Hadoop Assignment ---

Problem Statement 1:

1. Create a folder in HDFS by name “dir01” and move input1.txt , input2.txt and input3.txt into /dir01.

A screenshot of a computer screen

Description automatically generated

1. List only the file names present in “/dir01”
2. Change the replication factor for the content present in directory “/dir01” to 5 and display the replication factor for the files present in “/dir01”.

A screenshot of a computer

Description automatically generated

1. Create a folder in HDFS by name “scenario01” and create directory “level01” inside “scenario01” directory and create another directory “level02” inside directory “level01”. Once the required directories are created copy input1.txt to scenario01, input2.txt to level01 and input3.txt to level02 and finally recursively print only the file names present in directory scenario01

A screenshot of a computer

Description automatically generated

Problem statement 2:----

1. Run MapReduce Program and capture the application Id of the job

A screen shot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Re-run the MapReduce program and kill the application using the yarn command.

A computer screen shot of a computer screen

Description automatically generated

1. List all the applications which are RUNNING state  
   yarn application -appStates RUNNING -list | grep "applicationName"

A screenshot of a computer

Description automatically generated

1. View the logs of any of the jobs which are already completed

A screenshot of a computer

Description automatically generated

A screenshot of a computer screen

Description automatically generated