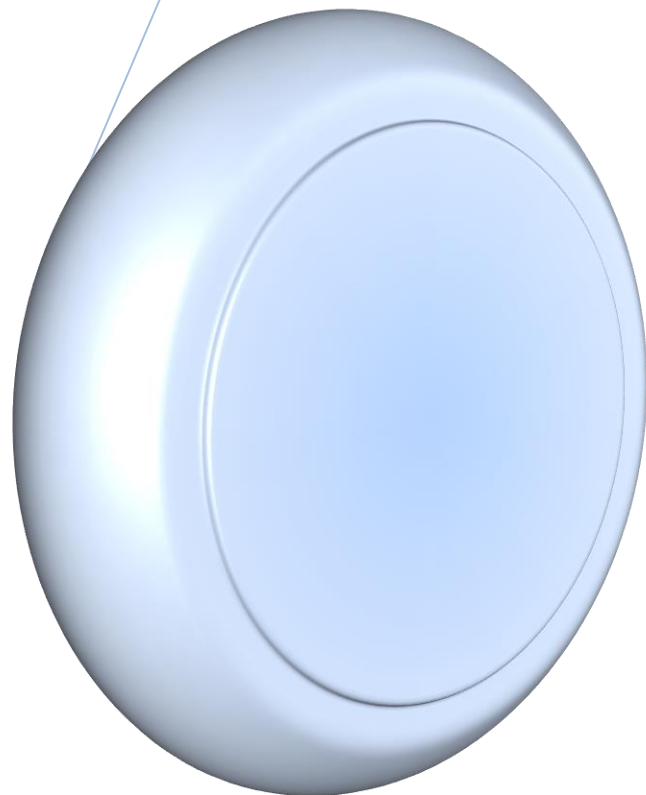# Big Data Hadoop & Spark Training
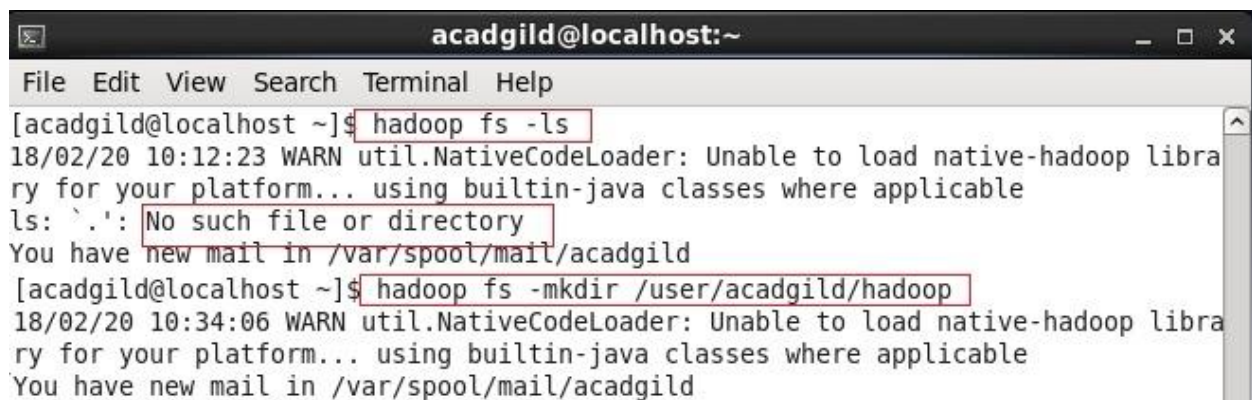
Assignment-2

Rashmi.K

Task 1: Check whether /user/acadgild directory exists or not in the HDFS.If it doesn't exist, then create this. Create a directory /user/acadgild/hadoop.

❖ Checked if /user/acadgild directory exists or not using the cmd: Hadoop fs –ls, directory did not exist. So created a directory /user/acadgild/Hadoop using the cmd: Hadoop fs –mkdir /user/acadgild/Hadoop, as shown in the below screenshot

Task 2: Create a file in HDFS under directory /user/acadgild/hadoop, with name word-count.txt. Whatever we type on screen should get appended to the file.

