



# Bike Rental Prediction Using Azure ML

Date- 28-09-2020

By- Vrisha Parekh  
parekh.vrisha@gmail.com

## Table of Contents

<i>1. Creating an Azure ML workspace .....</i>	<i>2</i>
<i>2. Creating Compute instance .....</i>	<i>3</i>
<i>3. Creating Compute cluster .....</i>	<i>4</i>
<i>4. Creating Inference cluster .....</i>	<i>5</i>
<i>5. Uploading bike data to Azure cloud .....</i>	<i>6</i>
<i>6. Setup and run Auto ML .....</i>	<i>7</i>
<i>7. Make ML pipeline on Azure Designer .....</i>	<i>9</i>
<i>8. Set up model for deployment .....</i>	<i>10</i>

# 1. Creating an Azure ML workspace

Firstly, a machine learning workspace was created using Azure ML.

This screenshot shows the 'Review + create' step in the Azure Machine Learning portal. The page title is 'Machine Learning' with a subtitle 'Create a machine learning workspace'. A green banner indicates 'Validation passed'. The 'Basics' tab is selected, showing project details. The 'Subscription' is 'Free Trial' and the 'Resource group' is '(New) Bike\_sharing\_prediction'. The 'Workspace details' section shows the 'Workspace name' as 'Bike\_sharing\_prediction\_workspace' and the 'Region' as 'Australia East'. At the bottom, there are navigation buttons: 'Review + create' (highlighted), '< Previous', and 'Next: Networking'.

Microsoft Azure Search resources, services, and docs (G+)

Home > New >

## Machine Learning

Create a machine learning workspace

Validation passed

Basics Networking Advanced Tags Review + create

**Project details**

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription \* Free Trial

Resource group \* (New) Bike\_sharing\_prediction [Create new](#)

**Workspace details**

Specify the name, region, and edition for the workspace.

Workspace name \* Bike\_sharing\_prediction\_workspace

Region \* Australia East

[Review + create](#) < Previous Next: Networking

This screenshot shows the 'Review + create' step in the Azure Machine Learning portal, specifically the 'Basics' tab. The page title is 'Machine Learning' with a subtitle 'Create a machine learning workspace'. A green banner indicates 'Validation passed'. The 'Basics' tab is selected, showing project details. The 'Subscription' is 'Free Trial' and the 'Resource group' is '(New) Bike\_sharing\_prediction'. The 'Region' is 'Australia East' and the 'Workspace name' is 'Bike\_sharing\_prediction\_workspace'. The 'Networking' tab shows the 'Connectivity method' as 'Public endpoint (all networks)'. The 'Advanced' tab shows the 'Encryption type' as 'Microsoft-managed keys' and the 'Enable HBI Flag' as 'Disabled'. At the bottom, there are navigation buttons: 'Create' (highlighted), '< Previous', 'Next >', and 'Download a template for automation'.

Microsoft Azure Search resources, services, and docs (G+)

Home > New >

## Machine Learning

Create a machine learning workspace

Validation passed

Basics Networking Advanced Tags **Review + create**

**Basics**

Subscription Free Trial

Resource group (New) Bike\_sharing\_prediction

Region Australia East

Workspace name Bike\_sharing\_prediction\_workspace

**Networking**

Connectivity method Public endpoint (all networks)

