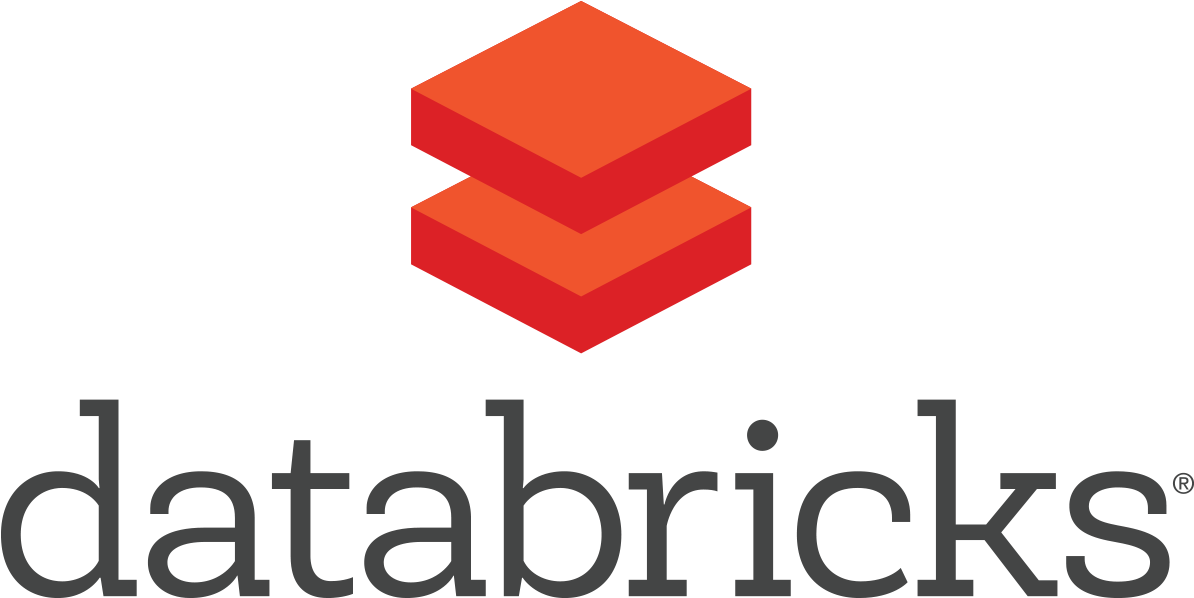
ASSESSMENT - 21 | Azure Databricks - 01

*“INTRO TO AZURE DATABRICKS”*



Submitted By

VAIBHAV PATIDAR

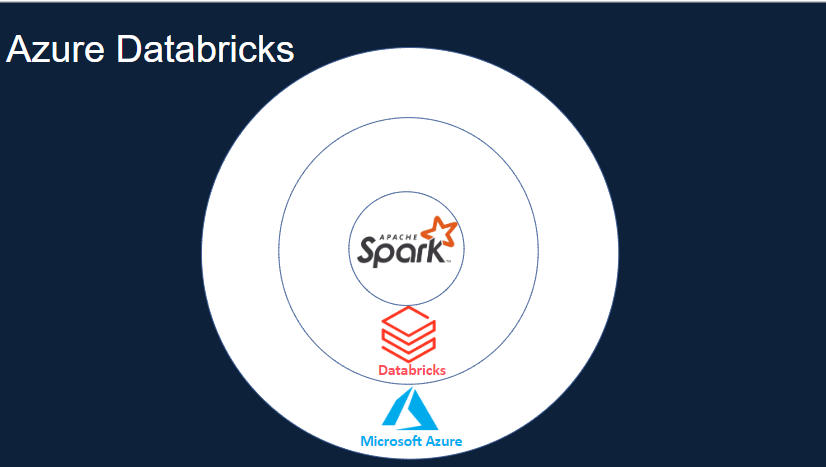
IPS ACADEMY, INDORE (M.P.)

**Date :** 28-12-2023 | **Day 21**

**Week 4** | **Day - 03**

**→ What is Azure Databricks?**

* Azure Databricks is a cloud-based big data analytics platform.
* It integrates Apache Spark and Azure services for fast data processing.
* Enables collaborative data science and engineering.
* Provides a unified environment for data preparation, exploration, and visualization.
* Scales resources dynamically to handle varying workloads.
* Streamlines machine learning workflows.



