**Setting up Azure Databricks**

### **1. Introduction to Azure Databricks**

**Azure Databricks** is a fast, easy, and collaborative Apache Spark-based analytics platform optimized for Microsoft Azure. It combines the best of Databricks and Azure to enable scalable big data processing, interactive notebooks, and machine learning on the cloud.

It is widely used for:

* Large-scale data engineering and analytics
* Machine learning model training
* Real-time data stream processing
* Interactive data science workflows

### **2. Objective**

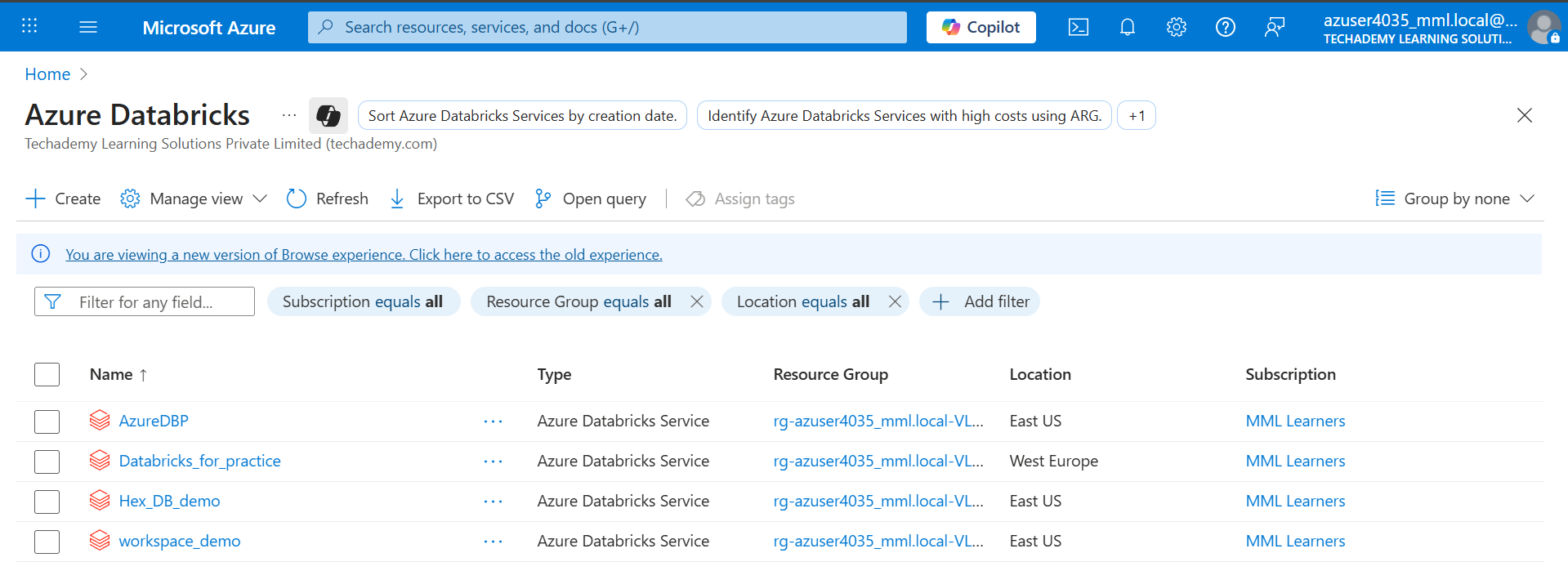
To set up and access Azure Databricks using the lab subscription provided by Hexaware, and create a workspace for running Spark-based analytics.

