**Project Report: Diabetes Prediction Using Classification in Azure ML**

**Name: Muhammad Munawar Khan  
Roll No: NUM-BSCS-2022-13  
Course: Coursera – Machine Learning with Microsoft Azure**

**Objective**

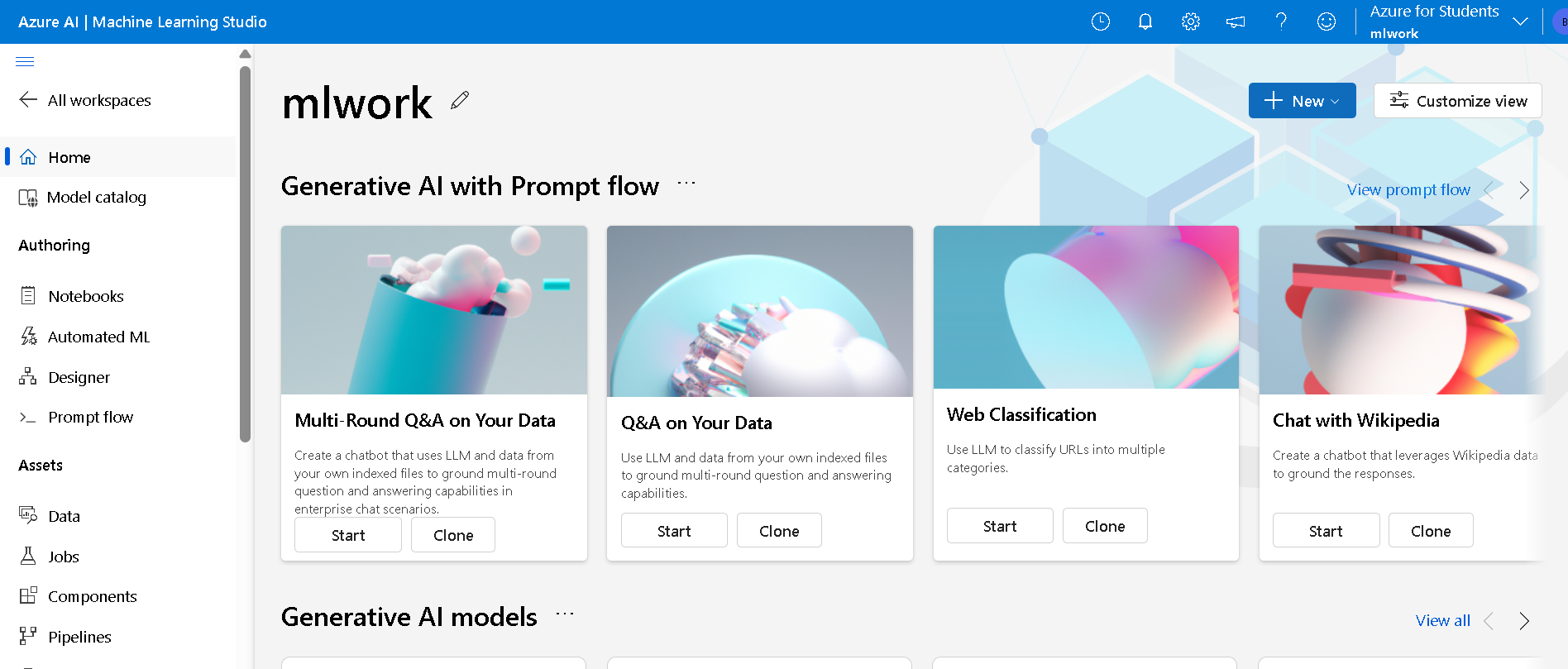
To build and deploy a classification machine learning model using Azure Machine Learning Designer that predicts whether a patient has diabetes based on health attributes.

**📘 Part 1: Create an Azure Machine Learning Workspace**

Created a new Azure Machine Learning workspace in the Azure portal.

Chose a globally unique workspace name, selected a region, and linked it to a resource group.

Verified the setup by accessing Azure ML Studio at ml.azure.com.

