

## CS 643 CLOUD COMPUTING PROGRAMMING ASSIGNMENT-2


-Ranjith Kumar Kotamraju rk939

Github link : [https://github.com/ranjit1706/cc\\_Assignment2](https://github.com/ranjit1706/cc_Assignment2)

Docker link : <https://hub.docker.com/r/ranjit1706/wine-quality-testing>

Training dataset:

← → ↻ 🔒 Not secure 3.91.134.163:8080

 **Spark Master at spark://ip-172-31-84-45.ec2.internal:7077**

URL: spark://ip-172-31-84-45.ec2.internal:7077  
Alive Workers: 4  
Cores in use: 8 Total, 8 Used  
Memory in use: 11.3 GiB Total, 5.0 GiB Used  
Resources in use:  
Applications: 1 Running, 0 Completed  
Drivers: 1 Running, 0 Completed  
Status: ALIVE

Workers (4)

Worker Id	Address	State	Cores	Memory	Resources
worker-20241210033025-172.31.81.139-40701	172.31.81.139:40701	ALIVE	2 (2 Used)	2.8 GiB (2.0 GiB Used)	
worker-20241210033025-172.31.84.119-35575	172.31.84.119:35575	ALIVE	2 (2 Used)	2.8 GiB (1024.0 MiB Used)	
worker-20241210033025-172.31.84.45-44499	172.31.84.45:44499	ALIVE	2 (2 Used)	2.8 GiB (1024.0 MiB Used)	
worker-20241210034945-172.31.92.27-44479	172.31.92.27:44479	ALIVE	2 (2 Used)	2.8 GiB (1024.0 MiB Used)	


Running Applications (1)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
app-20241210040231-0000	(kill) Wine Quality Prediction	7	1024.0 MiB		2024/12/10 04:02:31	ubuntu	RUNNING	1 s

Running Drivers (1)

Submission ID	Submitted Time	Worker	State	Cores	Memory	Resources	Main Class	Duration
driver-20241210040227-0000	(kill) 2024/12/10 04:02:27	worker-20241210033025-172.31.81.139-40701	RUNNING	1	1024.0 MiB		com.example.WineQualityPrediction	5 s

← → ↻ 🔒 Not secure 3.91.134.163:8080

 **Spark Master at spark://ip-172-31-84-45.ec2.internal:7077**

Applications: 0 Running, 1 Completed  
Drivers: 0 Running, 1 Completed  
Status: ALIVE

Workers (4)

Worker Id	Address	State	Cores	Memory	Resources
worker-20241210033025-172.31.81.139-40701	172.31.81.139:40701	ALIVE	2 (0 Used)	2.8 GiB (0.0 B Used)	
worker-20241210033025-172.31.84.119-35575	172.31.84.119:35575	ALIVE	2 (0 Used)	2.8 GiB (0.0 B Used)	
worker-20241210033025-172.31.84.45-44499	172.31.84.45:44499	ALIVE	2 (0 Used)	2.8 GiB (0.0 B Used)	
worker-20241210034945-172.31.92.27-44479	172.31.92.27:44479	ALIVE	2 (0 Used)	2.8 GiB (0.0 B Used)	

Running Applications (0)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
----------------	------	-------	---------------------	------------------------	----------------	------	-------	----------

Running Drivers (0)

Submission ID	Submitted Time	Worker	State	Cores	Memory	Resources	Main Class	Duration
---------------	----------------	--------	-------	-------	--------	-----------	------------	----------

Completed Applications (1)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
app-20241210040231-0000	Wine Quality Prediction	7	1024.0 MiB		2024/12/10 04:02:31	ubuntu	FINISHED	42 s

Completed Drivers (1)

Submission ID	Submitted Time	Worker	State	Cores	Memory	Resources	Main Class
driver-20241210040227-0000	2024/12/10 04:02:27	worker-20241210033025-172.31.81.139-40701	FINISHED	1	1024.0 MiB		com.example.WineQualityPrediction