**3. Steps to Set Up Azure Databricks**

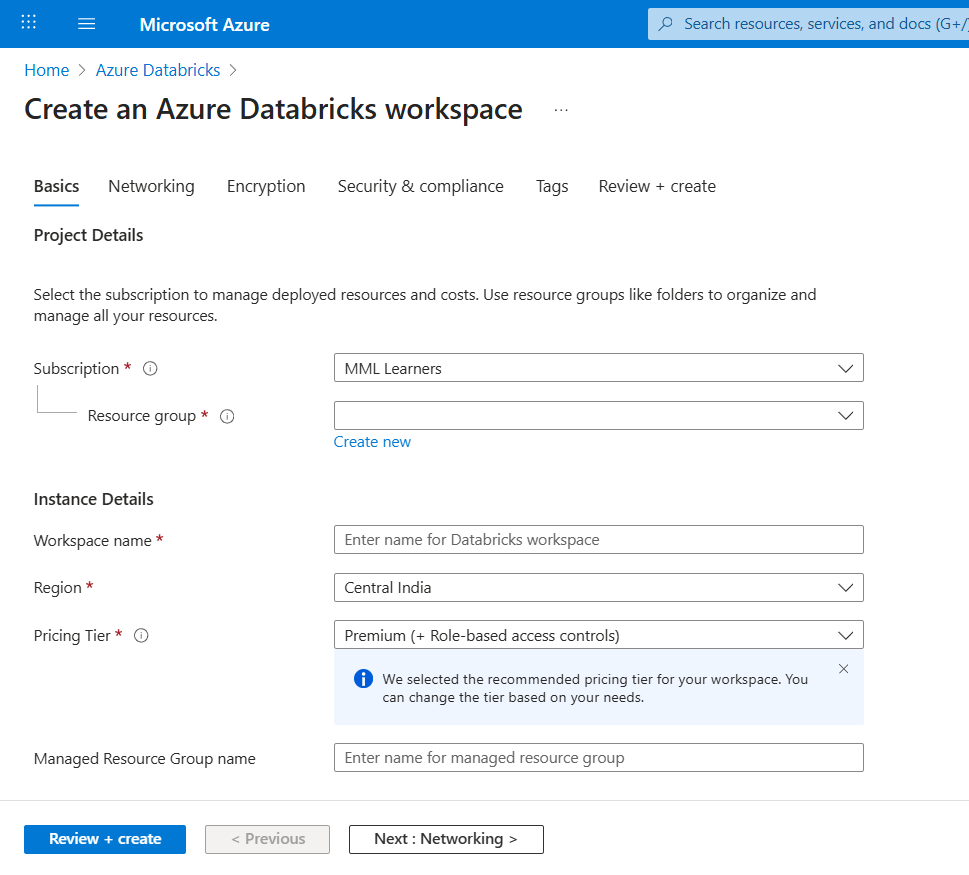
Step 1: Log in to Azure Portal

* Navigate to: <https://portal.azure.com>
* Sign in using the credentials provided by Hexaware

