CS643 CLOUD COMPUTING
PROGRAMMING ASSIGNMENT 2

Smit Hapani
(snh29)

Github Link :

Model training :
https://github.com/hapanismit/CloudComputing_pa2/blob/master/wineModelling.py

Prediction Application :
https://github.com/hapanismit/CloudComputing_pa2/blob/master/wineTesting.py

Docker hub Link :

https://hub.docker.com/repository/docker/snh29/pa2

Steps :

- **Model Training**
  - Create EMR Cluster on AWS using key pair
  - In Applications, choose Spark
  - Choose number of Instances as 6 (1 Master and 5 cores)
  - Enable SSH inbound rule for port 22 in security group of EMR Master
  - Connect to Master instance using command:
    - chmod 400 key.pem
    - ssh -i key.pem hadoop@master-public-dns
  - Write python code for model training
  - Run the python code using
    - spark-submit wineModelling.py

- **Prediction without Docker**
  - Create EMR Cluster on AWS using key pair
  - In Applications, choose Spark
  - Choose number of Instances as 1 (1 Master and 0 cores)
  - Enable SSH inbound rule for port 22 in security group of EMR Master
  - Connect to Master instance using command:
    - chmod 400 key.pem
    - ssh -i key.pem hadoop@master-public-dns

- Write python code for prediction application
- Run the python code using
  - spark-submit wineTesting.py

- **Prediction without Docker**
  - Sign up on Docker hub
  - Connect to EC2 using SSH command
  - Install docker on EC2
    - sudo yum update -y
    - sudo yum install docker
    - sudo service docker start
  - Create a docker file
  - Create docker image using command
    - sudo docker build . -f docker-file-name -t image-name
  - Run docker image using command
    - sudo docker run image-name
  - Log into the Docker Hub from the command line
    - docker login --username=yourusername --email=youremail@company.com
    - Enter your password when prompted
  - Tag your image
    - docker tag bb38976d03cf yourusername/repo-name
  - Push image to docker using command
    - sudo docker push yourusername/repo-name