

BIG DATA ENGINEERING FOR HADOOP & SPARK TRAINING

ASSIGNMENT 2

Task 1 Solution:

Q) Check whether /user/acadgild directory exists or not in the HDFS.

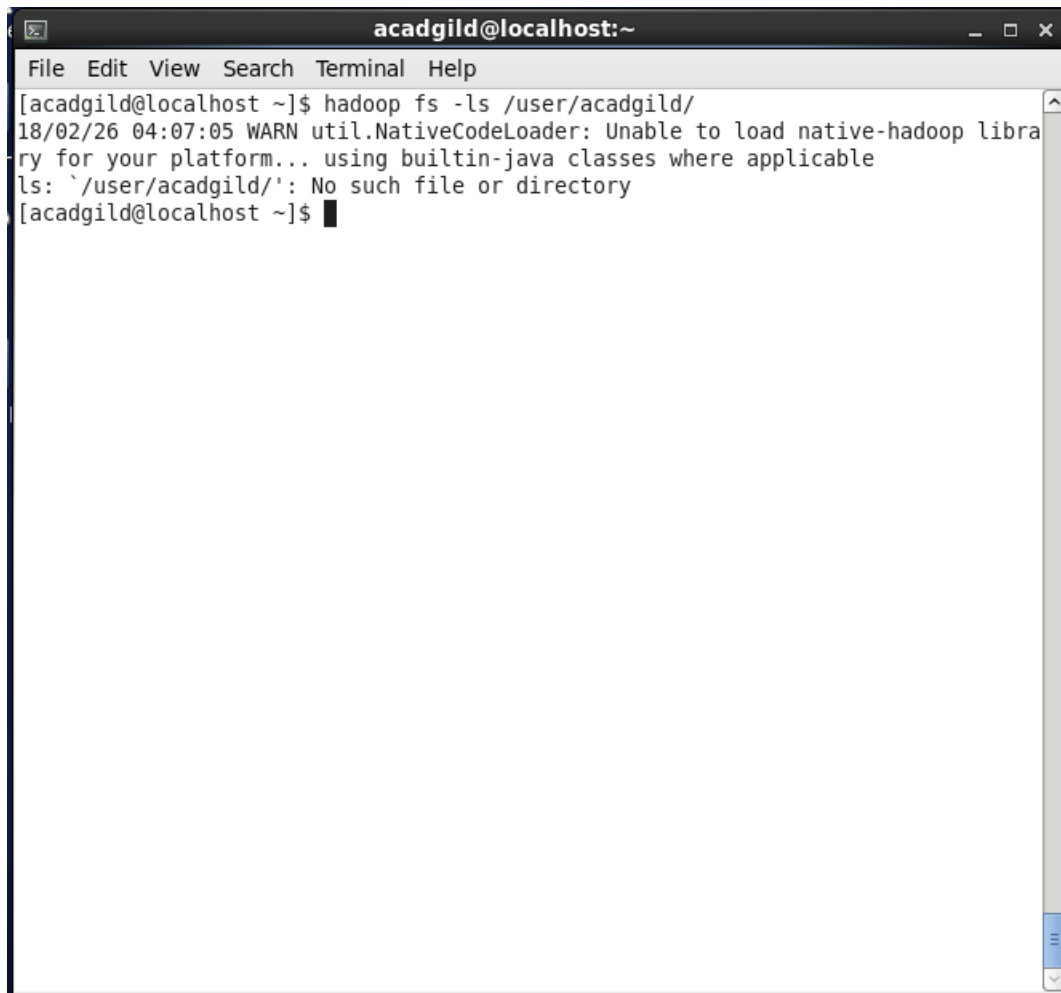
If it doesn't exist, then create this.

Create a directory /user/acadgild/hadoop.

Sol : `hadoop fs -ls /user/acadgild/hadoop` this command will list the files of the given directory.

If file or the directory doesn't exist then we get an error "No such file or directory".

Here we got the error so we have to create the whole directory.

