

# LLM Batch Processing Using Spark Distributed Systems

Tweet Sentiment Analysis

# Prerequisites

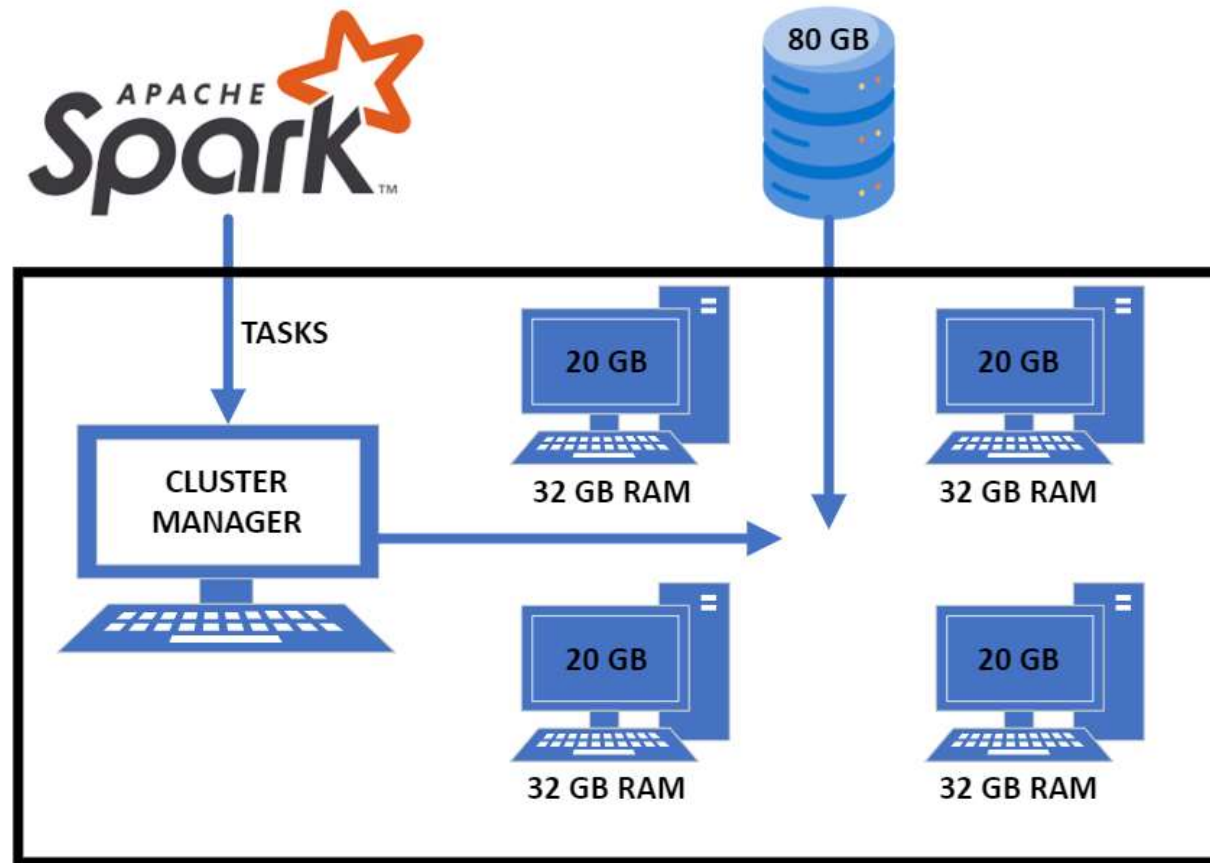
- Spark
- Multiple Computers (Single Computer too can work but Multi-node cannot be displayed)
- Transformers Library
- Google/flan-t5-small model (Will be downloaded automatically when you run the program once)

# Contents

- Tweets of political parties (CSV)
- Tweets extraction Guide
- Walkthrough Guide
- Code for Sentimental Analysis (`my_spark_app.py`)