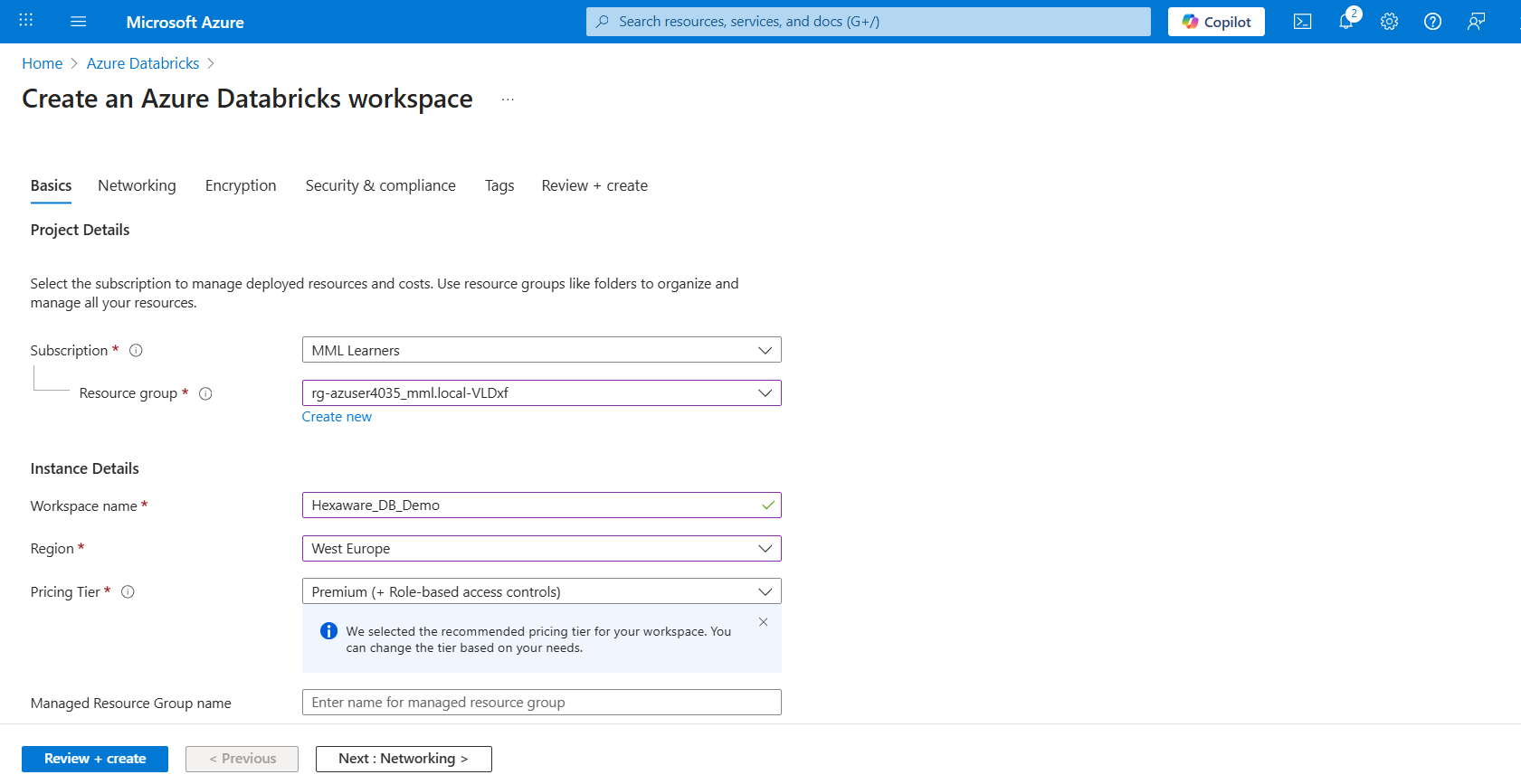
Step 2: Navigate to Azure Databricks

****