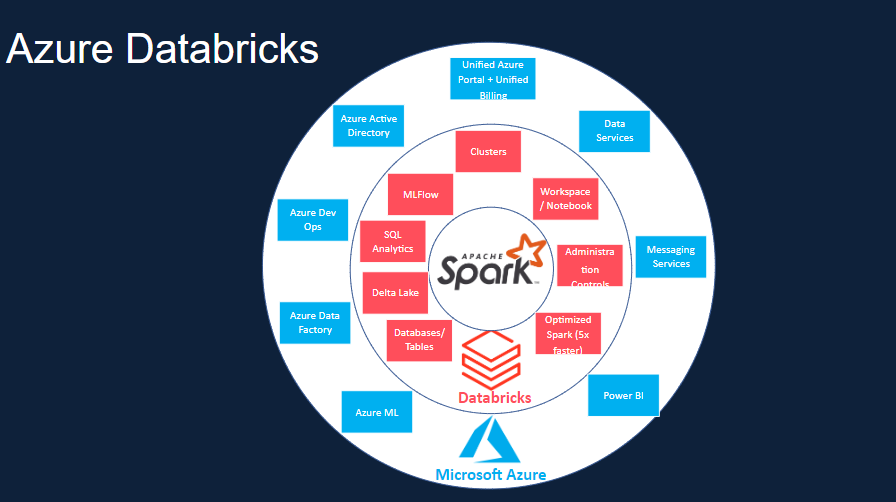
***→ Apache Spark*** is a lightning-fast unified analytics engine for big data processing and machine learning

→ **Services offered by Apache Spark :**

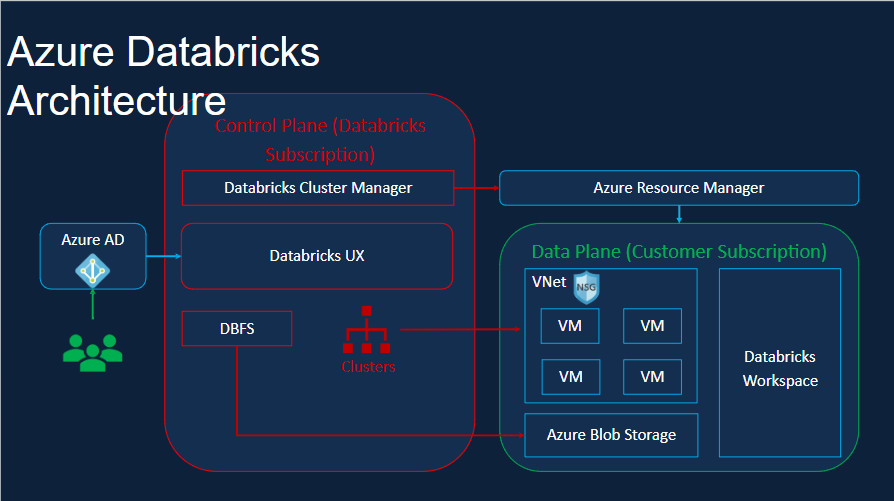
1. Clusters
2. Workspace / Notebook
3. Administration Controls
4. Optimized Spark (5x faster)
5. Databases/ Tables
6. Delta Lake
7. SQL Analytics
8. MLFlow

→ **Services offered by Databricks :**

1. Data Services
2. Messsaging Services
3. Power Bi
4. Azure ML
5. Azure Data Factory
6. Azure DevOps
7. Azure Active directory
8. Unified Billing