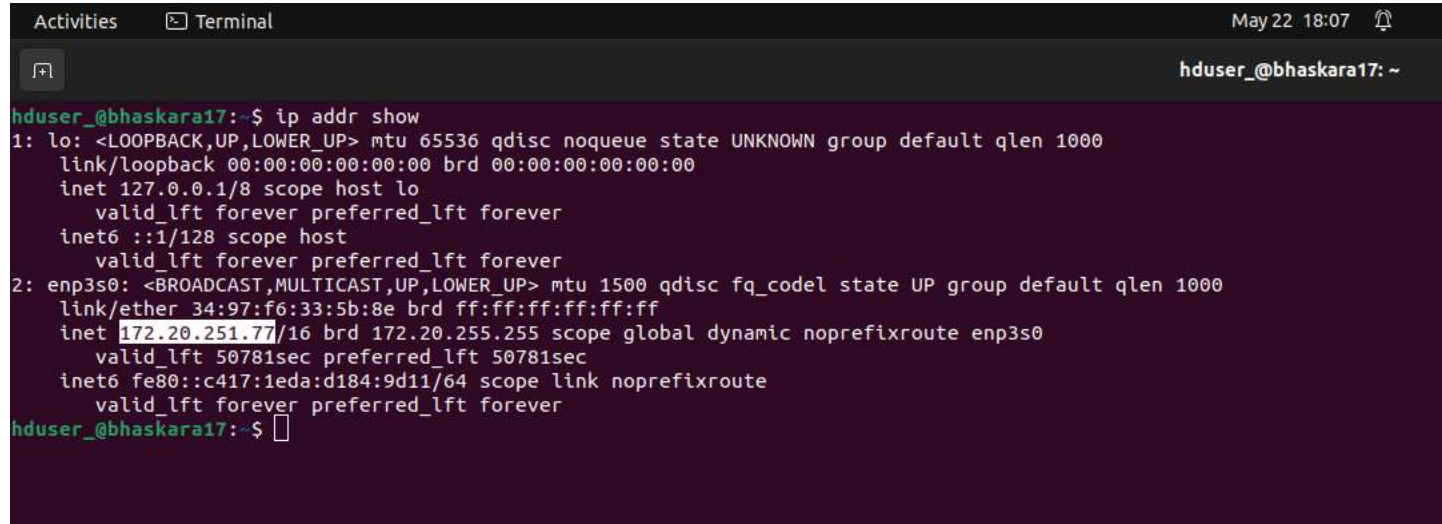
Note: You can also create your own Dataframe with own content (See Code)

# WorkFlow



# Setting Up the Master Node

- **ip addr show** – To Check the Master Node IP Address



```
Activities Terminal May 22 18:07
hduser_@bhaskara17: ~
hduser_@bhaskara17:~$ ip addr show
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: enp3s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 34:97:f6:33:5b:8e brd ff:ff:ff:ff:ff:ff
    inet 172.20.251.77/16 brd 172.20.255.255 scope global dynamic noprefixroute enp3s0
        valid_lft 50781sec preferred_lft 50781sec
    inet6 fe80::c417:1eda:d184:9d11/64 scope link noprefixroute
        valid_lft forever preferred_lft forever
hduser_@bhaskara17:~$
```

- **cd \$SPARK\_HOME/conf**
- **cp spark-env.sh.template spark-env.sh**
- **export SPARK\_MASTER\_HOST=<give your own ip address>**
- **export SPARK\_MASTER\_PORT=7077**
- **export SPARK\_WORKER\_CORES=3**
- **export SPARK\_WORKER\_MEMORY=3g**
- **\$SPARK\_HOME/sbin/start-master.sh**

```

hduser_@bhaskara17:/opt/spark/conf$ cp spark-env.sh.template spark-env.sh
hduser_@bhaskara17:/opt/spark/conf$ export SPARK_MASTER_HOST= 172.20.251.77
bash: export: `172.20.251.77': not a valid identifier
hduser_@bhaskara17:/opt/spark/conf$ export SPARK_MASTER_HOST= <172.20.251.77>
bash: syntax error near unexpected token `newline'
hduser_@bhaskara17:/opt/spark/conf$ export SPARK_MASTER_HOST=172.20.251.77
hduser_@bhaskara17:/opt/spark/conf$ export SPARK_MASTER_PORT=7077
hduser_@bhaskara17:/opt/spark/conf$ export SPARK_WORKER_CORES=3
hduser_@bhaskara17:/opt/spark/conf$ export SPARK_WORKER_MEMORY=3g
hduser_@bhaskara17:/opt/spark/conf$ $SPARK_HOME/sbin/start-master.sh
starting org.apache.spark.deploy.master.Master, logging to /opt/spark/logs/spark-hduser_-org.apache.spark.deploy.master.Master-1-bhaskara17.out
hduser_@bhaskara17:/opt/spark/conf$ █

