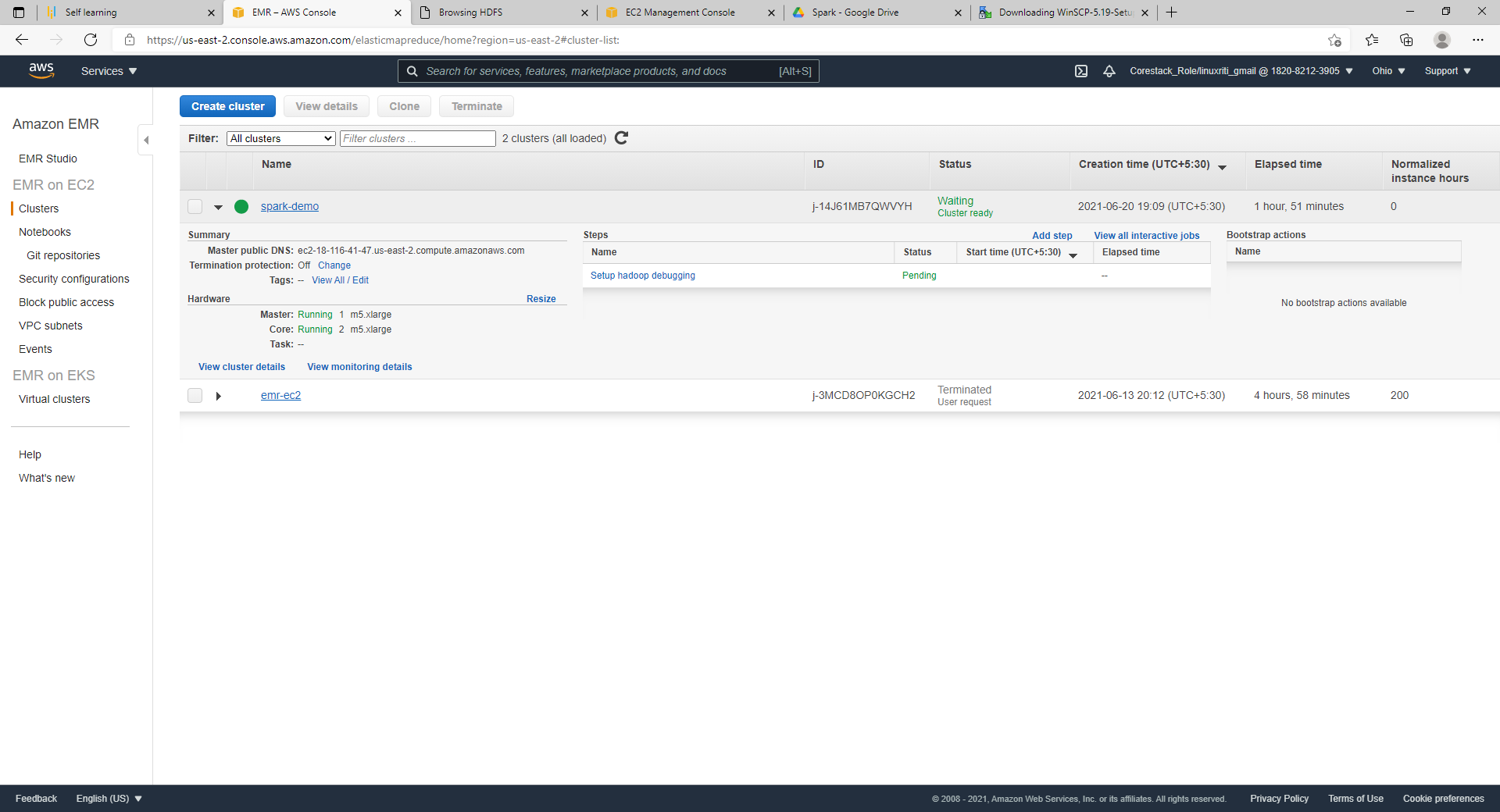
