

Lab 37

Classroom Activity: Getting started with Azure ML Designer

What is Azure ML Designer?

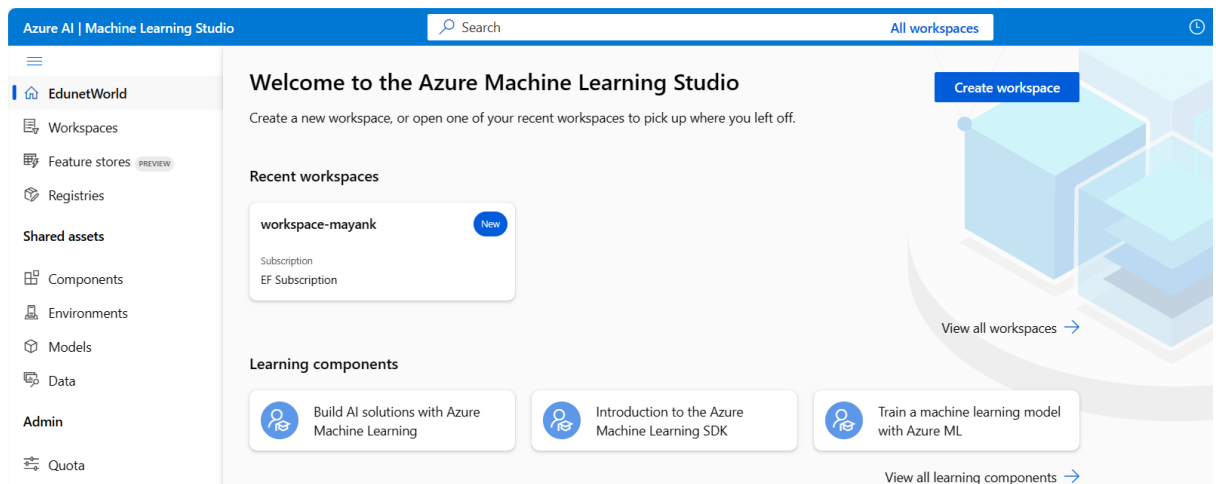
- Azure ML Designer is a visual interface for building, testing, and deploying machine learning models on Azure.
- It's designed for data scientists, analysts, and developers to collaborate and create ML pipelines without writing code.
- With a drag-and-drop interface, it simplifies the ML workflow and accelerates model development.

Key Features

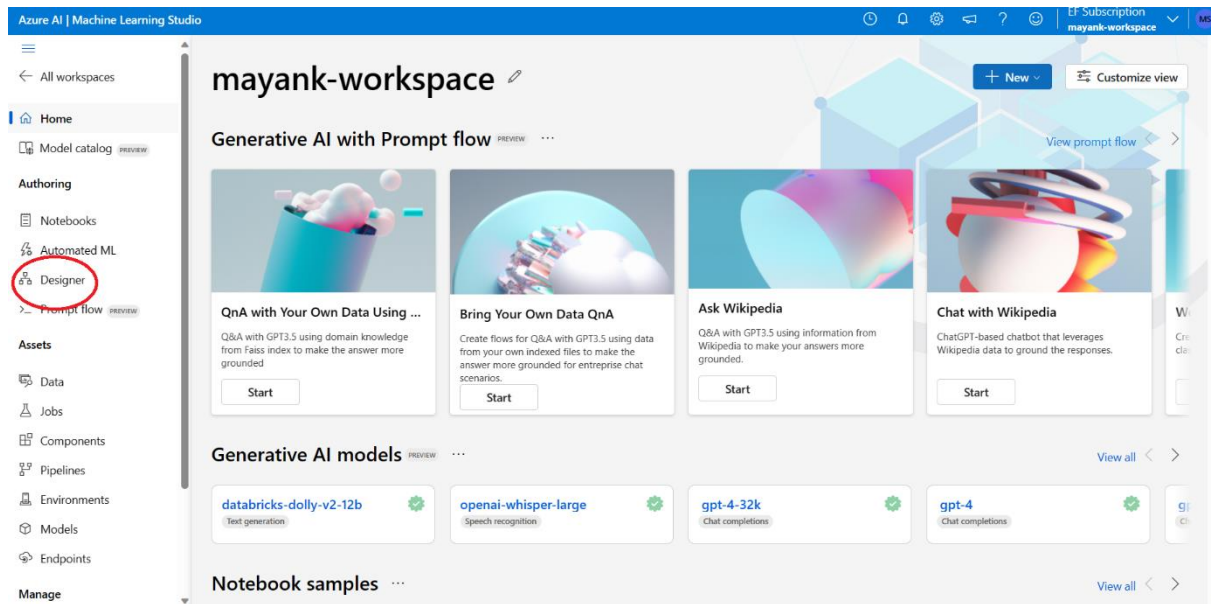
- Drag-and-Drop Interface: Easily connect and configure modules using a visual interface.
- Data Integration: Connect to various data sources, preprocess data, and explore it.
- Model Building: Build, train, and evaluate machine learning models.