```

- cat /opt/spark/logs/spark-hduser\_-org.apache.spark.deploy.master.Master-1-bhaskara17.out

```

hduser_@bhaskara17:/opt/spark/conf$ cat /opt/spark/logs/spark-hduser_-org.apache.spark.deploy.master.Master-1-bhaskara17.out
Spark Command: /usr/lib/jvm/java-1.17.0-openjdk-amd64/bin/java -cp /opt/spark/conf/:/opt/spark/jars/* -Xmx1g org.apache.spark.deploy.master.Master --host 172.20.251.77 --port 7077 --webui-port 8080
=====
Using Spark's default log4j profile: org/apache/spark/log4j2-defaults.properties
24/05/22 18:12:27 INFO Master: Started daemon with process name: 532633@bhaskara17
24/05/22 18:12:27 INFO SignalUtils: Registering signal handler for TERM
24/05/22 18:12:27 INFO SignalUtils: Registering signal handler for HUP
24/05/22 18:12:27 INFO SignalUtils: Registering signal handler for INT
24/05/22 18:12:27 WARN Utils: Your hostname, bhaskara17 resolves to a loopback address: 127.0.1.1; using 172.20.251.77 instead (on interface enp3s0)
24/05/22 18:12:27 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
24/05/22 18:12:27 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
24/05/22 18:12:28 INFO SecurityManager: Changing view acls to: hduser_
24/05/22 18:12:28 INFO SecurityManager: Changing modify acls to: hduser_
24/05/22 18:12:28 INFO SecurityManager: Changing view acls groups to:
24/05/22 18:12:28 INFO SecurityManager: Changing modify acls groups to:
24/05/22 18:12:28 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: hduser_; groups with view permissions: EMPTY; users with modify permissions: hduser_; groups with modify permissions: EMPTY
24/05/22 18:12:28 INFO Utils: Successfully started service 'sparkMaster' on port 7077.
24/05/22 18:12:28 INFO Master: Starting Spark master at spark://172.20.251.77:7077
24/05/22 18:12:28 INFO Master: Running Spark version 3.4.2
24/05/22 18:12:28 INFO JettyUtils: Start Jetty 0.0.0.0:8080 for MasterUI
24/05/22 18:12:28 INFO Utils: Successfully started service 'MasterUI' on port 8080.
24/05/22 18:12:28 INFO MasterWebUI: Bound MasterWebUI to 0.0.0.0, and started at http://172.20.251.77:8080
24/05/22 18:12:28 INFO Master: I have been elected leader! New state: ALIVE
hduser_@bhaskara17:/opt/spark/conf$ █


```

On accessing the highlighted section, it takes you to Spark UI which lets you see worker nodes status and Job application status.

Activities Firefox Web Browser May 22 18:15

Spark Master at spark:// x +

172.20.251.77:8080

 Spark Master at spark://172.20.251.77:7077

URL: spark://172.20.251.77:7077  
Alive Workers: 1  
Cores in use: 3 Total, 0 Used  
Memory in use: 3.0 GiB Total, 0.0 B Used  
Resources in use:  
Applications: 0 Running, 0 Completed  
Drivers: 0 Running, 0 Completed  
Status: ALIVE

Workers (1)

Worker Id	Address	State	Cores	Memory	Resources
worker-20240522181512-172.20.251.77-40823	172.20.251.77:40823	ALIVE	3 (0 Used)	3.0 GiB (0.0 B Used)	

Running Applications (0)

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

Completed Applications (0)