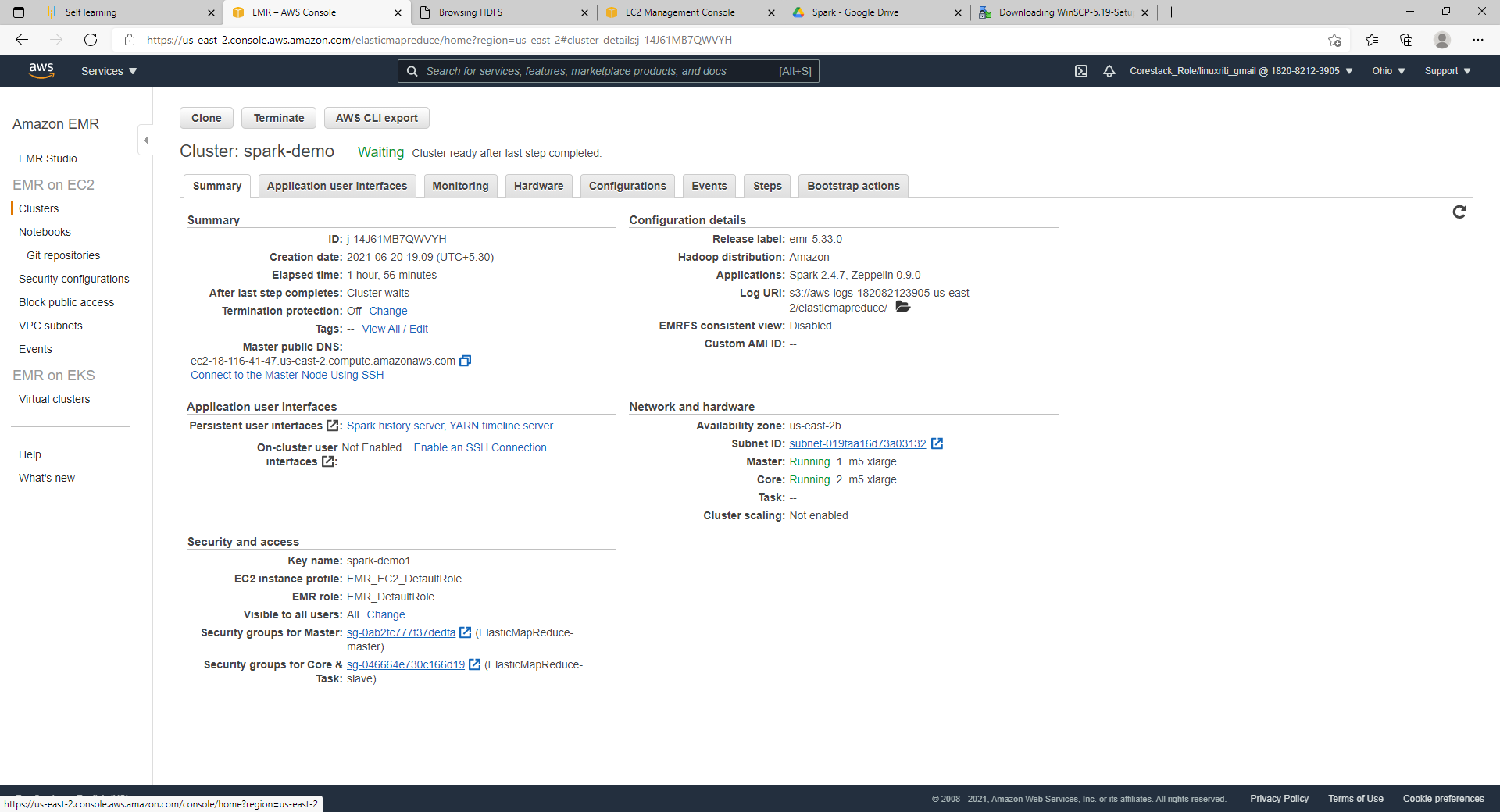
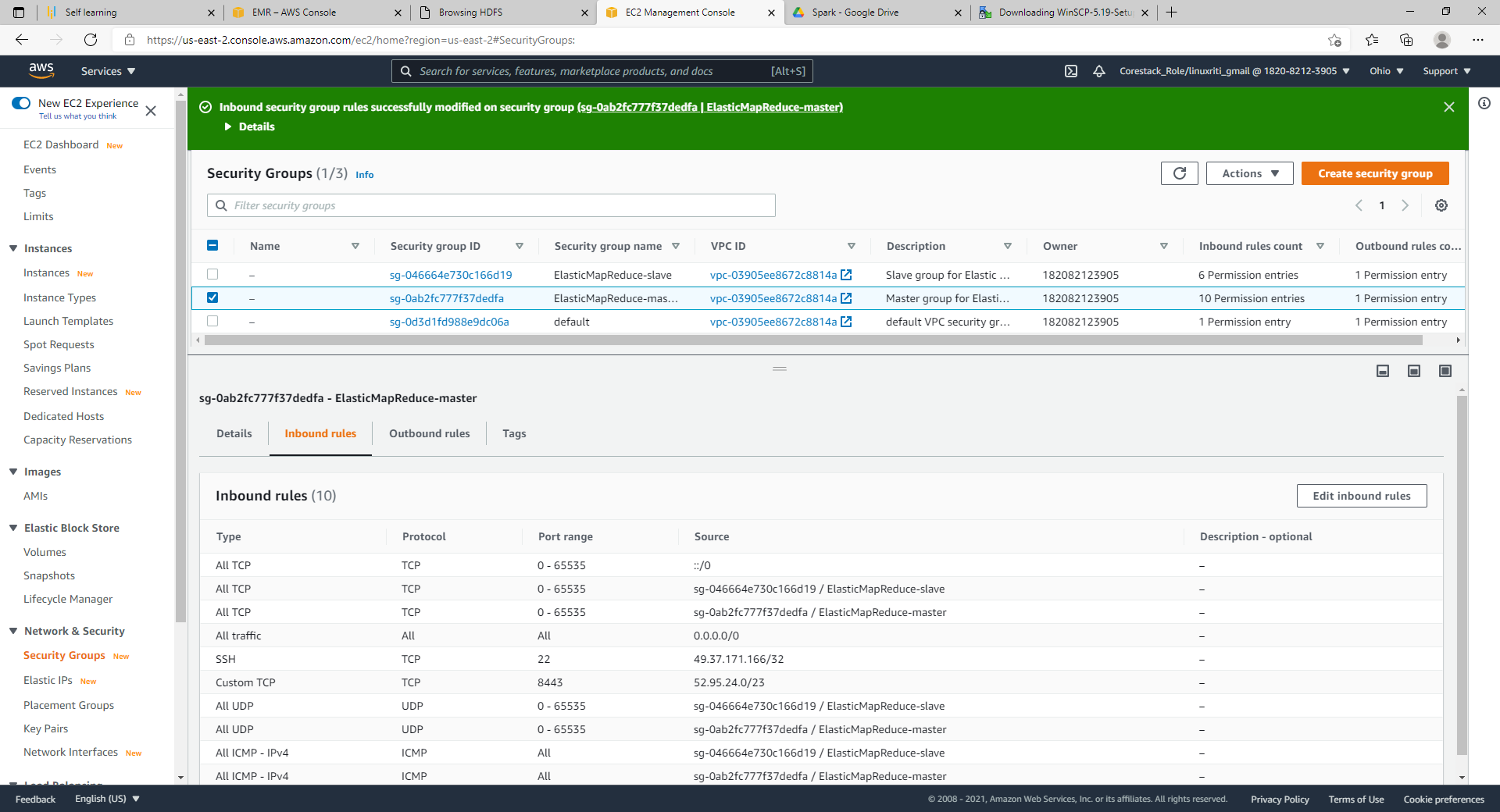