**Advanced**

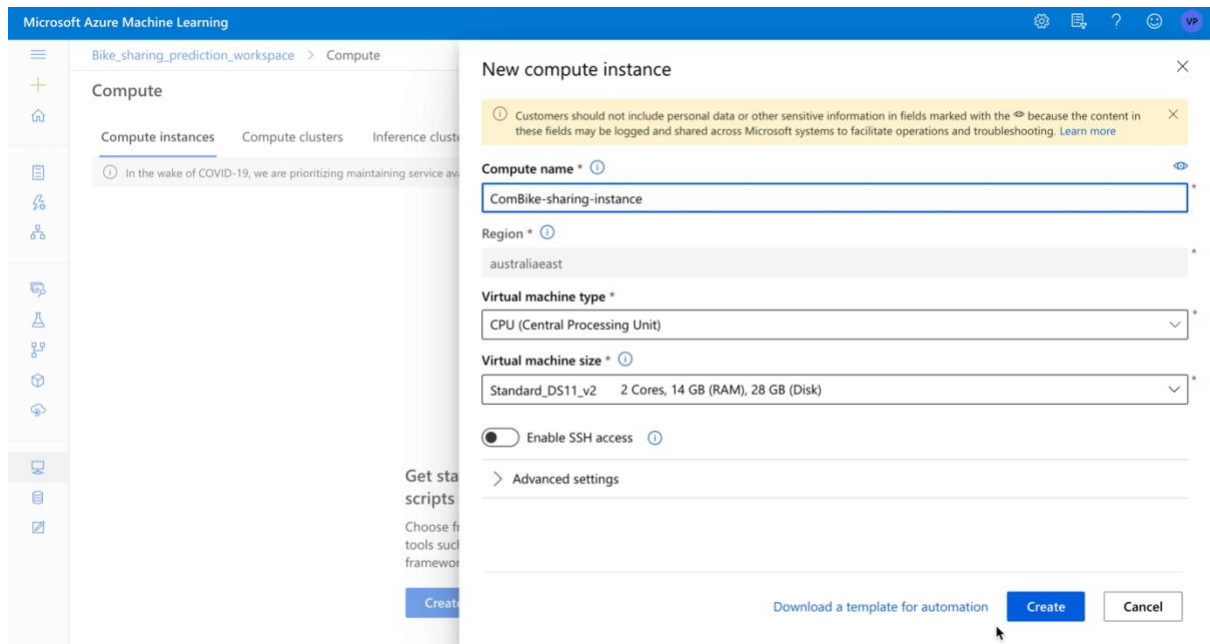
Encryption type Microsoft-managed keys

Enable HBI Flag Disabled

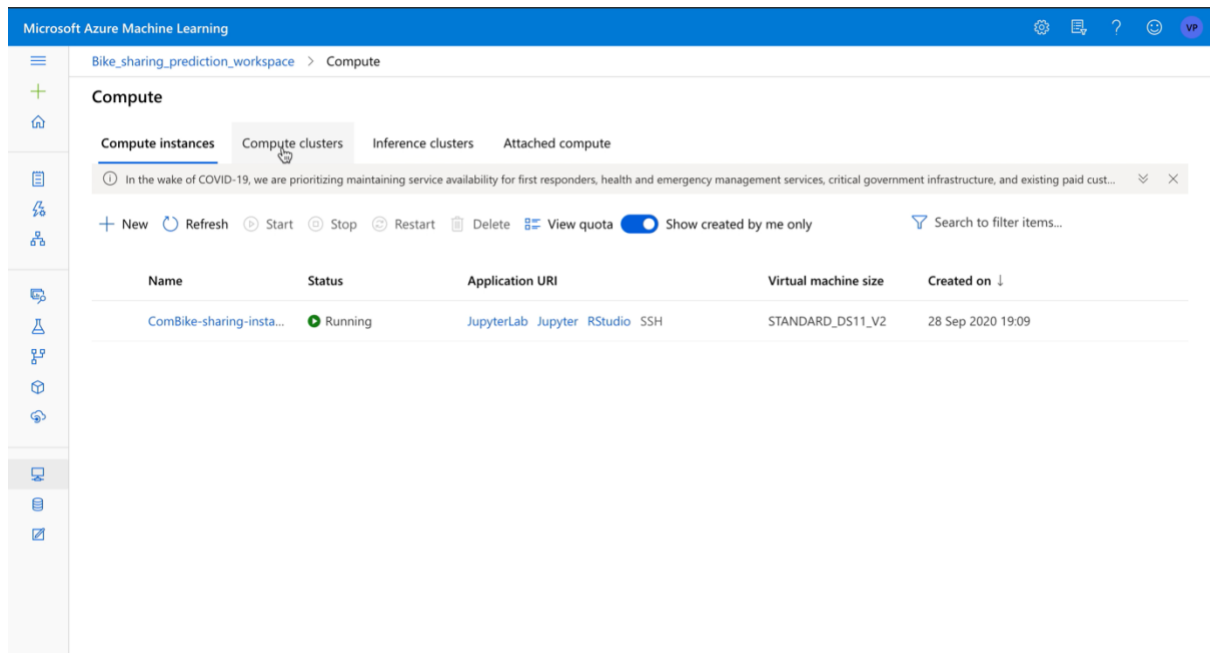
[Create](#) < Previous Next > [Download a template for automation](#)

## 2. Creating Compute instance

Secondly, compute instance was created to use as a workstation on Azure ML Studio.