**🖼 Screenshot: Workspace dashboard in Azure ML Studio**  


**📘 Part 2: Create Compute Resources**

Compute Cluster:

Name: diabetes-compute

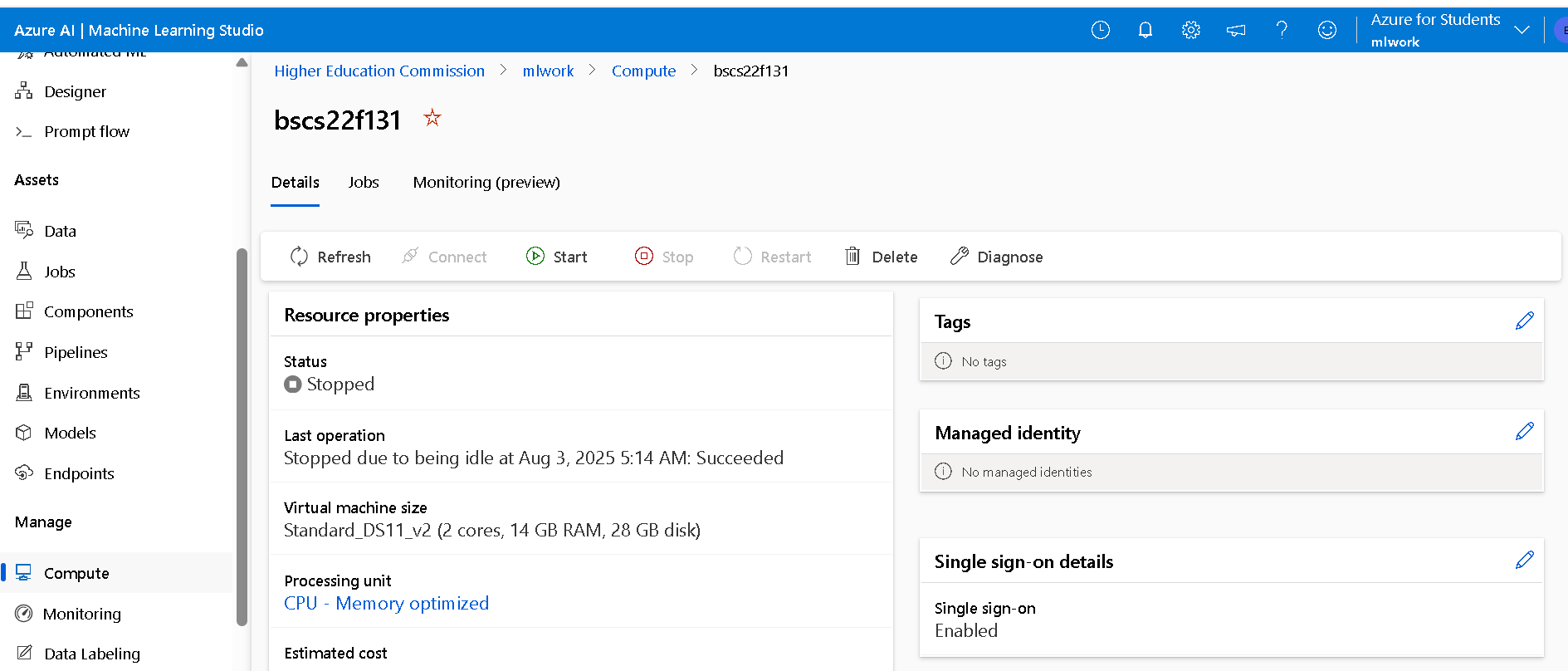
VM Type: Standard\_DS11\_v2

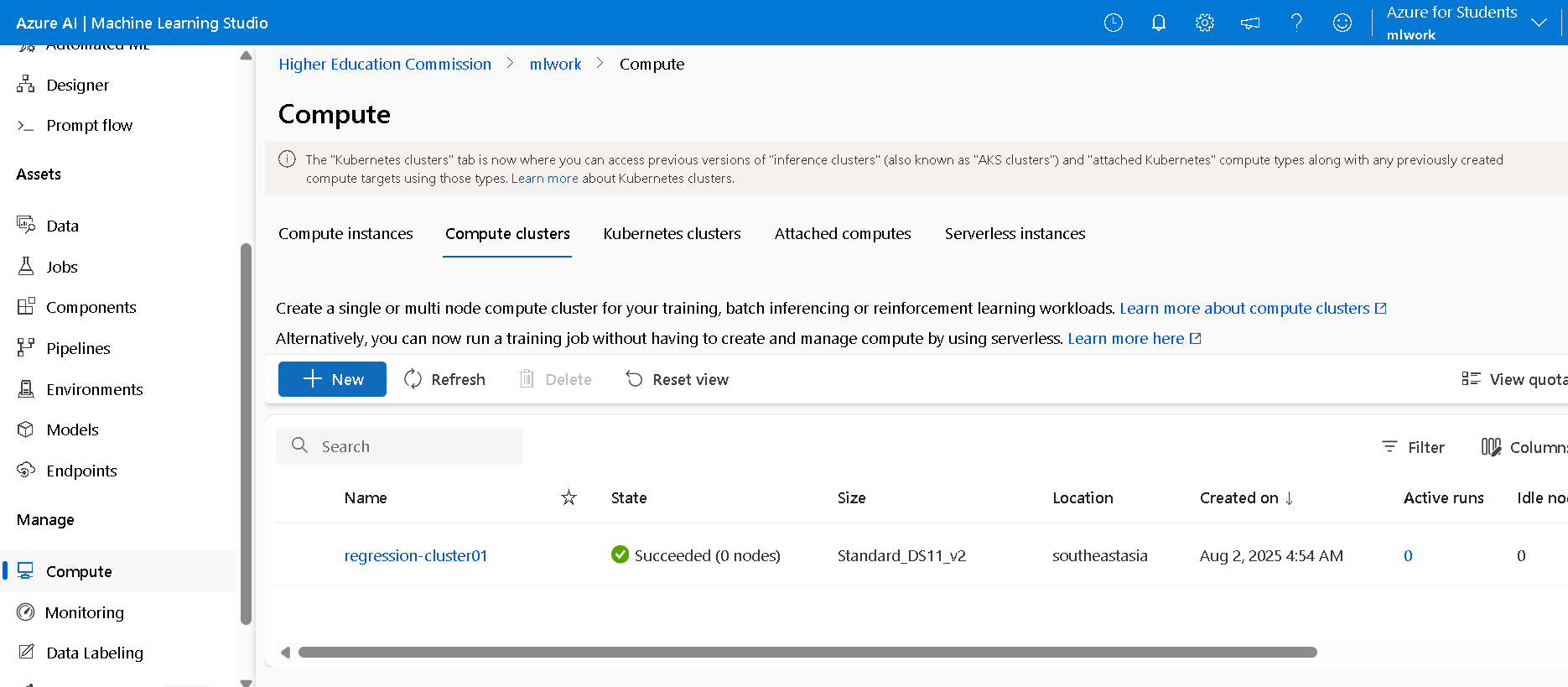
Minimum nodes: 0

