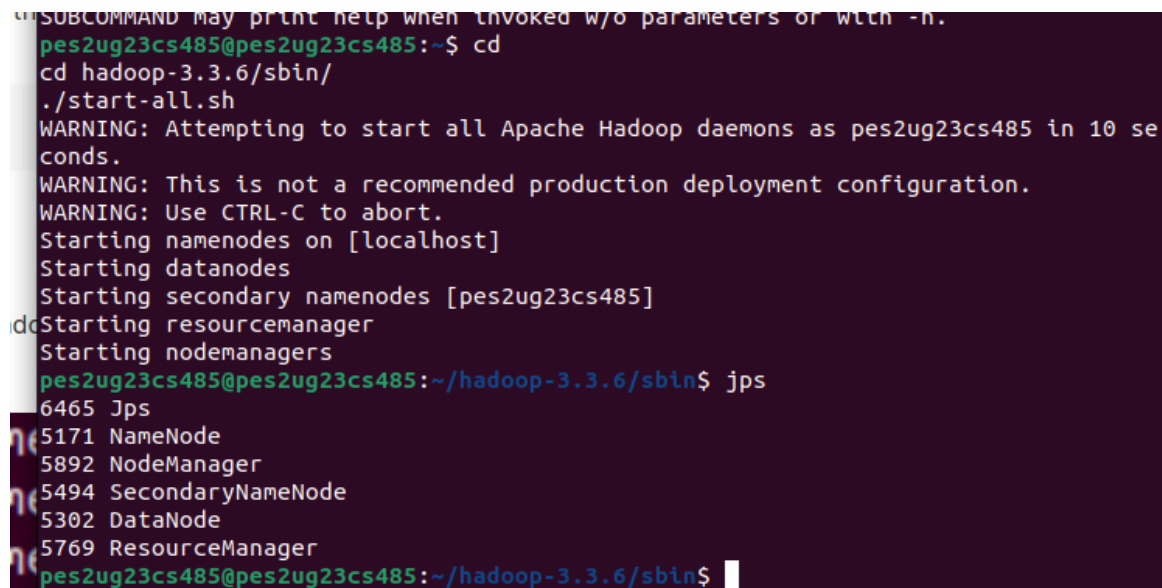# *Big Data assignment 1*

Name: Rithvik Rajesh Matta
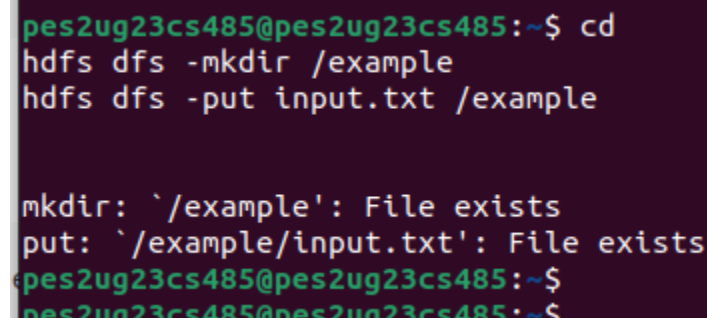**SRN: PES2UG23CS485**
Class: 5H

1.Hadoop installation **(run hadoop version)** and All daemons running **(run jps)**

```
SUBCOMMAND May print help when invoked w/o parameters or with -h.
pes2ug23cs485@pes2ug23cs485:~$ cd
cd hadoop-3.3.6/sbin/
./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as pes2ug23cs485 in 10 se
conds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [pes2ug23cs485]
Starting resourcemanager
Starting nodemanagers
pes2ug23cs485@pes2ug23cs485:~/hadoop-3.3.6/sbin$ jps
6465 Jps
5171 NameNode
5892 NodeManager
5494 SecondaryNameNode
5302 DataNode
5769 ResourceManager
pes2ug23cs485@pes2ug23cs485:~/hadoop-3.3.6/sbin$
```

2.Uploading the text file from local to HDFS **(hdfs dfs -mkdir .... & hdfs dfs -put ....)**

```
pes2ug23cs485@pes2ug23cs485:~$ cd
hdfs dfs -mkdir /example
hdfs dfs -put input.txt /example


mkdir: `/example': File exists
put: `/example/input.txt': File exists
pes2ug23cs485@pes2ug23cs485:~$
pes2ug23cs485@pes2ug23cs485:~$
```

3.Running MapReduce JAR commands **(hadoop jar $HADOOP_HOME ...)**

```
pes2ug23cs485@pes2ug23cs485:~$ hadoop jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.6.jar wordcount /example/input.txt /example/output
2025-09-02 11:58:54,091 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /127.0.0.1:8032
org.apache.hadoop.mapred.FileAlreadyExistsException: Output directory hdfs://127.0.0.1:9000/example/output already exists
        at org.apache.hadoop.mapreduce.lib.output.FileOutputFormat.checkOutputSpecs(FileOutputFormat.java:164)
        at org.apache.hadoop.mapreduce.JobSubmitter.checkSpecs(JobSubmitter.java:277)
        at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:143)
        at org.apache.hadoop.mapreduce.Job$11.run(Job.java:1678)
        at org.apache.hadoop.mapreduce.Job$11.run(Job.java:1675)
        at java.security.AccessController.doPrivileged(Native Method)
        at javax.security.auth.Subject.doAs(Subject.java:422)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1899)
        at org.apache.hadoop.mapreduce.Job.submit(Job.java:1675)
        at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1696)
        at org.apache.hadoop.examples.WordCount.main(WordCount.java:87)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.ProgramDriver$ProgramDescription.invoke(ProgramDriver.java:71)
        at org.apache.hadoop.util.ProgramDriver.run(ProgramDriver.java:144)
        at org.apache.hadoop.examples.ExampleDriver.main(ExampleDriver.java:74)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:328)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:241)
pes2ug23cs485@pes2ug23cs485:~$
```

4.MapReduce code output **(hdfs dfs -cat /example/output/part-r-00000)**

```
pes2ug23cs485@pes2ug23cs485:~$ hdfs dfs -cat /example/output/part-r-00000
hadoop! 1
hello,  1
i'm     2
rithvik,         1
using   1
pes2ug23cs485@pes2ug23cs485:~$
```

5.Simulating MapReduce in Local System **(cat input.txt | ./mapper.py | sort | ./reducer.py)**

```
pes2ug23cs485@pes2ug23cs485:~/hands_on$ cat input.txt | ./mapper.py
Hello,1
Hello,1
World,1
Big,1
Big,1
Big,1
pes2ug23cs485@pes2ug23cs485:~/hands_on$
```