Steps

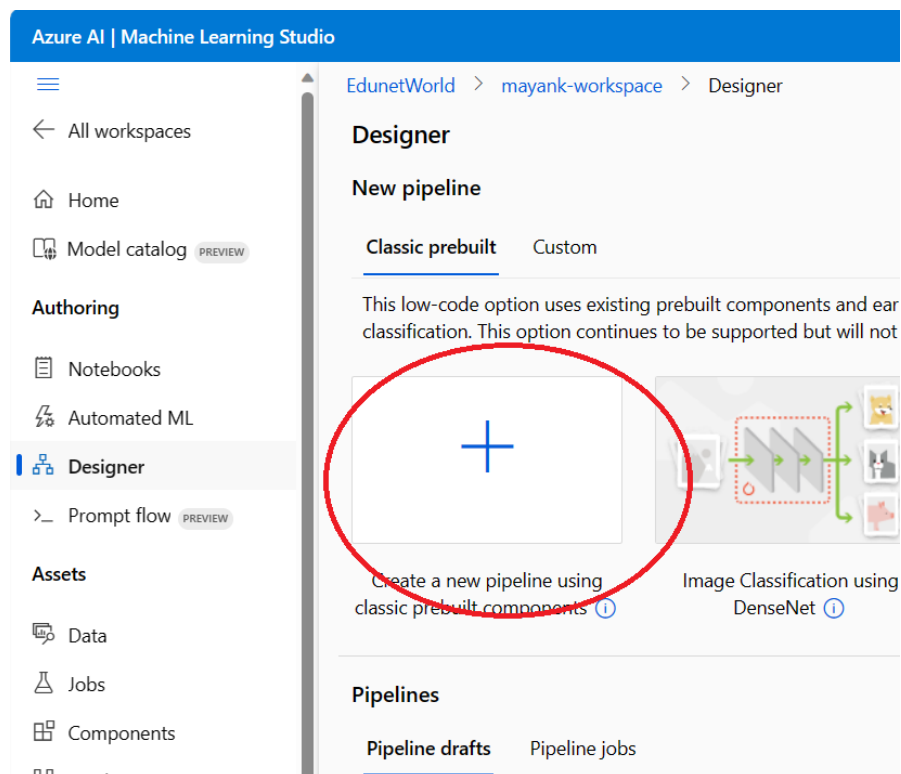
- **Step 1: Access Azure ML Studio:** Log in ml.azure.com
- **Step 2: Create a New Workspace:** Create an Azure Machine Learning workspace by clicking on **Create workspace**.



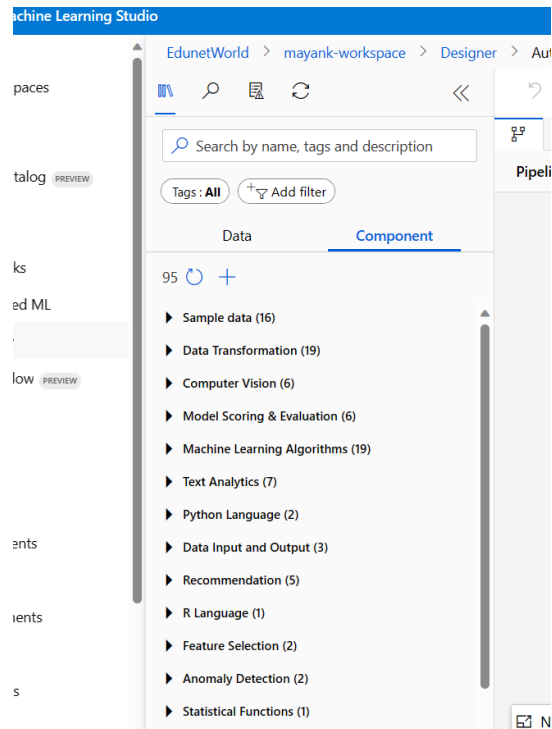
- **Step 3: Open Azure ML Designer:** In your workspace, navigate to Azure ML Designer from the left pane.



- **Step 4: Create a New Pipeline:** Start building your ML pipeline.

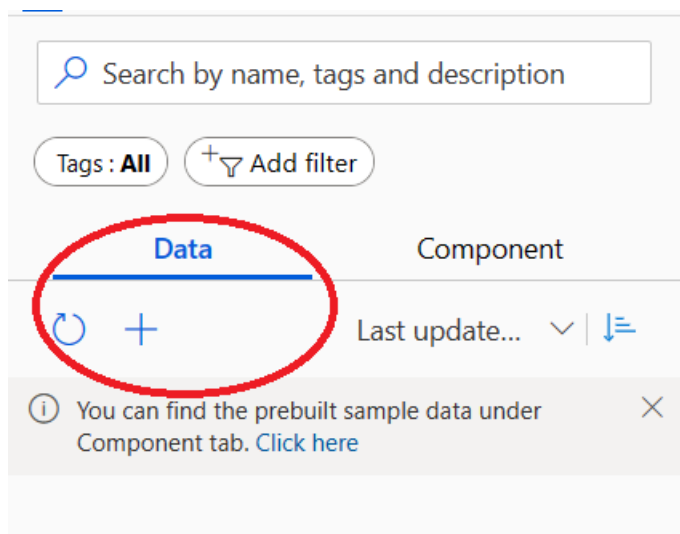


- **Step 5: Create a Classification model using prebuilt components:**



- **Step 6: Upload your dataset or use an existing dataset.**

To upload a local dataset, click on Data Tab and then +.



Provide the respective details. Click on next.

Create data asset

1 Data type

2 Data source

Set the name and type for your data asset

Name *

Diabetes-classification

Description

Data asset description

Type *

Tabular

Back Next

Select **From local files**. Click **Next**.

1 Data type

2 Data source

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, or from a publicly available web location.

From Azure storage

Create a data asset from registered data storage services including Azure Blob Storage, Azure file share, and Azure Data Lake.

From local files

Create a data asset by uploading files from your local drive.

From SQL databases

Create a dataset from Azure SQL database and Azure PostgreSQL database.

From web files

Create a data asset from a single file located at a public web URL.

From Azure Open Datasets

Create a dataset with one-click from pre-made data sets. These data sets are created by the general public and published as Azure Open Datasets.

Back Create

Select default **workspaceblobstore**.

Create data asset

✓ Data type

✓ Data source

3 Destination storage type

4 File or folder selection

5 Settings

6 Schema

7 Review

Select a datastore

Choose a storage type and a datastore to upload your data to in the

Datastore type *

Azure Blob Storage

Create new datastore

Search datastore

Name ↓	Storage
workspaceblobstore	mayar
workspaceartifactstore	mayar

Click on **Upload**. Select your file and click next.

Create data asset

✓ Data type

✓ Data source

✓ Destination storage type

4 File or folder selection

5 Settings

6 Schema

7 Review

Choose a file or folder

Choose files or folders to upload from your local drive. If you up stored in a containing folder.

Upload path

azureml://subscriptions/6a39ae4f-5e3b-4665-9057-6a2f927c4f

Upload

☐ Overwrite if already exists

Upload list

File Types supported are delimited (i.e. csv, Lines, and plain text.

After Uploading the file next screen looks like this.

Create data asset

✓ Data type

✓ Data source

✓ Destination storage type

✓ File or folder selection

5 Settings

6 Schema

7 Review

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

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedi...	Age	Outcome
6	148	72	35	0	33.6	0.627	50	1
1	85	66	29	0	26.6	0.351	31	0
8	183	64	0	0	23.3	0.672	32	1
1	89	66	23	94	28.1	0.167	21	0
0	137	40	35	168	43.1	2.288	33	1
5	116	74	0	0	25.6	0.201	30	0

Back

Next

Review

Cancel

Select usable columns. Click **Next**.

Create data asset

✓ Data type

✓ Data source

✓ Destination storage type

✓ File or folder selection

✓ Settings

6 Schema

7 Review

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 null-filled or replaced with error value. Any conversions preview errors are non-blocking and you can proceed.

Search column name

Include	Column name	Type	Example values	Date format	Properties
<input type="checkbox"/>	Path	String		Not applicable to select...	Not applicable t...
<input checked="" type="checkbox"/>	Pregnancies	Integer	6, 1, 8	Not applicable to select...	Not applicable t...
<input checked="" type="checkbox"/>	Glucose	Integer	148, 85, 183	Not applicable to select...	Not applicable t...
<input checked="" type="checkbox"/>	BloodPressure	Integer	72, 66, 64	Not applicable to select...	Not applicable t...
<input checked="" type="checkbox"/>	SkinThickness	Integer	35, 29, 0	Not applicable to select...	Not applicable t...
<input checked="" type="checkbox"/>	Insulin	Integer	0, 0, 0	Not applicable to select...	Not applicable t...
<input checked="" type="checkbox"/>	BMI	Decimal (dot '.')	33.6, 26.6, 23.3	Not applicable to select...	Not applicable t...