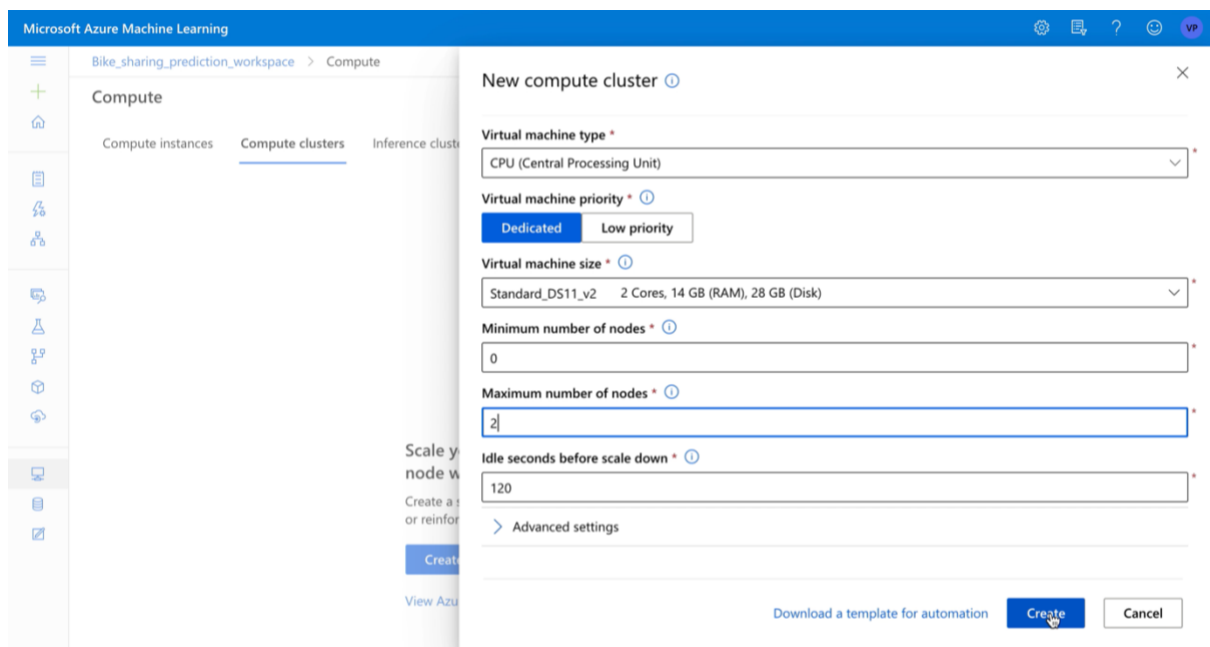


- Compute Instance was created

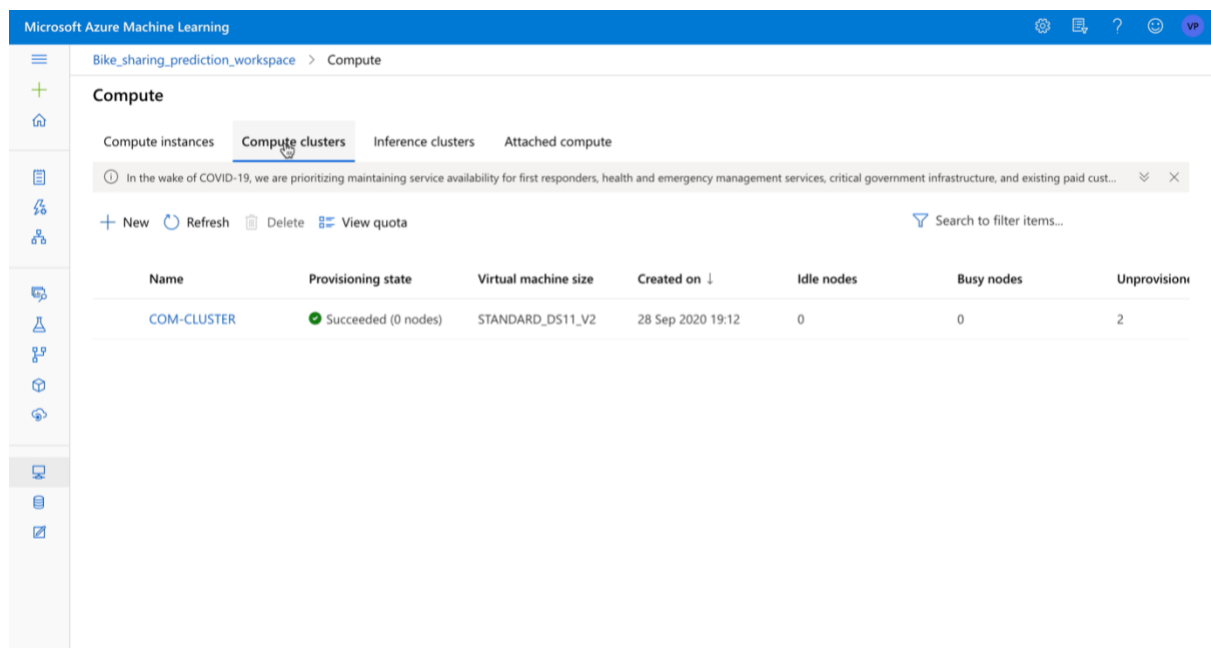


### 3. Creating Compute cluster

Then a compute cluster was created to process the on-demand code, make the Machine Learning (ML) pipeline and to train the model.



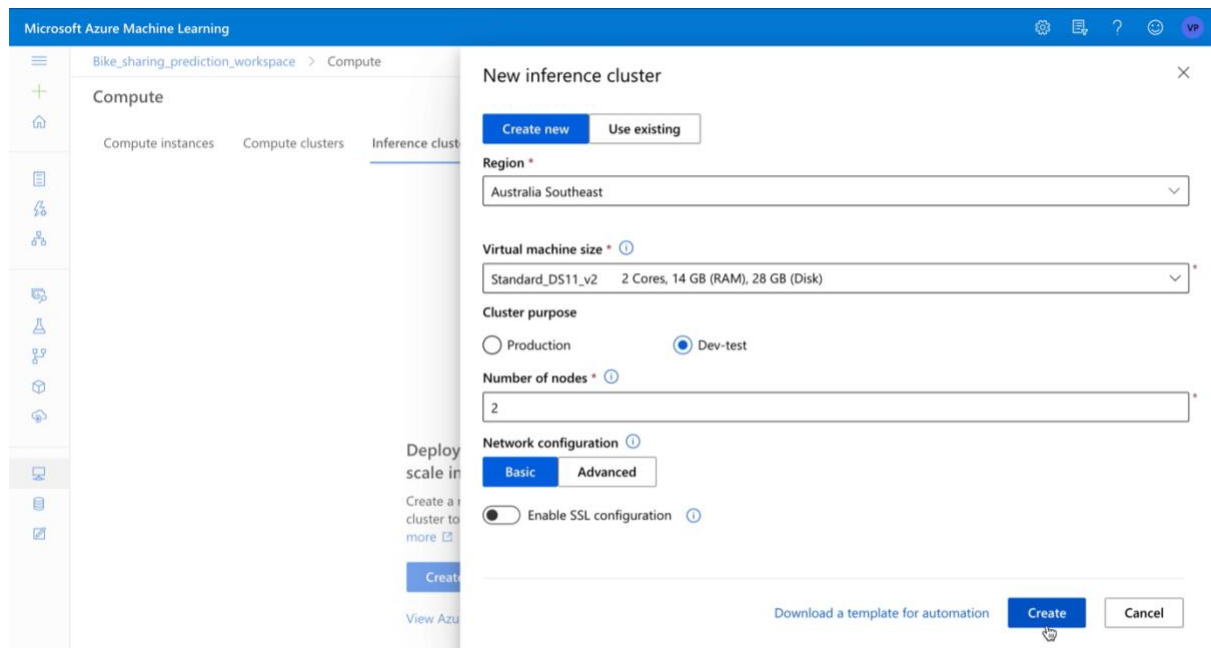
- Compute cluster was created



Name	Provisioning state	Virtual machine size	Created on ↓	Idle nodes	Busy nodes	Unprovisioned
COM-CLUSTER	Succeeded (0 nodes)	STANDARD_DS11_V2	28 Sep 2020 19:12	0	0	2

## 4. Creating Inference cluster

Lastly, an inference cluster was made to deploy the model for it to be used as predictive services.