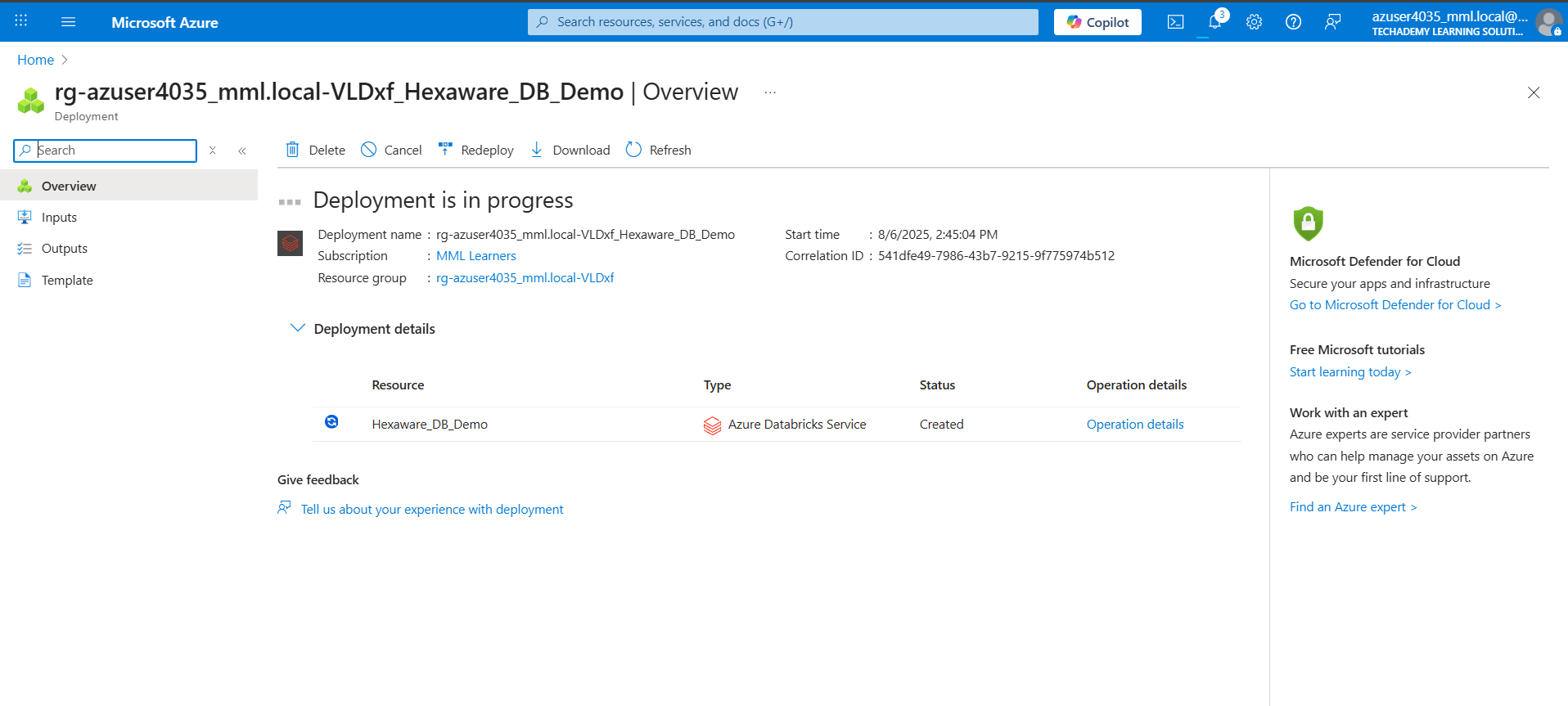
Step 3: Create a new Workspace

****

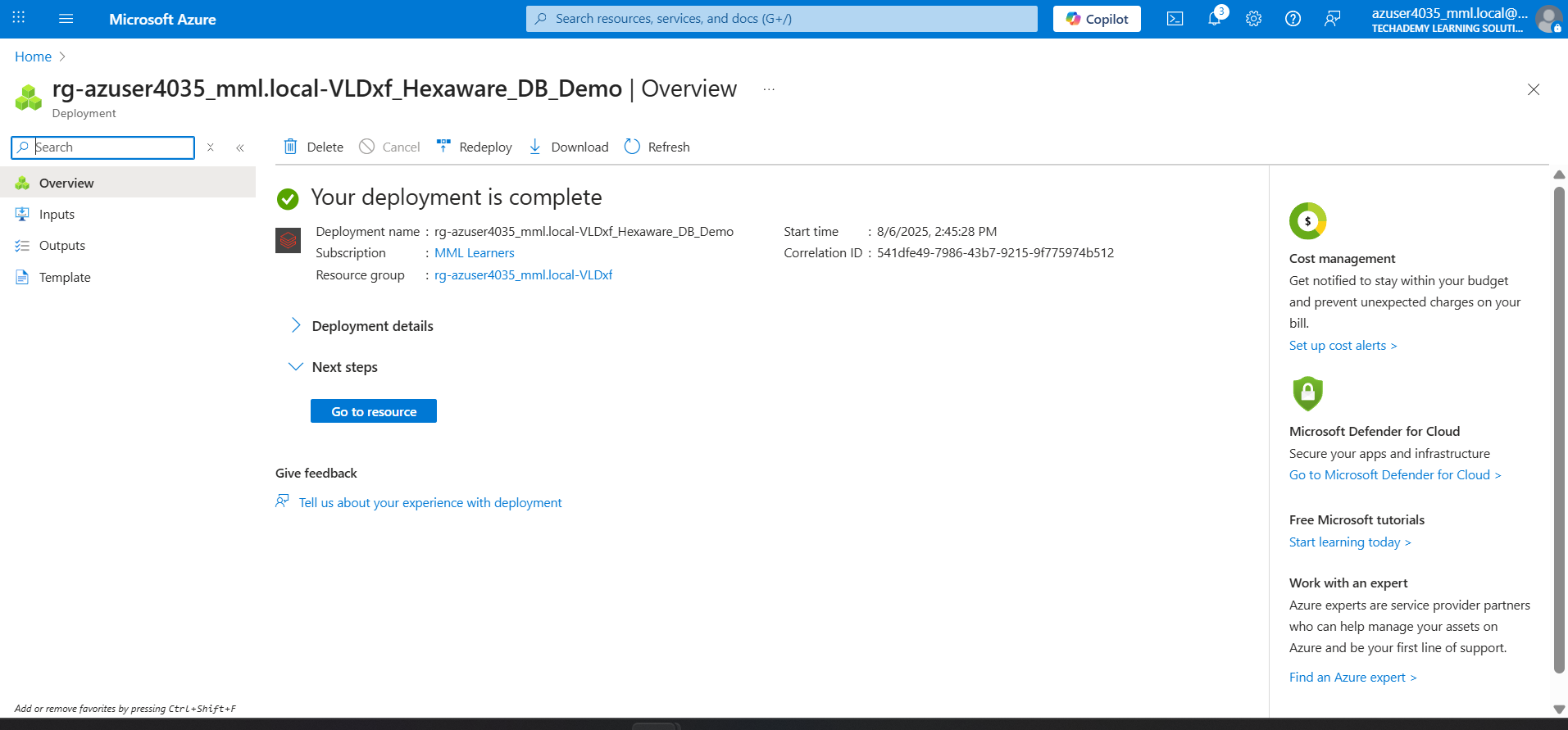
Step 4: Fill all the required fields and click the review + create button to deploy workspace