**→ Azure Databricks Archietecture :**

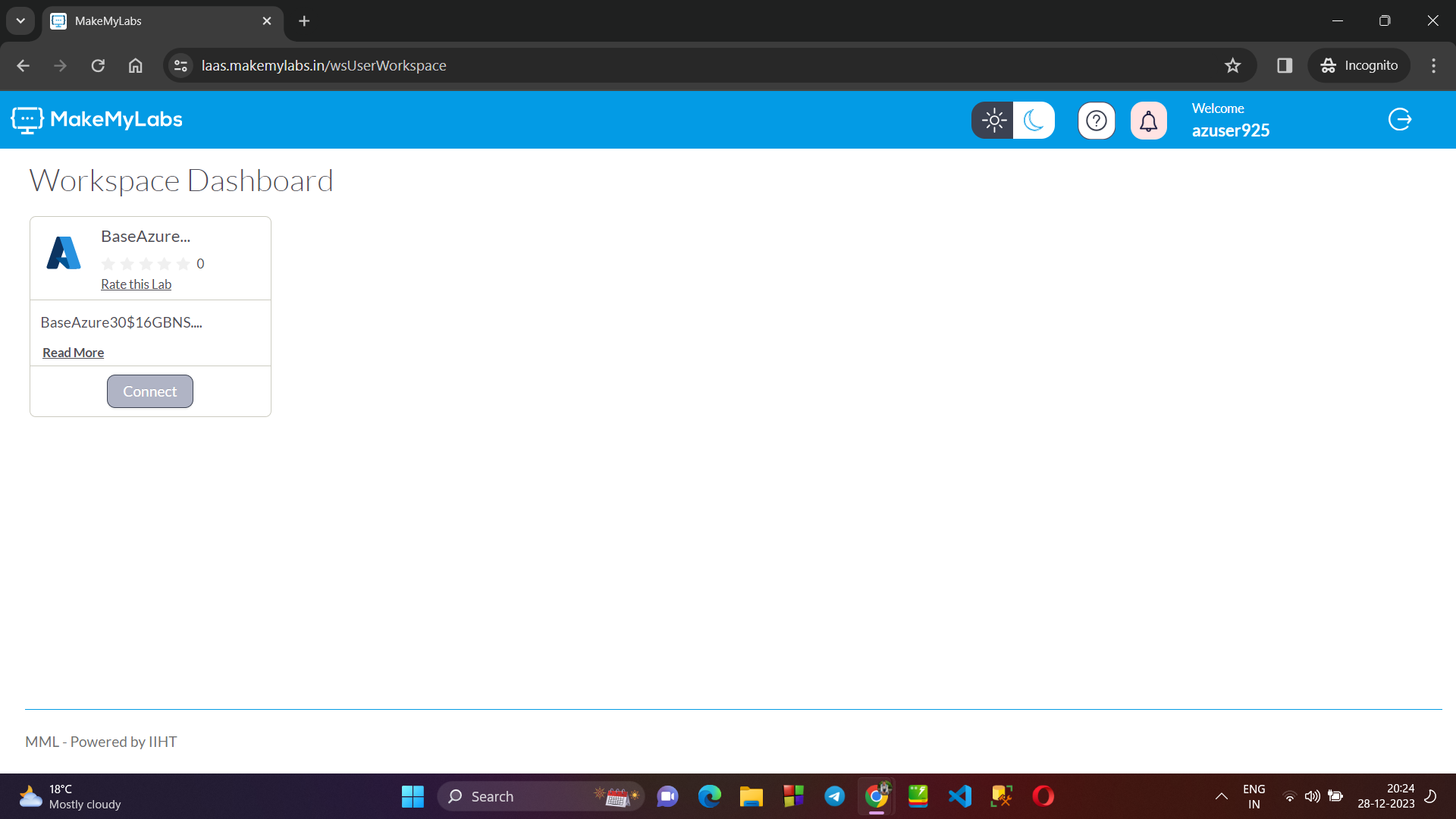
****

**→ How to setup Azure Databricks Account :**