New inference cluster

Create new Use existing

Region \*  
Australia Southeast

Virtual machine size \*  
Standard\_DS11\_v2 2 Cores, 14 GB (RAM), 28 GB (Disk)

Cluster purpose  
☐ Production ☒ Dev-test

Number of nodes \*  
2

Network configuration  
Basic Advanced

☐ Enable SSL configuration

Create

Download a template for automation Create Cancel

- Inference cluster was created

The screenshot shows the 'Compute' section of the Microsoft Azure Machine Learning interface. The breadcrumb path is 'Bike\_sharing\_prediction\_workspace > Compute'. The 'Compute' section has four tabs: 'Compute instances', 'Compute clusters', 'Inference clusters' (which is selected), and 'Attached compute'. Below the tabs, there are buttons for '+ New', 'Refresh', 'Delete', and 'Detach', along with a search bar 'Search to filter items...'. A table lists the inference clusters:

Name	Type	Created/Attached	Provisioning state	Created on ↓
Infer-cluster	Kubernetes service	Created	Creating	28 Sep 2020 19:15

## 5. Uploading bike data to Azure cloud

Day.csv was uploaded in the dataset section of the Azure ML studio.

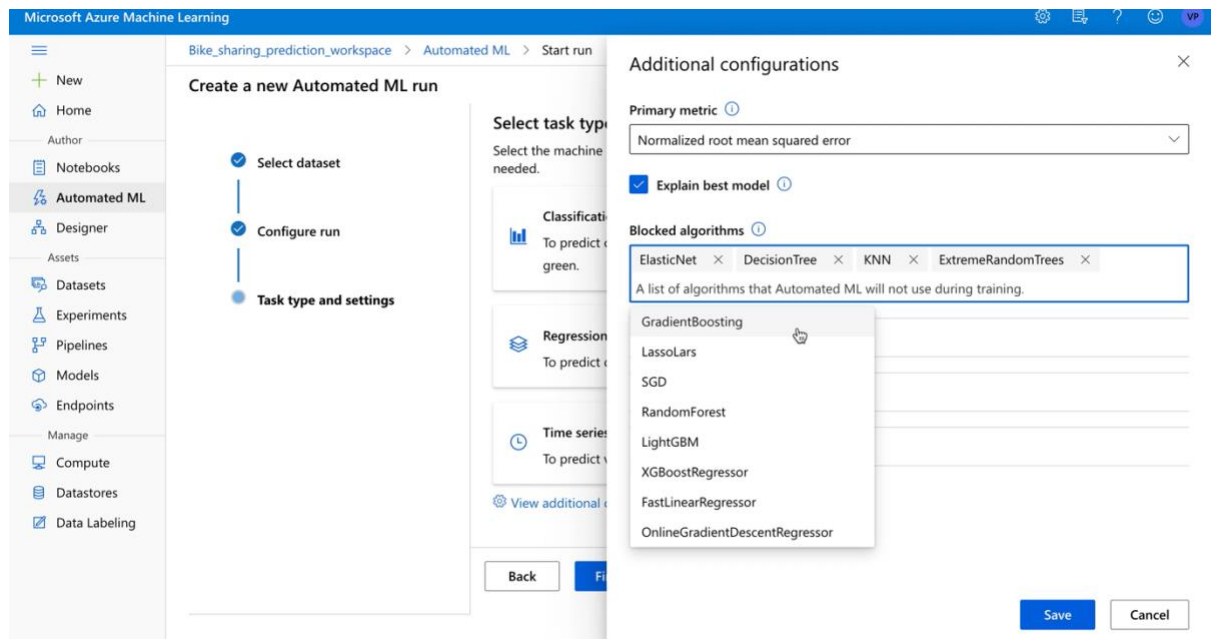
The screenshot shows the 'Datasets' section of the Microsoft Azure Machine Learning interface. The breadcrumb path is 'Bike\_sharing\_prediction\_workspace > Datasets'. A green notification bar at the top says 'Success: Bike\_sharing dataset created successfully'. Below this, the 'Datasets' section has two tabs: 'Registered datasets' (selected) and 'Dataset monitors (preview)'. There are buttons for '+ Create dataset', 'Refresh', and 'Unregister', along with a search bar 'Search to filter items...'. A table lists the registered datasets:

Name	Version	Created on	Modified on	Properties	Created by	Tags
Bike_sharing	1	28 Sep 2020 19:24	28 Sep 2020 19:24	Tabular	Vrisha Parekh	