****

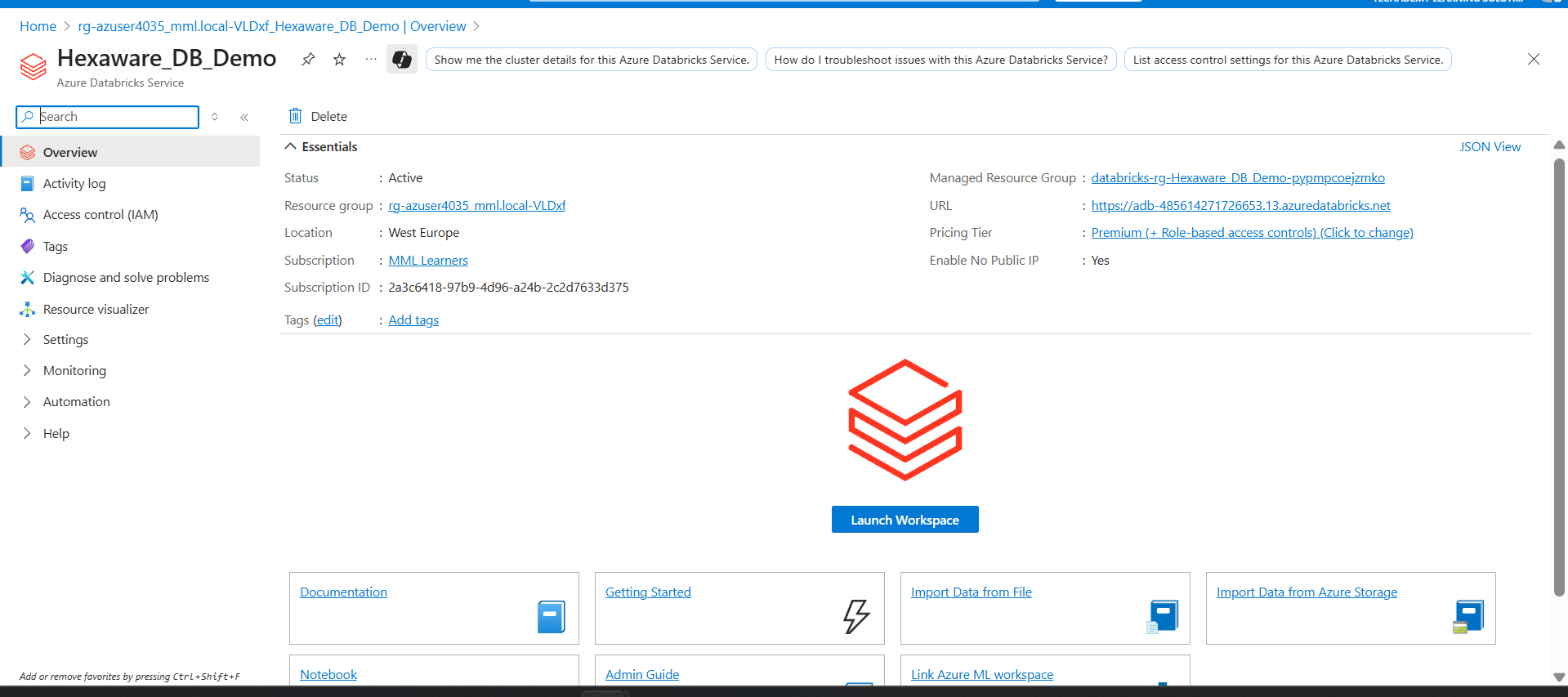
Step 5: The deployment is in progress

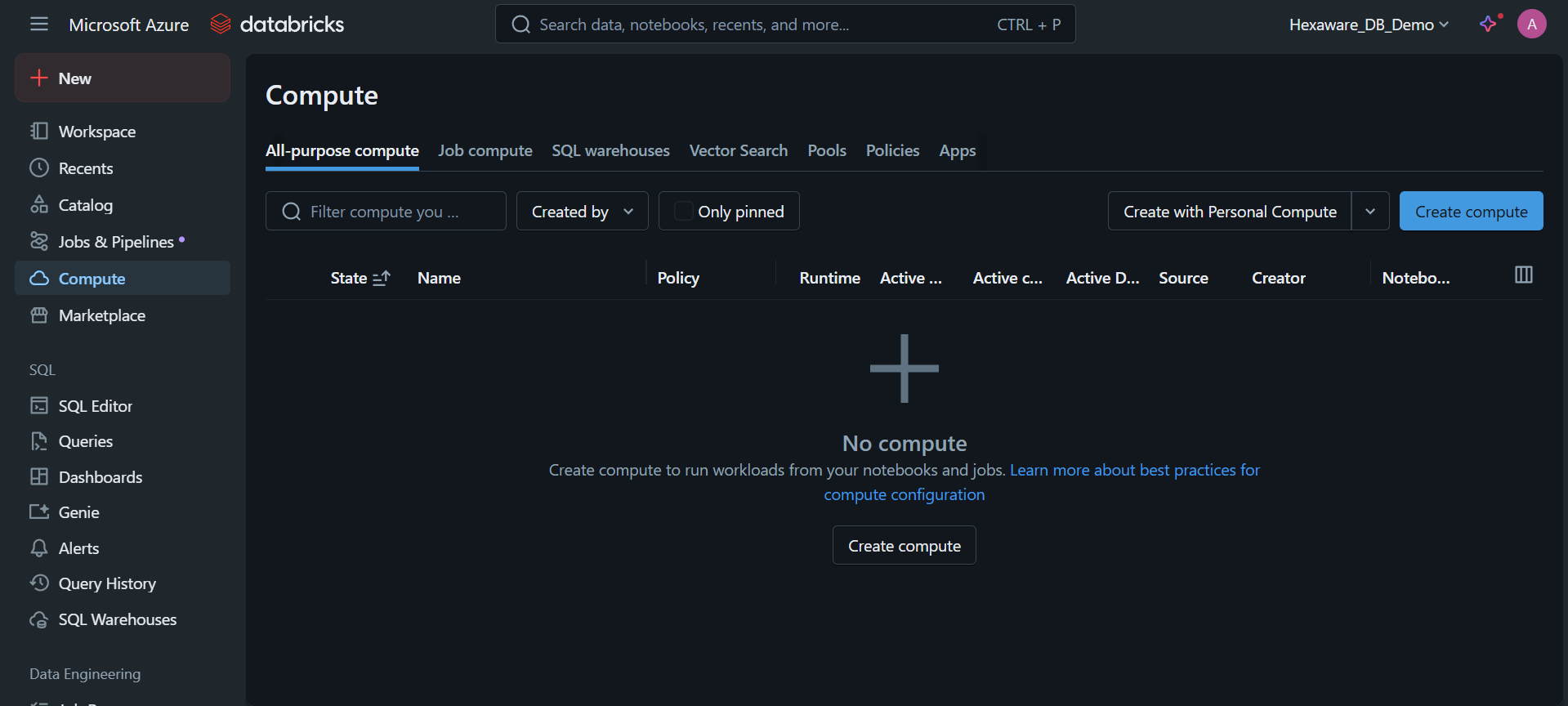