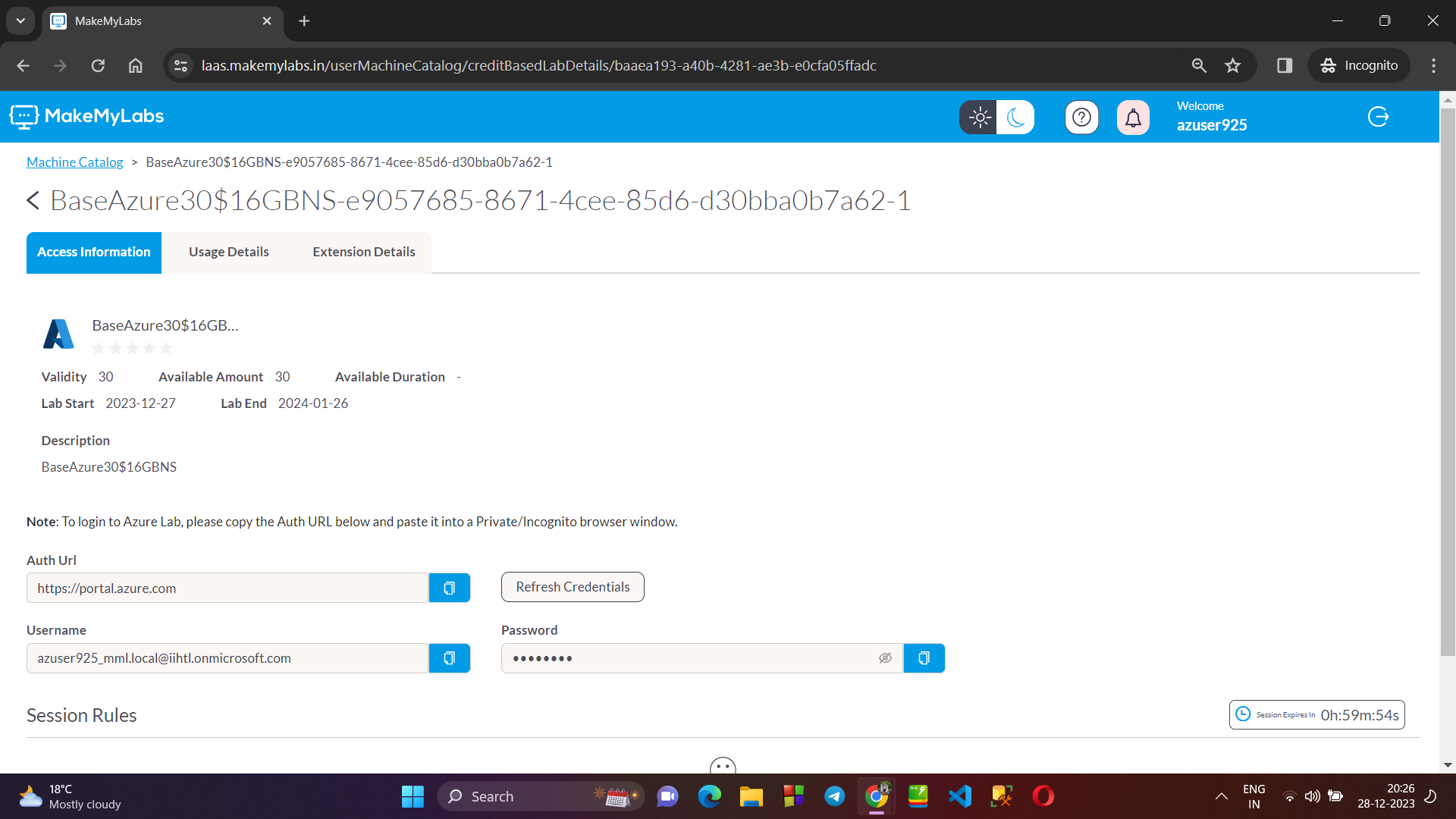
**STEP 1 :** Visit <https://laas.makemylabs.in/wsmLogin> and enter the username and password and login :

