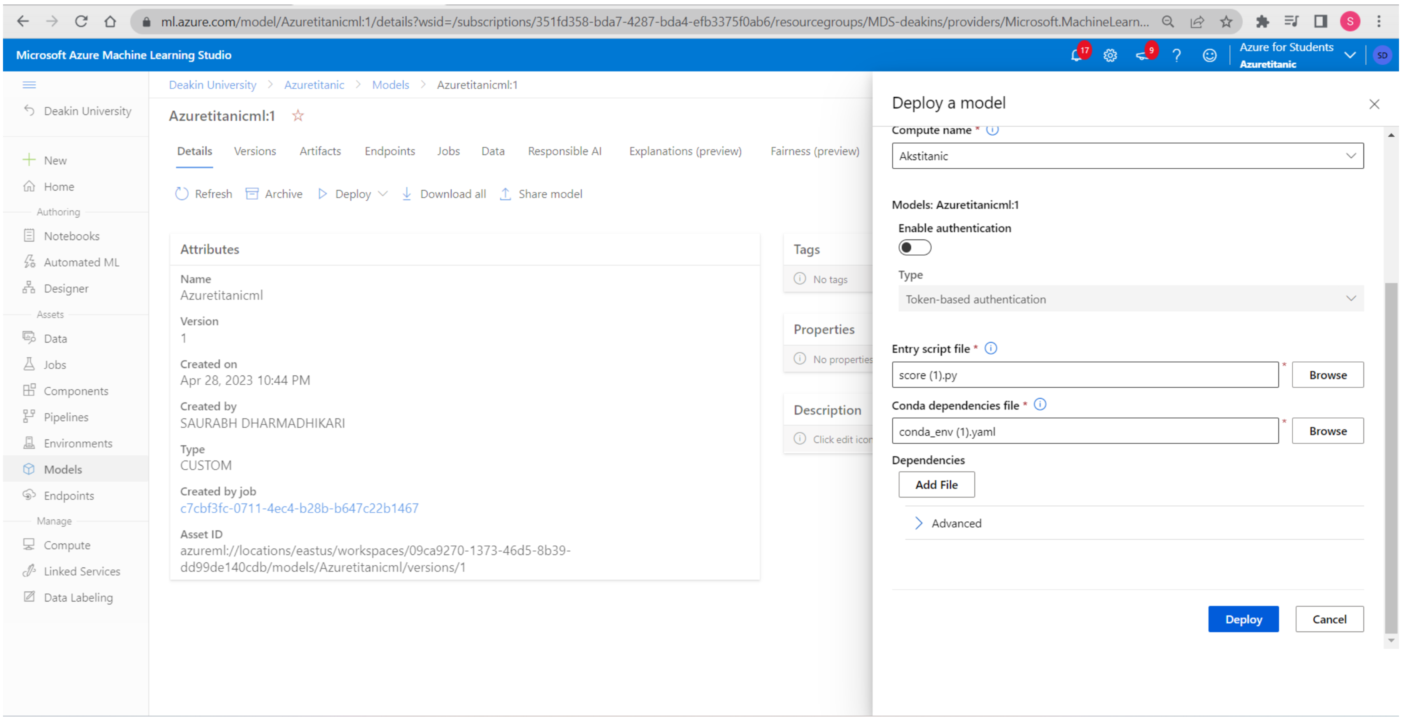
Asset ID
azureml://locations/eastus/workspaces/09ca9270-1373-46d5-8b39-dd99de140cdb/models/Azuretitanicml/versions/1

Tags
No tags

Properties
No properties

Description
Click edit icon to add a description

We put compute name as Akstitanic.



ml.azure.com/model/Azuretitanicml:1/details?wsid=/subscriptions/351fd358-bda7-4287-bda4-efb3375f0ab6/resourcegroups/MDS-deakins/providers/Microsoft.MachineLearn...