****

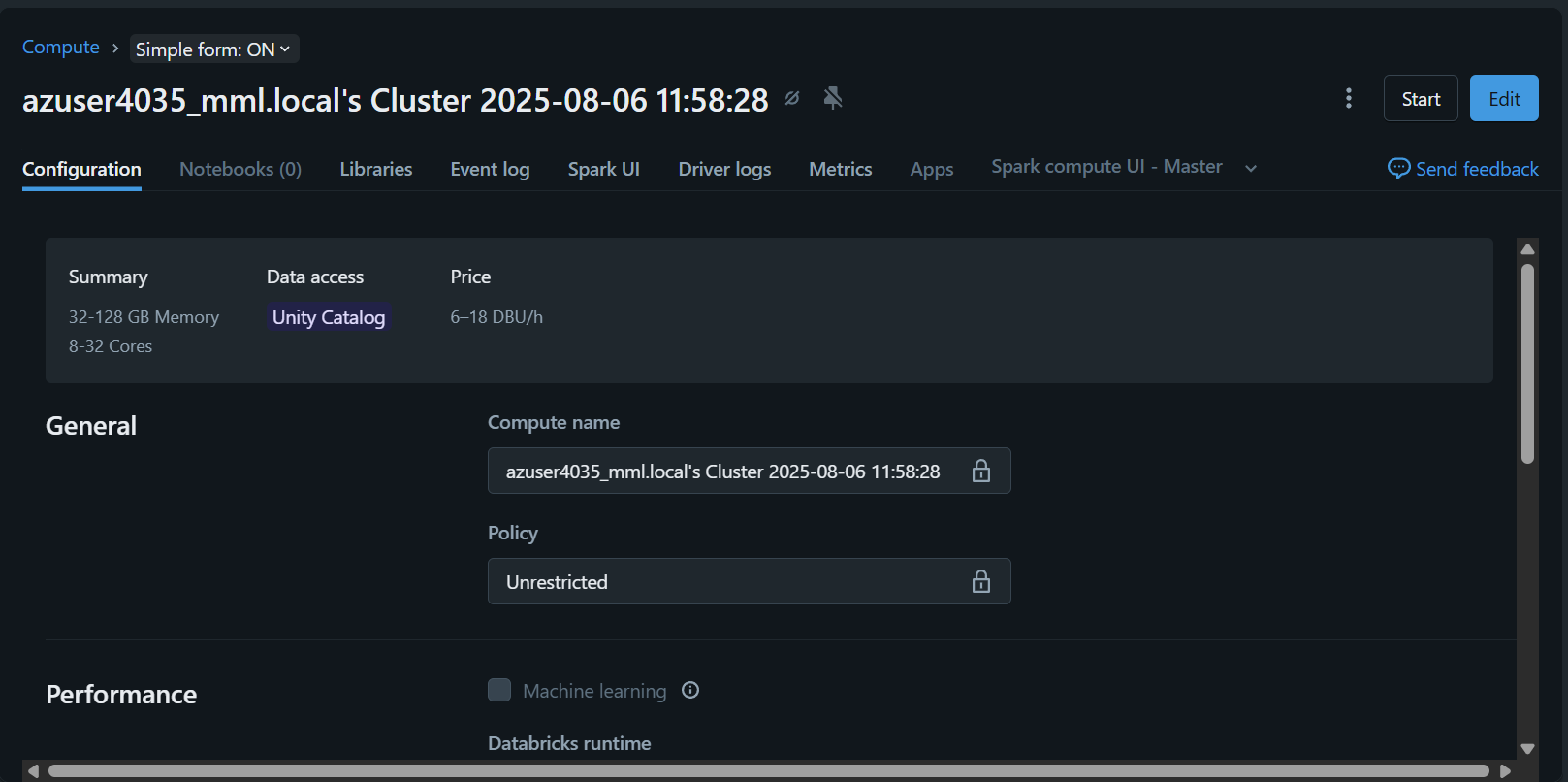
Step 6: After the deployment click “Get Resources”