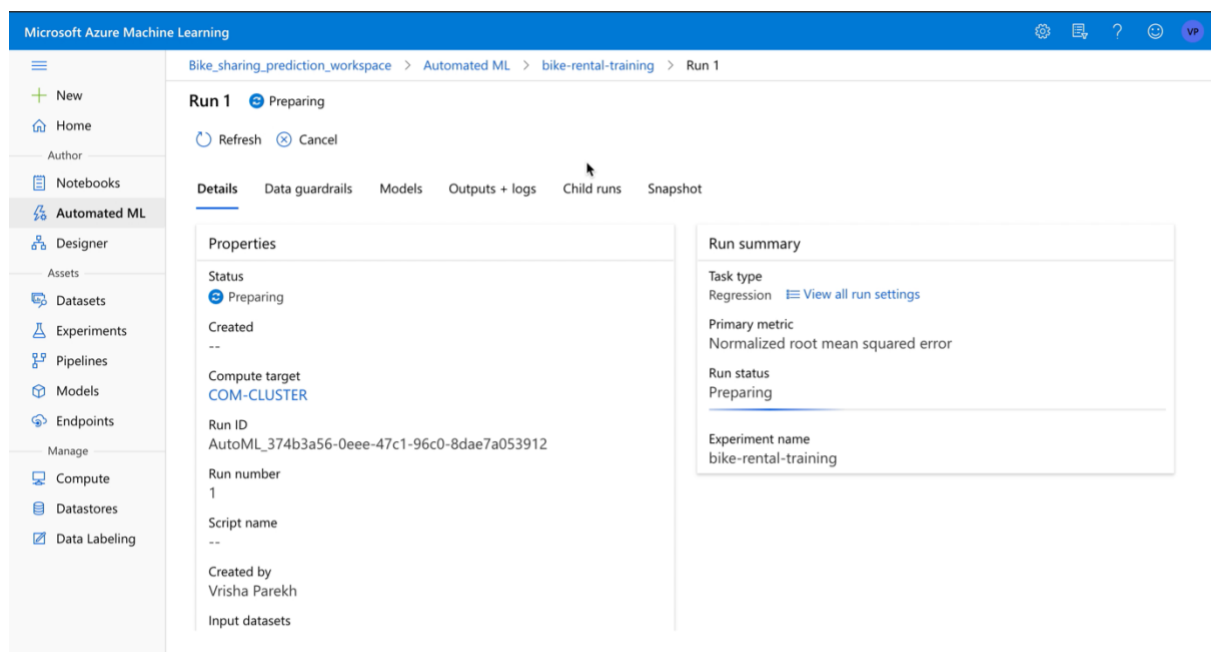
At the bottom of the page, there are navigation links: '< Prev' and 'Next >'.

## 6. Setup and run Auto ML

A mix of algorithms were used as a regression test to see which model performs the best.



- AutoML initiated



- Different regression models utilised



Microsoft Azure Machine Learning

Bike\_sharing\_prediction\_workspace > Automated ML > bike-rental-training > Run 1

Run 1 Running

Refresh Cancel

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

Deploy Download Explain model Search to filter items...

Algorithm name	Explained	Norma... ↑	Sampling	Run	Created	Duration	Status
MaxAbsScaler, ElasticNet		0.00252	100.00 %	<a>Run 17</a>	28 Sep 2020 20:01	51s	Completed
MaxAbsScaler, ElasticNet		0.00572	100.00 %	<a>Run 15</a>	28 Sep 2020 19:58	53s	Completed
StandardScalerWrapper, ElasticNet		0.00830	100.00 %	<a>Run 13</a>	28 Sep 2020 19:56	50s	Completed
MaxAbsScaler, XGBoostRegressor		0.01441	100.00 %	<a>Run 6</a>	28 Sep 2020 19:47	51s	Completed
MaxAbsScaler, LightGBM		0.01455	100.00 %	<a>Run 5</a>	28 Sep 2020 19:47	52s	Completed
MaxAbsScaler, DecisionTree		0.03388	100.00 %	<a>Run 9</a>	28 Sep 2020 19:51	55s	Completed
MaxAbsScaler, DecisionTree		0.03579	100.00 %	<a>Run 11</a>	28 Sep 2020 19:54	52s	Completed
MaxAbsScaler, DecisionTree		0.03698	100.00 %	<a>Run 10</a>	28 Sep 2020 19:52	52s	Completed

## - Best Model summary

Microsoft Azure Machine Learning

Bike\_sharing\_prediction\_workspace > Automated ML > bike-rental-training > Run 1

