

## Submission Details:

Q1: 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. Solution:

## Solution:

Design and deploy an AI-based solution on a cloud platform ‘Azure’ with Microsoft Azure Machine Learning Studio and finding key elements associated for integration into different applications.

## Steps:

Login to Azure portal and create a resource group dss223014687.



Validation passes.

Basics

Subscription: Azure for Students  
Resource group: ds223014687  
Region: Central India

Tags

None.

Activate Windows  
Go to Settings to activate Windows.

26°C Sunny 08:30 09-04-2023

ds223014687

Resource group

Search

+ Create Manage view Delete resource group Refresh Export to CSV Open query Assign tags Move Delete Export template JSON View

Activity log

Access control (IAM)

Tags

Resource visualizer

Events

Deployments

Security

Policies

Properties

Locks

Cost management

Cost analysis

Cost alerts (preview)

Budgets

Advisor recommendations

No resources match your filters  
Try changing or clearing your filters.  
Create resources Clear filters Learn more

Activate Windows  
Go to Settings to activate Windows.

26°C Sunny 08:30 09-04-2023

## Create Azure Machine Learning workspace dsstask3-1.

Activate Windows  
Go to Settings to activate Windows.

26°C Sunny 08:30 09-06-2023

Azure Machine Learning

ds223014687@deakin.edu.au

Home

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

No grouping List view

No workspaces to display

Workspaces are where you manage all the models, assets, and data related to your machine learning projects. Create one now to start using Azure Machine Learning.

Learn more

Activate Windows  
Go to Settings to activate Windows.

26°C Sunny 08:30 09-06-2023

[https://portal.azure.com/#view/Microsoft\\_Azure\\_MLTeamAccounts/CreateMachineLearningServicesBladeV2](https://portal.azure.com/#view/Microsoft_Azure_MLTeamAccounts/CreateMachineLearningServicesBladeV2)

**Azure Machine Learning**

Create a machine learning workspace

**Resource details**

Every workspace must be assigned to an Azure subscription, which is where billing happens. You use resource groups like folders to organize and manage resources, including the workspace you're about to create.

Learn more about Azure resource groups [? Help](#)

Subscription [Azure for Students](#) [Create new](#)

Resource group [dss223014687](#) [Create new](#)

**Workspace details**

Configure your basic workspace settings like its storage connection, authentication, container, and more. [Learn more](#) [? Help](#)

Workspace name [dsstask3-1](#)

Region [Central India](#)

Storage account [\(new\) dsstask313612875439](#) [Create new](#)

Key vault [\(new\) dsstask317823210246](#) [Create new](#)

Application insights [\(new\) dsstask319677879259](#) [Create new](#)

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

Activate Windows  
Go to Settings to activate Windows.

26°C Sunny 08:23 09-04-2023

[https://portal.azure.com/#view/Microsoft\\_Azure\\_MLTeamAccounts/CreateMachineLearningServicesBladeV2](https://portal.azure.com/#view/Microsoft_Azure_MLTeamAccounts/CreateMachineLearningServicesBladeV2)

**Azure Machine Learning**

Create a machine learning workspace

**Validation passed**

**Basics**

Subscription [Azure for Students](#)  
Resource group [dss223014687](#)  
Region [Central India](#)  
Workspace name [dsstask3-1](#)  
Storage account [\(new\) dsstask313612875439](#)  
Key vault [\(new\) dsstask317823210246](#)  
Application insights [\(new\) dsstask319677879259](#)  
Container registry [None](#)

**Networking**

Connectivity method [Enable public access from all networks](#)

**Advanced**

Identity type [System assigned](#)  
Encryption type [Microsoft managed keys](#)  
Enable HBI Flag [Disabled](#)

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

Activate Windows  
Go to Settings to activate Windows.

26°C Sunny 08:24 09-04-2023

<https://portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/~/overview/id/%2Fsubscriptions%2Fc5dd5bc-e6d1-4053-b114-9f70c1dew90e%2FresourceGroups%2Fdss223014687%2Fdeployments%2Fdsstask3-1>

**Microsoft.MachineLearningServices | Overview**

Deployment

[Search](#) [Delete](#) [Cancel](#) [Redeploy](#) [Download](#) [Refresh](#)

**Your deployment is complete**

Deployment name: [Microsoft.MachineLearningServices](#)  
Subscription: [Azure for Students](#)  
Resource group: [dss223014687](#)

Start time: 4/9/2023, 8:24:22 AM  
Correlation ID: 291def5c-3217-484f-a644-ddd1d9a5ebef

**Deployment details**

Resource	Type	Status	Operation details
<a href="#">dsstask3-1</a>	<a href="#">Microsoft.MachineLearningService</a>	<a href="#">Accepted</a>	<a href="#">Operation details</a>
<a href="#">dsstask319677879259</a>	<a href="#">Microsoft.Insights/components</a>	<a href="#">OK</a>	<a href="#">Operation details</a>
<a href="#">dsstask313612875439</a>	<a href="#">Microsoft.Storage/storageAccounts</a>	<a href="#">OK</a>	<a href="#">Operation details</a>
<a href="#">dsstask317823210246</a>	<a href="#">Microsoft.KeyVault/vaults</a>	<a href="#">OK</a>	<a href="#">Operation details</a>
<a href="#">DeployLogWorkspace-kkbqjgh65oz</a>	<a href="#">Microsoft.Resources/deployments</a>	<a href="#">OK</a>	<a href="#">Operation details</a>

**Next steps**

[Go to resource](#)

**Cost Management**  
Get notified to stay within your budget and prevent unexpected charges on your bill.  
[Set up cost alerts >](#)

**Microsoft Defender for Cloud**  
Secure your apps and infrastructure  
[Go to Microsoft Defender for Cloud >](#)

**Free Microsoft tutorials**  
Start learning today >

**Work with an expert**  
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.  
[Find an Azure expert >](#)

Activate Windows  
Go to Settings to activate Windows.

26°C Sunny 08:33 09-04-2023

The screenshot shows the Microsoft Azure Machine Learning workspace overview page for 'dsstask3-1'. The left sidebar includes sections for Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Events, Settings (Networking, Properties, Locks), Monitoring (Alerts, Metrics, Diagnostic settings, Logs), and Automation (Tasks (preview)). The main content area displays 'Essentials' information: Resource group (ds223014687), Location (Central India), Subscription (Azure for Students), Subscription ID (cc9dd5bc-e6d1-4053-b114-9f70c1dae90e), and Storage (dsstask3118612875439). It also lists Studio web URL, Container Registry, Key Vault, Application Insights, and Mflow tracking URL. A central icon is a grey 3D cube. Below it, a callout reads 'Work with your models in Azure Machine Learning Studio' with a 'Launch studio' button. The status bar at the bottom shows the date (09-04-2023) and time (08:26).

The screenshot shows the Microsoft Azure Machine Learning Studio Designer interface. The left sidebar has sections for Design (New pipeline, Classic pipeline, Custom) and Pipelines (Pipeline drafts, Pipeline jobs). The main area displays a 'New pipeline' section with various prebuilt components: Create a new pipeline using classic prebuilt components, Image Classification using DenseNet, Binary Classification using VGGish Model - A..., Wide & Deep based Recommendation - Retailu..., Regression - Automobile Price Prediction (Basic), Regression - Automobile Price Prediction (Compare algori..., and Binary Classification with Feature Selection - Income. Below this is a 'Pipelines' table with columns for Name, Pipeline type, Updated on, and Created by. The status bar at the bottom shows the date (09-04-2023) and time (08:26).

**Data Ingestion:** The ‘Admission\_Predict’ dataset is used, in which **Chance of Admit** (target) is classified using Decision Tree Model.

**Create data asset**

**Set the name and type for your data asset**

Name: Admission\_Predict  
Description: C:\Users\Admin\Desktop\DATASETS\Admission\_Predict  
Type: Tabular

**Use cases for data types**

The File type is recommended in most scenarios when you are working with a single data file or any type (including tabular data). This type allows you to specify a file location by URL in a storage location on your local computer, an attached Datastore, blob/ADLS storage, or a publicly available http(s) location. There are many ways to use this type in your machine learning training CLU v2 or Python SDK v2, this data type is called url\_file. Learn more about the url\_file type.

When should I use File type?

The Folder type has all the same capabilities and use cases as the File type, but is used when specifying a folder location. In the Azure Machine Learning CLU v2 or Python SDK v2, this data type is called url\_folder. Learn more about the url\_folder type.

When should I use Table type?

The Table type is most useful for advanced scenarios where you might need to abstract the schema definition for easier sharing. You should use it when you have complex transformations and schema mapping requirements. The Table type is recommended for simple tabular data, the File and Folder types are recommended. If you choose the Table type in the Azure Machine Learning CLU v2 or Python SDK v2, this data type is called mtable. Learn more about the mtable type.

Activate Windows  
Go to Settings to activate Windows  
Cancel

Back Next

**Create data asset**

**Choose a source for your data asset**

Choose the data source you want to create your asset from. A data source can be from a local storage location on your computer, from an attached datastore, from Azure storage, or from a publicly available web location.

From Azure storage  
From local files  
From SQL databases  
From web files  
From Azure Open Datasets

No data assets found

Activate Windows  
Go to Settings to activate Windows  
Cancel

Back Next

**Create data asset**

**Select a datastore**

Choose a storage type and a datastore from the list. You can also create a new datastore.

Datastore type: Azure Blob Storage  
Create new datastore

Search datastore

Name	Storage name	Created on
workspaceblobstore	dostask313612875439	Apr 9, 2023 8:25 AM
workspaceartifactstore	dostask313612875439	Apr 9, 2023 8:25 AM

All filters Clear all

No data assets found

Activate Windows  
Go to Settings to activate Windows  
Cancel

Back Next

**Create data asset**

**Choose a file or folder**

Choose files or folders to upload from your local drive. If you upload multiple folders or files, they will be stored in a containing folder.

**Upload path**  
AzureML/subscriptions/cchdd5bc-e6d1-4053-b114-9f70c1daef0e/resourcegroups/ds223014687/workspace

**Information**

**What file types can I use?**  
Supported file types include: delimited (such as csv or tsv), Parquet, JSON Lines, and plain text.

**Where are files uploaded?**  
Files will be uploaded to the selected datastore and made available in your workspace.

**Upload list**  
Admission\_Predict.csv 10.67 KB/10.67 KB ...

**File or folder selection**

**Settings**

**Schema**

**Review**

No data assets found

Back Next

Activate Windows  
Go to Settings to activate Windows

Cancel

**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	Delimiter	Example	Encoding
Delimited	Comma	Field1,Field2,Field3	UTF-8

**Column headers**  
Skip rows  
All files have same headers 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**

Serial No.	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Chance of Admit
1	337	118	4	4	4.5	9.65	1
2	324	107	4	4	4.5	8.87	1
3	316	104	3	3	3.5	8	0
4	322	110	3	3.5	2.5	8.67	1
5	314	103	2	2	3	8.21	0
6	330	115	5	4.5	3	9.34	1
7	321	109	3	3	4	8.2	1
8	308	101	2	3	4	7.9	0
9	302	102	1	2	1.5	8	0
10	323	108	3	3.5	3	8.6	0
11	325	106	3	3.5	4	8.87	1

Back Next

Activate Windows  
Go to Settings to activate Windows

Review Cancel

**Create data asset**

**Schema**

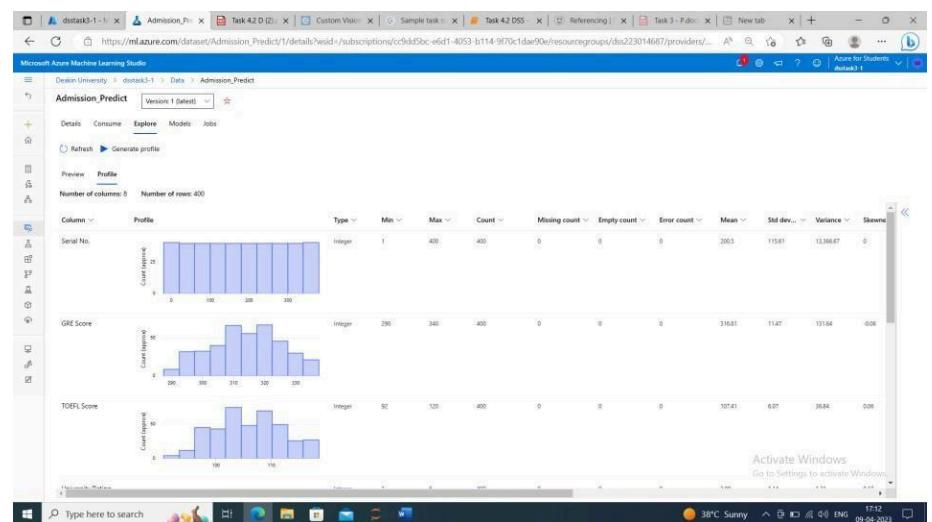
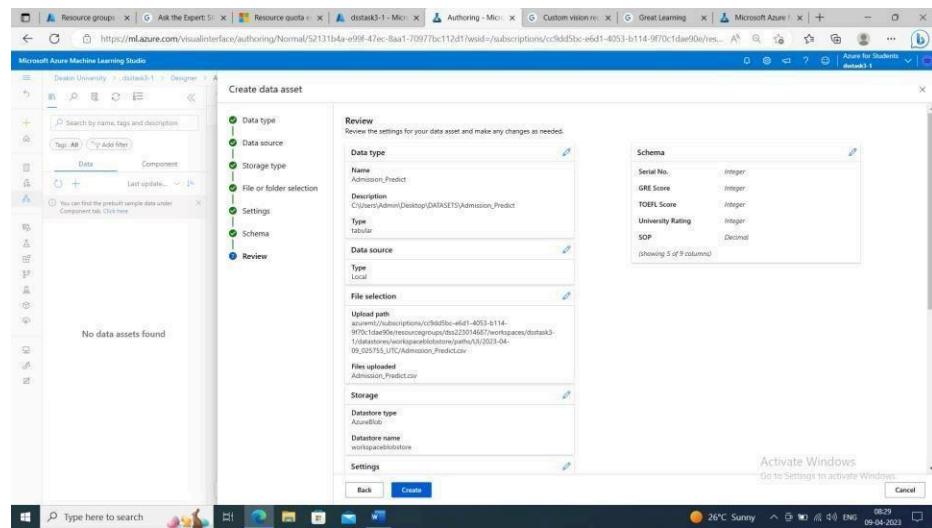
Column types are auto-detected based on the initial subset of the data and can be updated here. Values not aligning with the specified column type will fail conversion and would be either nullified or replaced with error value. Any conversions preview errors are non-blocking and you can proceed.