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



# Setting Up Worker Nodes

# Note: Master Node can also act as Worker Node !

- `cd $SPARK_HOME/conf`
- `cp spark-env.sh.template spark-env.sh`
- `export SPARK_WORKER_CORES=3`
- `export SPARK_WORKER_MEMORY=3g`
- `$SPARK_HOME/sbin/start-worker.sh spark://<master node ip>:7077`


```
hduser_@bhaskara18:/opt/spark/conf$ sudo cp spark-env.sh.template spark-env.sh
[sudo] password for hduser_:
hduser_@bhaskara18:/opt/spark/conf$ export SPARK_WORKER_CORES=3
hduser_@bhaskara18:/opt/spark/conf$ export SPARK_WORKER_MEMORY=3g
hduser_@bhaskara18:/opt/spark/conf$ $SPARK_HOME/sbin/start-worker.sh spark://172.20.251.77:7077
starting org.apache.spark.deploy.worker.Worker, logging to /opt/spark/logs/spark-hduser_@org.apache.spark.deploy.worker.Worker-1-bhaskara18.out
hduser_@bhaskara18:/opt/spark/conf$ cat /opt/spark/logs/spark-hduser_@org.apache.spark.deploy.worker.Worker-1-bhaskara18.out
Spark Command: /usr/lib/jvm/java-1.17.0-openjdk-amd64/bin/java -cp /opt/spark/conf:/opt/spark/jars/* -Xmx1g org.apache.spark.deploy.worker.Worker --webui-port 8081 spark://172.20.251.77:7077
=====
Using Spark's default log4j profile: org/apache/spark/log4j2-defaults.properties
24/05/22 18:24:08 INFO Worker: Started daemon with process name: 3288@bhaskara18
24/05/22 18:24:08 INFO SignalUtils: Registering signal handler for TERM
24/05/22 18:24:08 INFO SignalUtils: Registering signal handler for HUP
24/05/22 18:24:08 INFO SignalUtils: Registering signal handler for INT
24/05/22 18:24:08 WARN Utils: Your hostname, bhaskara18 resolves to a loopback address: 127.0.1.1; using 172.20.251.181 instead (on interface enp3s0)
24/05/22 18:24:08 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
24/05/22 18:24:08 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
24/05/22 18:24:08 INFO SecurityManager: Changing view acls to: hduser_
24/05/22 18:24:08 INFO SecurityManager: Changing modify acls to: hduser_
24/05/22 18:24:08 INFO SecurityManager: Changing view acls groups to:
24/05/22 18:24:08 INFO SecurityManager: Changing modify acls groups to:
24/05/22 18:24:08 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: hduser_; groups with view permissions: EMPTY; users with modify permissions: hduser_; groups with modify permissions: EMPTY
24/05/22 18:24:09 INFO Utils: Successfully started service 'sparkWorker' on port 35127.
24/05/22 18:24:09 INFO Worker: Worker decommissioning not enabled.
24/05/22 18:24:09 INFO Worker: Starting Spark worker 172.20.251.181:35127 with 3 cores, 3.0 GiB RAM
24/05/22 18:24:09 INFO Worker: Running Spark version 3.4.2
24/05/22 18:24:09 INFO Worker: Spark home: /opt/spark
24/05/22 18:24:09 INFO ResourceUtils: =====
24/05/22 18:24:09 INFO ResourceUtils: No custom resources configured for spark.worker.
24/05/22 18:24:09 INFO ResourceUtils: =====
24/05/22 18:24:09 INFO JettyUtils: Start Jetty 0.0.0.0:8081 for WorkerUI
24/05/22 18:24:09 INFO Utils: Successfully started service 'WorkerUI' on port 8081.
24/05/22 18:24:09 INFO WorkerWebUI: Bound WorkerWebUI to 0.0.0.0, and started at http://172.20.251.181:8081
24/05/22 18:24:09 INFO Worker: Connecting to master 172.20.251.77:7077...
24/05/22 18:24:09 INFO TransportClientFactory: Successfully created connection to /172.20.251.77:7077 after 31 ms (0 ms spent in bootstraps)
24/05/22 18:24:09 INFO Worker: Successfully registered with master spark://172.20.251.77:7077
hduser_@bhaskara18:/opt/spark/conf$
```

# Refresh Spark UI

Activities Firefox Web Browser May 22 18:31

New Tab Spark Master at spark://1 x +

172.20.251.77:8080

 **Spark Master at spark://172.20.251.77:7077**

URL: spark://172.20.251.77:7077  
Alive Workers: 2  
Cores in use: 6 Total, 0 Used  
Memory in use: 6.0 GiB Total, 0.0 B Used  
Resources in use:  
Applications: 0 Running, 0 Completed  
Drivers: 0 Running, 0 Completed  
Status: ALIVE

Workers (3)

Worker Id	Address	State	Cores	Memory	Resources
worker-20240522181512-172.20.251.77-40823	172.20.251.77-40823	DEAD	3 (0 Used)	3.0 GiB (0.0 B Used)	
<a href="#">worker-20240522181659-172.20.251.77-37019</a>	172.20.251.77-37019	ALIVE	3 (0 Used)	3.0 GiB (0.0 B Used)	
<a href="#">worker-20240522182409-172.20.251.181-35127</a>	172.20.251.181-35127	ALIVE	3 (0 Used)	3.0 GiB (0.0 B Used)	

Running Applications (0)