Microsoft Azure Machine Learning Studio

Deakin University > Azuretitanic > Models > Azuretitanicml:1

Azuretitanicml:1

Attributes

Name
Azuretitanicml

Version
1

Created on
Apr 28, 2023 10:44 PM

Created by
SAURABH DHARMADHIKARI

Type
CUSTOM

Created by job
c7cbf3fc-0711-4ec4-b28b-b647c22b1467

Asset ID
azureml://locations/eastus/workspaces/09ca9270-1373-46d5-8b39-dd99de140cdb/models/Azuretitanicml/versions/1

Deploy a model

Compute name *
Akstitanic

Models: Azuretitanicml:1

Enable authentication

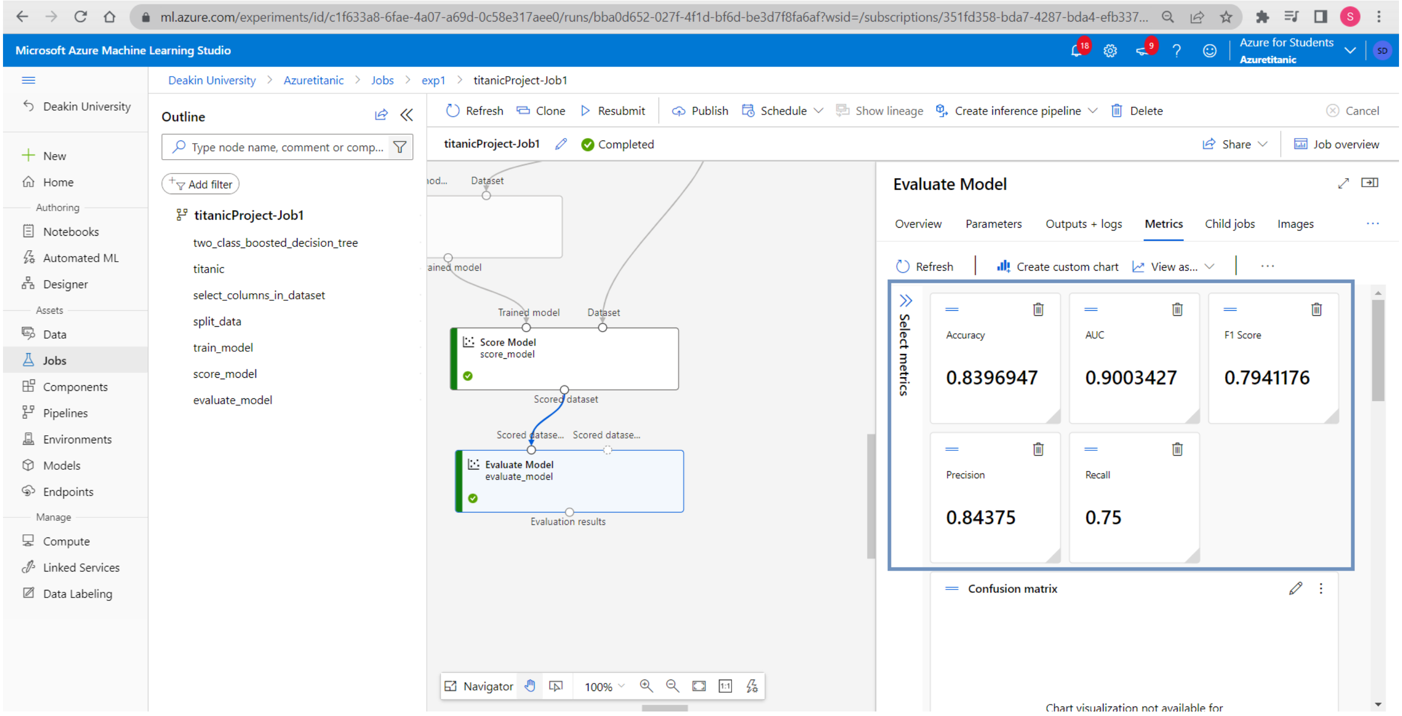
Type
Token-based authentication

Entry script file *
score (1).py

Conda dependencies file *
conda_env (1).yaml

Dependencies

Once deployed we can now evaluate and check the performance of the model that we created.



ml.azure.com/experiments/id/c1f633a8-6fae-4a07-a69d-0c58e317aee0/runs/bba0d652-027f-4f1d-bf6d-be3d7f8fa6af?wsid=/subscriptions/351fd358-bda7-4287-bda4-efb337...