❖ Created a wordcount.txt file in local system in nano editor.
❖ Moved the file from local system to HDFS file under the directory /user/acadgild/Hadoop using Put command.
❖ Outputs are as shown in the below screenshots

```
                        acadgild@localhost:~                    _ □ x

 File  Edit  View  Search  Terminal  Help

You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ nano wordcount.txt ——creating wordcount.txt file
[acadgild@localhost ~]$ cat wordcount.txt
Hello All,

This is training on Big Data Hadoop & Spark.
This is a sample file to count the number of words in this file.
This file is saved in HDFS file system.
File path where this data is saved is /user/acadgild.
                                        contents of wordcount.txt
Thank you all!

[acadgild@localhost ~]$ hadoop fs -put wordcount.txt /user/acadgild/hadoop
18/02/20 10:35:01 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/hadoop/wordcount.txt
18/02/20 10:37:38 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
Hello All,

This is training on Big Data Hadoop & Spark.
This is a sample file to count the number of words in this file.
This file is saved in HDFS file system.
File path where this data is saved is /user/acadgild.

Thank you all!      ouput of wordcount.txt after moving the file from local system to HDFS

You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ █
```

❖ Created a file append.txt in local system and *appended this file to wordcount.txt* in HDFS, output is as shown below-

```
[acadgild@localhost ~]$ nano append.txt    created a file append.txt
[acadgild@localhost ~]$ cat append.txt
Welcome All,        This is the contents of the file append.txt

This file is to append file to the existing wordcount file in hdfs file system
[acadgild@localhost ~]$ hadoop fs -appendToFile append.txt /user/acadgild/hadoop
/wordcount.txt
18/02/20 10:41:04 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
/wordcount.txt
[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/hadoop/wordcount.txt
18/02/20 10:41:45 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
Hello All,

This is training on Big Data Hadoop & Spark.
This is a sample file to count the number of words in this file.
This file is saved in HDFS file system.
File path where this data is saved is /user/acadgild.
                    This is the output after appending the file append.txt to
Thank you all!      wordcount.txt.

Welcome All,

This file is to append file to the existing wordcount file in hdfs file system
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ █
```
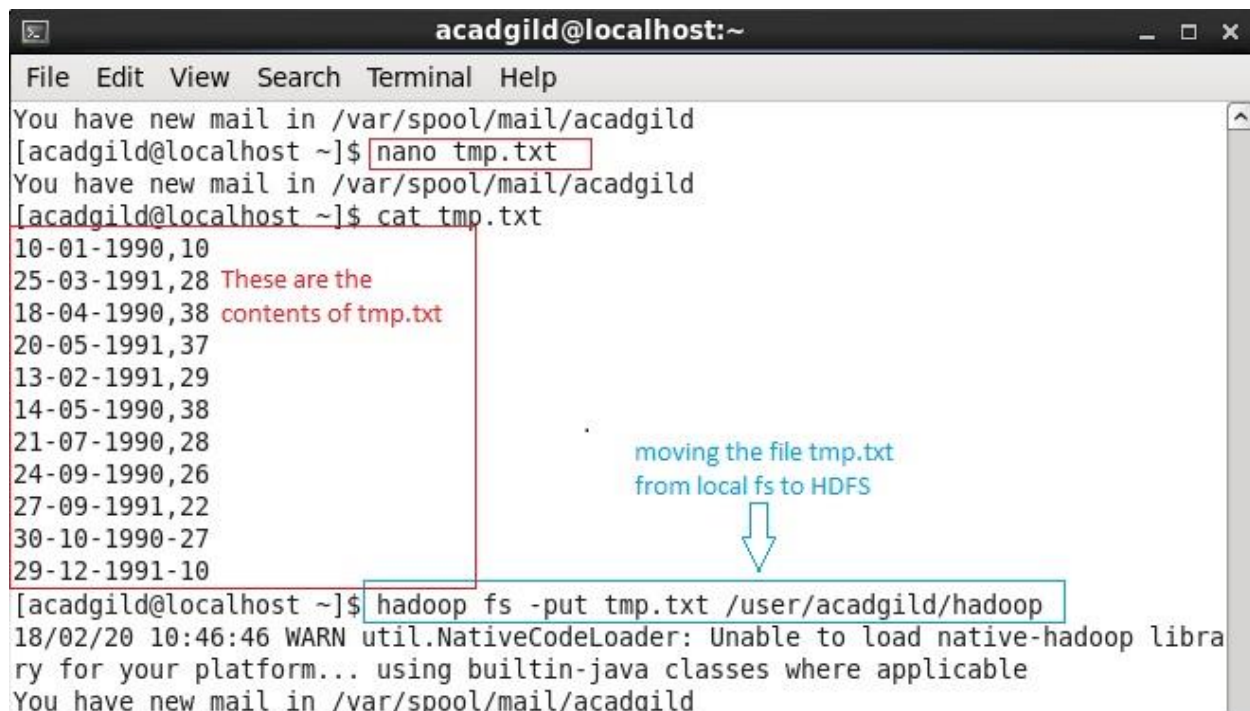
Task 3: Create a file max-temp.txt in local FS.Put some 10-15 records of date and temperature example:dd-mm-yyyy,temperature.Move this file to HDFS at /user/acadgild/hadoop.