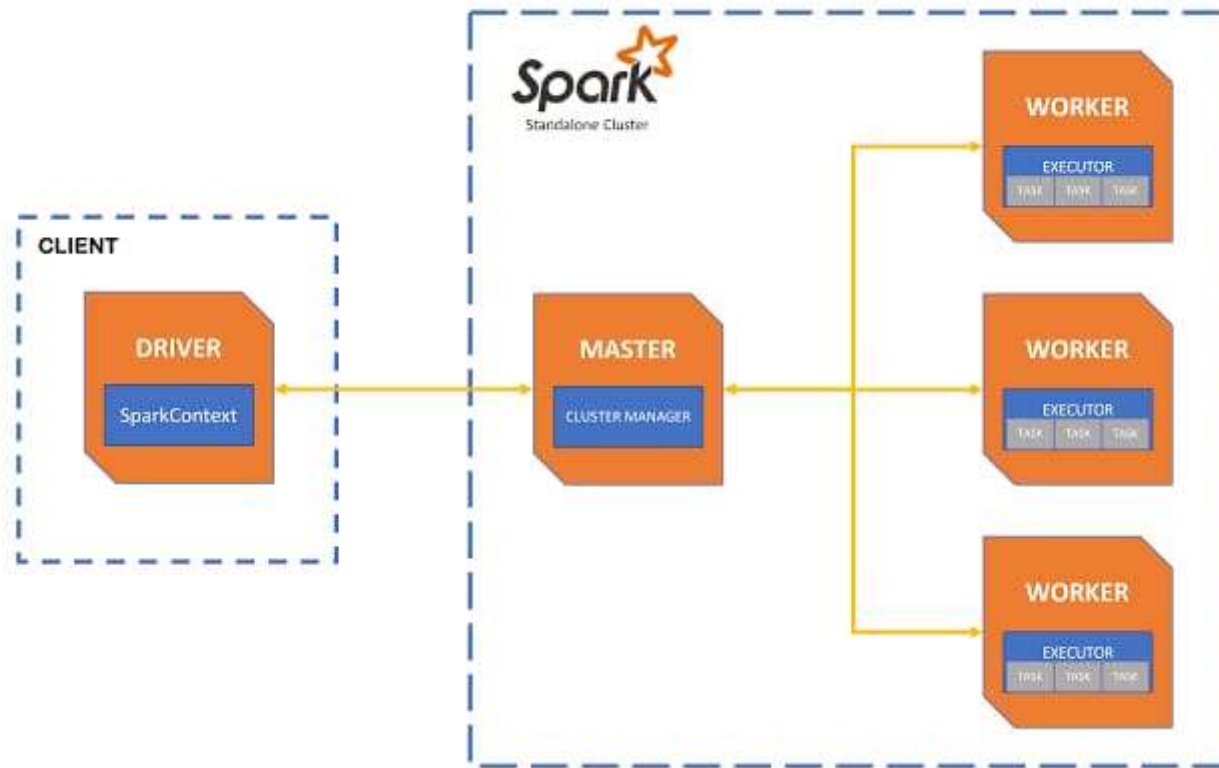
Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
----------------	------	-------	---------------------	------------------------	----------------	------	-------	----------

Completed Applications (0)

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

# Tweet Sentiment Analysis

# Workflow (Again)





`spark-submit <your .py file name>.py`

(Check the my\_spark\_app.py: Change the SparkSession Builder master address to your respective master node IP address)