Back

Next

Cancel

Review and click on **Create**.

Create data asset

✓ Data type

✓ Data source

✓ Destination storage type

✓ File or folder selection

✓ Settings

✓ Schema

ⓘ Review

Review

Review the settings for your data asset and make any changes as needed.

Data type

Name

Diabetes-classification

Description

--

Type

tabular

Data source

Type

Local

File selection

Upload path

azurerml:/subscriptions/6a39ae4f-5e3b-4665-9057-6a2f927c406b/resourcegroups/mayank-rg/workspaces/mayank-workspace/datastores/workspaceblobstore/paths/UI/2023-09-06_055418_UTC/diabetes.csv

Files uploaded

Schema

Pregnancies

Integer

Glucose

Integer

BloodPressure

Integer

SkinThickness

Integer

Insulin

Integer

(showing 5 of 10 columns)

Back

Create

Cancel

- **Step 7: Search components from search bar and develop a model as follows:**

Pipeline-Created-on-09-06-2023

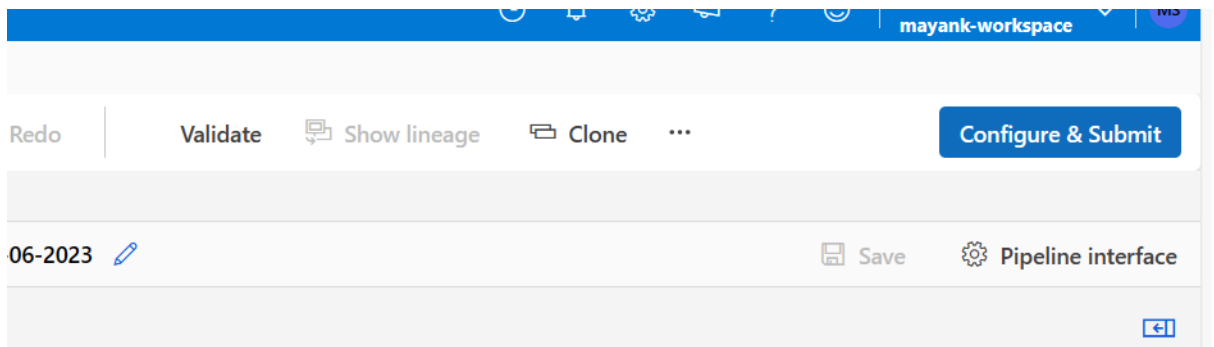
Save

Pipeline interface

```

graph TD
    A[Diabetes-classification  
Diabetes-classification] --> B[Split Data  
split_data]
    B --> C[Results datas...]
    C --> D[Train Model  
train_model]
    D --> E[Trained model]
    E --> F[Score Model  
score_model]
    F --> G[Scored dataset]
    G --> H[Evaluate Model  
evaluate_model]
    H --> I[Evaluation results]
  
```

- **Step 8: Click on Configure & Submit.**



Fill up the required details.

Set up pipeline job

1 Basics

2 Inputs & outputs

3 Runtime settings

4 Review + Submit

Basics

Experiment name

Select existing

Create new

New experiment name *

Classification-example

Job display name

Pipeline-Created-on-09-06-2023

Job description

Pipeline created on 20230906

Job tags

Name

:

Value

Add

Review + Submit

Back

Next

Close

In the Runtime settings. Fill up the required details. Click on **Review + submit** and then **Submit**. If you don't have any Compute instance then create one. Steps are provided in the last.

Set up pipeline job

✓ Basics

✓ Inputs & outputs

3 Runtime settings

4 Review + Submit

Runtime settings

Default compute ⓘ

Select compute type

Compute instance

Select Azure ML compute instance

mayank1 - Running

Create Azure ML compute instance

Refresh Compute

Default datastore ⓘ

Select datastore *

workspaceblobstore

Advanced settings

☒ Continue on step failure ⓘ

Review + Submit

Back

Next

Close

After successful submission. Pipeline job will be submitted. Click on **View Details**.