Run 1 Canceled

Refresh Cancel

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

Properties

Status Canceled

Created 28 Sep 2020 19:47

Duration 20m 28.12s

Compute target COM-CLUSTER

Run ID AutoML\_374b3a56-0eee-47c1-96c0-8dae7a053912

Run number 1

Script name --

Created by

Best model summary

Algorithm name MaxAbsScaler, ElasticNet

Normalized root mean squared error 0.00252 View all other metrics

Sampling 100.00 %

Registered models No registration yet

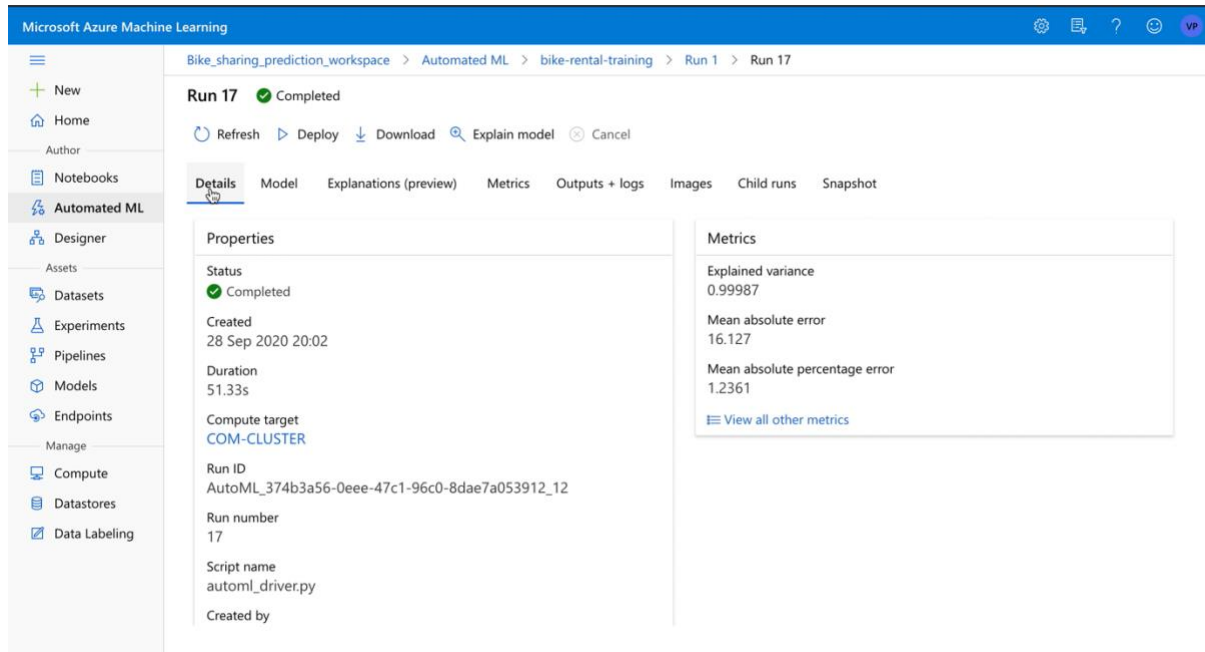
Deployment status No deployment yet

Run summary

Task type Regression View all run settings

<https://ml.azure.com/visualinterface?wsid=/subscriptions/1886747b-7f...>

## - Metrics of Best model

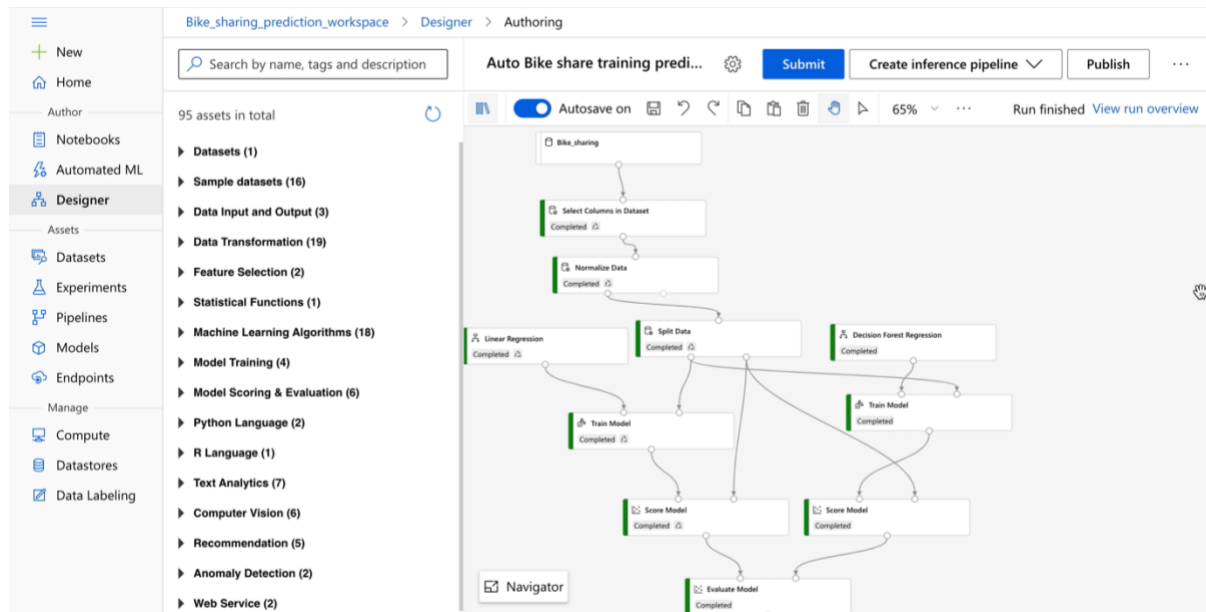


## 7. Make ML pipeline on Azure Designer

The following steps were performed to build the pipeline below:

- Assign the compute cluster to this pipeline.
- Load the dataset in the designer section and observe the stats and plots of each feature in the visualise option of the data box.
- Drag the 'Select columns in Dataset' box and select all the important features for the model.
- Drag the 'Normalize Data' box to the designer and using the Minmax scaling to scale the numeric features as per the maximum and minimum feature of the columns.
- Use the 'Split Data' module to split the data into train and test sets (Train-70% and test- 30%)
- Use the Linear regression module to train the data.
- Utilise the Train data module which trains the model. Assign the label 'cnt' in this case to this module. This module receives input from the Machine learning model (Linear Regression) and Split Data module.

- Next, use the score module to give the scores of the trained model. This module receives one input from trained model and the other input as test set from the Split data module to predict the scores.
- Lastly, evaluate module is put to use to evaluate the efficiency of the model.