Maximum nodes: 2

Purpose: Used for training, scoring, and deployment.

**🖼 Screenshot: Created compute instance and compute cluster**





**📘 Part 3: Explore and Prepare Data Using Azure ML Designer**

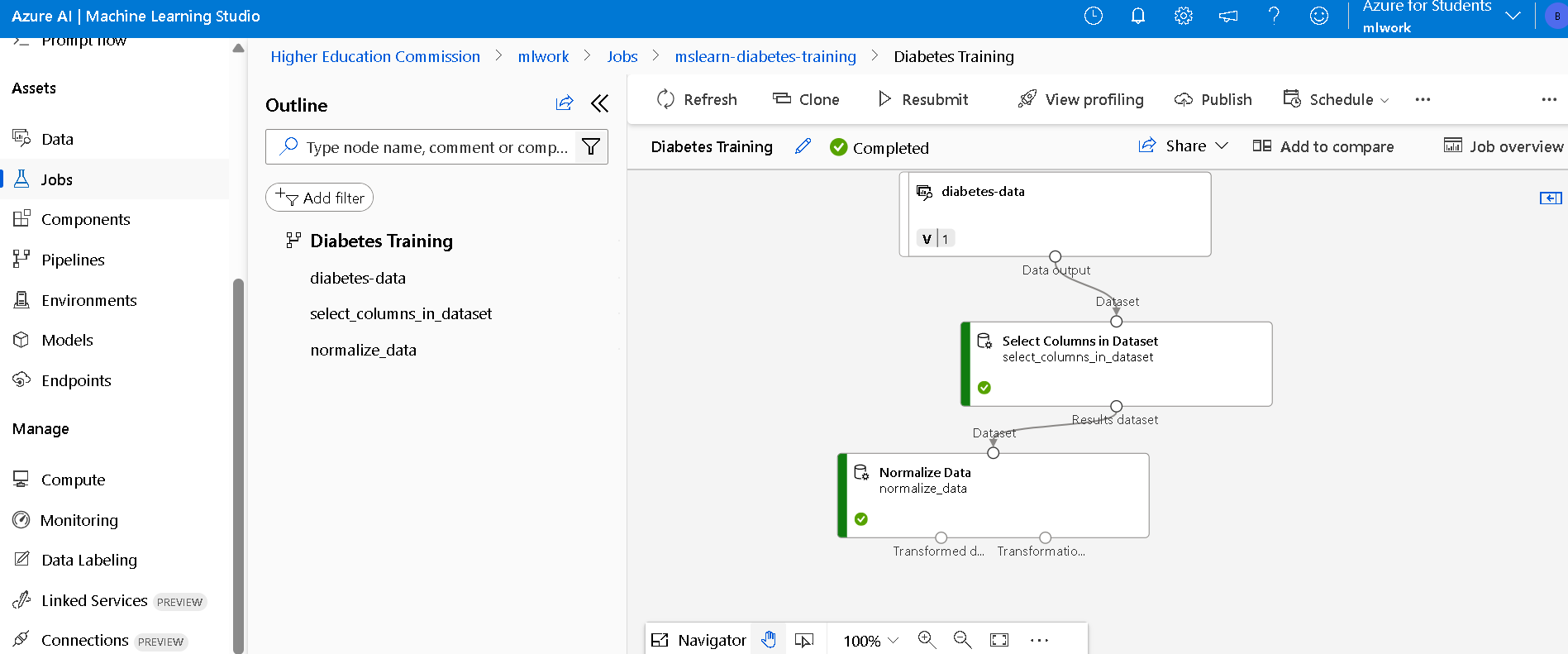
Imported the Diabetes dataset.

