Programming Assignment-2

Cloud Spark Prediction:

This project aims to develop a Python application leveraging the PySpark framework, operating within an Amazon Web Services (AWS) Elastic MapReduce (EMR) cluster. The primary objective is to concurrently train a machine learning model on EC2 instances using publicly available data to predict wine quality. Following training, the model is utilized to make predictions on wine quality. Docker is employed to construct a container image for the trained model, simplifying deployment processes.

Github: https://github.com/Yeswanth234/Programming-Assignment-

Docker: https://hub.docker.com/repository/docker/yn78/qulpred

EMR Cluster Steps:

1. Access the AWS Management Console by visiting https://aws.amazon.com/ and logging in with your credentials.

2. Navigate to the EC2 Dashboard by typing "EC2" into the search bar within the AWS services and selecting it.

3. Choose "Key Pairs" from the "Network & Security" section on the left-hand side of the EC2 Dashboard.

4. Generate a new Key Pair by clicking on the "Create Key Pair" button.

5. Provide a name for your Key Pair.

6. Download the Key Pair and save it as "predkey.pem".

7. Proceed to the EMR console and initiate the creation of an EMR cluster.

8. Utilize the EMR console to configure the Spark cluster and create four instances. Name and application:

Amazon EMR release: emr-5.33.0

Application bundle: Hadoop 2.10.1 Spark 2.4.7, Zippeline 0.9.0, and Yarn

By following images, we need to create EMR Cluster:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Running AWS in CMD:

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer screen

Description automatically generated

Docker:

A screen shot of a computer

Description automatically generated

A screen shot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated