**Links**

[**https://hub.docker.com/repository/docker/christopherc0923/wine-app/general**](https://hub.docker.com/repository/docker/christopherc0923/wine-app/general)

**https://github.com/njitchrisc0923/CS643-Assignment-2/**

**EC2**

https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#Home:

Launch 5 Instance (4 Worker & 1 Master Node)

Initialize EC2 Name

Amazon Linux

Amazon Linux 2023 AMI, 64-bit (x86) architecture

T2.medium

Add security group rule

All traffic

Anywhere

**EC2 Setup on both Instances**

sudo yum update -y

sudo yum install -y java-1.8.0-openjdk-devel python3

wget<https://archive.apache.org/dist/spark/spark-2.4.7/spark-2.4.7-bin-hadoop2.7.tgz>

tar xvf spark-2.4.7-bin-hadoop2.7.tgz

sudo mv spark-2.4.7-bin-hadoop2.7 /opt/spark

nano ~/.bashrc

export SPARK\_HOME=/opt/spark

export PATH=$PATH:/opt/spark/bin

export PYSPARK\_PYTHON=/usr/bin/python3

export PYSPARK\_DRIVER\_PYTHON=/usr/bin/python3

source ~/.bashrc

sudo pip3 install numpy pandas

**Master Node Only**

cd /opt/spark/conf

cp spark-env.sh.template spark-env.sh

nano /opt/spark/conf/spark-env.sh

export SPARK\_MASTER\_HOST='172.31.26.109’

export SPARK\_WORKER\_MEMORY=2g

export SPARK\_WORKER\_CORES=2

/opt/spark/sbin/start-master.sh

/opt/spark/sbin/stop-master.sh

**Worker Node Only**

cd /opt/spark/conf

cp spark-env.sh.template spark-env.sh

nano /opt/spark/conf/spark-env.sh

export SPARK\_MASTER\_HOST='172.31.22.109'

/opt/spark/sbin/start-slave.sh spark://172.31.26.109:7077

/opt/spark/sbin/stop-slave.sh

**Running Application**

**Parallel Training**

/opt/spark/bin/spark-submit --packages org.apache.hadoop:hadoop-aws:2.7.7,com.amazonaws:aws-java-sdk:1.7.4 --master spark:// 172.31.26.109:7077 TrainSimple.py

A screenshot of a computer

Description automatically generated

**Dockerless Prediction**

/opt/spark/bin/spark-submit --packages org.apache.hadoop:hadoop-aws:2.7.7,com.amazonaws:aws-java-sdk:1.7.4 --master local[4]  TestSimple.py







**Docker Prediction**

sudo yum install -y docker

sudo systemctl start docker

sudo usermod -aG docker $(whoami)

sudo docker build -t wine-app .

docker login

docker tag c569f32916d6 christopherc0923/wine-app:latest

docker push christopherc0923/wine-app:latest

docker pull christopherc0923/wine-app:latest

docker run -it christopherc0923/wine-app:latest

A screenshot of a computer

Description automatically generated

/opt/spark/bin/spark-submit --packages org.apache.hadoop:hadoop-aws:2.7.7,com.amazonaws:aws-java-sdk:1.7.4 --master local[4]  TrainSimple.py

/opt/spark/bin/spark-submit --packages org.apache.hadoop:hadoop-aws:2.7.7,com.amazonaws:aws-java-sdk:1.7.4 --master spark://172.31.25.16:7077 TrainSimple.py