Include	Column name	Type	Example values	Date format	Properties
Path	String		Not applicable to selected type	Not applicable to selected type	
Serial No.	Integer	1, 2, 3	Not applicable to selected type	Not applicable to selected type	
GRE Score	Integer	337, 324, 316	Not applicable to selected type	Not applicable to selected type	
TOEFL Score	Integer	118, 107, 104	Not applicable to selected type	Not applicable to selected type	
University Rating	Integer	4, 4, 3	Not applicable to selected type	Not applicable to selected type	
SOP	Decimal (dot:.)	4.5, 4.3	Not applicable to selected type	Not applicable to selected type	
LOR	Decimal (dot:.)	4.5, 4.5, 3.5	Not applicable to selected type	Not applicable to selected type	
CGPA	Decimal (dot:.)	9.65, 8.87, 8	Not applicable to selected type	Not applicable to selected type	
Chance of Admit	Integer	1, 1, 0	Not applicable to selected type	Not applicable to selected type	

Back Next

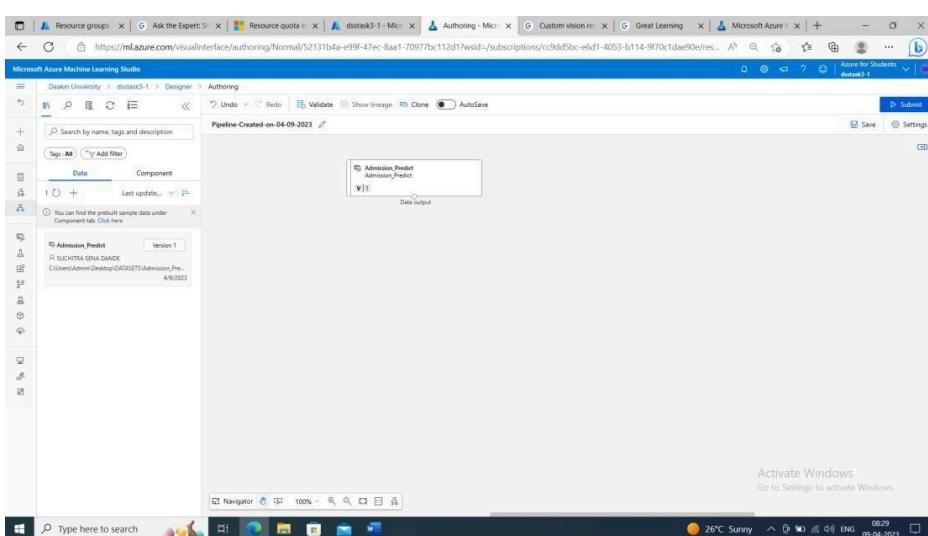
Activate Windows  
Go to Settings to activate Windows

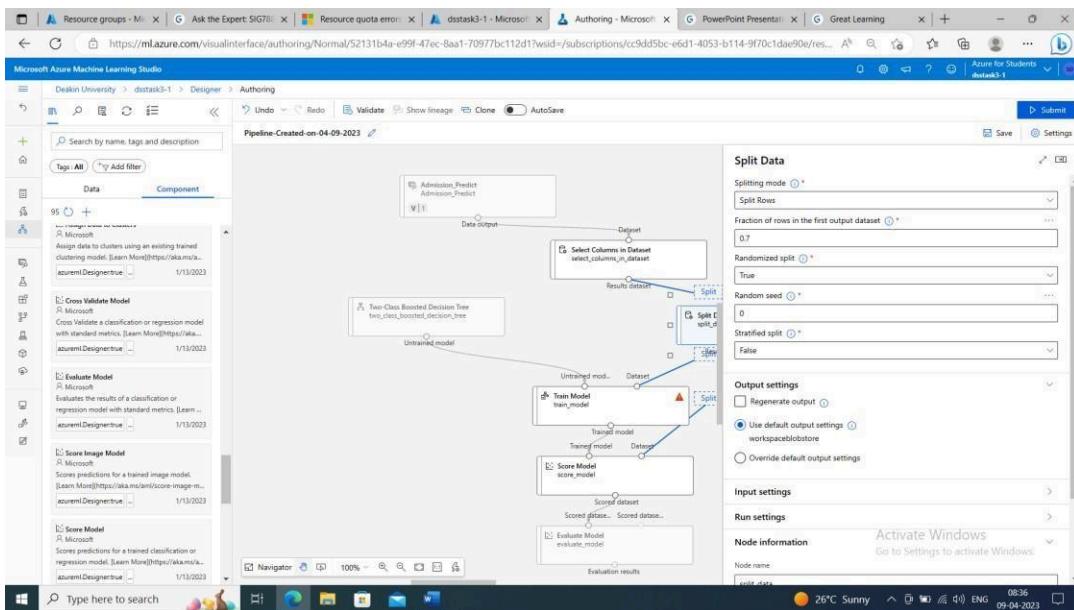
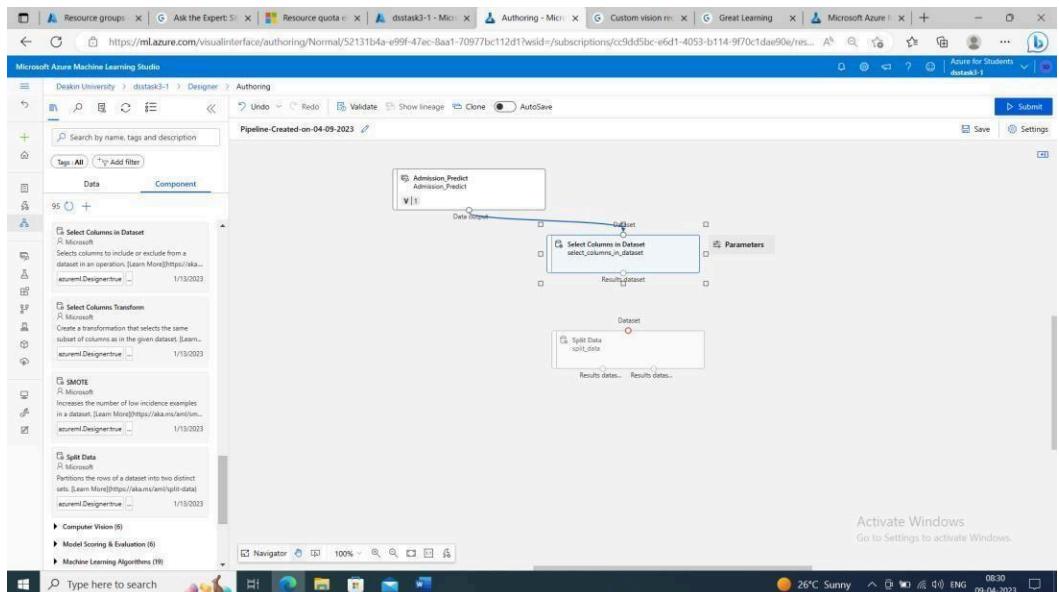
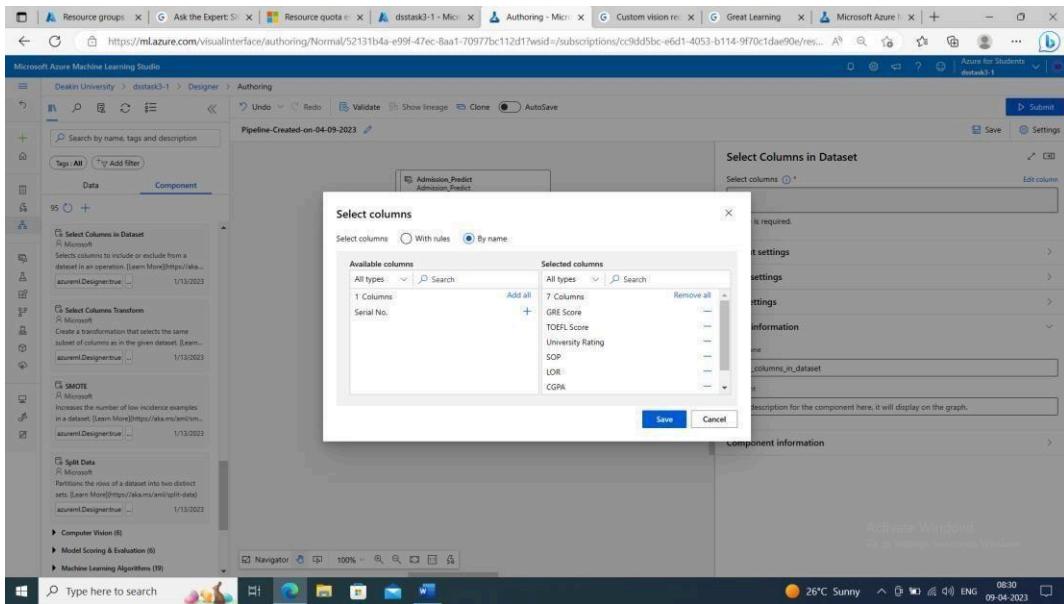
Cancel

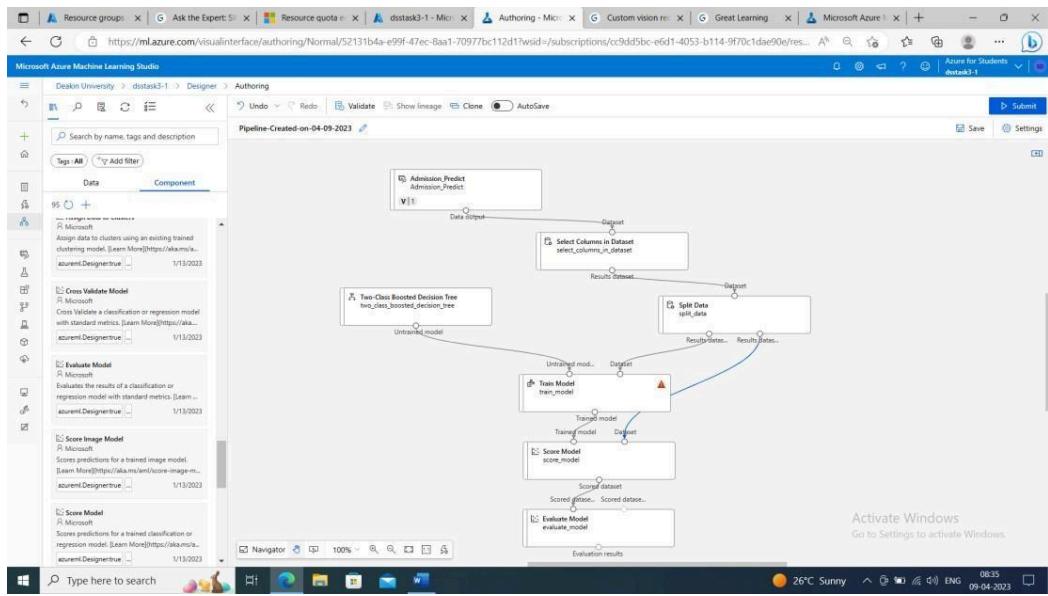


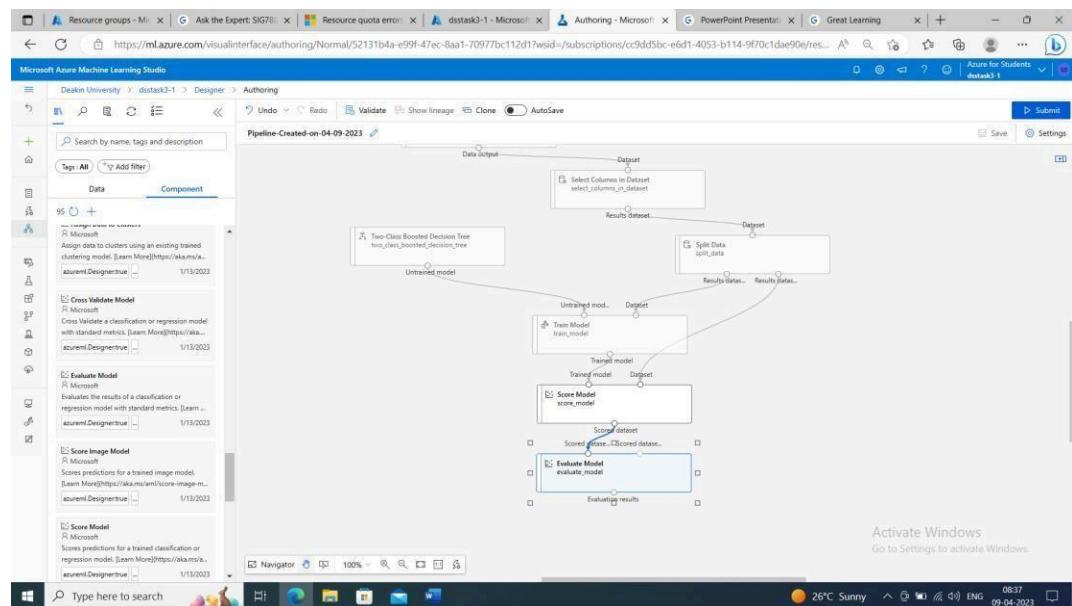
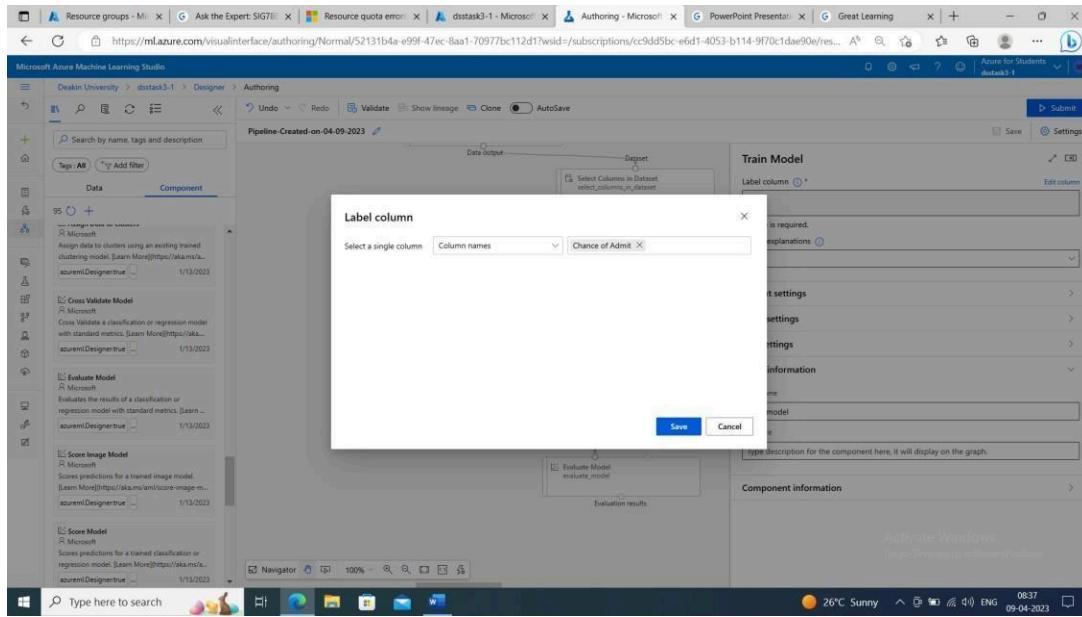
**Output:** The dataset has 400 rows, 8 columns and no missing values.

**Building the Pipeline:** Add data and required components in the pipeline. The model is **two class decision tree model**.









Create Compute Cluster: After the model is saved, a compute cluster ‘senacc’ is created for model submission.

Microsoft Azure Machine Learning Studio

Compute instances Compute clusters Kubernetes clusters Attached computes

Scale your compute cluster from a single node to a multi node workload

Create a single or multi node compute cluster for your training, batch inferencing or reinforcement learning workloads. Learn more

+ New

View Azure Machine Learning tutorials View available quotas

Activate Windows Go to Settings to activate Windows.

26°C Sunny 08:38 09-04-2023

Microsoft Azure Machine Learning Studio

Create compute cluster

Virtual Machine Advanced Settings

Select virtual machine Select the virtual machine size you would like to use for your compute cluster.

Location Central India

Virtual machine tier Dedicated Low priority

Virtual machine type CPU GPU

Virtual machine size Select from recommended options Select from all options

Name	Category	Workload types	Available quota	Cost
Standard_DS11_v2	Memory optimized	Development on Notebooks (or other IDE) and light weight testing	6 cores	\$0.19/hr
Standard_DS2_v2	General purpose	Classical ML, model training on small datasets	6 cores	\$0.34/hr
Standard_DS2_v2	Memory optimized	Data manipulation and training on medium-sized datasets (1-10GB)	6 cores	\$0.30/hr
Standard_F4s_v2	Compute optimized	Data manipulation and training on large datasets (> 10 GB)	16 cores	\$0.17/hr

Activate Windows Go to Settings to activate Windows.

Back Next Download a template for automation Cancel

26°C Sunny 08:39 09-04-2023

**Create compute cluster**

**Configure Settings**

Configure compute cluster settings for your selected virtual machine size.

Name	Category	Cores	Available quota	RAM	Storage	Cost/Node
Standard_DS11_v2	Memory optimized	2	6 cores	14 GB	28 GB	\$0.19/hr

Compute name \*

Minimum number of nodes \*

Maximum number of nodes \*

Idle seconds before scale down \*

Enable SSH access

Advanced settings

Add tags

No tags

Back Create Download a template for automation.

Activate Windows  
Go to Settings to activate Windows.

**Compute**

Compute instances Compute clusters Kubernetes clusters Attached computes

New Refresh Delete Edit columns Reset view View quota

Search

Name	State	Size	Location	Created on	Active runs	Idle nodes	Busy nodes	Unprovisioned nodes
senacc	Succeeded (0 nodes)	STANDARD_DS11_V2	centralindia	Apr 9, 2023 8:39 AM	0	0	0	3

Activate Windows  
Go to Settings to activate Windows.

Kubernetes AKS compute Cluster: The AKS compute cluster ‘senacc’ is created which is used in model deployment.

Microsoft Azure Machine Learning Studio

Compute

Kubernetes clusters

Train or deploy models with your self-managed Kubernetes clusters anywhere

+ New

View Azure Machine Learning tutorials

View available quota

Activate Windows

Go to Settings to activate Windows.

Microsoft Azure Machine Learning Studio

Create AksCompute

Virtual Machine

Select virtual machine

Select the virtual machine size you would like to use for your inference cluster:

Kubernetes Service

Create new  Use existing

Location \*

Select or search by name

Activate Windows

Go to Settings to activate Windows.

Microsoft Azure Machine Learning Studio

Create AksCompute

Virtual Machine

Select virtual machine

Select the virtual machine size you would like to use for your inference cluster:

Kubernetes Service

Create new  Use existing

Location \*

Central India

Add filter

Search by VM name...

Showing 477 VM sizes | Current selection: Standard\_A2\_v2.

Name	Category	Available quota
Standard_A2_v2 2 cores, 4GB RAM, 20GB storage	General purpose	4 cores
Standard_A2m_v2 2 cores, 16GB RAM, 20GB storage	General purpose	4 cores
Standard_A4_v2 4 cores, 8GB RAM, 40GB storage	General purpose	4 cores
Standard_Adv_v2 4 cores, 32GB RAM, 40GB storage	General purpose	4 cores
Standard_A8_v2 8 cores, 16GB RAM, 80GB storage	General purpose	4 cores
Standard_A8m_v2 8 cores, 32GB RAM, 80GB storage	General purpose	4 cores
Standard_B12H8 12 cores, 40GB RAM, 95GB storage	General purpose	4 cores

Activate Windows

Go to Settings to activate Windows.

## Run the Model and Create Pipeline Job:

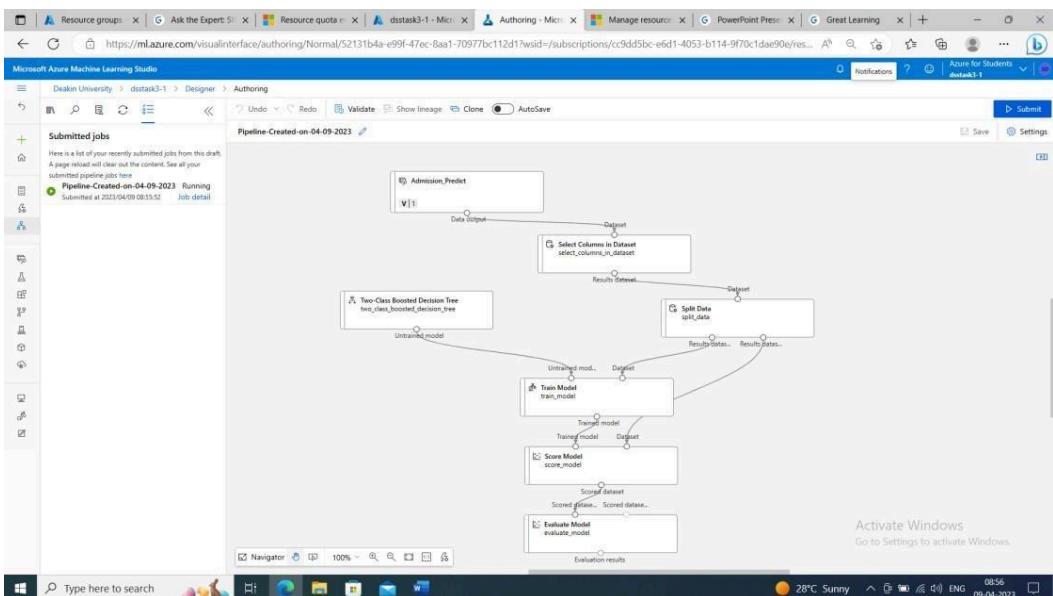
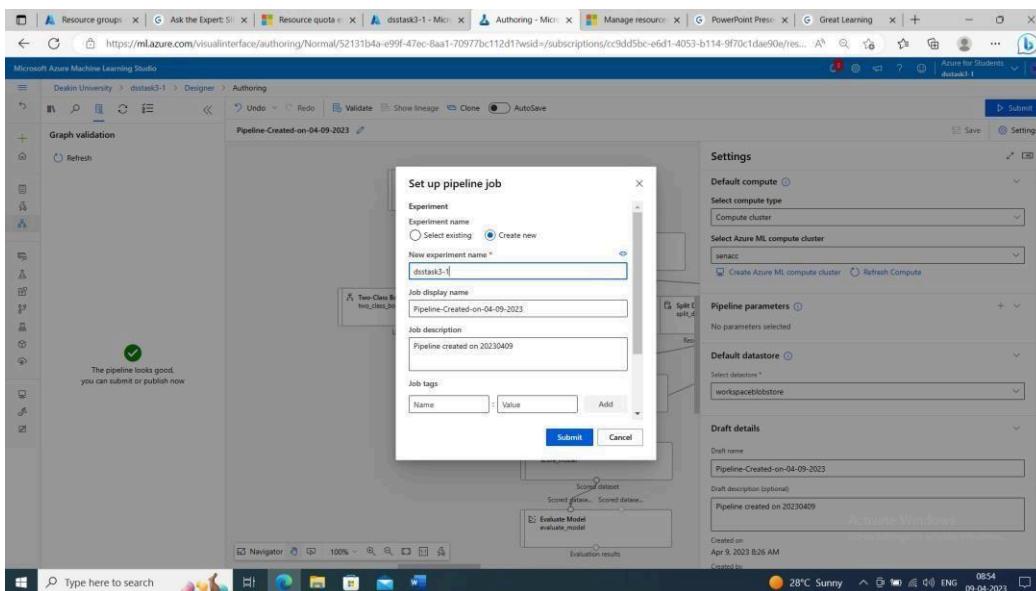
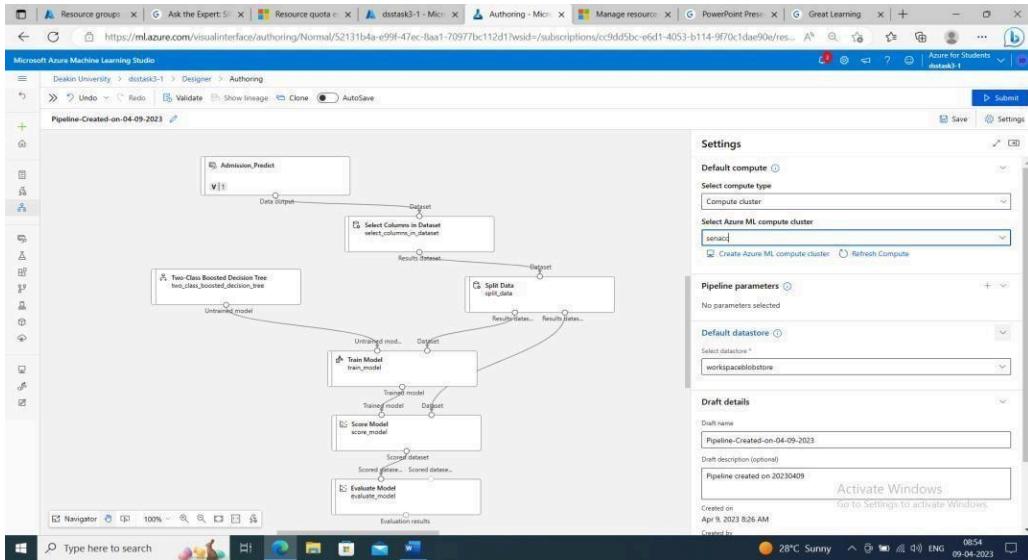
The model selected is Two Class Boosted Decision Tree, which has following advantages and disadvantages.