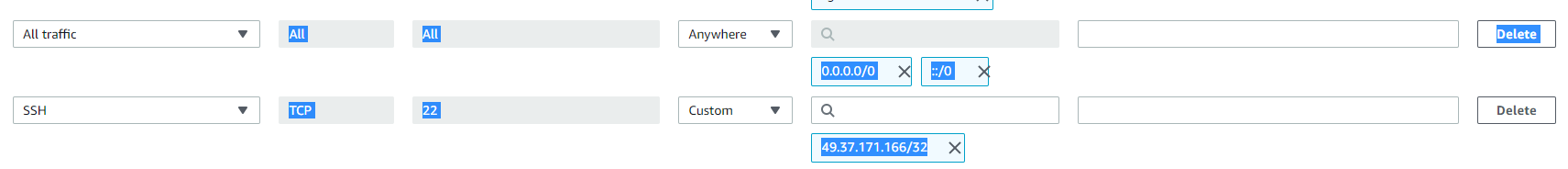
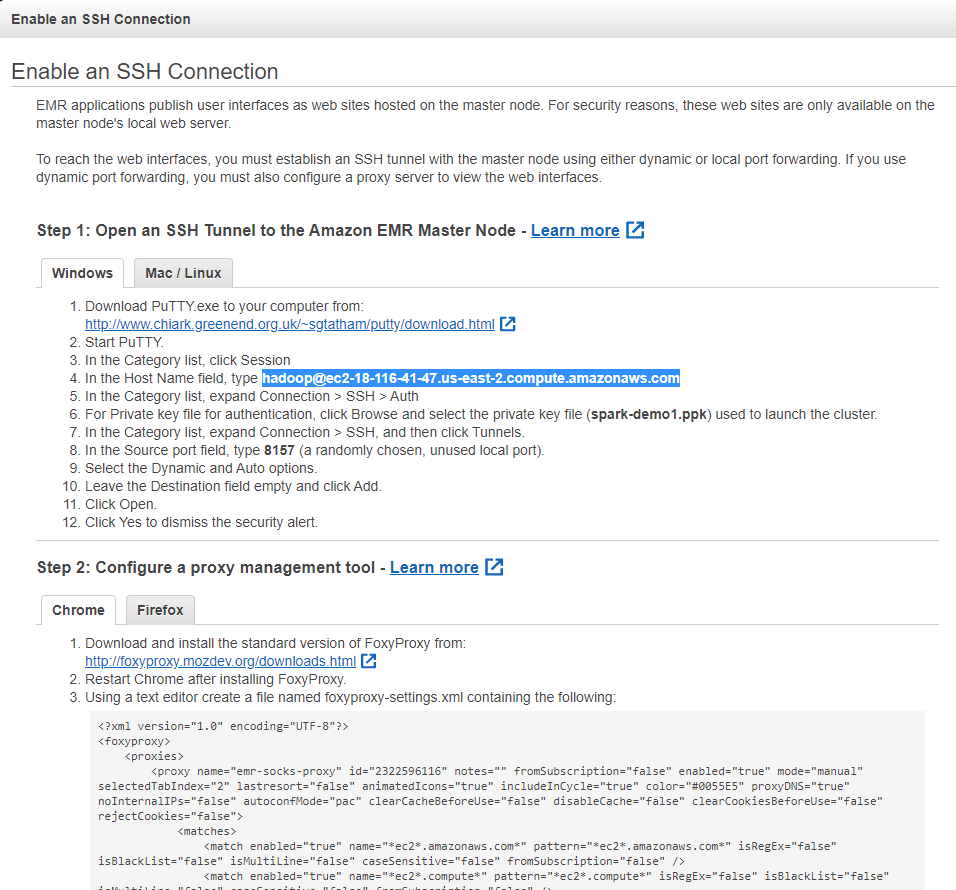
**AWS EMR -> Spark On EC2 HDFS**