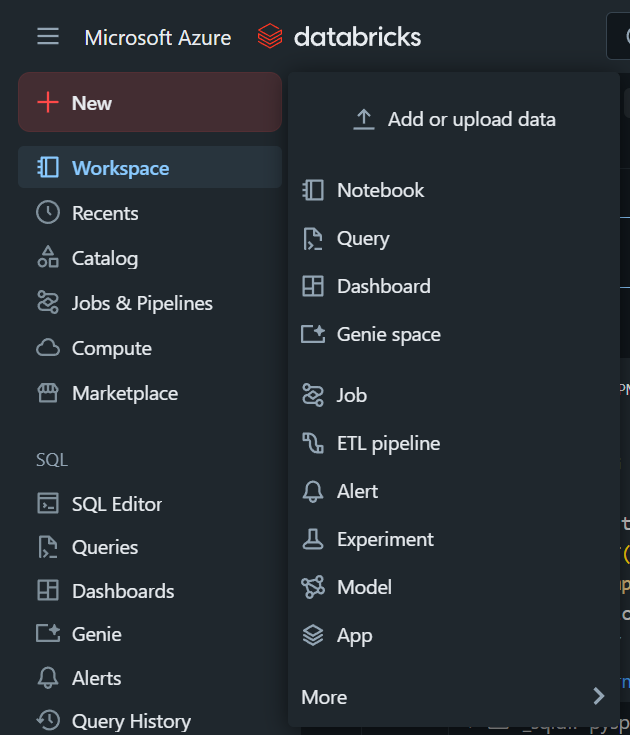
****

Step 7: Launch the workspace

****

Step 8: And you will be redirected to the databricks website,here navigate to “Compute”

Step 9: Create a Compute and start the cluster with default entries.



Step 10: Now you can create a new

Notebook and write you PySpark

Programs.

**Conclusion:**

The lab exercise involved attempting to deploy **Azure Databricks workspace** via Azure Portal. It helped in understanding:

* How to navigate Azure services.
* Resource group and workspace creation.
* Common deployment errors and diagnostics.