(Incase of Error : unset PYSPARK\_DRIVER\_PYTHON then above code)

```
hduser@bhaskara17: /opt/spark/conf$ unset PYSPARK_DRIVER_PYTHON
hduser@bhaskara17: /opt/spark/conf$ spark-submit /home/hduser/Desktop/my_spark_app.py
24/05/22 18:31:59 WARN Utils: Your hostname, bhaskara17 resolves to a loopback address: 127.0.1.1; using 172.20.251.77 instead (on interface enp3s0)
24/05/22 18:31:59 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
24/05/22 18:32:02 INFO SparkContext: Running Spark version 3.4.2
24/05/22 18:32:02 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
24/05/22 18:32:02 INFO ResourceUtils: =====
24/05/22 18:32:02 INFO ResourceUtils: No custom resources configured for spark.driver.
24/05/22 18:32:02 INFO ResourceUtils: =====
24/05/22 18:32:02 INFO SparkContext: Submitted application: FlanT5Seq2SeqExample
24/05/22 18:32:02 INFO ResourceProfile: Default ResourceProfile created, executor resources: Map(memory -> name: memory, amount: 1024, script: , vendor: , offHeap -> name: offHeap, amount: 0, script: , vendor: )
, task resources: Map(cpus -> name: cpus, amount: 1.0)
24/05/22 18:32:02 INFO ResourceProfile: Limiting resource is cpu
24/05/22 18:32:02 INFO ResourceProfileManager: Added ResourceProfile id: 0
24/05/22 18:32:02 INFO SecurityManager: Changing view acls to: hduser_
24/05/22 18:32:02 INFO SecurityManager: Changing modify acls to: hduser_
24/05/22 18:32:02 INFO SecurityManager: Changing view acls groups to:
24/05/22 18:32:02 INFO SecurityManager: Changing modify acls groups to:
24/05/22 18:32:02 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: hduser_; groups with view permissions: hduser_
; groups with modify permissions: EMPTY
24/05/22 18:32:02 INFO Utils: Successfully started service 'sparkDriver' on port 38053.
24/05/22 18:32:02 INFO SparkEnv: Registering MapOutputTracker
24/05/22 18:32:02 INFO SparkEnv: Registering BlockManagerMaster
24/05/22 18:32:02 INFO BlockManagerMasterEndpoint: Using org.apache.spark.storage.DefaultTopologyMapper for getting topology information
24/05/22 18:32:02 INFO BlockManagerMasterEndpoint: BlockManagerMasterEndpoint up
24/05/22 18:32:02 INFO SparkEnv: Registering BlockManagerMasterHeartbeat
24/05/22 18:32:02 INFO DiskBlockManager: Created local directory at /tmp/blockmgr-121eef02-4674-4ee2-86a4-e177a0dfe8db
24/05/22 18:32:02 INFO MemoryStore: MemoryStore started with capacity 434.4 MiB
24/05/22 18:32:02 INFO SparkEnv: Registering OutputCommitCoordinator
24/05/22 18:32:02 INFO JettyUtils: Start Jetty 0.0.0.0:4040 for SparkUI
24/05/22 18:32:02 INFO Utils: Successfully started service 'SparkUI' on port 4040.
24/05/22 18:32:02 INFO StandaloneAppClientSClientEndpoint: Connecting to master spark://172.20.251.77:7077...
24/05/22 18:32:02 INFO TransportClientFactory: Successfully created connection to /172.20.251.77:7077 after 26 ms (0 ms spent in bootstraps)
24/05/22 18:32:03 INFO StandaloneSchedulerBackend: Connected to Spark cluster with app ID app-20240522183203-0000
24/05/22 18:32:03 INFO Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port 38359.
24/05/22 18:32:03 INFO NettyBlockTransferService: Server created on 172.20.251.77:38359
24/05/22 18:32:03 INFO BlockManager: Using org.apache.spark.storage.RandomBlockReplicationPolicy for block replication policy
24/05/22 18:32:03 INFO BlockManagerMaster: Registering BlockManager BlockManagerId(driver, 172.20.251.77, 38359, None)
24/05/22 18:32:03 INFO BlockManagerMasterEndpoint: Registering block manager 172.20.251.77:38359 with 434.4 MiB RAM, BlockManagerId(driver, 172.20.251.77, 38359, None)
24/05/22 18:32:03 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 172.20.251.77, 38359, None)
24/05/22 18:32:03 INFO BlockManager: Initialized BlockManager: BlockManagerId(driver, 172.20.251.77, 38359, None)
24/05/22 18:32:03 INFO StandaloneAppClientSClientEndpoint: Executor added: app-20240522183203-0000/0 on worker-20240522182409-172.20.251.181-35127 (172.20.251.181:35127) with 3 core(s)
24/05/22 18:32:03 INFO StandaloneSchedulerBackend: Granted executor ID app-20240522183203-0000/0 on hostPort 172.20.251.181:35127 with 3 core(s), 1024.0 MiB RAM
24/05/22 18:32:03 INFO StandaloneAppClientSClientEndpoint: Executor added: app-20240522183203-0000/1 on worker-20240522181659-172.20.251.77-37019 (172.20.251.77:37019) with 3 core(s)
24/05/22 18:32:03 INFO StandaloneSchedulerBackend: Created executor ID app-20240522183203-0000/1 on hostPort 172.20.251.77:37019 with 3 core(s), 1024.0 MiB RAM
```

# Update Spark UI

Activities Firefox Web Browser May 22, 18:32

New Tab Spark Master at spark://172.20.251.77:7077

172.20.251.77:7077

**Spark Master at spark://172.20.251.77:7077**

URL: spark://172.20.251.77:7077  
Alive Workers: 2  
Cores in use: 6 Total, 6 Used  
Memory in use: 6.0 GiB Total, 2.0 GiB Used  
Resources in use:  
Applications: 1 [Running](#), 0 [Completed](#)  
Drivers: 0 Running, 0 Completed  
Status: ALIVE

Workers (3)

Worker Id	Address	State	Cores	Memory	Resources
worker-20240522181512-172.20.251.77-40823	172.20.251.77:40823	DEAD	3 (0 Used)	3.0 GiB (0.0 B Used)	
<a href="#">worker-20240522181659-172.20.251.77-37019</a>	172.20.251.77:37019	ALIVE	3 (3 Used)	3.0 GiB (1024.0 MiB Used)	
<a href="#">worker-20240522182409-172.20.251.181-35127</a>	172.20.251.181:35127	ALIVE	3 (3 Used)	3.0 GiB (1024.0 MiB Used)	

Running Applications (1)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
<a href="#">app-20240522183203-0000</a>	(kill) <a href="#">FlanT5Seq2SeqExample</a>	6	1024.0 MiB		2024/05/22 18:32:03	hduser_	RUNNING	7 s

Completed Applications (0)

