

Akilesh K

k.akilesh123@gmail.com

Data engineering - Batch 1

Date: 12-02-24

DAY 16 – Azure –Databricks-Setup

- Databricks + Apache Spark + enterprise cloud = Azure Databricks
- It is a fully-managed version of the open-source Apache Spark data analytics and it features optimized connectors to storage platforms for the quickest possible data access.
- It offers a notebook-oriented Apache Spark as-a-service workspace environment which makes it easy to explore data interactively and manage clusters.
- It is secure cloud-based machine learning and big data platform.
- It is supporting multiple languages such as Scala, Python, R, Java, and SQL.

Login to labs



Workspace Login

Username

azuser1079@mml.local


Password

Login

Connect to azure



Workspace Dashboard



BaseAzure...

★★★★★ 0

[Rate this Lab](#)


BaseAzure30\$16GBNS....

[Read More](#)

Connect

Generate link

Access InformationUsage DetailsExtension Details



BaseAzure30\$16GB...

★★★★★

Validity 45Available Amount 30Available Duration -

Lab Start 2024-02-05Lab End 2024-03-21

DescriptionBaseAzure30\$16GBNS

Note: To login to Azure Lab, please copy the Auth URL below and paste it into a Private/Incognito browser window.

Auth Url

https://portal.azure.com

Refresh Credentials

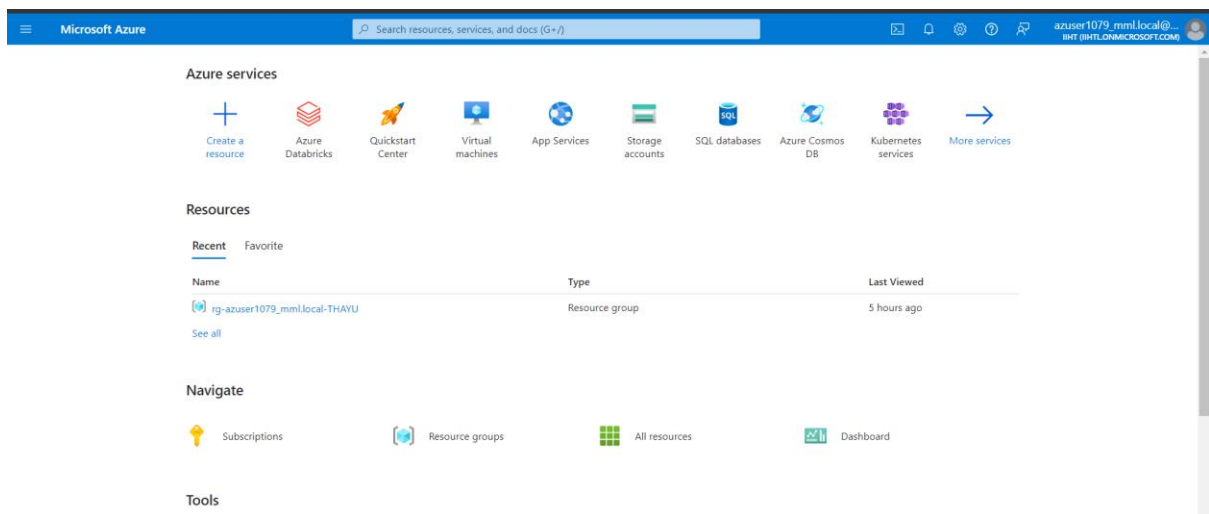
Usernameazuser1079_mml.local@iimt.onmicrosoft.com

Password

Session Rules

Session Expires In 0h:59m:54s

Login to azure cloud



Create a data bricks workspace

Create an Azure Databricks workspace ...

Basics Networking Encryption Tags Review + create

Project Details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.


Subscription *	<input type="text" value="Azure subscription 1"/>
Resource group *	<input type="text" value="rg-azuser1079_mm1.local-THAYU"/>

[Create new](#)

Instance Details

Workspace name *	<input type="text" value="hexa-deb-1079"/>
Region *	<input type="text" value="North Central US"/>
Pricing Tier *	<input type="text" value="Standard (Apache Spark, Secure with Microsoft Entra ID)"/>
Managed Resource Group name	<input type="text" value="Enter name for managed resource group"/>

Deployed along resource

 **rg-azuser1079_mml.local-THAYU_hexa-deb-1079** | Overview ✨ ...
Deployment

<< Delete Cancel Redeploy Download Refresh

Overview

Inputs

Outputs

Template

✔ Your deployment is complete

Deployment name : rg-azuser1079_mml.local-THAYU_hexa-deb-1079

Subscription : [Azure subscription 1](#)

Resource group : [rg-azuser1079_mml.local-THAYU](#)

Start time : 2/12/2024, 10:23:28 PM

Correlation ID : d33ede8d-527e-4e38-b392-c2c66039d075

> Deployment details


< Next steps

[Go to resource](#)

Give feedback

[Tell us about your experience with deployment](#)

Deployed

 **hexa-deb-1079** ✨ ☆ ...
Azure Databricks Service

<< Delete

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Settings

Virtual Network Peerings

Encryption

Networking

Properties

Locks

Automation

CLI / PS

Tasks (review)

Essentials

Status : Active

Resource group : [rg-azuser1079_mml.local-THAYU](#)

Location : North Central US

Subscription : [Azure subscription 1](#)


Subscription ID : 984f097c-963c-4eb6-a20d-839457ae9f08

Tags [\(edit\)](#) : [Add tags](#)

Managed Resource Group : [databricks-rg-hexa-deb-1079-gzm57dkk2qfau](#)

URL : [https://adb-5212341692865180.0.azuredatabricks.net](#)




Pricing Tier : [Standard \(Apache Spark, Secure with Microsoft Ent](#)



[Launch Workspace](#)

[Upgrade to Premium](#)

Create notebook

Microsoft Azure  **databricks** CTRL + P hexa-deb-1079   azuser1079_mml.local@iihtl.onmicr...

New

Workspace

Recents

Catalog

Workflows

Compute

Data Engineering

Job Runs

Machine Learning

Playground

Experiments

Features

Models

Serving

Workspace

Home

Workspace

Repos

Favorites

Trash

Workspace > Users >


azuser1079_mml.local@iihtl.onmicrosoft.com ☆

Name

Type

Owner

Created



This folder is empty

Create a cluster

Compute > Preview [Send feedback](#)

azuser1079_mml.local's Cluster

More Terminate

Configuration Notebooks (0) Libraries Event log Spark UI Driver logs Metrics Apps Spark compute UI - Master

Multi node Single node

Access mode Single user access

Single user azuser1079_mml.local

Performance

Databricks Runtime Version

13.3 LTS (includes Apache Spark 3.4.1, Scala 2.12)

Use Photon Acceleration

Worker type

Workers Current

Standard_DS3_v2 14 GB Memory, 4 Cores

8

0

Spot instances

Driver type

Standard_DS3_v2 14 GB Memory, 4 Cores

Enable autoscaling

Summary

8 Workers 112 GB Memory 32 Cores
1 Driver 14 GB Memory, 4 Cores
Runtime 13.3.x-scala2.12

Photon Standard_DS3_v2 13.5 DBU/h

Azure data bricks

Untitled Notebook 2024-02-12 22:33:43

Python

☆

File Edit View Run Help Last edit was now

New cell UI: OFF

Run all

azuser1079_mml.local's...

Cmd 1

1

Start typing or generate with AI (Ctrl + I)...

[Shift+Enter] to run

[Shift+Ctrl+Enter] to run selected text