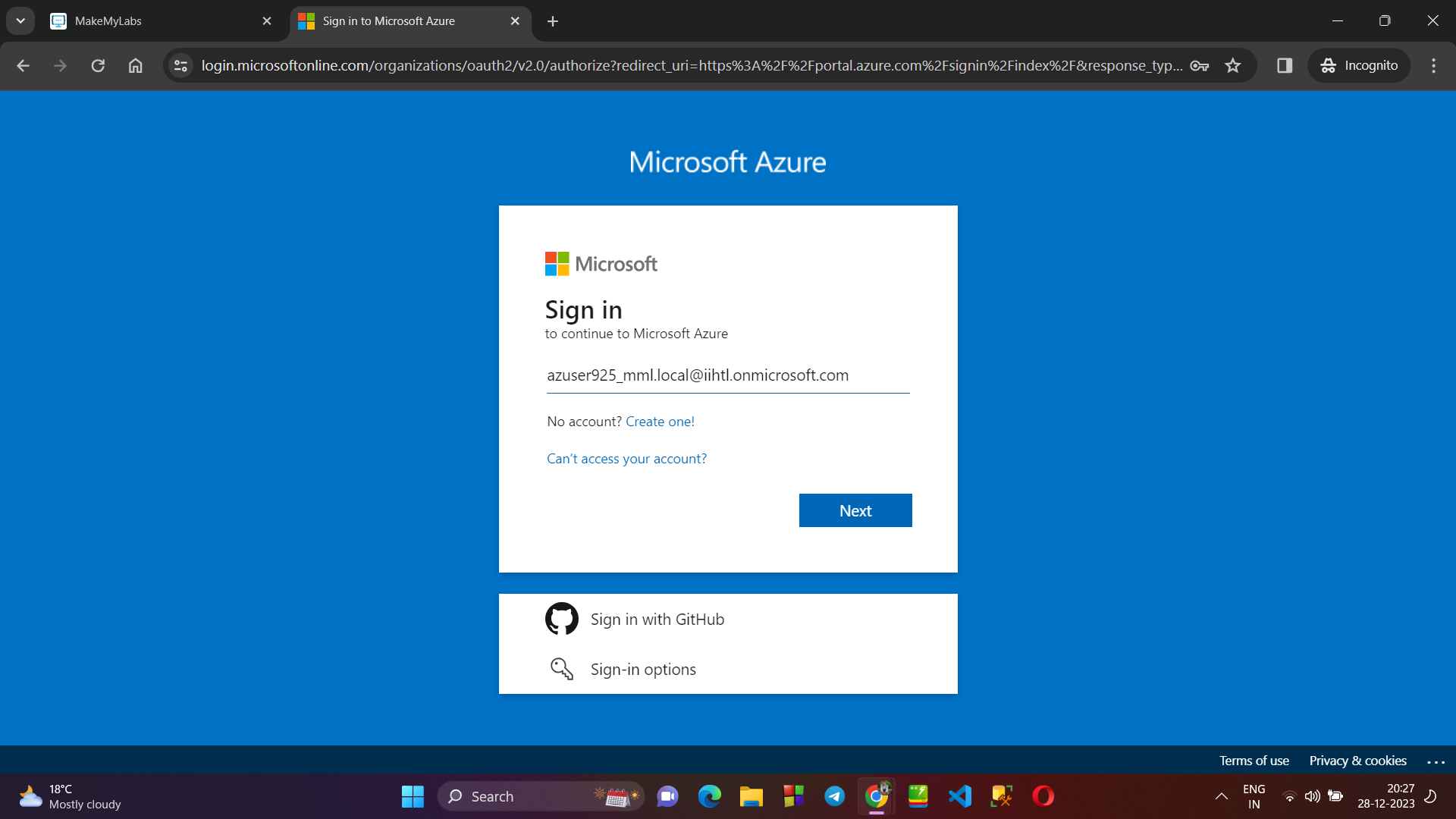


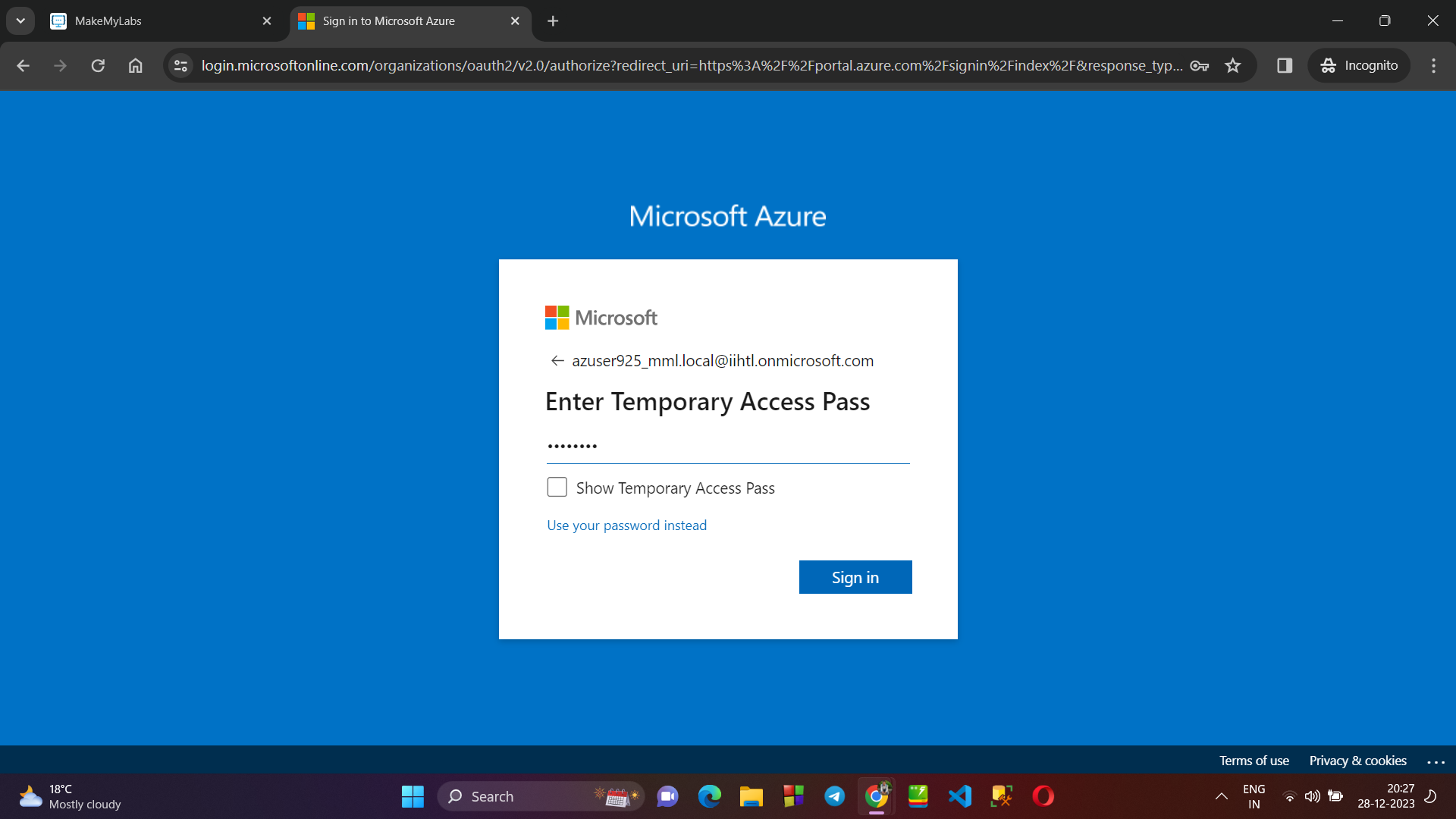
**STEP 2 :** Click on CONNECT :