Applied the following data transformation modules:

Select Columns in Dataset – Chose features relevant to prediction.

Clean Missing Data – Removed rows with missing values.

Normalize Data – Applied MinMax normalization to numerical columns (excluding target).

**🖼 Screenshot: Data preprocessing pipeline**

**📘 Part 4: Create and Run a Training Pipeline**

Pipeline structure:

Select Columns in Dataset

Clean Missing Data

Normalize Data

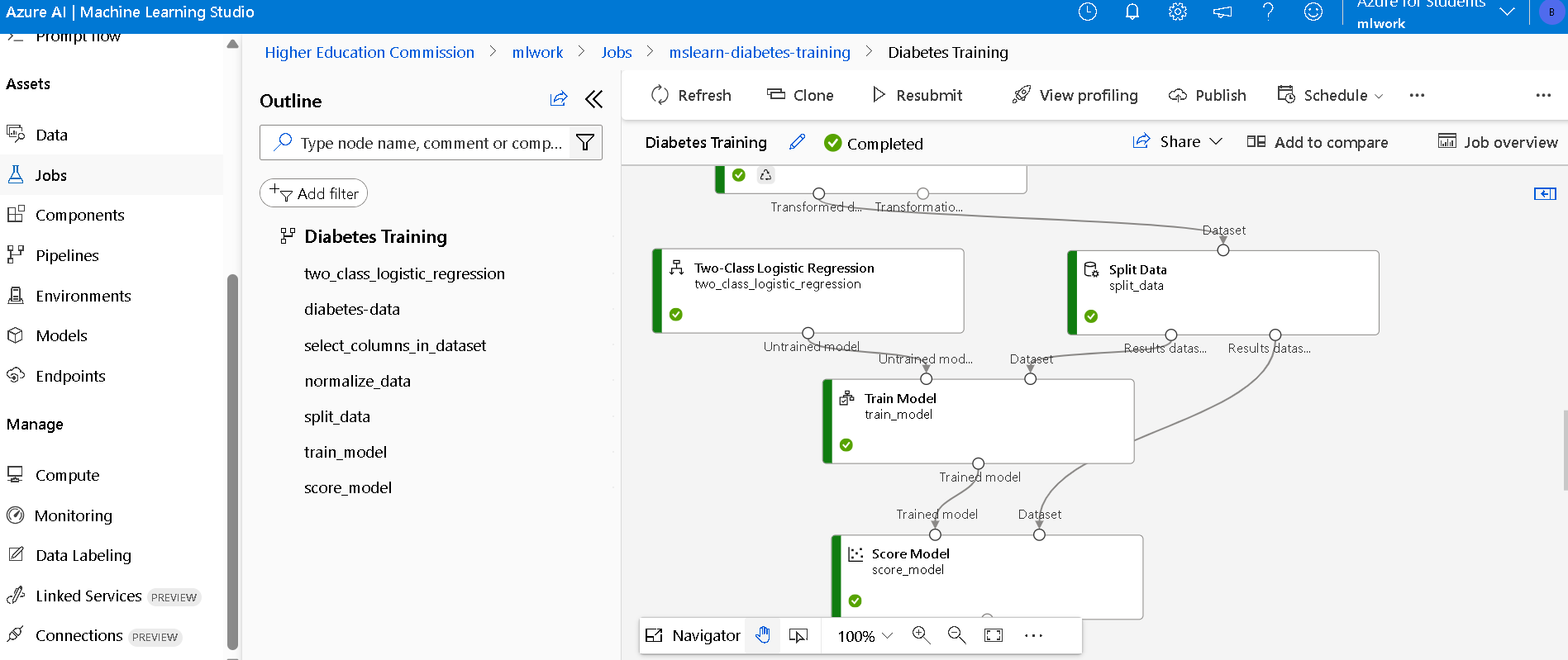
Split Data (70% training / 30% testing)

Train Model (target column: Outcome)

Score Model

Evaluate Model

Used classification algorithm (e.g., Logistic Regression).