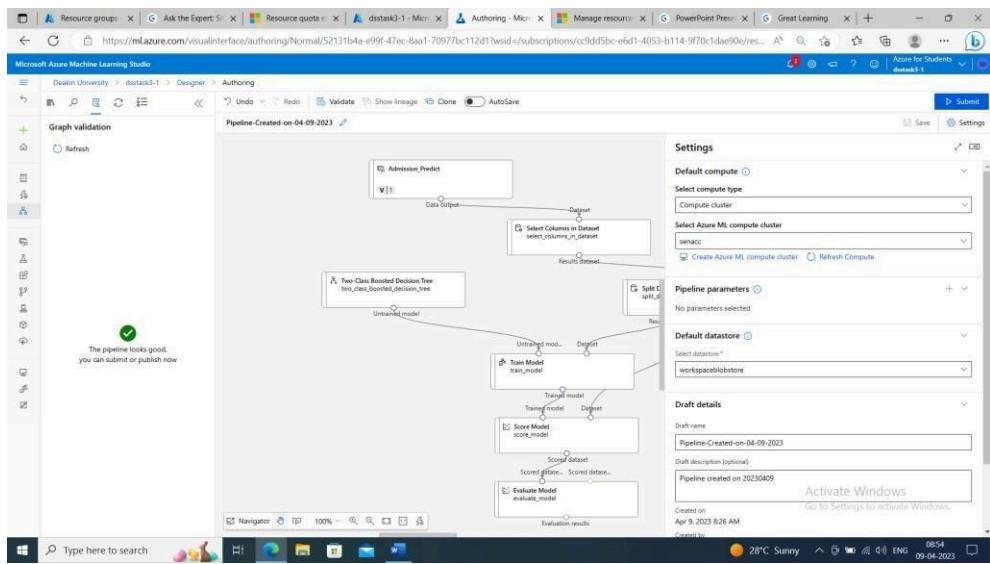
### Advantages of Decision Tree:

- Easy to understand and is robust to outliers and missing values in data.
- Very fast at testing time (Big O (depth of tree)) and highly compact after pruning.

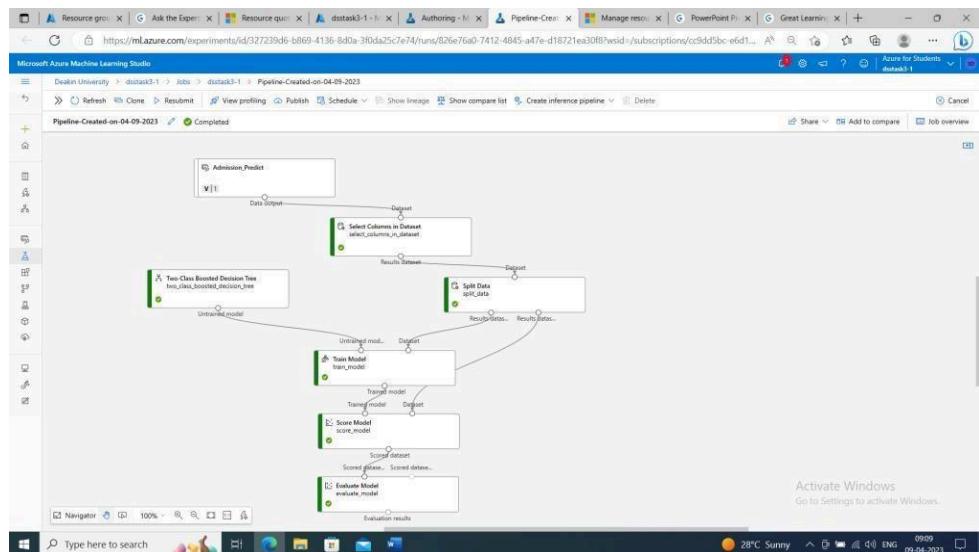
### Disadvantages of Decision Tree:

- Greedy and generates only axis-aligned splits of data.

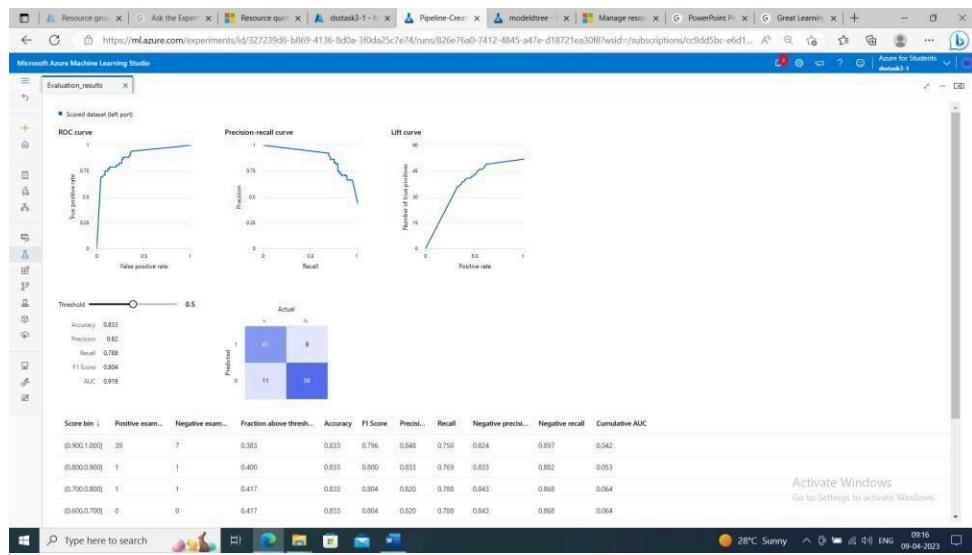




The model is run completely, and all components turn green.



Evaluation Results:



**Accuracy – 0.83, Precision – 0.82, Recall – 0.788, F1 Score – 0.804, AUC – 0.919**

```

PY scorepy x conda_env.yaml x _sample.json x
1 import os
2 import json
3
4 from azureml.studio.core.io.model_directory import ModelDirectory
5 from pathlib import Path
6 from azureml.studio.core.io.model_directory.scorer import generic_module.scorer.generic_module import ScoreModuleModule
7 from azureml.designer.serving.dagengine.converter import create_dfd_from_dict
8 from collections import defaultdict
9 from azureml.designer.serving.dagengine.utils import decode_nan
10 from azureml.designer.serving.dagengine.utils import decode_nan
11
12 model_path = os.path.join(os.getcwd(), 'AZUREML_MODEL_DIR'), 'trained_model_outputs')
13 schema_file_path = Path(model_path) / '_schema_json'
14 with open(schema_file_path) as fp:
15     schema_data = json.load(fp)
16
17 schema_data = json.loads(schema_data)
18
19 def init():
20     global model
21     model = ModelDirectory.load(model_path).model
22
23
24 def run(data):
25     data = json.loads(data)
26     input_entry = defaultdict(list)
27     for row in data:
28         for key, val in row.items():
29             input_entry[key].append(decode_nan(val))
30
31     data_frame_directory = create_dfd_from_dict(input_entry, schema_data)
32
33     score_module = ScoreModuleModule()

```

The trained model component is observed. The Output + Logs are viewed.

Name	Description	Tags	Source
AzureML-ACPI-pytorch-1.1-py38...	Recommended environment for Deep Learning in public preview with PyT...	GPU: Cuda11 OS: Ubuntu20.04 PyTorch: 1.11	Microsoft (curated)
responsibleai-vision-ubuntu20.04...	AzureML Responsible AI Vision environment.	OS: Ubuntu20.04 Preview Training	azurent
responsibleai-ubuntu20.04-py38-cu...	AzureML Responsible AI environment.	OS: Ubuntu20.04 Preview Training	azurent
responsibleai-test-ubuntu20.04-py3...	AzureML Responsible AI test environment.	OS: Ubuntu20.04 Preview Training	azurent
tensorflow-2.9-cuda11-tf	An environment for deep learning with Tensorflow containing the Azure...	GPU: Cuda11 OS: Ubuntu20.04 Preview Pyt	azurent
tensorflow-2.8-cuda11-tf	An environment for deep learning with Tensorflow containing the Azure...	GPU: Cuda11 OS: Ubuntu20.04 Preview Pyt	azurent
mlDesigner-minimal5	CPU based environment for pipelines (ML Designer) minimal version.	OS: Ubuntu20.04 Python3.8 Private Windows	azurent

Create Model: The files \_Sample.Json, Conda-env.Yaml, Score.py are downloaded in the local machine.

The screenshot shows the Microsoft Azure Machine Learning Studio interface. The main window title is "Train Model". The left sidebar lists various outputs: logs, module\_statistics, system\_logs, trained\_model\_outputs (selected), meta.yaml, samples.json (highlighted in blue), schema.json, conda\_env.yaml, datalearner, modelspec.yaml, PY score.py, and user\_logs. The right pane contains three code editors. The top one is titled "PY score.py" and contains the following Python code:

```
1  {
2      "GPA Score": 3.1,
3      "TOEFL Score": 107,
4      "University Rating": 3,
5      "SOP": 4,
6      "LOR": 3,
7      "CGPA": 3.69,
8      "Chance of Admit": 0
9  },
10 },
11 },
12 "GPA Score": 3.0,
13 "TOEFL Score": 105,
14 "University Rating": 2,
15 "SOP": 3,
16 "LOR": 3,
17 "CGPA": 3.69,
18 "Chance of Admit": 0
19 },
20 },
21 "GPA Score": 3.0,
22 "TOEFL Score": 112,
23 "University Rating": 4,
24 "SOP": 4,
25 "LOR": 4.5,
26 "CGPA": 3.8,
27 "Chance of Admit": 1
28 }
```

The middle editor is titled "conda\_env.yaml" and the bottom one is titled "samples.json". At the bottom of the screen, there is a Windows taskbar with icons for Start, Search, Task View, File Explorer, Edge, Mail, Photos, and File Explorer. The system tray shows the date and time as "09-04-2023 09:12", the temperature as "28°C", and the weather as "Sunny".

The model is registered as modeldtree.

Microsoft Azure Machine Learning Studio

Train Model

Overview Parameters Outputs + logs Metrics Child jobs Images

Refresh Register model Debug and monitor Download all Enable log streaming Word wrap

Data outputs Hide data outputs

Trained model +

Other outputs

logs module\_statistics system\_logs trained\_model\_outputs \_metayaml \_samplesjson \_schema.json conda\_env.yaml data.learner model\_spec.yaml PY score.py user\_logs

PY score.py

```
1  [
2   {
3     "GRE Score": 313,
4     "TOEFL Score": 10,
5     "University Rating": 4,
6     "SOP": 4,
7     "LOR": 4.5,
8     "CGPA": 8.69,
9     "Chance of Admit": 0
10    },
11   {
12     "GRE Score": 308,
13     "TOEFL Score": 10,
14     "University Rating": 3,
15     "SOP": 3,
16     "LOR": 3.5,
17     "CGPA": 8.49,
18     "Chance of Admit": 0
19    },
20   {
21     "GRE Score": 328,
22     "TOEFL Score": 118,
23     "University Rating": 4,
24     "SOP": 4,
25     "LOR": 4.5,
26     "CGPA": 9.1,
27     "Chance of Admit": 1
28   }
29 ]
```

Register model

Model name:   
Description (Optional):

Save Cancel

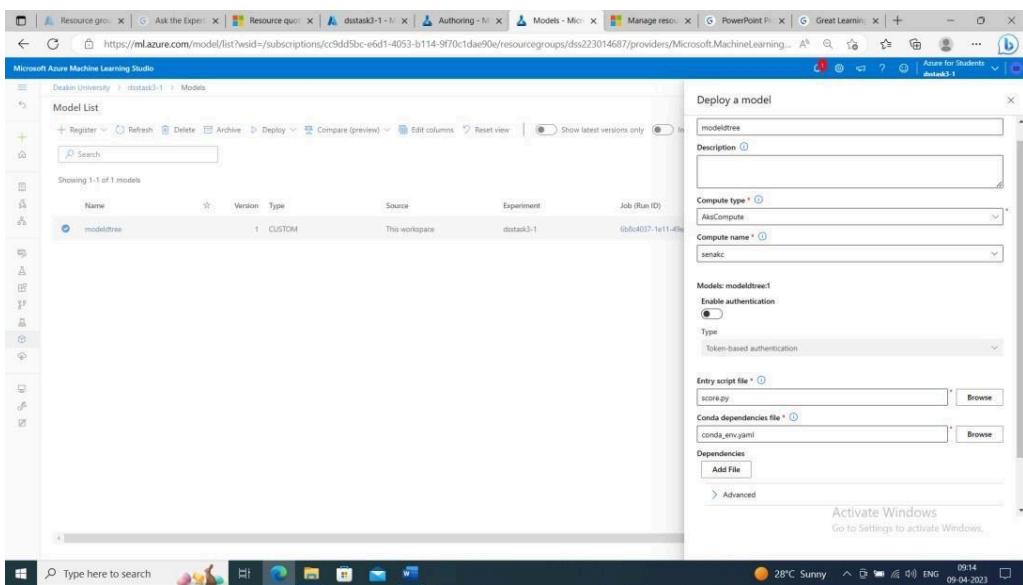
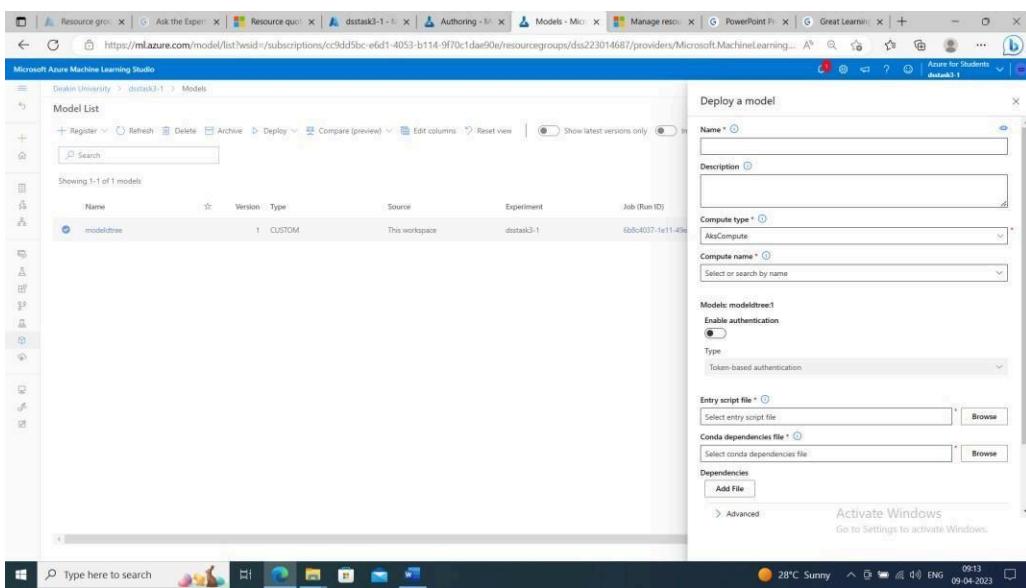
Activate Windows  
Get to Synergy & activate Windows

Type here to search 28°C Sunny ENG 09:13 09-04-2023

The screenshot shows the Microsoft Azure Machine Learning Studio interface. The browser tab is titled "Microsoft Azure Machine Learning Studio". The main content area displays a "Model List" for the workspace "dssatck3-1". A single model named "modeltree" is listed. The table columns include Name, Version, Type, Source, Experiment, Job (Run ID), Created on, Tags, and Properties. The "Created on" column is sorted in descending order. The "Properties" column for the model shows three dots (...). The status bar at the bottom right indicates "Activate Windows Go to Settings to activate Windows".

Name	Version	Type	Source	Experiment	Job (Run ID)	Created on	Tags	Properties
modeltree	1	CUSTOM	This workspace	dssatck3-1	6b8c4037-1e11-4fe4-adff-a54...	Apr 9, 2023 9:13 AM		...

## Model Deployment:



The **modeldtree** is deployed and deployment state changed from Transitioning to Healthy.

This screenshot shows the Microsoft Azure Machine Learning Studio interface. The URL in the address bar is <https://ml.azure.com/endpoints/lists/realmineendpoints/modeldtree/detail?wsid=/subscriptions/c9dd5bc-e6d1-4053-b114-9f70c1dae90e/resourcegroups/ds22301...>. The main content area displays the 'modeldtree' endpoint details. Under 'Endpoint attributes', the 'Deployment state' is shown as 'Transitioning'. After deployment, it would change to 'Healthy'. Other attributes include Service ID (modeldtree), Compute type (AksCompute), Created by (SUCHITRA SENA DANDE), Model ID (modeldtree1), and REST endpoint (<http://20.194.7.210:80/api/v1/service/modeldtree/score>). The 'Properties' section shows 'hasInferenceSchema' as True, 'hasImage' as False, and 'authenticated' as False. A status bar at the bottom right indicates '28°C Sunny' and the date '09-08-2023'.

The REST Endpoint is obtained which is consumed in C#, Python or R [1].

This screenshot shows the Microsoft Azure Machine Learning Studio interface. The URL in the address bar is <https://ml.azure.com/endpoints/lists/realmineendpoints/modeldtree/detail?wsid=/subscriptions/c9dd5bc-e6d1-4053-b114-9f70c1dae90e/resourcegroups/ds22301...>. The main content area displays the 'modeldtree' endpoint details. Under 'Endpoint attributes', the 'Deployment state' is now 'Healthy'. Other attributes remain the same. The 'Properties' section shows 'hasInferenceSchema' as True, 'hasImage' as False, and 'authenticated' as False. A status bar at the bottom right indicates '31°C Sunny' and the date '09-04-2023'.

This screenshot shows the Microsoft Azure Machine Learning Studio interface. The URL in the address bar is <https://ml.azure.com/endpoints/lists/realmineendpoints/modeldtree/consume?wsid=/subscriptions/c9dd5bc-e6d1-4053-b114-9f70c1dae90e/resourcegroup/ds22301...>. The main content area displays the 'Basic consumption info' section, which includes the REST endpoint (<http://20.194.7.210:80/api/v1/service/modeldtree/score>). Below this, the 'Consumption option' section shows 'Consumption types' with tabs for C#, Python, and R. The C# tab is selected. A code editor window shows a C# script for calling the endpoint:

```
1 // This code requires the NuGet package Microsoft.AspNetCore.Client to be installed.
2 // Instructions for doing this in Visual Studio:
3 // Tools -> NuGet Package Manager -> Package Manager Console
4 // Install-Package Microsoft.AspNetCore.Client
5 // .NET Framework 4.7.1 or greater must be used
6
7 using System;
8 using System.Collections.Generic;
9 using System.IO;
10 using System.Net.Http;
11 using System.Net.Http.Headers;
12 using System.Text;
13 using System.Threading.Tasks;
14 using Newtonsoft.Json;
15
16 namespace CallRequestResponseService
17 {
18     class Program
19     {
20         static void Main(string[] args)
21         {
22             InvokeRequestResponseService().Wait();
23         }
24     }
25 }
```

A status bar at the bottom right indicates '31°C Sunny' and the date '09-04-2023'.

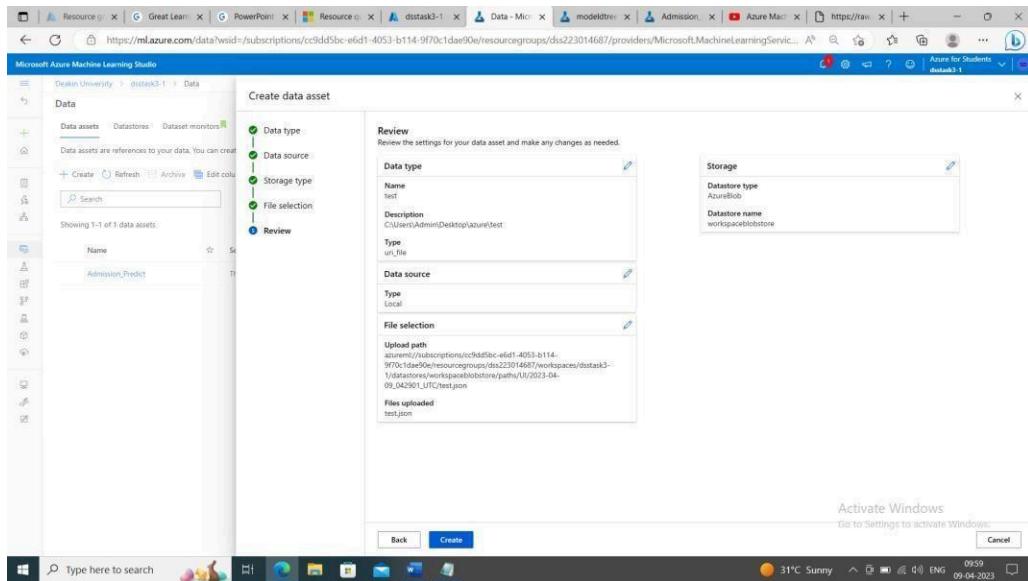
**Rest Endpoint URL:** <http://20.219.47.210:80/api/v1/service/modeldtree/score>

Q2: You need to use Azure Machine learning Python SDK to train and deploy your best built model (Decision tree or Random Forest) in part1. To complete this task, you need to login to your Azure Portal and create your workspace for machine learning and then train your model (you can use the model you built in week 2 or train a new model using your selected dataset) and then deploy the model on Azure. You need to provide a screen shot of your built workspace, the built model using Python and how you have deployed the model using Azure Python SDK and dependencies. Every step of the development and deployment should be explained.

Solution:

Data Ingestion:

The screenshot shows the Microsoft Azure Machine Learning Studio interface. A modal window titled "Create data asset" is open, divided into four main sections: Data type, Data source, Storage type, and File selection. The "Storage type" section is currently active, with "Azure Blob Storage" selected. Below it, a list of existing datastores is shown, including "workspaceblobstore" (storage name: dsstask313612875439, created on: Apr 9, 2023 8:25 AM), "workspaceartifactstore" (storage name: dsstask313612875439, created on: Apr 9, 2023 8:25 AM), and "azurerm" (storage name: dsstask313612875439, created on: Apr 9, 2023 9:41 AM). The "File selection" section shows a file named "test.json" has been uploaded to the "azurerm" datastore. The "Information" section provides details about supported file types (CSV, JSON Lines, Plain Text) and the upload process. At the bottom of the modal, there are "Back" and "Next" buttons, along with an "Activate Windows" watermark.



Building Two-class Decision Forest Pipeline and Run it using Compute Cluster.

Microsoft Azure Machine Learning Studio

Designer

New pipeline

Classic prebuilt Custom

This low-code option uses existing prebuilt components and earlier dataset types (tabular, file) and is best suited for data processing and traditional machine learning tasks like regression and classification. This option continues to be supported but will not have any new components added.

Pipelines

Pipeline drafts Pipeline jobs

Search

View my only Created by All filters Clear all

Select a pipeline from list to preview

Showing 1-2 of 2 pipeline drafts

Name	Pipeline type	Updated on	Created by
Pipeline-Created-on-04-09-2023	N/A	Apr 9, 2023 10:10 AM	SUCHTRA SE...
Pipeline-Created-on-04-09-2023	Training	Apr 9, 2023 8:55 AM	SUCHTRA SE...

Page size 10

Activate Windows Go to Settings to activate Windows.

31°C Sunny 10:10 ENG 09-04-2023

Microsoft Azure Machine Learning Studio

Designer University > dstack3-1 > Designer > Authoring

Search by name, tags and description

Top AB Add filter

Data Component

95 +

Create an averaged perceptron binary classification model. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Boosted Decision Tree A Microsoft Creates a binary classifier using a boosted decision tree algorithm. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Decision Forest A Microsoft Creates a two-class classification model using the decision forest algorithm. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Logistic Regression A Microsoft Creates a two-class logistic regression model. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Neural Network A Microsoft Creates a binary classifier using a neural network algorithm. [Learn More] [https://aka.ms/autorem/Designer...]

Authoring

Pipeline-Created-on-04-09-2023

Undo Redo Validate Show imageage Clone AutoSave

Select Columns in Dataset

Select columns Select columns with rules By name

All available columns Selected columns

1. Columns Add all Remove all

Serial No. GRE Score TOEFL Score University Rating SOP LOR CGPA

Save Cancel

Component information

Activate Windows Go to Settings to activate Windows.

31°C Sunny 10:12 ENG 09-04-2023

Microsoft Azure Machine Learning Studio

Designer University > dstack3-1 > Designer > Authoring

Search by name, tags and description

Top AB Add filter

Data Component

95 +

Create an averaged perceptron binary classification model. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Boosted Decision Tree A Microsoft Creates a binary classifier using a boosted decision tree algorithm. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Decision Forest A Microsoft Creates a two-class classification model using the decision forest algorithm. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Logistic Regression A Microsoft Creates a two-class logistic regression model. [Learn More] [https://aka.ms/autorem/Designer...]

Two-Class Neural Network A Microsoft Creates a binary classifier using a neural network algorithm. [Learn More] [https://aka.ms/autorem/Designer...]

Authoring

Pipeline-Created-on-04-09-2023

Undo Redo Validate Show imageage Clone AutoSave

Split Data

Splitting mode Split Rows Fraction of rows in the first output dataset 0.7 Randomized split True Random seed 0 Stratified split False

Output settings

Input settings

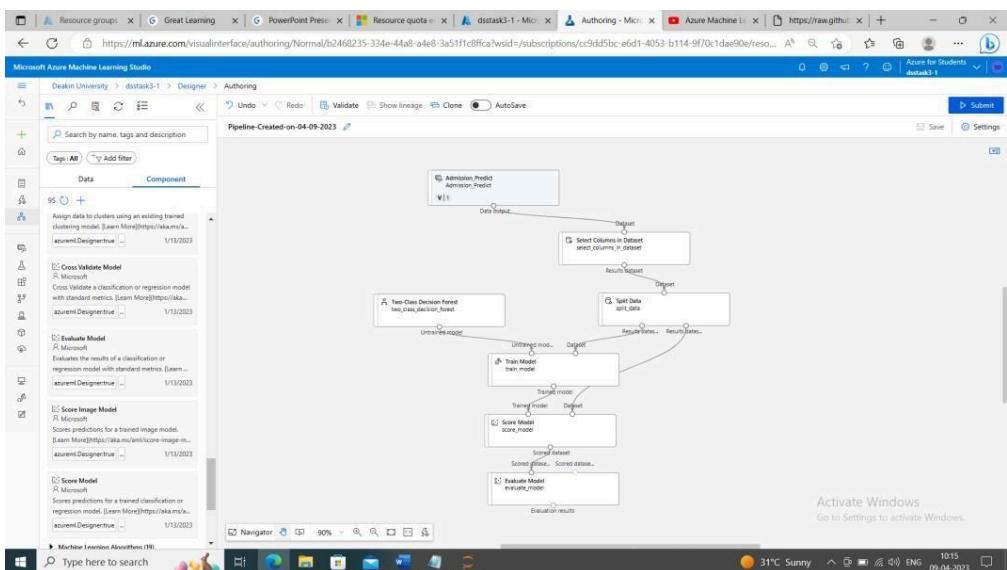
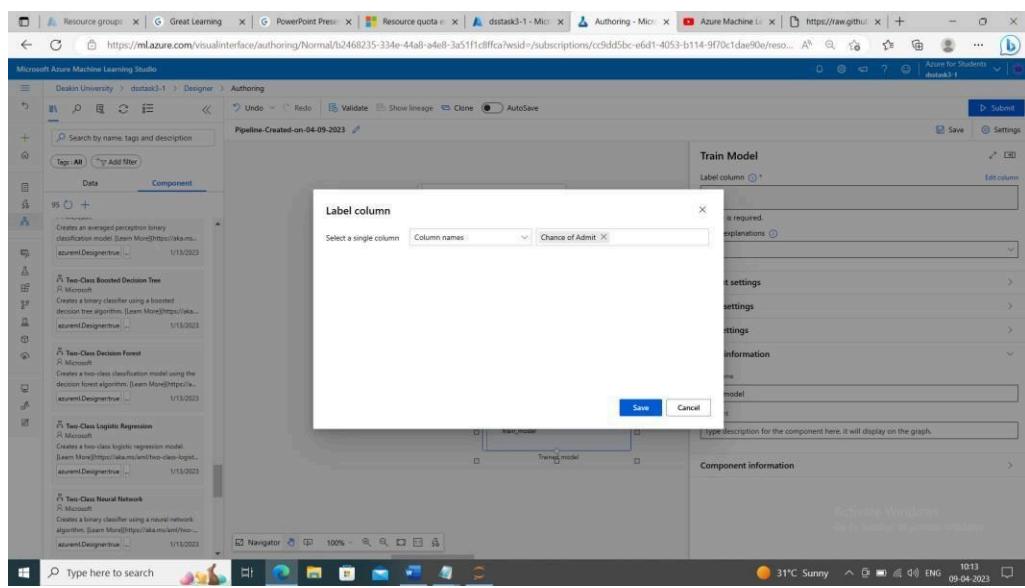
Run settings

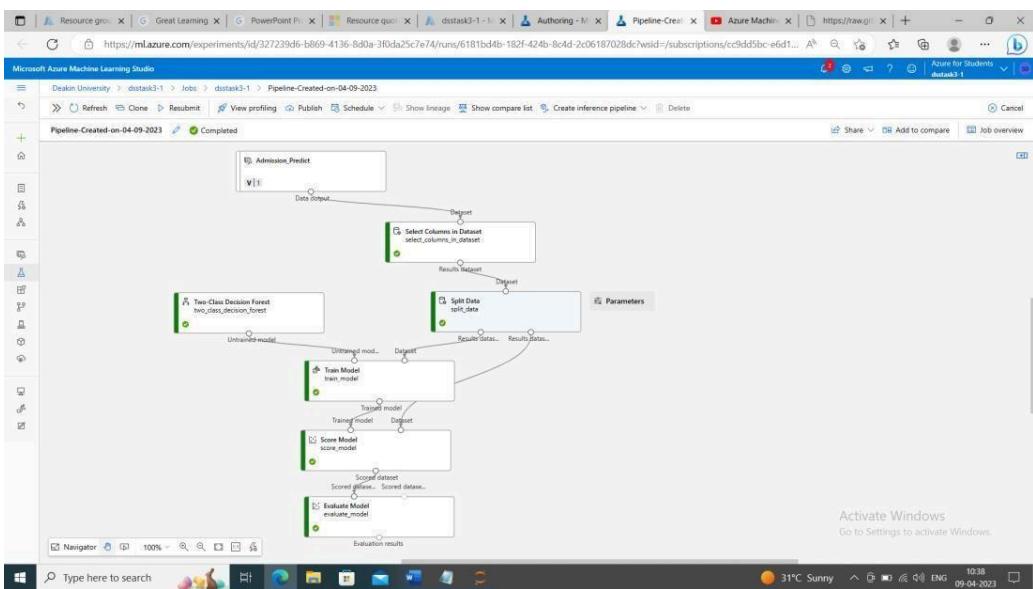
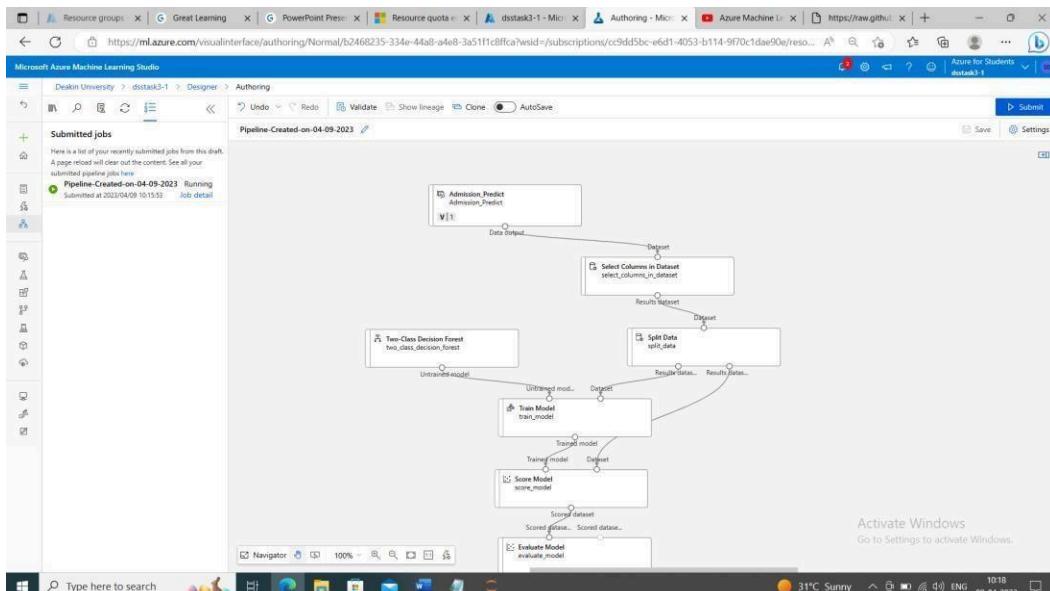
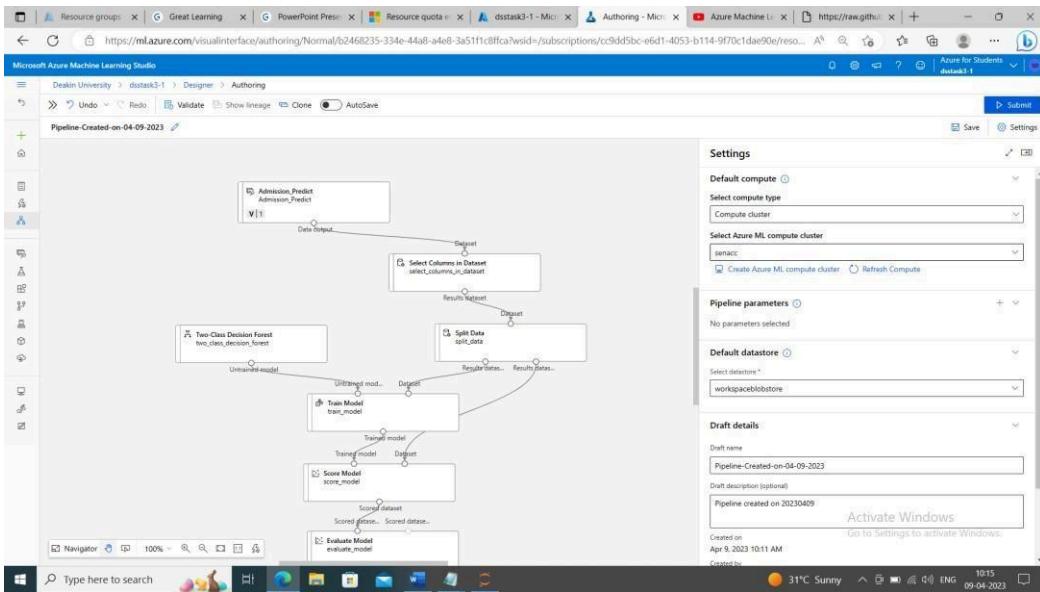
Node information Node name split\_data Comment Description for the component here. It will display on the graph. Go to Settings to activate Windows.

Component information

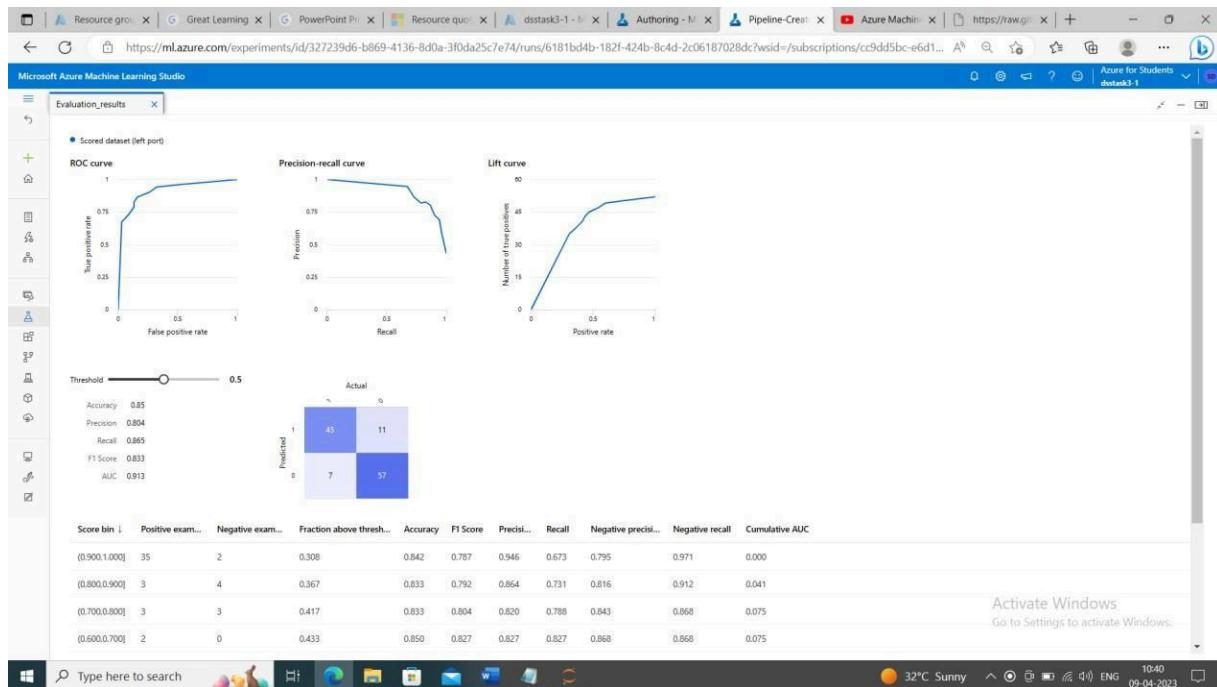
Activate Windows Go to Settings to activate Windows.

31°C Sunny 10:12 ENG 09-04-2023





## Evaluation Results for Train Data:



**Accuracy: 0.85, Precision: 0.804, Recall: 0.865, F1 Score: 0.833, AUC: 0.913**

The trained model component is observed. The Output + Logs are viewed, and we download the files.

Create Model: The files - Sample.Json, Conda-env.Yaml, Score.py are downloaded in the local machine.

The screenshot shows the Microsoft Azure Machine Learning Studio interface. The title bar has tabs for Resource group, Great Learning, PowerPoint Pres, Resource quota, dss task3-1 - M, Authoring - M, Pipeline-Creat..., Azure Machine L, and https://raw.githubusercontent.com/... . The main area is titled "Train Model". Under "Outputs + logs", there are tabs for Overview, Parameters, Outputs + logs, Metrics, Child jobs, Images, Code, Explanations (preview), Fairness (preview), and Monitoring. The "Outputs + logs" tab is selected. It shows a list of outputs: std\_log.txt, \_samples.json, PY score.py, and conda\_env.yaml. The "conda\_env.yaml" file is selected and its contents are displayed in a code editor window:

```

1 name: project_environment
2 channels:
3   - conda-forge
4 dependencies:
5   - python3.8.10
6   - pip=20.2
7   - pip:
8     - azureml-defaults
9     - azureml-designer-classic-modules==0.0.177
10    - azureml-designer-serving==0.0.12
11

```

The bottom right corner of the screen shows a Windows taskbar with icons for Start, Search, Task View, File Explorer, Edge, Mail, Photos, and Settings, along with system status like 32°C Sunny, ENG, and 09-04-2023.

The model is created as ‘dsstaskrf’, which is then deployed using Aks Compute cluster.

The screenshot shows the Microsoft Azure Machine Learning Studio interface. The title bar indicates the URL is <https://ml.azure.com/model/list?wsid=...>. The main content area is titled "Model List" and shows a table of two models:

Name	Version	Type	Source	Experiment	Job (Run ID)	Created on	Tags	Properties
dstatkef	1	CUSTOM	This workspace	dstatk3-1	503f159c-862f-4c2c-bd1c-02a9...	Apr 9, 2023 10:05 AM	...	
modelstree	1	CUSTOM	This workspace	dstatk3-1	60b04037-1e11-49e4-ad00-a34...	Apr 9, 2023 9:13 AM	...	

The interface includes a search bar, filter options (Registry, Created on, Created by, Tags), and a toolbar with various icons. The status bar at the bottom shows "Activate Windows" and the system tray shows the date and time as "09/04/2023".

## Model Deployment as Web Service:

This model is not deployed to any endpoints

Deploy model

Name: dsstaskrf

Description:

Compute type: AksCompute

Compute name: senakc

Model: dsstaskrf1

Enable authentication:

Type: Key-based authentication

Entry script file: score\_rtpy

Conda dependencies file: conda\_env\_rf.yaml

Dependencies: Add File

dsstaskrf

Details Test Consume Deployment logs

Endpoint attributes

Service ID: dsstaskrf

Description: ..

Deployment state: Healthy

Compute type: AksCompute

Created by: SUCHITRA SENA DANDE

Model ID: dsstaskrf1

Created on: Apr 9, 2023 10:47 AM

Last updated on: Apr 9, 2023 10:47 AM

Compute target: senakc

Image ID: ..