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



## Spark Worker at 172.20.251.181:35127

ID: worker-20240522182409-172.20.251.181-35127

Master URL: spark://172.20.251.77:7077

Cores: 3 (0 Used)

Memory: 3.0 GiB (0.0 B Used)

Resources:

[Back to Master](#)

### Running Executors (0)

ExecutorID	State	Cores	Memory	Resources	Job Details	Logs
------------	-------	-------	--------	-----------	-------------	------

### Finished Executors (1)

ExecutorID	State	Cores	Memory	Resources	Job Details	Logs
0	KILLED	3	1024.0 MiB		ID: app-20240522183203-0000 Name: FlanT5Seq2SeqExample User: hduser_	<a href="#">stdout stderr</a>



## Spark Worker at 172.20.251.77:37019

ID: worker-20240522181659-172.20.251.77-37019

Master URL: spark://172.20.251.77:7077

Cores: 3 (0 Used)

Memory: 3.0 GiB (0.0 B Used)

Resources:

[Back to Master](#)

### Running Executors (0)

ExecutorID	State	Cores	Memory	Resources	Job Details	Logs
------------	-------	-------	--------	-----------	-------------	------

### Finished Executors (1)

ExecutorID	State	Cores	Memory	Resources	Job Details	Logs
1	KILLED	3	1024.0 MiB		ID: app-20240522183203-0000 Name: FlanT5Seq2SeqExample User: hduser_	<a href="#">stdout stderr</a>



# Workings

[illegible]