A terminal window titled 'acadgild@localhost:~' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows the command 'hadoop fs -ls /user/acadgild/' being executed. The output includes a warning message: '18/02/26 04:07:05 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable' followed by the error 'ls: '/user/acadgild/': No such file or directory'. The prompt returns to 'acadgild@localhost ~]' with a cursor.

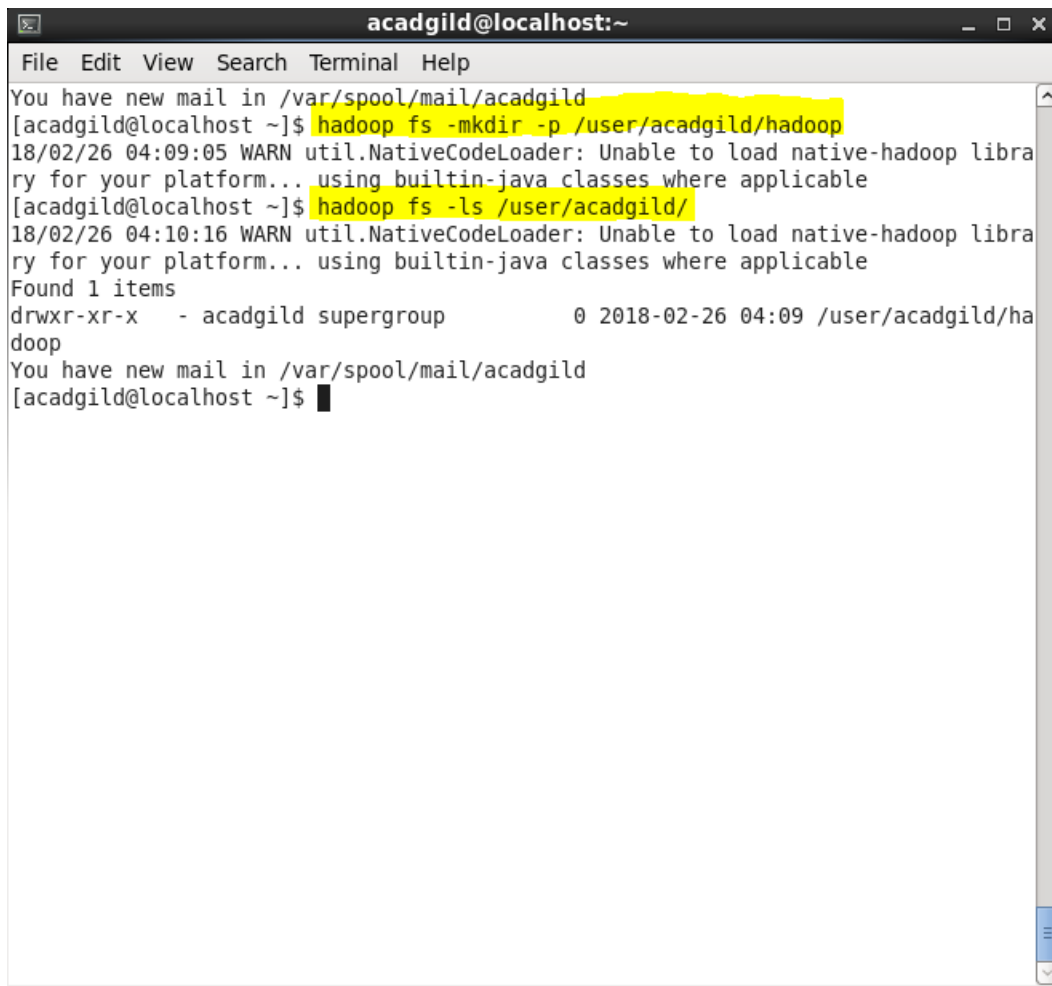
```
acadgild@localhost:~  
File Edit View Search Terminal Help  
[acadgild@localhost ~]$ hadoop fs -ls /user/acadgild/  
18/02/26 04:07:05 WARN util.NativeCodeLoader: Unable to load native-hadoop libra  
ry for your platform... using builtin-java classes where applicable  
ls: '/user/acadgild/': No such file or directory  
[acadgild@localhost ~]$
```

“hadoop fs -mkdir -p /user/acadgild/hadoop” command is used to create.