**🖼 Screenshot: Full training pipeline  
**

**📘 Part 5: Evaluate the Classification Model**

Evaluation Metrics:

Accuracy

Precision

Recall

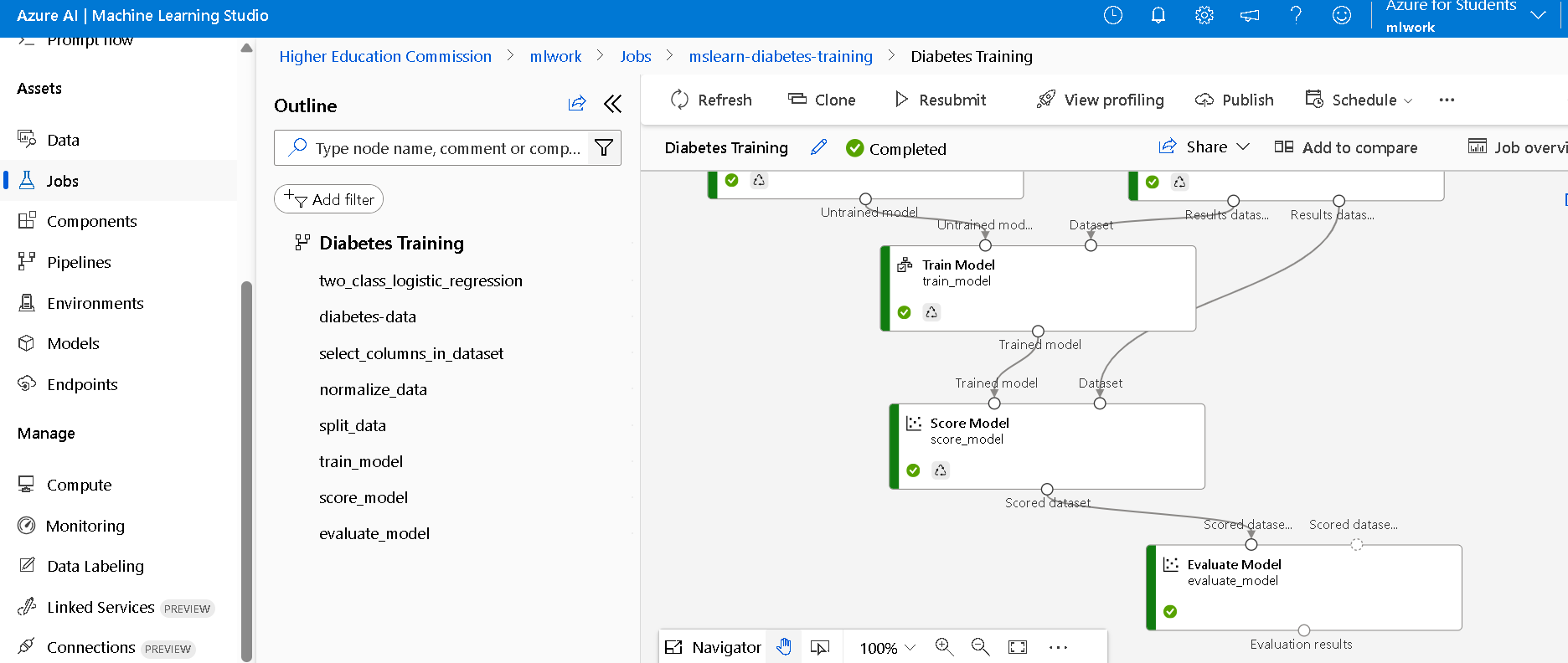
F1 Score

AUC

Observations:

Model shows acceptable predictive performance.

Potential improvements may include feature engineering or handling class imbalance.

**🖼 Screenshot: Evaluation results  
**

**📘 Part 6: Create an Inference Pipeline**

Cloned the training pipeline to convert it into an inference pipeline.

Replaced the dataset input with Enter Data Manually (excluding target column).

Removed the Evaluate Model module.

Used Execute Python Script to return only prediction and probability scores.