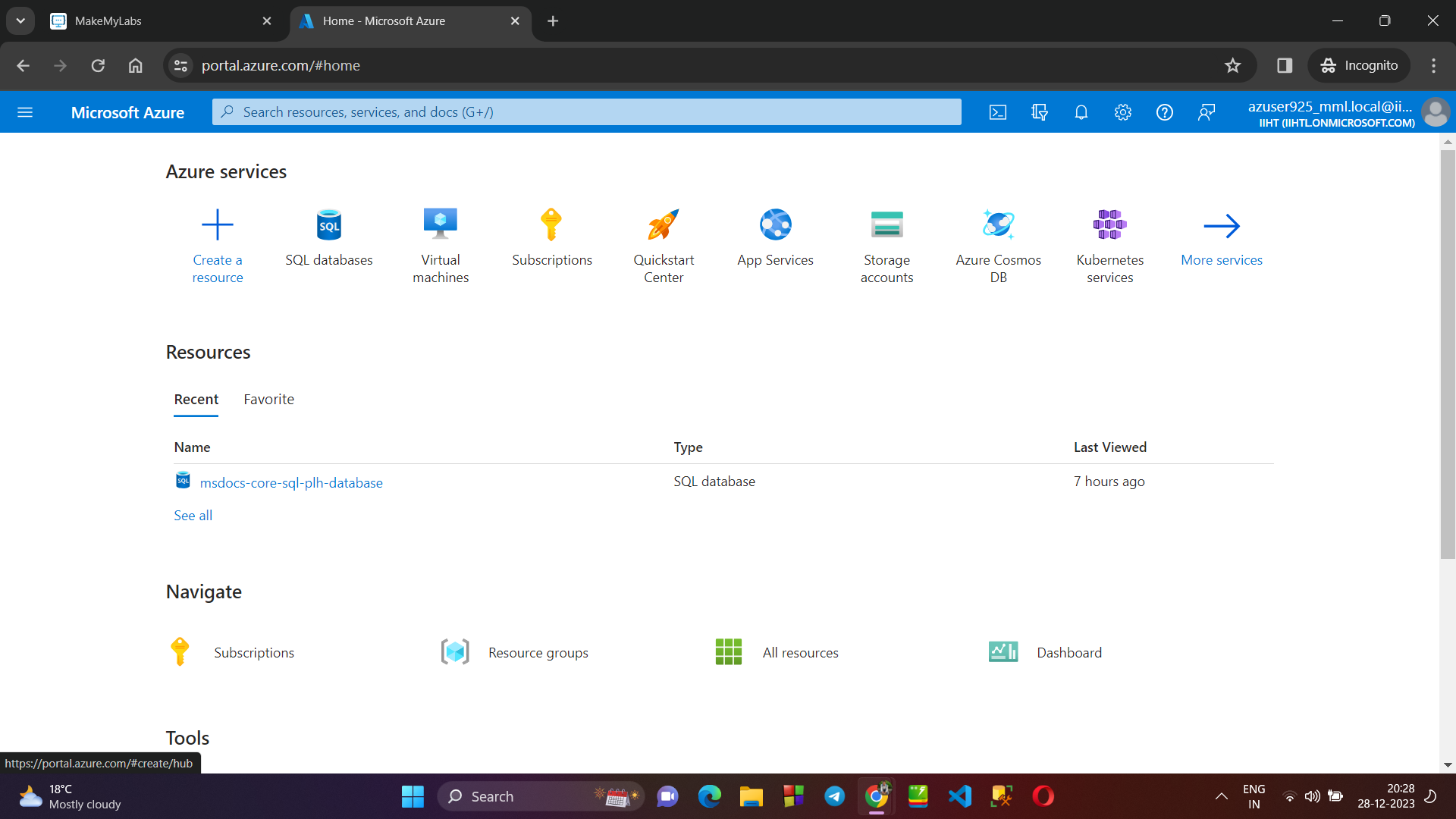
**STEP 3 :** Copy the Auth URL and open it new tab, copy username and password and enter :