“/user/acadgild/hadoop” directory recursively in hadoop file system as we have used mkdir with option -p.

Now we can check the directory path created or not by using the command

“hadoop fs -ls /user/acadgild”.



```
acadgild@localhost:~  
File Edit View Search Terminal Help  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ hadoop fs -mkdir -p /user/acadgild/hadoop  
18/02/26 04:09:05 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
[acadgild@localhost ~]$ hadoop fs -ls /user/acadgild/  
18/02/26 04:10:16 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
Found 1 items  
drwxr-xr-x - acadgild supergroup          0 2018-02-26 04:09 /user/acadgild/hadoop  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$
```

Task 2 Solution:

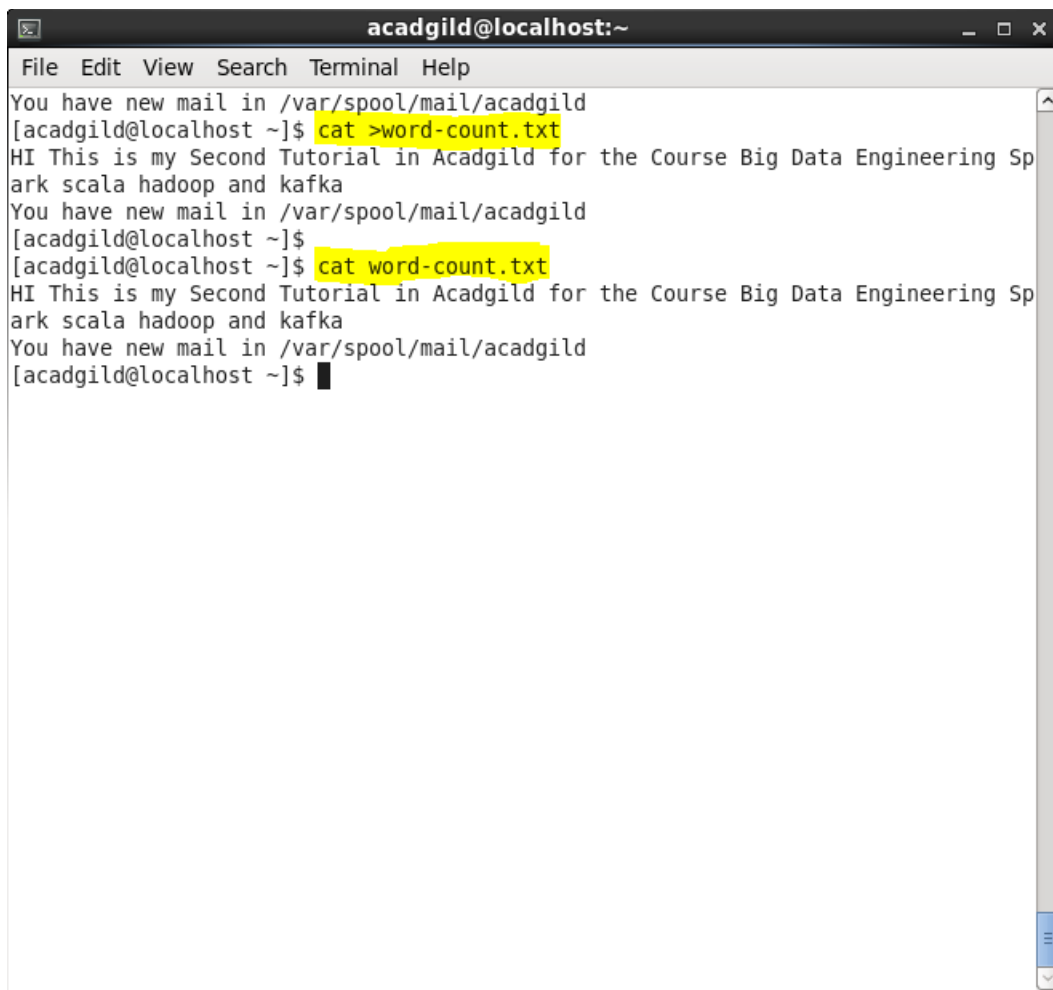
Q) Create a file in HDFS under directory /user/acadgild/hadoop, with name word-count.txt.

Whatever we type on screen should get appended to the file.

Try to type (on screen) few lines from any online article or textbook.

Sol: Create a file in the local file system and add some data to it using "cat >word-count.txt"

Then check the file content using "cat word-count.txt" this will display the content of the File if the file exists.

A terminal window titled 'acadgild@localhost:~' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows a series of commands and their outputs. First, the user runs 'cat >word-count.txt', which creates a file. Then, they type three lines of text: 'HI This is my Second Tutorial in Acadgild for the Course Big Data Engineering Sp', 'ark scala hadoop and kafka', and 'You have new mail in /var/spool/mail/acadgild'. Finally, they run 'cat word-count.txt', which displays the three lines of text that were just entered. The terminal has a scrollbar on the right side.

```
acadgild@localhost:~  
File Edit View Search Terminal Help  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ cat >word-count.txt  
HI This is my Second Tutorial in Acadgild for the Course Big Data Engineering Sp  
ark scala hadoop and kafka  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$  
[acadgild@localhost ~]$ cat word-count.txt  
HI This is my Second Tutorial in Acadgild for the Course Big Data Engineering Sp  
ark scala hadoop and kafka  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$
```

Now copy the file “word-count.txt” to hadoop file system using the command.

“hadoop fs -put /home/user/word-count.txt /user/acadgild/hadoop/word-count.txt”.

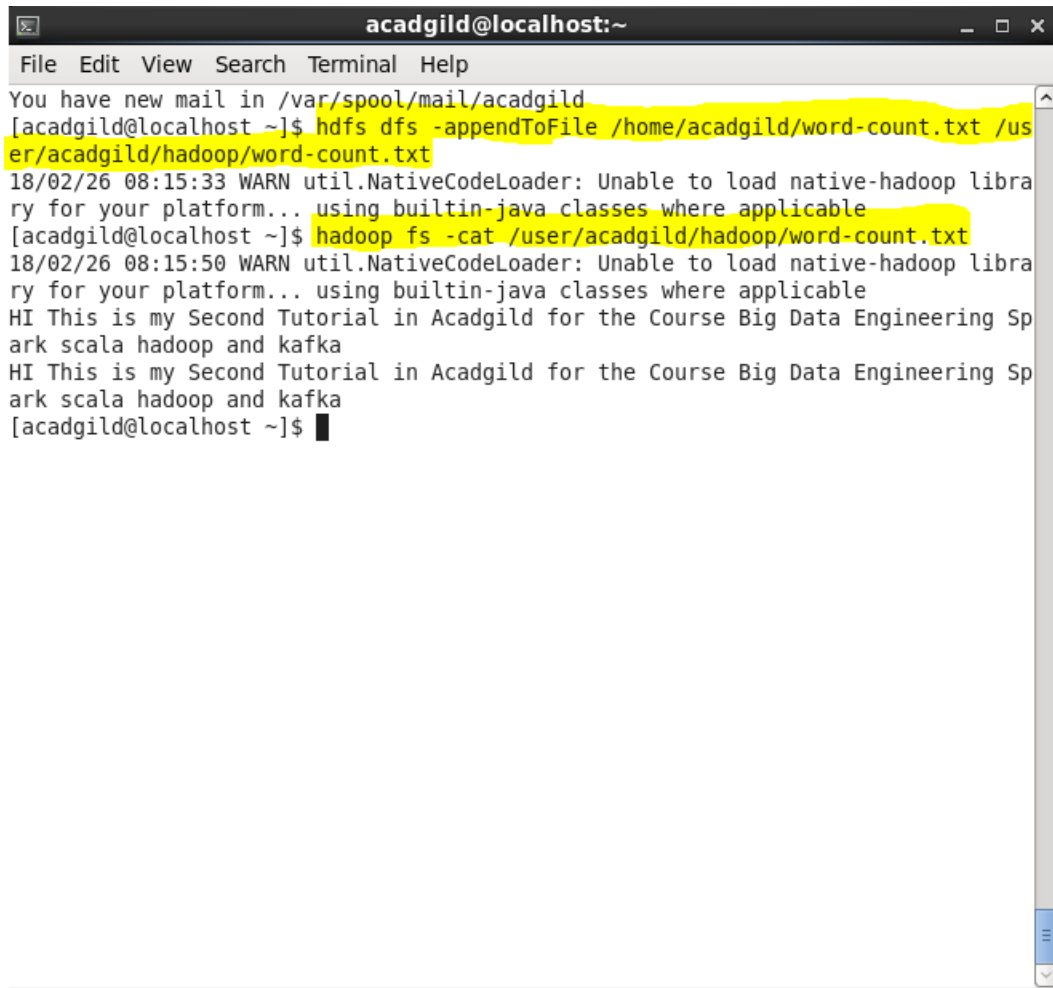
This command will copy the local file to hadoop file system.

To Append the file with the existing file which is available in the local directory use the

Following command

“hdfs dfs -appendToFile /home/acadgild/word-count.txt /user/acadgild/hadoop/word-count.txt”.

Now check the file in hdfs is updated with the contents of local file .

A terminal window titled 'acadgild@localhost:~' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows a notification 'You have new mail in /var/spool/mail/acadgild'. The user enters the command 'hdfs dfs -appendToFile /home/acadgild/word-count.txt /user/acadgild/hadoop/word-count.txt', which is highlighted in yellow. The output shows a warning from 'util.NativeCodeLoader' and then the user enters 'hadoop fs -cat /user/acadgild/hadoop/word-count.txt', also highlighted in yellow. The output displays the contents of the file: 'HI This is my Second Tutorial in Acadgild for the Course Big Data Engineering Spark scala hadoop and kafka' repeated twice.

```
acadgild@localhost:~  
File Edit View Search Terminal Help  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ hdfs dfs -appendToFile /home/acadgild/word-count.txt /user/acadgild/hadoop/word-count.txt  
18/02/26 08:15:33 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/hadoop/word-count.txt  
18/02/26 08:15:50 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
HI This is my Second Tutorial in Acadgild for the Course Big Data Engineering Spark scala hadoop and kafka  
HI This is my Second Tutorial in Acadgild for the Course Big Data Engineering Spark scala hadoop and kafka  
[acadgild@localhost ~]$
```

Task 3 Solution :

Q) Create a file max-temp.txt in local FS.

Put some 10-15 records of date and temperature example:

dd-mm-yyyy,temperature

Example:

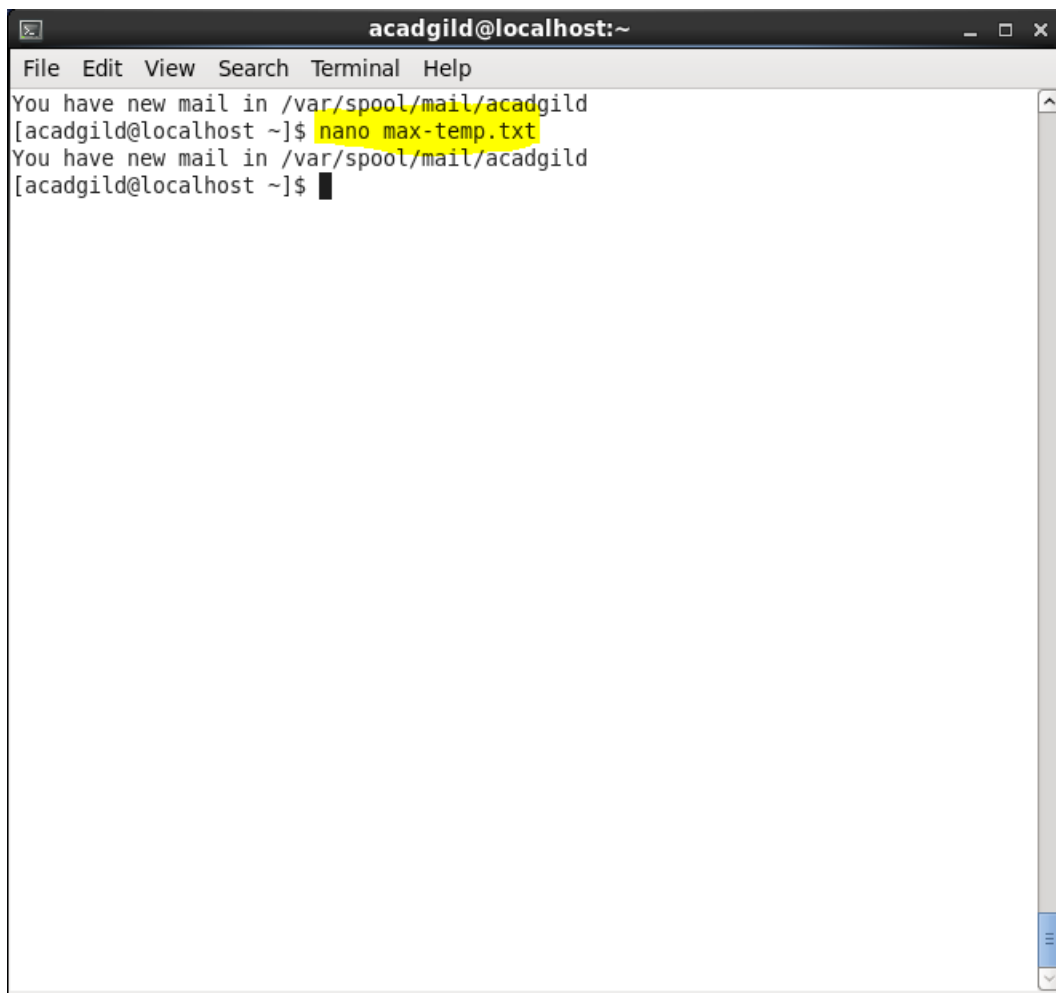
10-01-1990,10

10-02-1991,20

Move this file to HDFS at /user/acadgild/hadoop.

Sol) Use nano editor to create a file and update the file with data specified in the above format

“nano max-temp.txt” command will create a file in the present working directory.



```
acadgild@localhost:~  
File Edit View Search Terminal Help  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ nano max-temp.txt  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$
```

Update the file with the data of 15 records in the given format and save the file.

Now copy "max-temp.txt" file to hdfs file system using the command

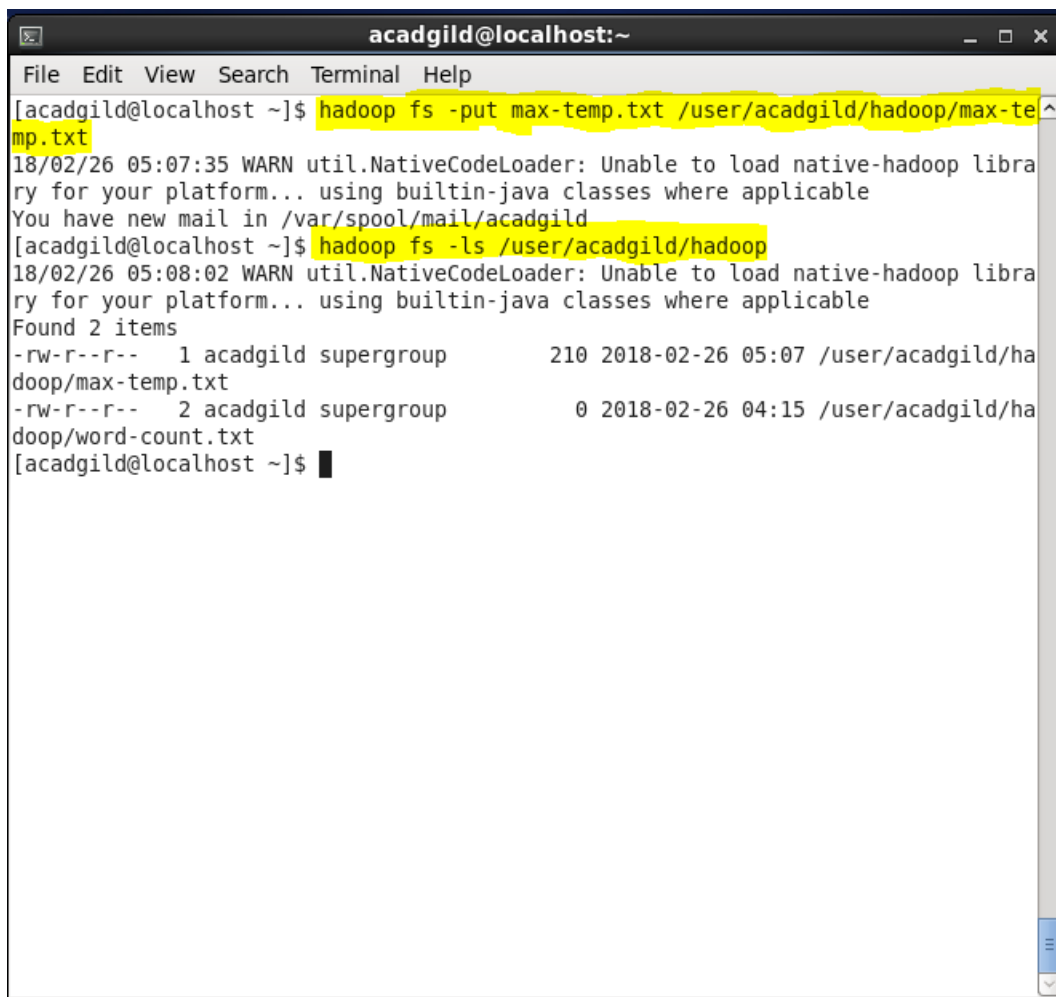
"hadoop fs -put max-temp.txt /user/acadgild/hadoop/max-temp.txt"