## 8. Set up model for deployment

As the pipeline was submitted and it ran successfully, this pipeline can now be used for real time predictions. For this, the model was deployed as a web-service. For this, following steps were taken:

- The 'Create inference pipeline' tab was clicked to deploy the model as a web service using the inference cluster.
- Some modifications to the above model were made like removing one model. Secondly, the transformations like splitting data and training model were all part of the 'Apply Transformation' module. Moreover, removing the output variable from the 'Select columns in Dataset' module as that is something which the model will predict in real time.

Now this model, just needs to be connected to the client side for real-time predictions

Microsoft Azure Machine Learning

Bike\_sharing\_prediction\_workspace > Designer > Authoring

Training pipeline Real-time inference pipeline

Search by name, tags and description

Auto Bike share training prediction-real time inference

Submit Deploy

Autosave on 80% Draft autosaved on 9/28/2020, 9:25

95 assets in total

**Designer**

Assets

- Datasets
- Experiments
- Pipelines
- Models
- Endpoints

Manage

- Compute
- Datastores
- Data Labeling

**Datasets (1)**

- Sample datasets (16)
- Data Input and Output (3)
- Data Transformation (19)
- Feature Selection (2)
- Statistical Functions (1)
- Machine Learning Algorithms (18)
- Model Training (4)
- Model Scoring & Evaluation (6)
- Python Language (2)
- R Language (1)
- Text Analytics (7)
- Computer Vision (6)
- Recommendation (5)
- Anomaly Detection (2)

Web Service Input

Bike\_sharing

TD-Auto\_Bike\_share\_training\_predi...

Select Columns in Dataset

MD-Auto\_Bike\_share\_training\_pre...

Apply Transformation

Score Model

Web Service Output

Evaluate Model

```
graph TD; WSInput[Web Service Input] --> TD[TD-Auto_Bike_share_training_predi...]; WSInput --> SCDS[Select Columns in Dataset]; TD --> MD[MD-Auto_Bike_share_training_pre...]; SCDS --> MD; MD --> AT[Apply Transformation]; AT --> SM[Score Model]; SM --> WSO[Web Service Output]; SM --> EM[Evaluate Model];
```