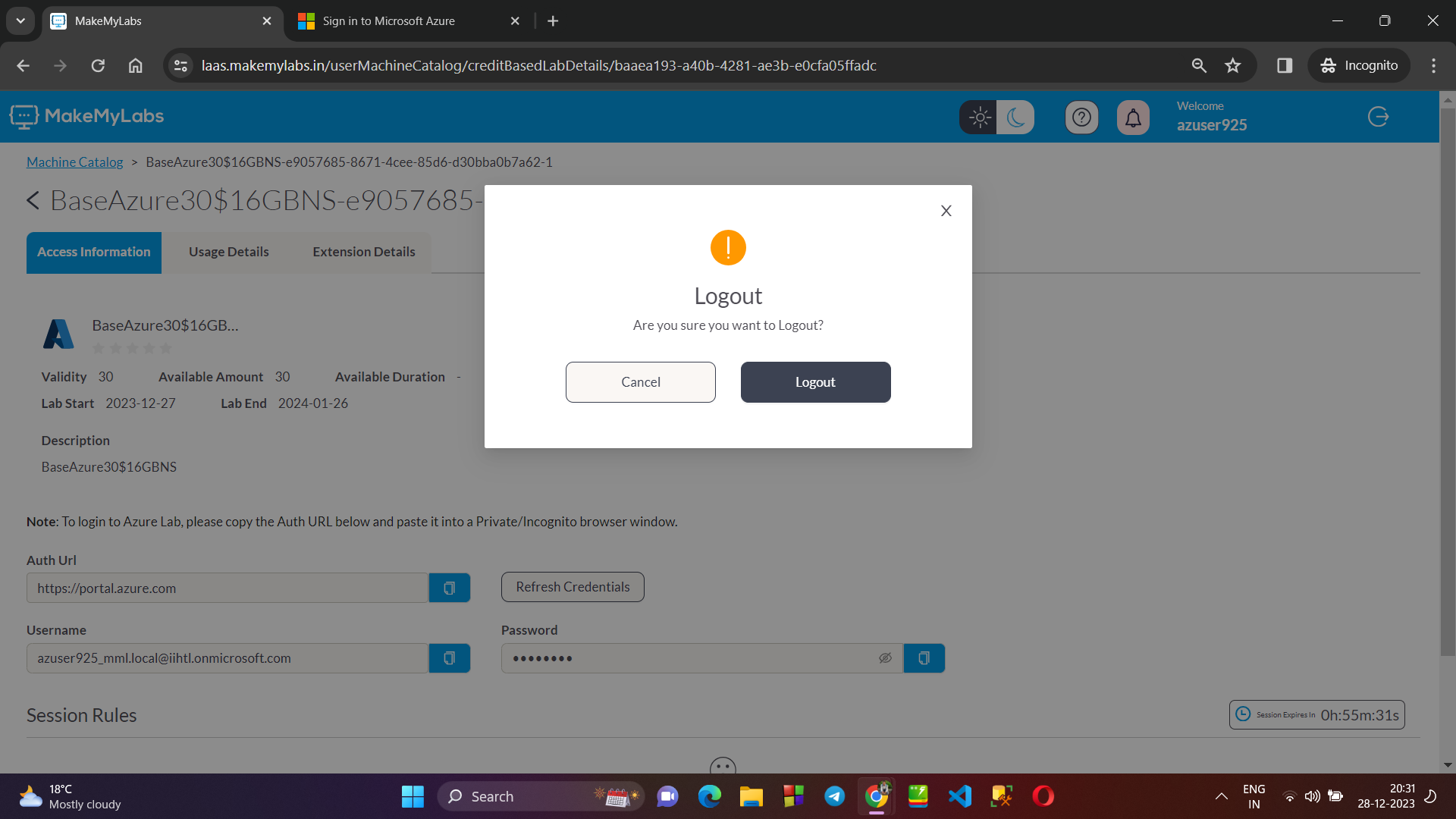
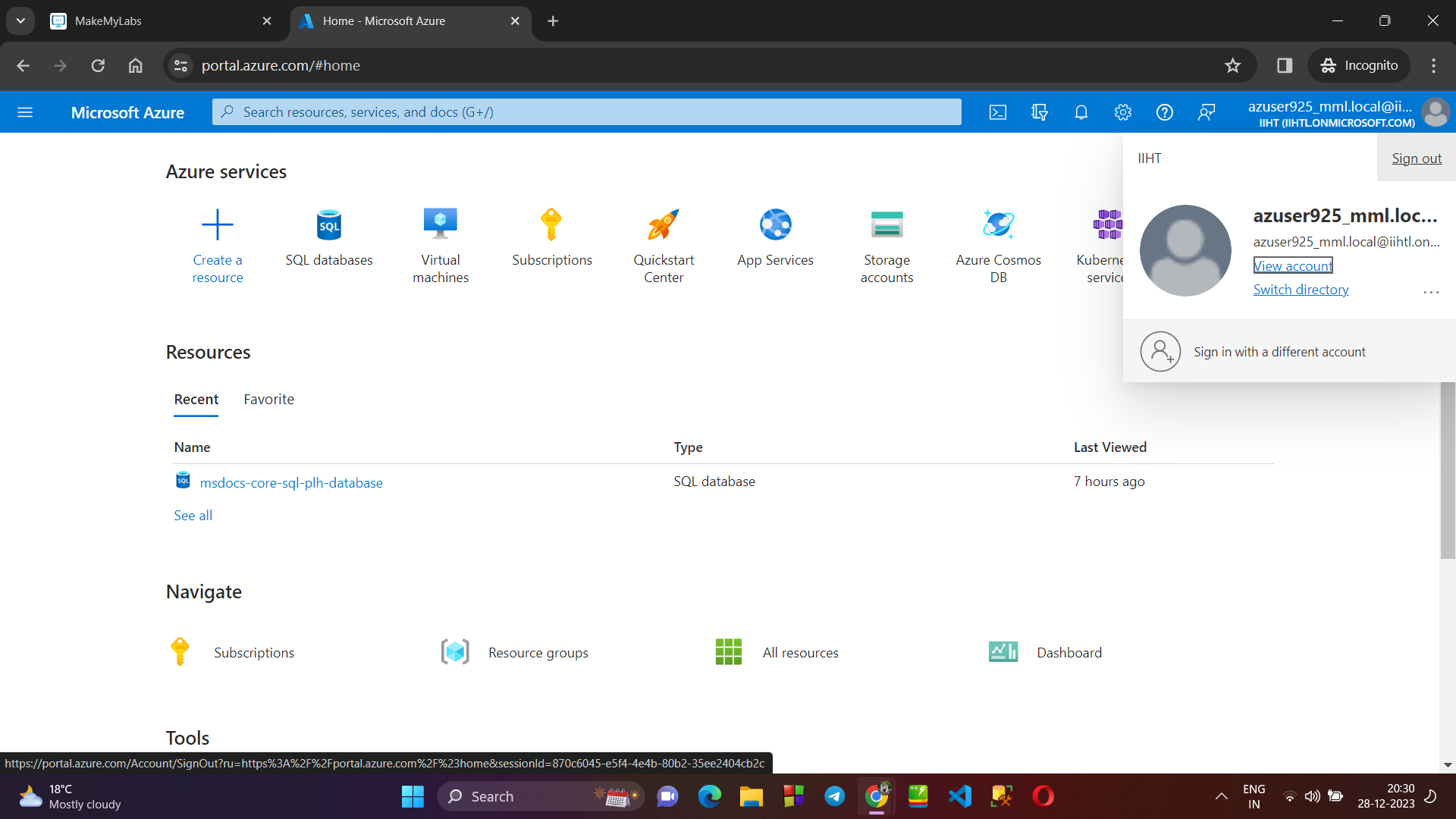