## Result :

**F1 Score: 0.8104636591478698**

```
24/12/10 21:29:46 INFO Executor: Finished task 0.0 in stage 7.0 (TID 6). 1916 bytes result sent to driver
24/12/10 21:29:46 INFO TaskSetManager: Finished task 0.0 in stage 7.0 (TID 6) in 353 ms on b03ae48739d7 (executor driver) (1/1)
24/12/10 21:29:46 INFO TaskSchedulerImpl: Removed TaskSet 7.0, whose tasks have all completed, from pool
24/12/10 21:29:46 INFO DAGScheduler: ShuffleMapStage 7 (map at MulticlassMetrics.scala:52) finished in 0.399 s
24/12/10 21:29:46 INFO DAGScheduler: looking for newly runnable stages
24/12/10 21:29:46 INFO DAGScheduler: running: Set()
24/12/10 21:29:46 INFO DAGScheduler: waiting: Set(ResultStage 8)
24/12/10 21:29:46 INFO DAGScheduler: failed: Set()
24/12/10 21:29:46 INFO DAGScheduler: Submitting ResultStage 8 (ShuffledRDD[35] at reduceByKey at MulticlassMetrics.scala:61), which has no missing parents
24/12/10 21:29:46 INFO MemoryStore: Block broadcast_13 stored as values in memory (estimated size 4.7 MiB, free 433.8 MiB)
24/12/10 21:29:46 INFO MemoryStore: Block broadcast_13_piece0 stored as bytes in memory (estimated size 2.8 MiB, free 433.8 MiB)
24/12/10 21:29:46 INFO BlockManagerInfo: Added broadcast_13_piece0 in memory on b03ae48739d7:34111 (size: 2.8 MiB, free: 434.3 MiB)
24/12/10 21:29:46 INFO SparkContext: Created broadcast 13 from broadcast at DAGScheduler.scala:1535
24/12/10 21:29:46 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 8 (ShuffledRDD[35] at reduceByKey at MulticlassMetrics.scala:61) (first 15
tasks are for partitions Vector(0))
24/12/10 21:29:46 INFO TaskSchedulerImpl: Adding task set 8.0 with 1 tasks resource profile 0
24/12/10 21:29:46 INFO TaskSetManager: Starting task 0.0 in stage 8.0 (TID 7) (b03ae48739d7, executor driver, partition 0, NODE_LOCAL, 7181 bytes)
24/12/10 21:29:46 INFO Executor: Running task 0.0 in stage 8.0 (TID 7)
24/12/10 21:29:46 INFO ShuffleBlockFetcherIterator: Getting 1 (304.0 B) non-empty blocks including 1 (304.0 B) local and 0 (0.0 B) host-local and 0 (0.0 B)
push-merged-local and 0 (0.0 B) remote blocks
24/12/10 21:29:46 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 3 ms
24/12/10 21:29:46 INFO Executor: Finished task 0.0 in stage 8.0 (TID 7). 2068 bytes result sent to driver
24/12/10 21:29:46 INFO TaskSetManager: Finished task 0.0 in stage 8.0 (TID 7) in 29 ms on b03ae48739d7 (executor driver) (1/1)
24/12/10 21:29:46 INFO TaskSchedulerImpl: Removed TaskSet 8.0, whose tasks have all completed, from pool
24/12/10 21:29:46 INFO DAGScheduler: ResultStage 8 (collectAsMap at MulticlassMetrics.scala:61) finished in 0.052 s
24/12/10 21:29:46 INFO DAGScheduler: Job 6 is finished. Cancelling potential speculative or zombie tasks for this job
24/12/10 21:29:46 INFO TaskSchedulerImpl: Killing all running tasks in stage 8: Stage finished
24/12/10 21:29:46 INFO DAGScheduler: Job 6 finished: collectAsMap at MulticlassMetrics.scala:61, took 0.468697 s
F1 Score: 0.8104636591478698
24/12/10 21:29:46 INFO SparkContext: SparkContext is stopping with exitCode 0.
24/12/10 21:29:46 INFO SparkUI: Stopped Spark web UI at http://b03ae48739d7:4040
24/12/10 21:29:46 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
24/12/10 21:29:46 INFO MemoryStore: MemoryStore cleared
24/12/10 21:29:46 INFO BlockManager: BlockManager stopped
24/12/10 21:29:46 INFO BlockManagerMaster: BlockManagerMaster stopped
24/12/10 21:29:46 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
24/12/10 21:29:46 INFO SparkContext: Successfully stopped SparkContext
24/12/10 21:29:46 INFO ShutdownHookManager: Shutdown hook called
24/12/10 21:29:46 INFO ShutdownHookManager: Deleting directory /tmp/spark-d454d585-1bb7-495f-ae2d-430264deb5d0
24/12/10 21:29:46 INFO ShutdownHookManager: Deleting directory /tmp/spark-524569fd-4d4a-48bd-afee-e1dcf7fb2c0b
ubuntu@ip-172-31-84-45:~$
```