🔍 Python Script Used:

import pandas as pd

def azureml\_main(dataframe1 = None, dataframe2 = None):

scored\_results = dataframe1[['PatientID', 'Scored Labels', 'Scored Probabilities']]

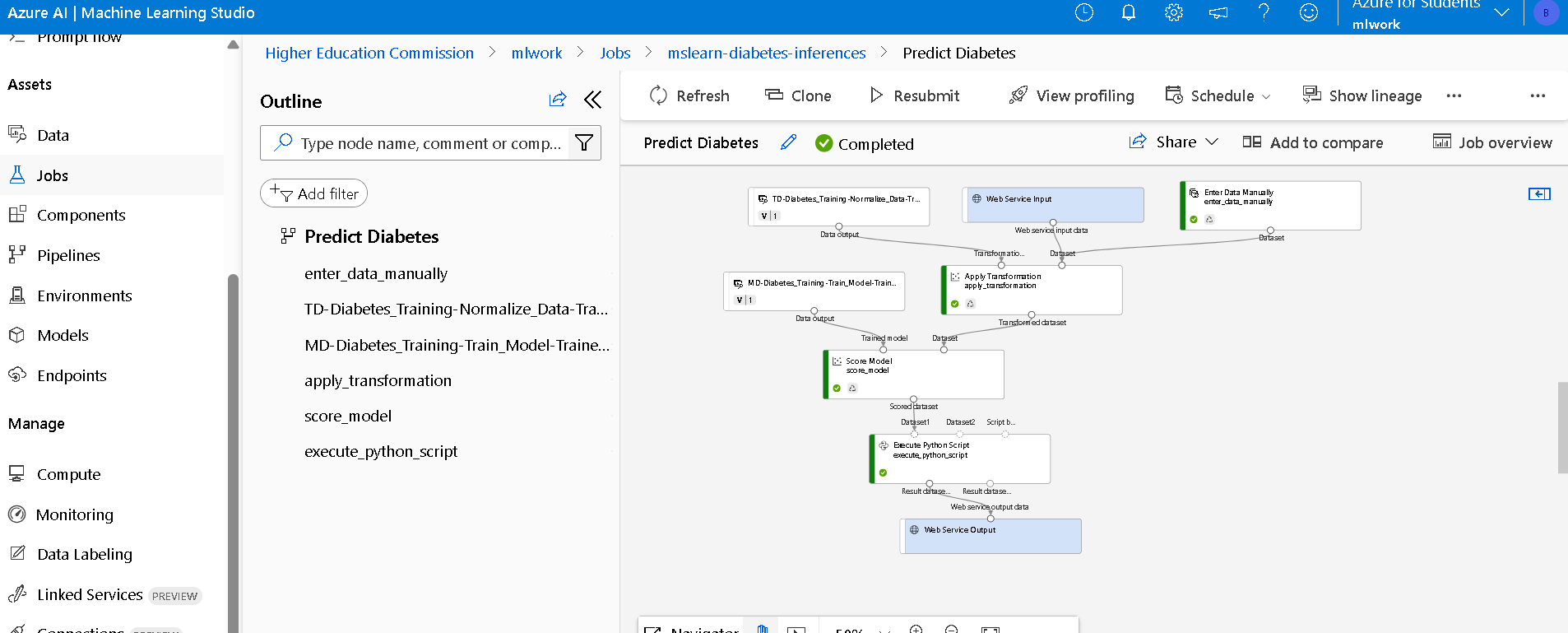
scored\_results.rename(columns={

'Scored Labels': 'DiabetesPrediction',

'Scored Probabilities': 'Probability'

}, inplace=True)

return scored\_results

**🖼 Screenshot: Inference pipeline view**  


**📘 Part 7: Deploy a Predictive Service**

Deployed the inference pipeline as a real-time web service using Azure Container Instance (ACI).

Encountered the following issue during deployment:

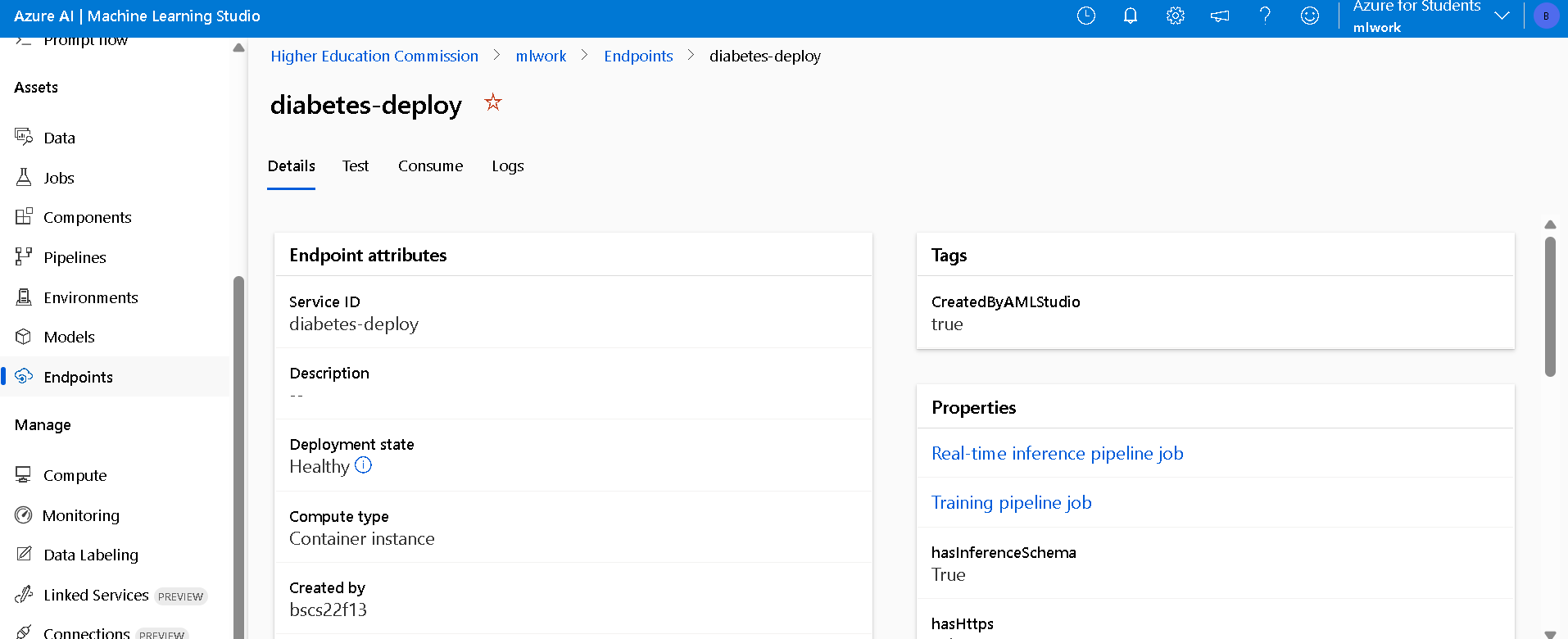
"User does not have authorization to perform action 'Microsoft.ContainerInstance/containerGroups/write'"

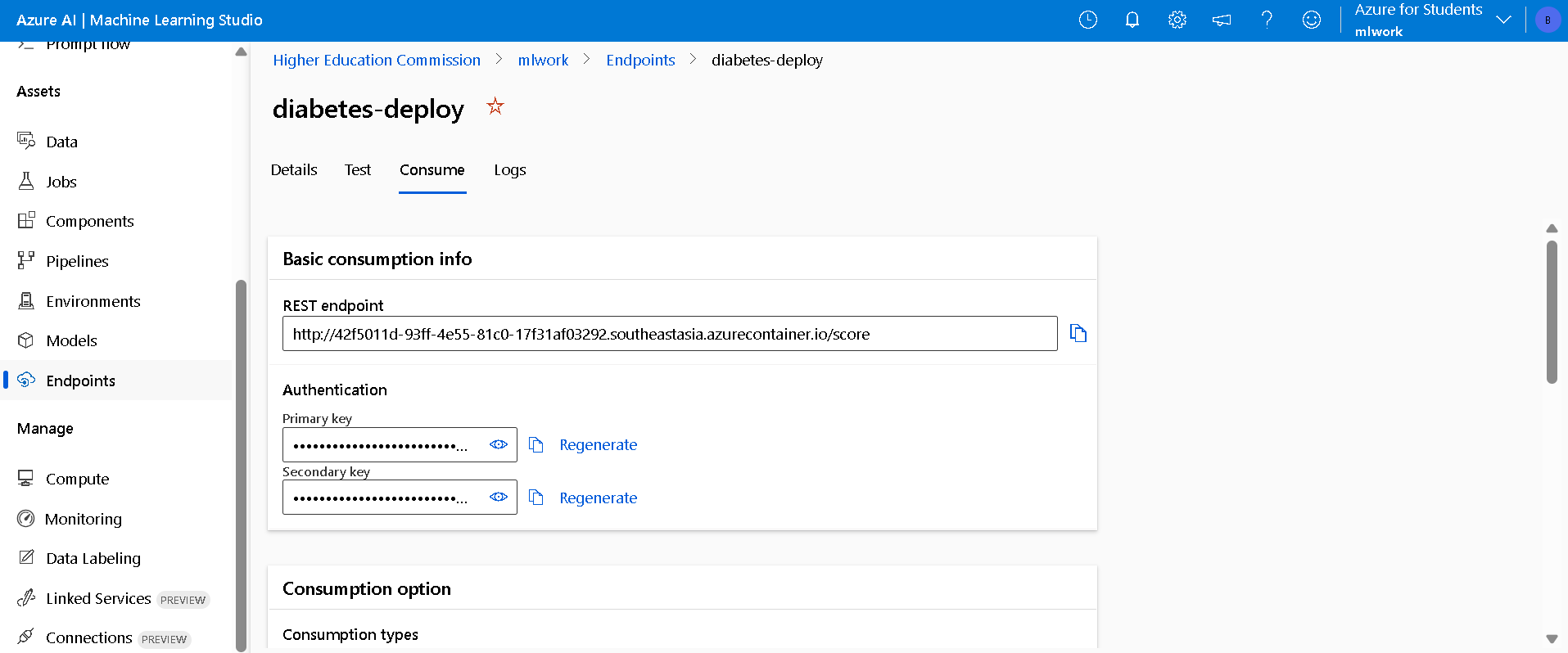
Resolution:

Requested Contributor access from Azure admin

Alternatively, deployed to the available diabetes-compute cluster instead of ACI

Web service was successfully deployed and tested using REST endpoint.