# Output


```
24/05/22 18:32:42 INFO StandaloneSchedulerBackend: Shutting down all executors
24/05/22 18:32:42 INFO CodeGenerator: Code generated in 11.715961 ms
+-----+-----+
|sentence|output_column|
+-----+-----+
|BJP National President Shri @JPNadda addressed the Cluster Meeting of Malkajgiri, Secunderabad, Hyderabad and Chevella Lok Sabha constituencies in Hyderabad, Telangana. https://t.co/HMO3fUPmvL|positive|
|Telangana is all set to vote for BJP-led NDA.Visuals from BJP National President Shri @JPNadda's mega roadshow in Nizampet. https://t.co/Ki4VtVQRyV|positive|
|The law and order situation in Karnataka is deteriorating. The state is on the brink of bankruptcy.Congress made tall promises but they applied terms and conditions.- PM @narendramodi Watch the complete video:https://t.co/96nanptoD4 https://t.co/5athSNT6re|negative|
|मैं विकास और विरासत की बात करता हूँ और वो विरासत को लूने की बात करते हैं। सुनिए, पीएम श्री @narendramodi ने और क्या कहा...पूरा वीडियो देखें https://t.co/96nanptoD4 https://t.co/jRQICegd2y|positive|
|Under the TMC, atrocities against women have increased. The incident in Sandeshkhali has shaken the people of West Bengal.People will vent their anger through their votes in West Bengal.- PM @narendramodi Watch the complete video:https://t.co/96nanptoD4 https://t.co/gPiZcC0aL0|negative|
|मुस्लिम अल्पसंख्यक पर मोदी जी ने बताया... कांग्रेस का इतिहास। पूरा वीडियो देखें https://t.co/96nanptoD4 https://t.co/9fmTECzici|negative|
|विपक्ष पहले राउंड के चुनाव में पस्त और दूसरे राउंड में ध्वस्त हो चुका है। - पीएम श्री @narendramodi पूरा वीडियो देखें https://t.co/96nanptoD4 https://t.co/uk1xm0AohZ|positive|
|Manifestos of political parties aren't showpieces, and it's the work of the media to analyse them. I had said that the manifesto of Congress has a reflection of the Muslim League. When the media failed to expose the drawbacks of the manifesto of Congress, I had to step in!- PM @narendramodi Watch the complete video:https://t.co/fch7HjcCsk|negative|
|The people flooded the streets and roads of Guwahati during Union Home Minister Shri @AmitShah's roadshow in Assam today. https://t.co/Ba2wMsrP68|negative|
|अब देश के लोगों में एक कॉन्फिडेंस लेवल आया है। अब ओ मुझे देश को अस्थिर की तरीक़ा सबसे बड़ी ताक़त बनाना है। मैंने देश को 11वें नंबर की अर्थव्यवस्था से 5वें नंबर की दुनिया की सबसे बड़ी अर्थव्यवस्था बनाया... अब भारत को तीसरी सबसे बड़ी अर्थव्यवस्था बनाना है। ओ आने वाले 5 साल में हमें क्या करना है... ये स्पष्ट है।- पीएम @narendramodi|positive|
|Under the TMC, the atrocities against women have increased. The incident in Sandeshkhali has shaken the people of West Bengal. The people will answer through their votes. In the last year, huge sums of money were unearthed in West Bengal. No matter how hard TMC may try to suppress it, people will not forget. We will do a clean sweep in West Bengal.- PM @narendramodi|negative|
|जहाँ तक बंगाल की बात है तो टीएमसी के शासन में जो वहाँ महिलाओं पर अत्याचार हो रहे हैं, संसदवादी की घटना ने तो पूरे देश को हिला दिया है। इसे ज़माना बर्बाद नहीं करेगा... इसका गुस्सा तो ज़माना निकालेगी ही। जो वहाँ से नोटों के बंडल के ढेर मिले हैं, पूरे देश ने देखा है। इस बार हम क्लीन स्वीप करेंगे।- पीएम @narendramodi|positive|
|Not even a single case has been dropped. The court will deliver verdicts on the cases. Only 3% of cases pertain to political leaders.The investigative agencies have been formed to work against the corrupt. The court isn't giving bail to the accused. We should not take the cases of corruption lightly...- PM @narendramodi|negative|
|इंडी अलायंस में चाहे समाजवादी पार्टी हो या कोई और... उनकी मजबूरी है कि उन्हें परिवार के लोगों को उत्तारना पड़ रहा है। ये उनका मूल कैरेक्टर है, क्योंकि इनके सिवा उनके पास कुछ है ही नहीं। पैमिली मेंबर्स के लिए लड़ना और पैमिली मेंबर्स को लड़ाना... यही उनका काम है।- पीएम @narendramodi|positive|
|हमने किसी काम में कभी कोई भेदभाव नहीं किया। हमने पूरे परिवार मुख़जी, नरसिंम्हा राव, चौधरी चरण सिंह, कर्पूरी ठाकुर को भारत रत्न देने का काम किया, किसी ने कोई प्रश्न नहीं उठाया। वैसे ही पद्म पुरस्कार मुलायम सिंह जी, तरुण गोगोई जी, पीए संमा जी, एसएम कृष्णा जी को दिया... ये सारे दूसरी पार्टियों के हैं। ये आरुंधत देश का है, किसी दल का नहीं है... ये मेरी सोच है।- पीएम @narendramodi|positive|
|The people of Karnataka are repenting for electing a Congress govt. The law and order situation in Karnataka is deteriorating. The economic situation is bad. Congress made tall promises but they applied terms and conditions to it. Therefore, they are betraying the people. Bengaluru was known as a tech hub, but they turned it into a tanker hub...- PM @narendramodi|negative|
|2014 में इनके पास इंडी थी, सीबीआई थी तो फिर हार क्यों गए। मेरे होम मिनिस्टर तक को जेल में डाल दिया था। आर इंडी, सीबीआई से चुनाव जीते जाते तो इतने इंडी, सीबीआई के साथ तो कांग्रेस ने काम किया है तो फिर वो जीत ही जाते।- पीएम @narendramodi https://t.co/TrKJ8dmyQh|positive|
+-----+-----+
24/05/22 18:32:42 INFO SparkContext: SparkContext is stopping with exitCode 0.
24/05/22 18:32:42 INFO SparkUI: Stopped Spark web UI at http://172.20.251.77:4040
24/05/22 18:32:42 INFO StandaloneSchedulerBackend: Shutting down all executors
24/05/22 18:32:42 INFO StandaloneSchedulerBackend: Asking each executor to shut down
```

# Update Spark UI

Activities Firefox Web Browser May 22 18:35

New Tab Spark Master at spark://172.20.251.77:7077

172.20.251.77:8080

 **Spark Master at spark://172.20.251.77:7077**

URL: spark://172.20.251.77:7077  
Alive Workers: 2  
Cores in use: 6 Total, 0 Used  
Memory in use: 6.0 GiB Total, 0.0 B Used  
Resources in use:  
Applications: 0 Running, 1 Completed  
Drivers: 0 Running, 0 Completed  
Status: ALIVE

Workers (2)

Worker Id	Address	State	Cores	Memory	Resources
<a href="#">worker-20240522181659-172.20.251.77-37019</a>	172.20.251.77:37019	ALIVE	3 (0 Used)	3.0 GiB (0.0 B Used)	
<a href="#">worker-20240522182409-172.20.251.181-35127</a>	172.20.251.181:35127	ALIVE	3 (0 Used)	3.0 GiB (0.0 B Used)	

Running Applications (0)

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

Completed Applications (1)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
<a href="#">app-20240522183203-0000</a>	FlanT5Seq2SeqExample	6	1024.0 MiB		2024/05/22 18:32:03	hduser_	FINISHED	40 s

# To Know more about the Model used

- <https://huggingface.co/google/flan-t5-small>