REST endpoint: http://20.219.47.210:80/api/v1/service/dsstaskrf/score

Key-based authentication enabled: true

Token-based authentication enabled: false

Tags: No data

Properties

hasInferenceSchema: True

hasHttps: False

authEnabled: True

Activate Windows

Go to Settings to activate Windows.

The Rest End Point URL is obtained, which is used in C#, Python or R [1].

The screenshot shows the Azure Machine Learning Studio interface. In the center, there's a card for the 'dsstaskrf' endpoint. The 'Basic consumption info' section displays the REST endpoint URL as <http://20.219.47.210:80/api/v1/service/dsstaskrf/score>. Below it, the 'Consumption types' section contains code snippets for C#, Python, and R. The Python code is as follows:

```
19 body = str.encode(json.dumps(data))
20
21 url = "http://20.219.47.210:80/api/v1/service/dsstaskrf/score"
22 # Replace this with the primary/secondary key or API Token for the endpoint
23 api_key = ''
24 if not api_key:
25     raise Exception("A key should be provided to invoke the endpoint")
26
27 headers = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + api_key}
28
29 req = urllib.request.Request(url, body, headers)
30
31 try:
32     response = urllib.request.urlopen(req)
```

At the bottom right of the interface, there's a message: 'Activate Windows' and 'Go to Settings to activate Windows.'

Rest Endpoint URL: <http://20.219.47.210:80/api/v1/service/dsstaskrf/score>

Hyperparameter Tuned Random Forest Model:

The screenshot shows the Azure Machine Learning Studio interface with the 'HyperTuned DTree' model selected. The 'Designer' tab is active, displaying a flowchart of the model architecture. The flow starts with an 'Admission\_Pred' component, which feeds into a 'Two-Class Decision Forest' component. This is followed by a 'Your Model Hyperparameters' component, which then splits into 'Train' and 'Test' paths. The 'Train' path leads to a 'Training' component, and the 'Test' path leads to a 'Score Model' component. Finally, the 'Score Model' component outputs 'Scored Data' and 'Scored Labels'. The interface includes various toolbars and a status bar at the bottom.

Microsoft Azure Machine Learning Studio

Task 3 - P: SIG788 - Engineering / HyperTuned DTree - Microsoft A / Authoring - Microsoft Azure M / +

https://ml.azure.com/visualinterface/authoring/Normal/8fbcd9a1-45c5-4f71-90c1-4f1c5bbabdc?wsid=/subscriptions/c9dd5bc-e6d1-4053-b114-9f70c1d... A? Not syncing ... b) Azure for Students SenaTask3-1

Search by name, tags and description

Tags: All Add filter

Data Component

1 + Last update... v9 Undo Redo Validate Show lineage Clone AutoSave

HyperTuned RandomForest

Admission\_Predict Data Output

Select Columns in Dataset

Two-Class Decision Forest

Number of decision trees: 8

Maximum depth of the decision trees: 32

Minimum number of samples per leaf node: 1

Resampling method: Bagging Resampling

Output settings

Input settings

Run settings

Node information

Activate Windows

Node name: two\_class\_decision\_forest

Microsoft Azure Machine Learning Studio

Task 3 - P: SIG788 - Engineering / HyperTuned DTree - Microsoft A / Authoring - Microsoft Azure M / +

https://ml.azure.com/visualinterface/authoring/Normal/8fbcd9a1-45c5-4f71-90c1-4f1c5bbabdc?wsid=/subscriptions/c9dd5bc-e6d1-4053-b114-9f70c1d... A? Not syncing ... b) Azure for Students SenaTask3-1

Graph validation Refresh

HyperTuned RandomForest

The pipeline looks good, you can submit or publish now

Set up pipeline job

Experiment: Select existing

Existing experiment: dtask3-1

Job display name: HyperTunedRandomForest

Job description: HyperTuned RandomForest

Job tags: Name: Value: Add

Submit Cancel

Two-Class Decision Forest

Create trainer mode: SingleParameter

Number of decision trees: 8

Maximum depth of the decision trees: 32

Minimum number of samples per leaf node: 1

Resampling method: Bagging Resampling

Output settings

Input settings

Run settings

Node information

Activate Windows

Node name: two\_class\_decision\_forest

Microsoft Azure Machine Learning Studio

Task 3 - P: SIG788 - Engineering / HyperTuned DTree - Microsoft A / Authoring - Microsoft Azure M / +

https://ml.azure.com/visualinterface/authoring/Normal/8fbcd9a1-45c5-4f71-90c1-4f1c5bbabdc?wsid=/subscriptions/cc9dd5bc-e6d1-4053-b114-9f70c1d... A? Not syncing ... b) Azure for Students SenaTask3-1

Search by name, tags and description

Tags: All Add filter

Data Component

1 + Last update... v9 Undo Redo Validate Show lineage Clone AutoSave

HyperTuned RandomForest

Admission\_Predict Data Output

Select Columns in Dataset

Two-Class Decision Forest

Number of decision trees: 8

Maximum depth of the decision trees: 32

Minimum number of samples per leaf node: 1

Resampling method: Bagging Resampling

Output settings

Input settings

Run settings

Node information

Activate Windows

Node name: two\_class\_decision\_forest

Specify parameter sweeping mode: Random sweep

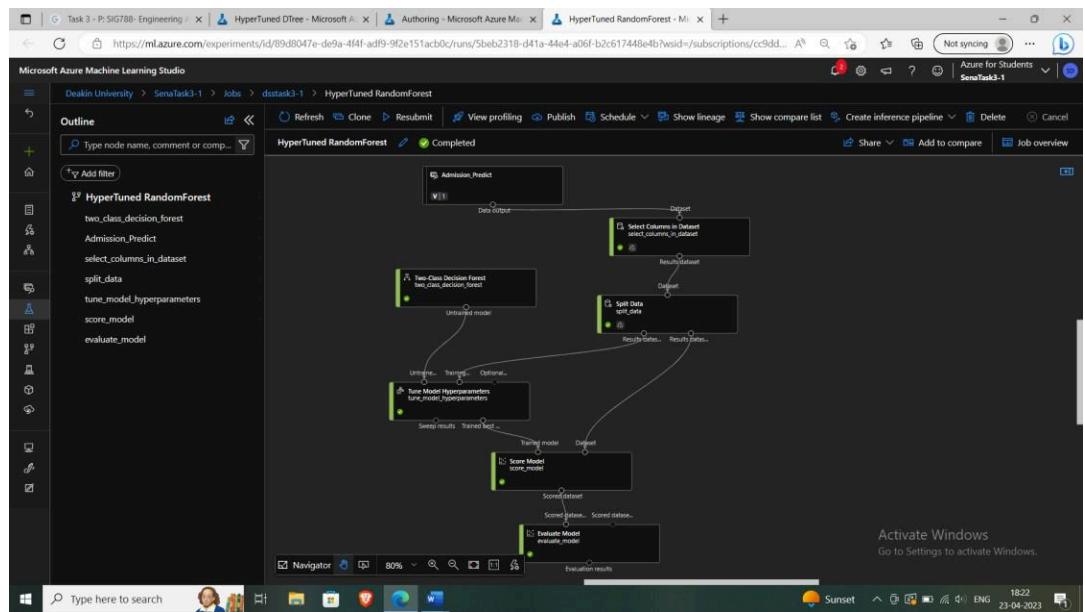
Maximum number of runs on random sweep: 5

Random seed: 1

Label column: Column names: Chance of Admit

Metric for measuring performance for classification: Precision

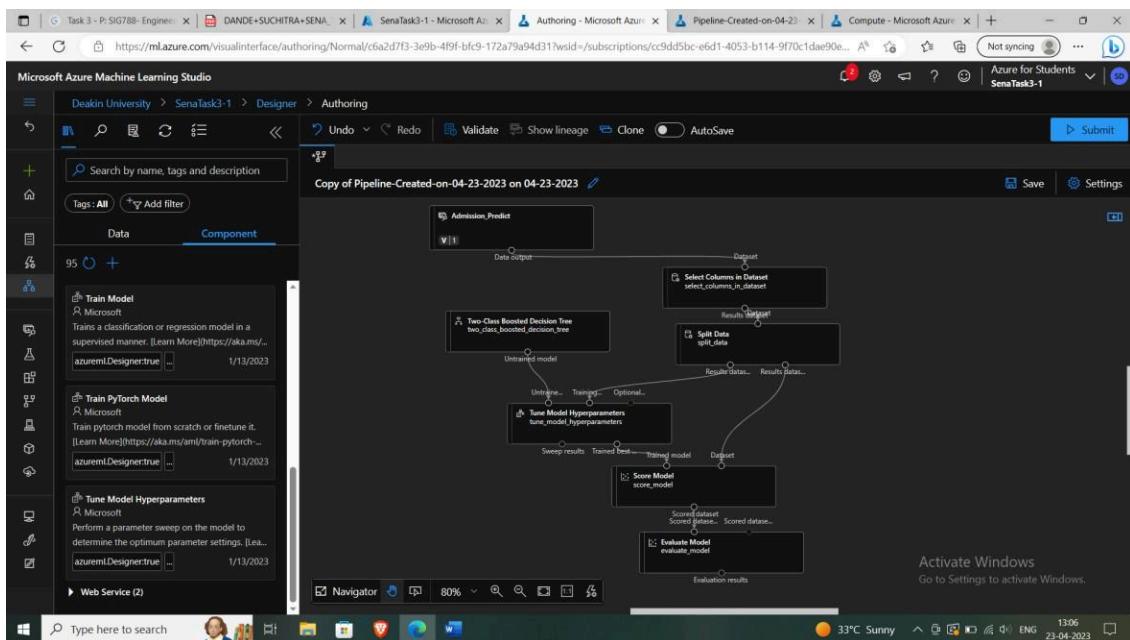
Metric for measuring performance for regression: Mean absolute error

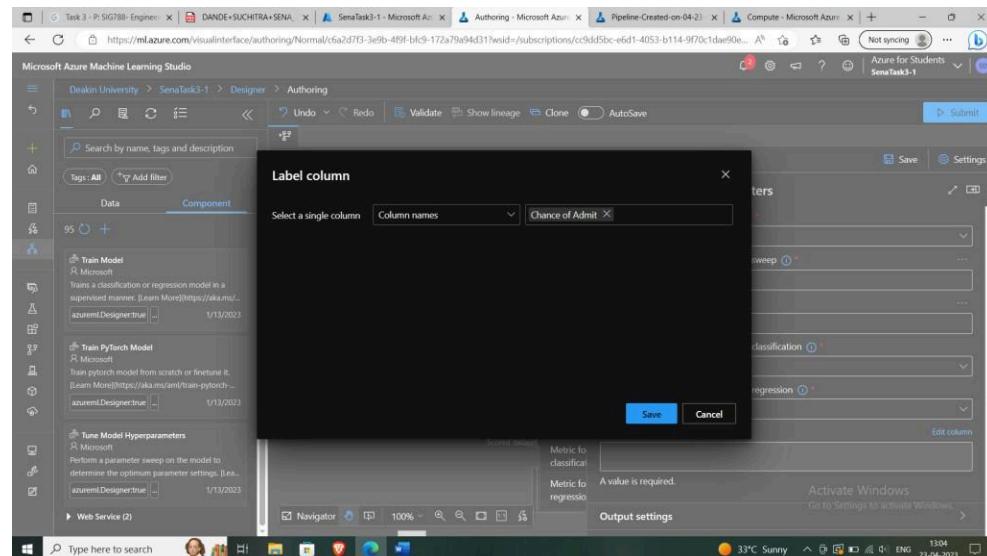
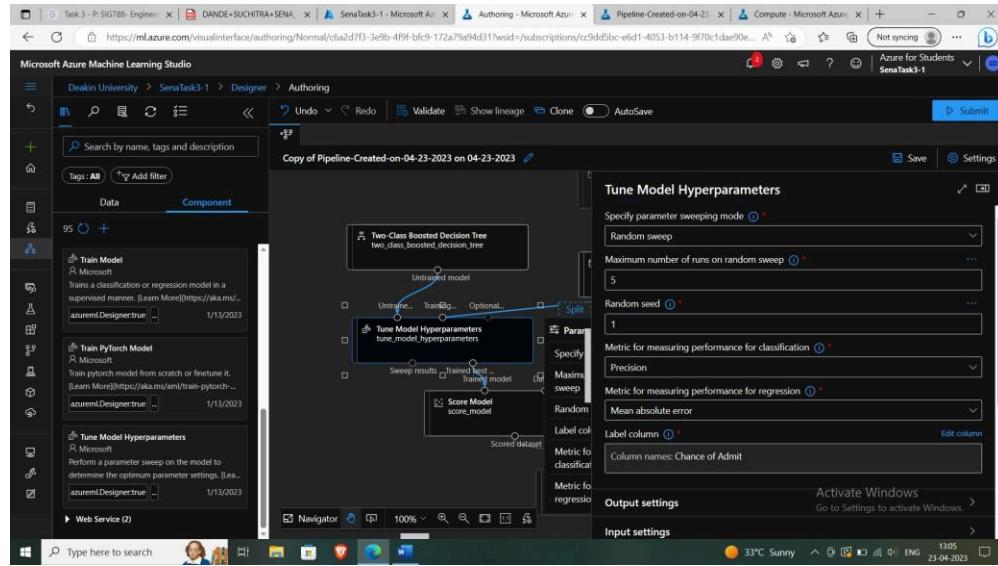


**Accuracy - 0.867, Precision - 0.833, Recall - 0.865, F1 Score - 0.849 and AUC - 0.953**

## Hyperparameter Tuned Decision Classifier Model:

The pipeline of the Hyperparameter Tuned Decision Tree model is shown below.