The above command will copy the max-temp.txt file to hdfs directory

"/user/acadgild/hadoop/"

Now to confirm whether the file is copied or not list the files in hdfs directory

Using the command **"hadoop fs -ls /user/acadgild/hadoop"**

A terminal window titled 'acadgild@localhost:~' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows two commands being executed. The first command is 'hadoop fs -put max-temp.txt /user/acadgild/hadoop/max-temp.txt', which is highlighted in yellow. It produces a warning message about native-hadoop library and a mail notification. The second command is 'hadoop fs -ls /user/acadgild/hadoop', also highlighted in yellow. It produces a similar warning and then lists two files in the directory: 'max-temp.txt' and 'word-count.txt'.

```
acadgild@localhost:~  
File Edit View Search Terminal Help  
[acadgild@localhost ~]$ hadoop fs -put max-temp.txt /user/acadgild/hadoop/max-temp.txt  
18/02/26 05:07:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ hadoop fs -ls /user/acadgild/hadoop  
18/02/26 05:08:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
Found 2 items  
-rw-r--r--  1 acadgild supergroup      210 2018-02-26 05:07 /user/acadgild/hadoop/max-temp.txt  
-rw-r--r--  2 acadgild supergroup       0 2018-02-26 04:15 /user/acadgild/hadoop/word-count.txt  
[acadgild@localhost ~]$
```

Task 4 Solution :

Q) Change the permission of the file /user/acadgild/hadoop/max-temp.txt, such that only the owner and

the group members have full control over the file.

Others do not have any control over it.

Sol) we have chmod command to change the read write and execute permissions for the

User,group user and others.

Ex : `chmod u=rwx,g=rw,o=r <filename>`

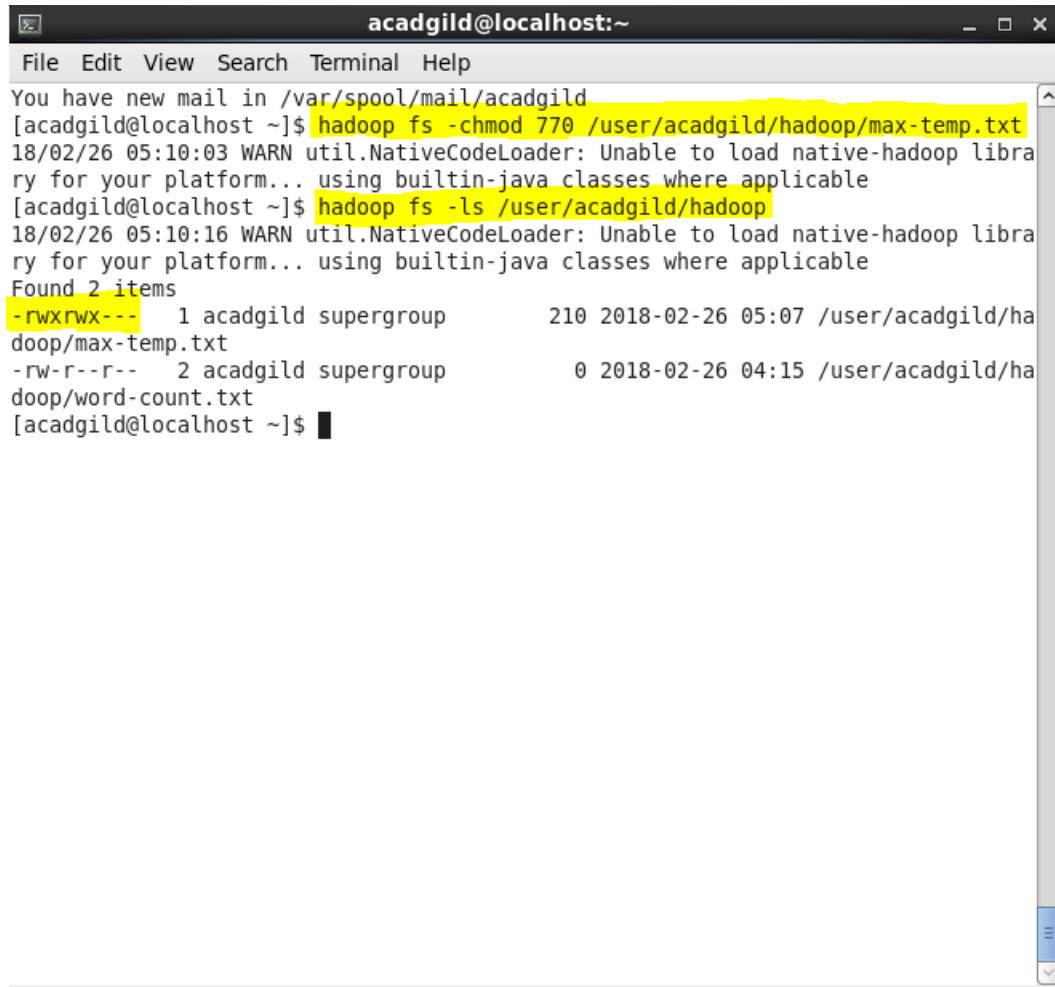
The letters **u**, **g**, and **o** stand for "**user**", "**group**", and "**other**". The equals sign ("=") means "set the permissions exactly like this," and the letters "**r**", "**w**", and "**x**" stand for "read", "write", and "execute", respectively. The commas separate the different classes of permissions, and there are no spaces in between them.

This is equivalent to "`chmod 754 <filename>`"

Here the digits **7**, **5**, and **4** each individually represent the permissions for the user, group, and others, in that order. Each digit is a combination of the numbers **4**, **2**, **1**, and **0**:

Now here we have to provide full permissions for user and group and no permissions to Others.

We have to use the command “`hadoop fs -chmod 770 /user/acadgild/hadoop/max-temp.txt`”.



```
acadgild@localhost:~  
File Edit View Search Terminal Help  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ hadoop fs -chmod 770 /user/acadgild/hadoop/max-temp.txt  
18/02/26 05:10:03 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
[acadgild@localhost ~]$ hadoop fs -ls /user/acadgild/hadoop  
18/02/26 05:10:16 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
Found 2 items  
-rwxrwx--- 1 acadgild supergroup      210 2018-02-26 05:07 /user/acadgild/hadoop/max-temp.txt  
-rw-r--r--  2 acadgild supergroup       0 2018-02-26 04:15 /user/acadgild/hadoop/word-count.txt  
[acadgild@localhost ~]$
```