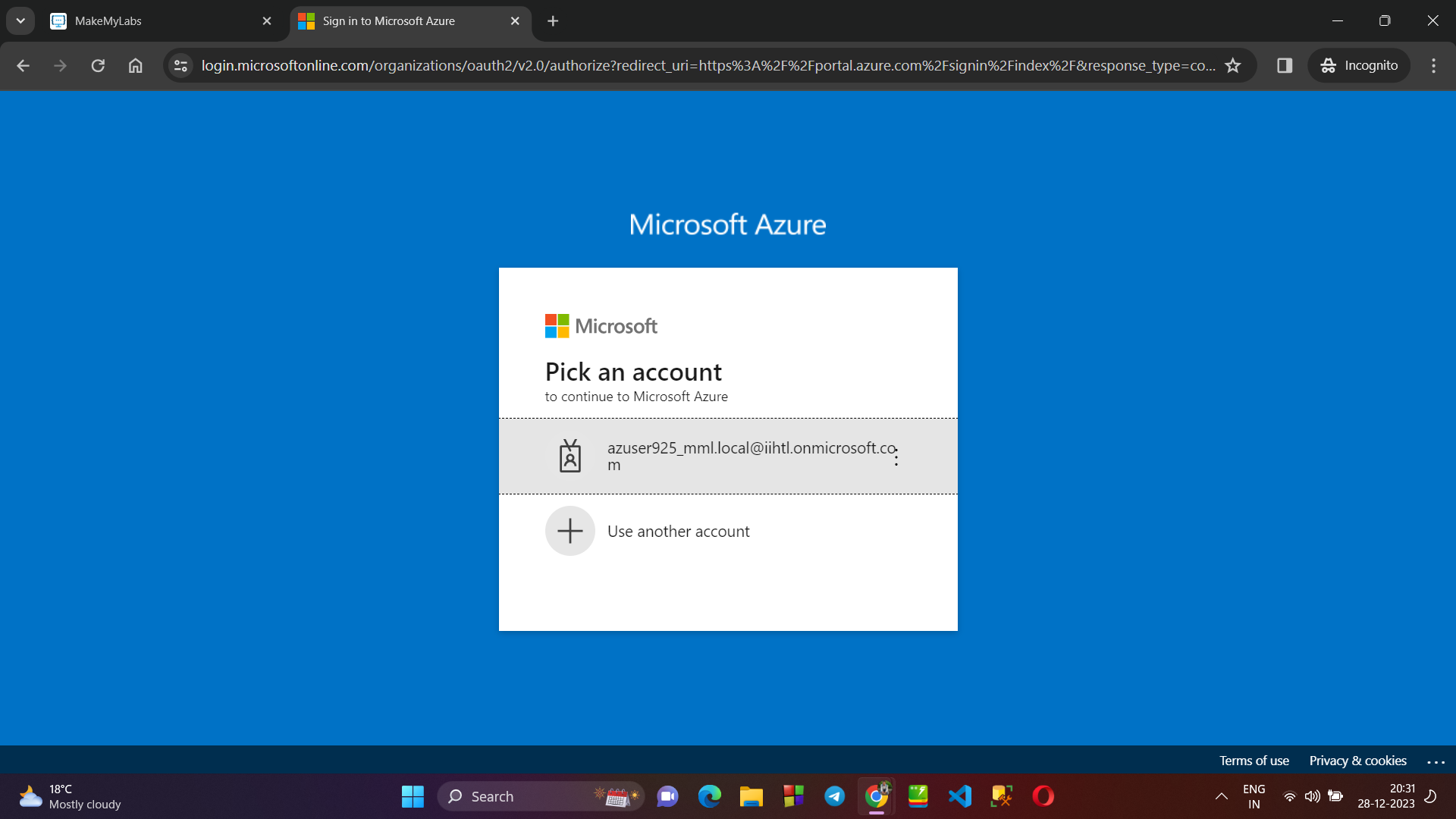
**STEP 3** : Now , the databricks setup is complete , now we can create resources using different services provided by Azure Databricks.



**STEP 4 :** As soon as we finish our task, we have to log out from the azure portal as well as the makemylabs.



**STEP 5 :** Next time, we can login back as and when required :



**→ Data Centers :**

* A data center is a facility that houses computer systems and related components.
* It stores, manages, and processes data for organizations.
* Includes servers, storage, networking equipment.
* Maintains a controlled environment for optimal performance.
* Provides security, power, and cooling to ensure data reliability and availability.