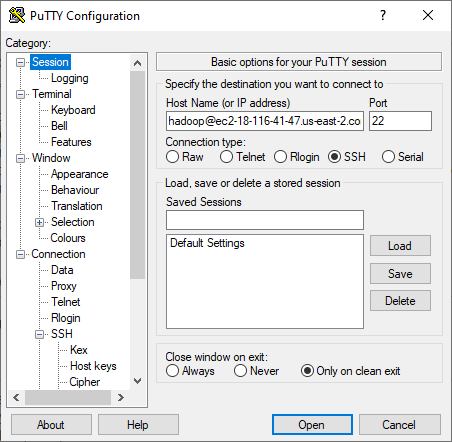


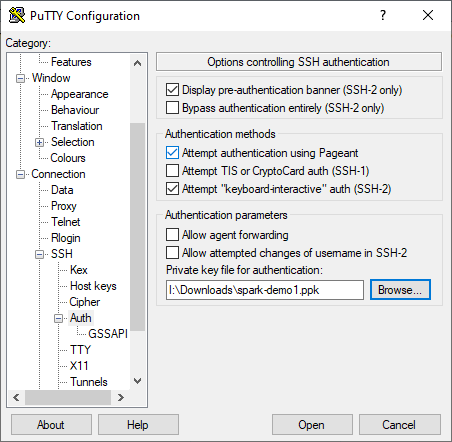












https://drive.google.com/drive/folders/1jq\_ihORB9Z-sCixpT9nbAg-944vOpu0p

Download below 2 files from above link

[hadoop@ip-172-31-27-251 ~]$ ll

total 36

-rw-rw-r-- 1 hadoop hadoop 3069 Jun 20 14:59 als\_example.py

-rw-rw-r-- 1 hadoop hadoop 32363 Jun 20 15:01 sample\_movielens\_ratings.txt

[hadoop@ip-172-31-27-251 ~]$

hdfs dfs -mkdir -p /user/hadoop/data/mllib/als

hdfs dfs -put sample\_movielens\_ratings.txt data/mllib/als

spark-submit als\_example.py

Root-mean-square error = 1.7119948432743584

+------+--------------------+

|userId| recommendations|

+------+--------------------+

| 12|[[46, 6.3836966],...|

| 1|[[29, 3.6966667],...|

| 6|[[25, 4.6179533],...|

| 3|[[51, 4.4775157],...|

| 4|[[83, 4.4641266],...|

| 8|[[29, 5.5296984],...|

| 11|[[46, 5.270946], ...|

| 19|[[90, 3.8684716],...|

| 23|[[7, 5.1011376], ...|

| 21|[[29, 4.8047805],...|

| 14|[[46, 6.4863787],...|

| 16|[[94, 5.6871014],...|

| 20|[[51, 3.64208], [...|

| 5|[[25, 5.201289], ...|

| 15|[[30, 5.2814465],...|

| 25|[[23, 4.3291316],...|

| 24|[[75, 5.1583447],...|

| 2|[[8, 5.008992], [...|

| 18|[[83, 5.0603013],...|

| 28|[[92, 5.02911], [...|

+------+--------------------+

only showing top 20 rows

+-------+--------------------+

|movieId| recommendations|

+-------+--------------------+

| 81|[[28, 4.933389], ...|

| 65|[[23, 4.8755374],...|

| 76|[[14, 4.8879633],...|

| 1|[[12, 3.2287781],...|

| 12|[[26, 6.4800053],...|

| 6|[[26, 2.8247209],...|

| 92|[[26, 6.004135], ...|

| 64|[[12, 4.726378], ...|

| 3|[[14, 2.57447], [...|

| 61|[[10, 2.9048662],...|

| 35|[[12, 4.8630567],...|

| 37|[[11, 2.0760088],...|

| 4|[[26, 3.6528506],...|

| 87|[[9, 3.7570255], ...|

| 8|[[2, 5.008992], [...|

| 60|[[24, 3.1212084],...|

| 90|[[17, 5.257891], ...|

| 11|[[6, 2.1886692], ...|

| 33|[[25, 4.1074724],...|

| 66|[[11, 4.1396217],...|

+-------+--------------------+

only showing top 20 rows

+------+--------------------+

|userId| recommendations|

+------+--------------------+

| 21|[[29, 4.8047805],...|

| 25|[[23, 4.3291316],...|

| 27|[[18, 3.5621157],...|

+------+--------------------+

+-------+--------------------+

|movieId| recommendations|

+-------+--------------------+

| 65|[[23, 4.8755374],...|

| 94|[[16, 5.6871014],...|

| 77|[[22, 5.5942373],...|

+-------+--------------------+