EdunetWorld > mayank-workspace > Designer > Authoring

✓ Success: Pipeline job has been submitted. [View Details](#)

evaluate

Tags: All

+ Add filter

Data

Component

7

↺

+

Most relevant

⌵

⌵

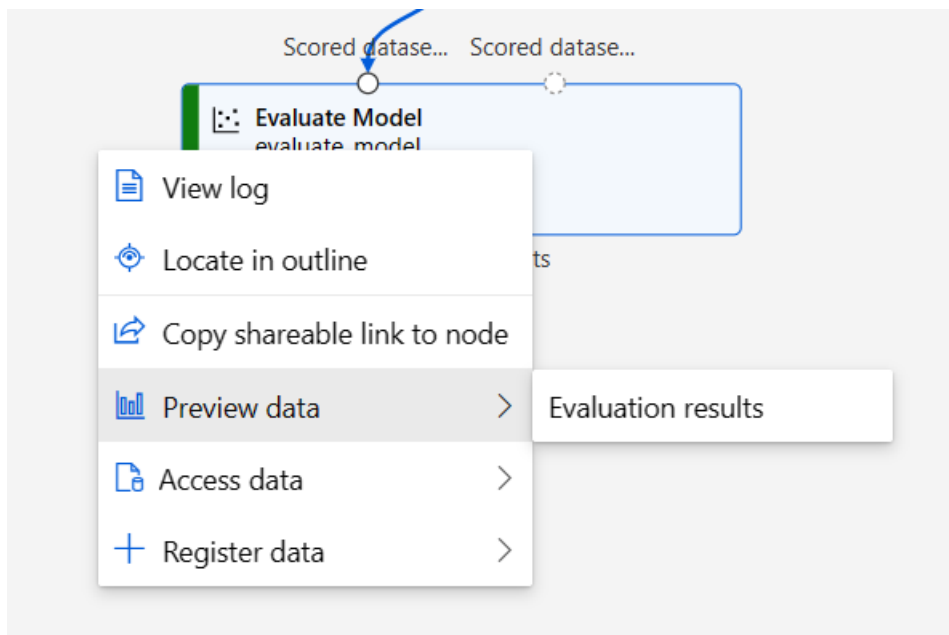
💡 Evaluate Recommender

Undo

Redo

Pipeline-Created-on-09-06-2023

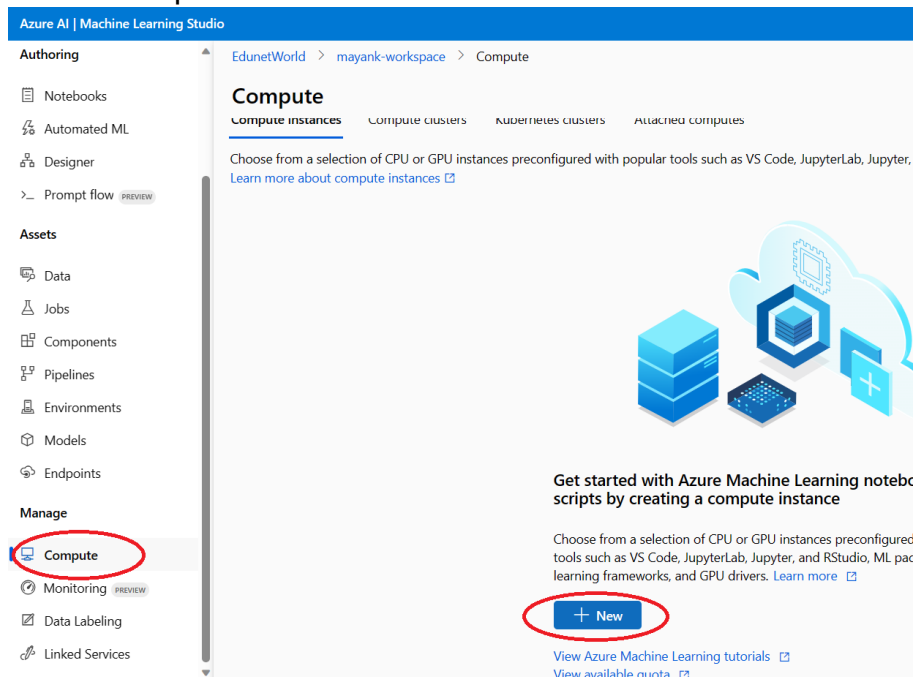
Once the run is successful, you can see the evaluation results by right-click on **Evaluate Model** -> **Preview Data** -> **Evaluation Results**.



Evaluation Results will be shown.



To create a compute instance.
Select Compute.



Select appropriate details. For our implementation, we will use a Memory optimized CPU : **Standard_E4ds_v4 4 cores, 32GB RAM, 150GB storage**

Click on **Create**.
Now, provision will take a couple of minutes.