❖ Created a file tmp.txt using nano editor in local file system and moved this file to HDFS using *Put* command, output is as shown in the below screenshot-
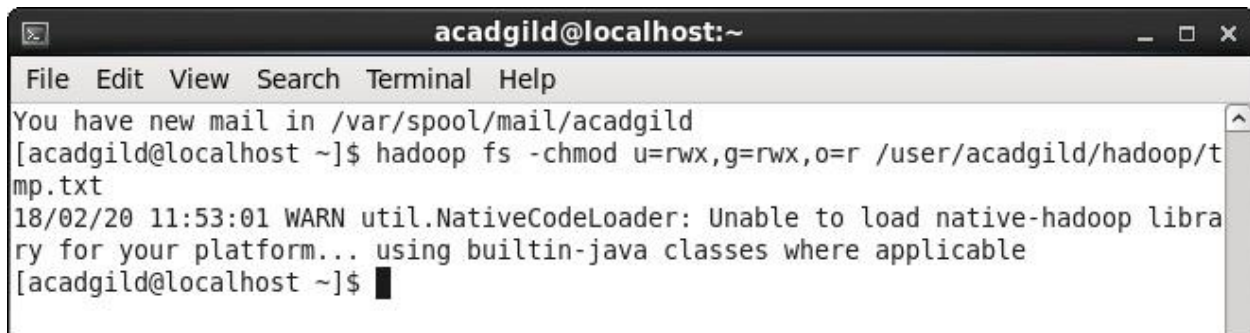
```
acadgild@localhost:~                                  _  □  ×
File  Edit  View  Search  Terminal  Help
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ nano tmp.txt
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ cat tmp.txt
10-01-1990,10
25-03-1991,28   These are the
18-04-1990,38   contents of tmp.txt
20-05-1991,37
13-02-1991,29
14-05-1990,38
21-07-1990,28              .
24-09-1990,26          moving the file tmp.txt
27-09-1991,22          from local fs to HDFS
30-10-1990-27
29-12-1991-10                ⇩
[acadgild@localhost ~]$ hadoop fs -put tmp.txt /user/acadgild/hadoop
18/02/20 10:46:46 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
You have new mail in /var/spool/mail/acadgild
```

Task 4: Change the permission of the file /user/acadgild/hadoop/max-temp.txt, such that only the owner andthe group members have full control over the file. Others do not have any control over it.

❖ Changed the permission of the file tmp.txt, such that only owner and group members can read,write and execute, while others can only read the file. Output is as shown below-