**🖼 Screenshot: Web service endpoint and deployment status**  




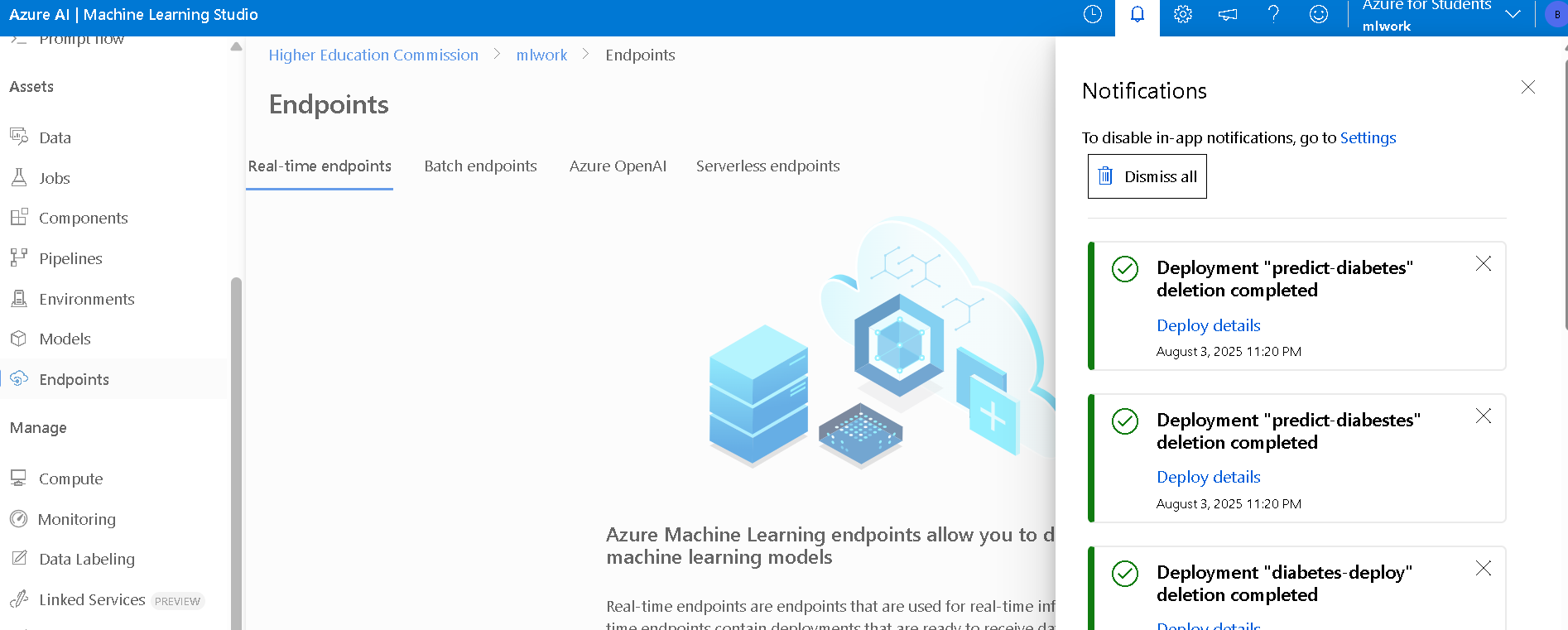
**📘 Part 8: Clean-Up**

Deleted the following to avoid unnecessary charges:

Real-time endpoint

Compute cluster

Workspace (optional)

**🖼 Screenshot: Deleted resources confirmation**  


**✅ Conclusion**

Successfully built, trained, and deployed a classification model to predict diabetes.

The project demonstrated the full ML lifecycle using Azure ML Designer, from preprocessing to real-time deployment.

The deployed service allows real-time predictions via an API, making it useful in clinical dashboards or health apps.

**📎 Technologies Used**

Microsoft Azure ML Studio (Designer)

Azure Compute Instances and Clusters

Azure Container Instances

Python (Execute Script module)

Logistic Regression (Azure ML built-in)