Microsoft Azure Machine Learning Studio

Deakin University > Azuretitanic > Jobs > exp1 > titanicProject-Job1

titanicProject-Job1 Completed

Outline

titanicProject-Job1
- two_class_boosted_decision_tree
- titanic
- select_columns_in_dataset
- split_data
- train_model
- score_model
- evaluate_model

Evaluate Model

Metrics

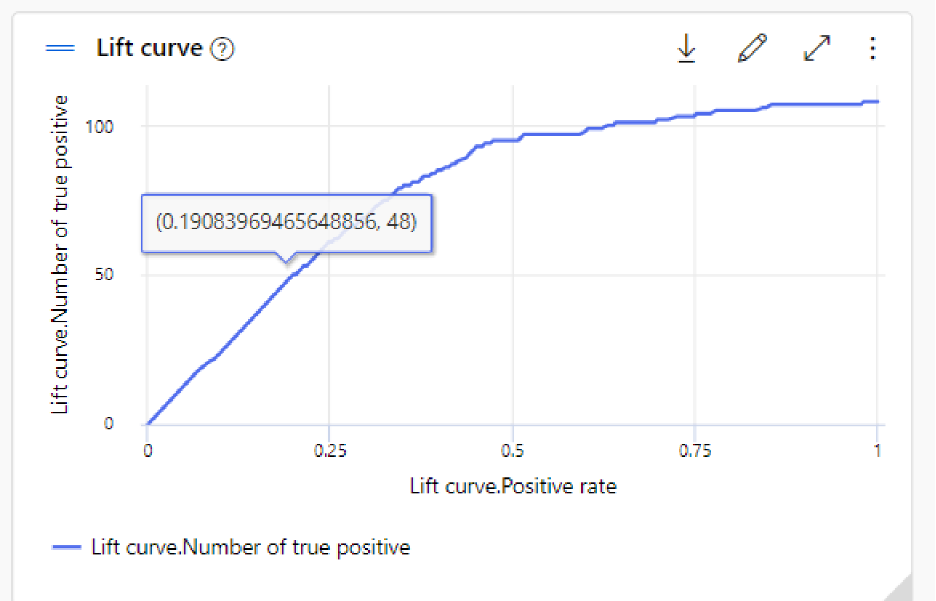
Accuracy	AUC	F1 Score
0.8396947	0.9003427	0.7941176