The label column is Chance of Admit with no. of epochs = 5 and random seed = 1. The optimal classification parameter is set to precision.

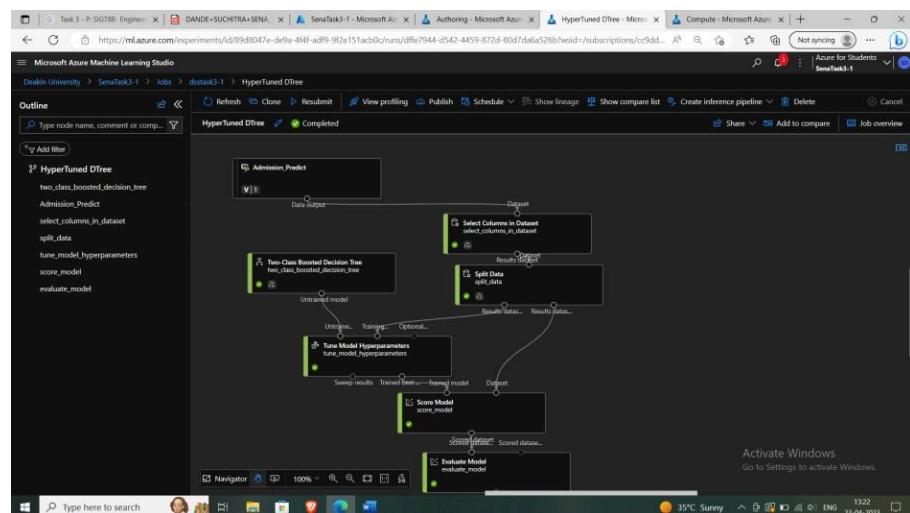
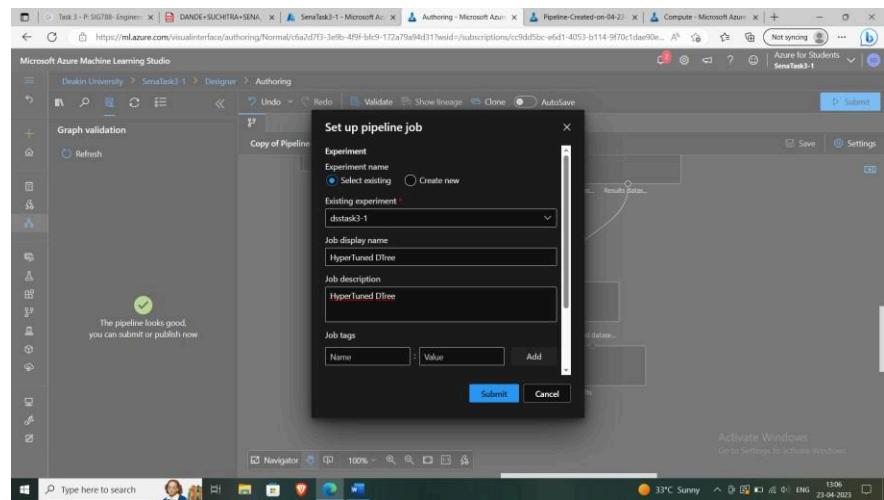
The model is then submitted to pipeline job which is run to get the evaluation metrics.

#### Advantages of Random Forest Model:

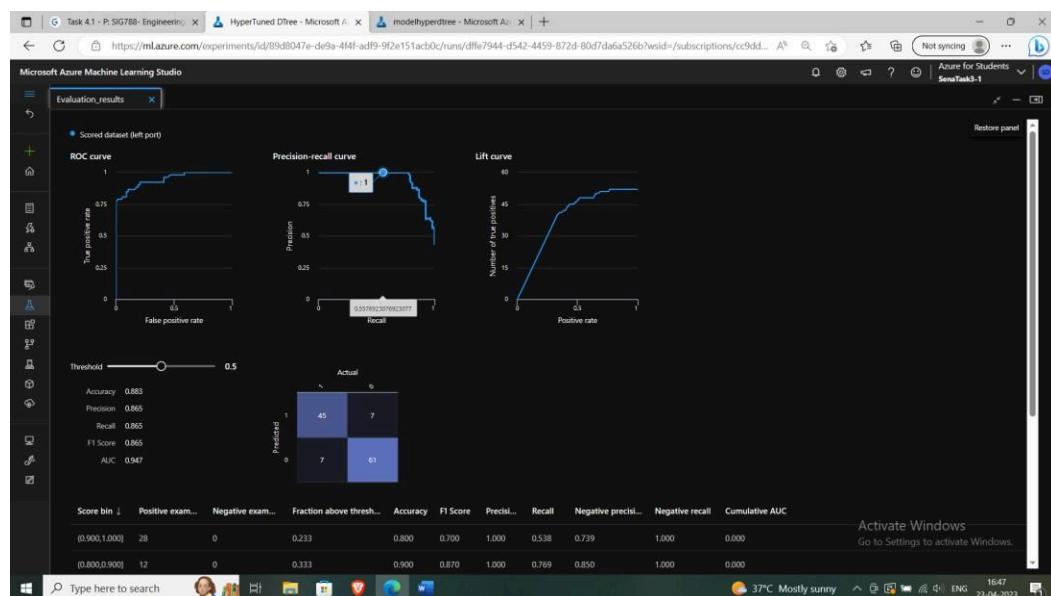
- Perform feature selection implicitly to generate decision trees, which are not correlated.
- Provide higher accuracy by balancing bias-variance trade off.
- Outliers have little influence on random forests. Scaling the data is not required.

#### Disadvantages of Random Forest:

- Can't be interpreted easily and are intensive computationally for large datasets.
- Little control on the model by the user.



The pipeline turns green indicating that the model is run successfully.  
**Output Reuslts and Logs:**



**Accuracy - 0.883, Precision - 0.865, Recall - 0.865, F1 Score - 0.865 and AUC - 0.947**

## Model Comparison:

Model	Accuracy	Precision	Recall	F1 - Score	AUC
Two-Class Boosted Decision Tree	0.83	0.82	0.788	0.804	0.919
Two-Class Decision Forest	0.85	0.804	0.865	0.833	0.913
Hyperparameter Tuned Two-Class Decision Forest	0.867	0.833	0.865	0.849	0.953
Hyperparameter Tuned Two-Class Boosted Decision Tree	0.883	0.865	0.865	0.865	0.947

The best model obtained is **Hyperparameter Tuned Decision Tree** which is registered and deployed as web service.

```

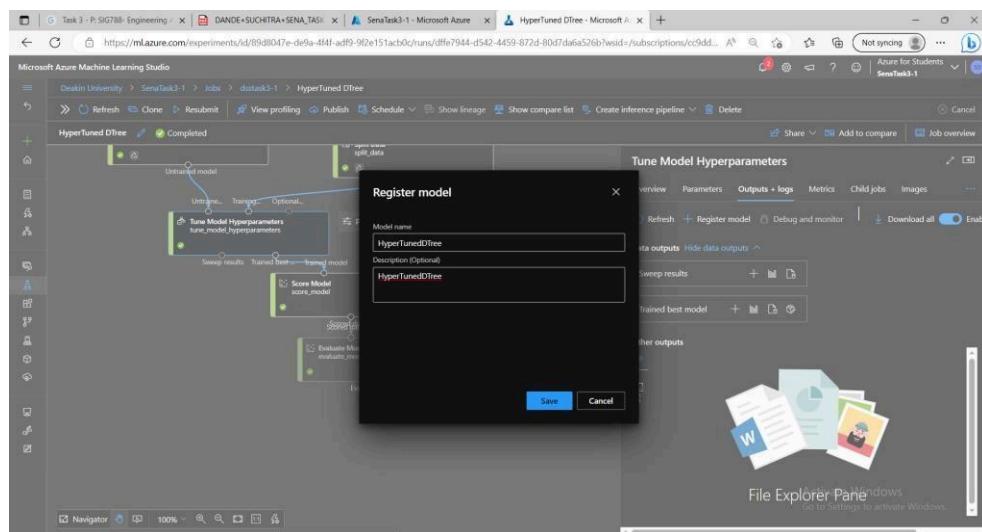
PY score.py
O_schema.json
O_samples.json

1 import os
2 import json
3
4 from azureml.studio.core.io.model_directory import ModelDirectory
5 from pathlib import Path
6 from azureml.studio.modules.ml.score.score_generic_module import ScoreModuleModule
7 from azureml.designer.serving.dagengine.converter import create_dfd_from_dict
8 from collections import defaultdict
9 from azureml.designer.serving.dagengine.utils import decode_name
10 from azureml.studio.common.datatable.DataTable import DataTable
11
12
13 model_path = os.path.join(os.getenv('AZUREML_MODEL_DIR'), 'trained_model_outputs')
14 schema_file_path = Path(model_path) / 'schema.json'
15 with open(schema_file_path) as fp:
16     schema_data = json.load(fp)
17
18
19 def init():
20     global model
21     model = ModelDirectory.load(model_path).model
22
23
24 def run(data):
25     data = json.loads(data)
26     input_entry = defaultdict(list)
27     for row in data:

```

The model is then registered as HyperTunedDTree. Advantages of Hyperparameter Tuning:

- More efficient (better evaluation metrics)



Microsoft Azure Machine Learning Studio

Model List

Showing 1-1 of 1 models

Name	Version	Type	Source	Experiment	Job (Run ID)	Created on	Tags
HyperTunedDTree	1	CUSTOM	This workspace	dtask3-1	0f0d48eb-05cc-4ae7-8736-71d...	Apr 23, 2023 1:29 PM	

Finally, the registered model is deployed as web service.

Microsoft Azure Machine Learning Studio

Model List

Showing 1-1 of 1 models

Deploy a model

Name: HyperTunedDTree

Description:

Compute type: AksCompute

Compute name: aksaks

Models: HyperTunedDTree:1

Enable authentication:

Type: Token-based authentication

Entry script file: score\_hyperDTree.py

Conda dependencies file: conda\_env\_hyperDTree.yaml

Microsoft Azure Machine Learning Studio

Model List

Showing 1-1 of 1 models

Deploy a model

Name: modelhyperDtree

Description:

Compute type: AksCompute

Compute name: semtask3-1

Models: HyperTunedDTree:1

Enable authentication:

Type: Token-based authentication

Entry script file: score\_hyperDTree.py

Conda dependencies file: conda\_env\_hyperDTree.yaml

The model `HyperTunedDTree` is being deployed. Deployment state: `Transitioning`. Compute target: `senakc`.

Compute name: `senakc`

Models: `HyperTunedDTree1`

Type: Token-based authentication

Entry script file: `score_hyperDtree.py`

Conda dependencies file: `conda_env_hyperDtree.yaml`

Dependencies: Add File

Deploy Cancel

Activate Windows  
Go to Settings to activate Windows.

The model `HyperTunedDTree` is deployed and deployment state changes from transitioning to healthy.

Deployment state: `Healthy`

Tags: No data

Properties: No data

REST endpoint: `http://20.204.203.177:80/api/v1/service/modelhyperdtree/score`

Activate Windows  
Go to Settings to activate Windows.

Deployment state: `Healthy`

Tags: No data

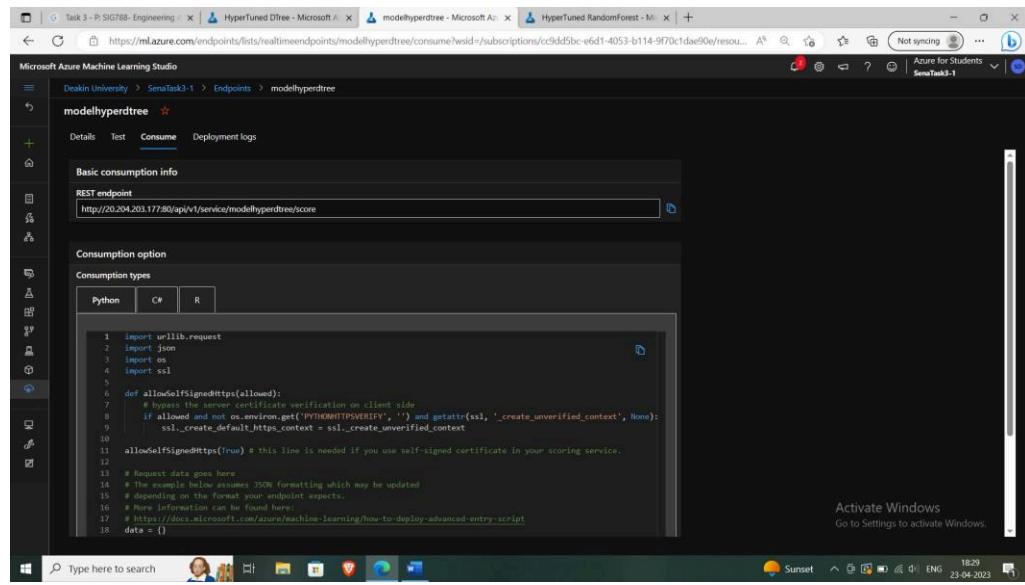
Properties:

- `hasInferenceSchema`: `True`
- `hasHttps`: `False`
- `authenticated`: `False`

REST endpoint: `http://20.204.203.177:80/api/v1/service/modelhyperdtree/score`

Activate Windows  
Go to Settings to activate Windows.

The REST Endpoint URL is obtained which is consumed in C#, Python or R [1].



Rest Endpoint: <http://20.204.203.177:80/api/v1/service/modelhyperdtree/score>

The best model obtained is **Hyperparameter Tuned Decision Tree** model with the rest endpoint URL <http://20.204.203.177:80/api/v1/service/modelhyperdtree/score>