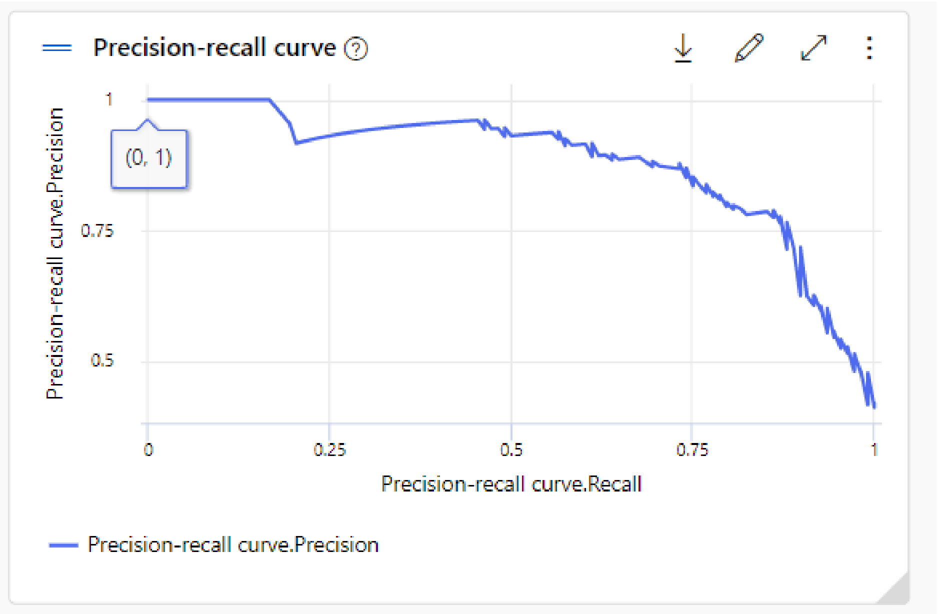
Precision	Recall
0.84375	0.75

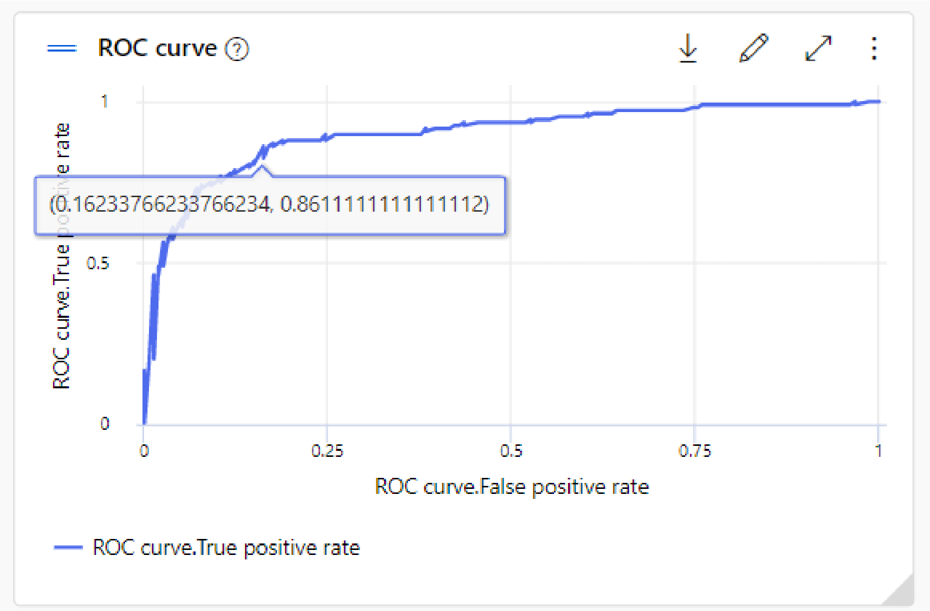
Confusion matrix

Chart visualization not available for

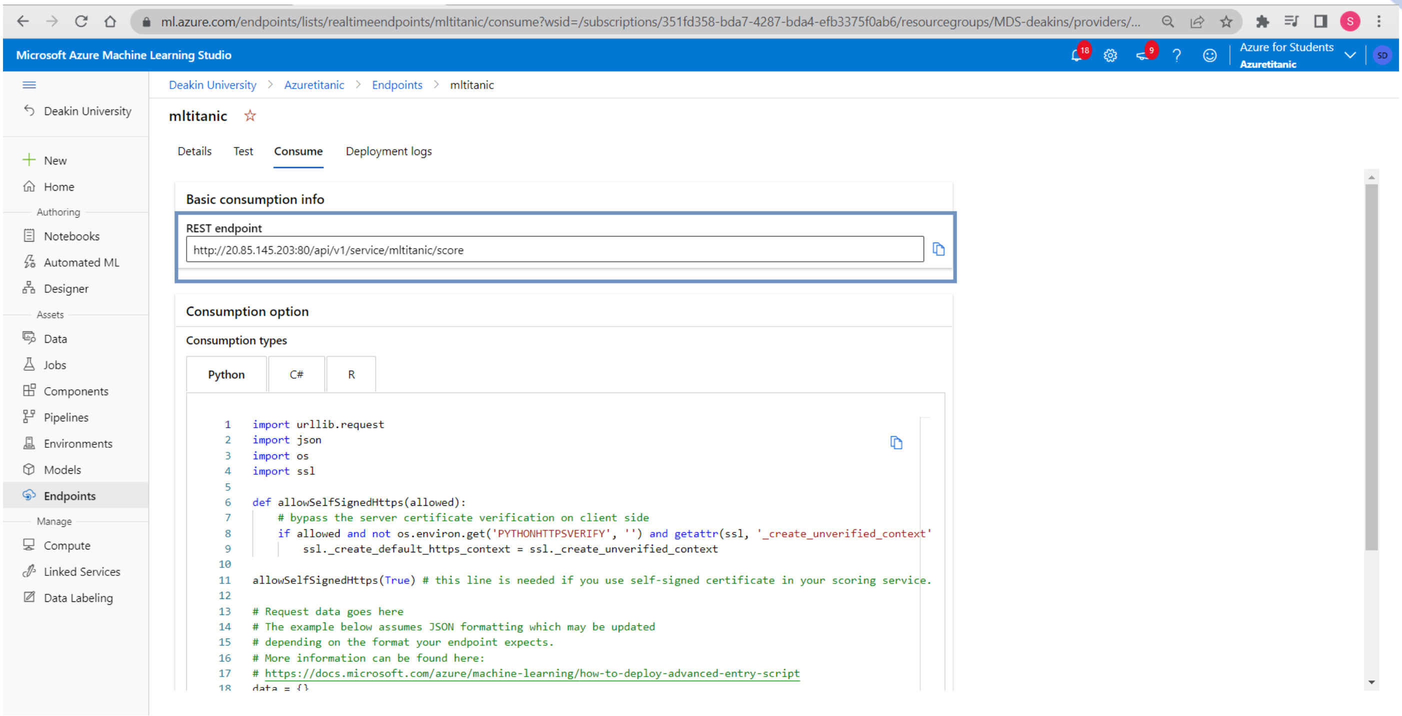
Double click evaluate model and after we click metrics we can understand the performance of the model that we created.


Lift curve
Lift curve.Number of true positive
100
50
0
(0.19083969465648856, 48)
0
0.25
0.5
0.75
1
Lift curve.Positive rate
Lift curve.Number of true positive


Precision-recall curve
Precision-recall curve.Precision
1
0.75
0.5
(0, 1)
0
0.25
0.5
0.75
1
Precision-recall curve.Recall
Precision-recall curve.Precision


ROC curve
ROC curve.True positive rate
1
0.5
0
(0.16233766233766234, 0.8611111111111112)
0
0.25
0.5
0.75
1
ROC curve.False positive rate
ROC curve.True positive rate

In the endpoint option we can find REST API endpoint required by the developer.



REST endpoint: http://20.85.145.203:80/api/v1/service/mltitanic/score

```python
import urllib.request
import json
import os
import ssl

def allowSelfSignedHttps(allowed):
    # bypass the server certificate verification on client side
    if allowed and not os.environ.get('PYTHONHTTPSVERIFY', '') and getattr(ssl, '_create_unverified_context'
        ssl._create_default_https_context = ssl._create_unverified_context

allowSelfSignedHttps(True) # this line is needed if you use self-signed certificate in your scoring service.

# Request data goes here
# The example below assumes JSON formatting which may be updated
# depending on the format your endpoint expects.
# More information can be found here:
# https://docs.microsoft.com/azure/machine-learning/how-to-deploy-advanced-entry-script
data = {}
```

We can also find the codes in different languages here.

References

No references were taken.