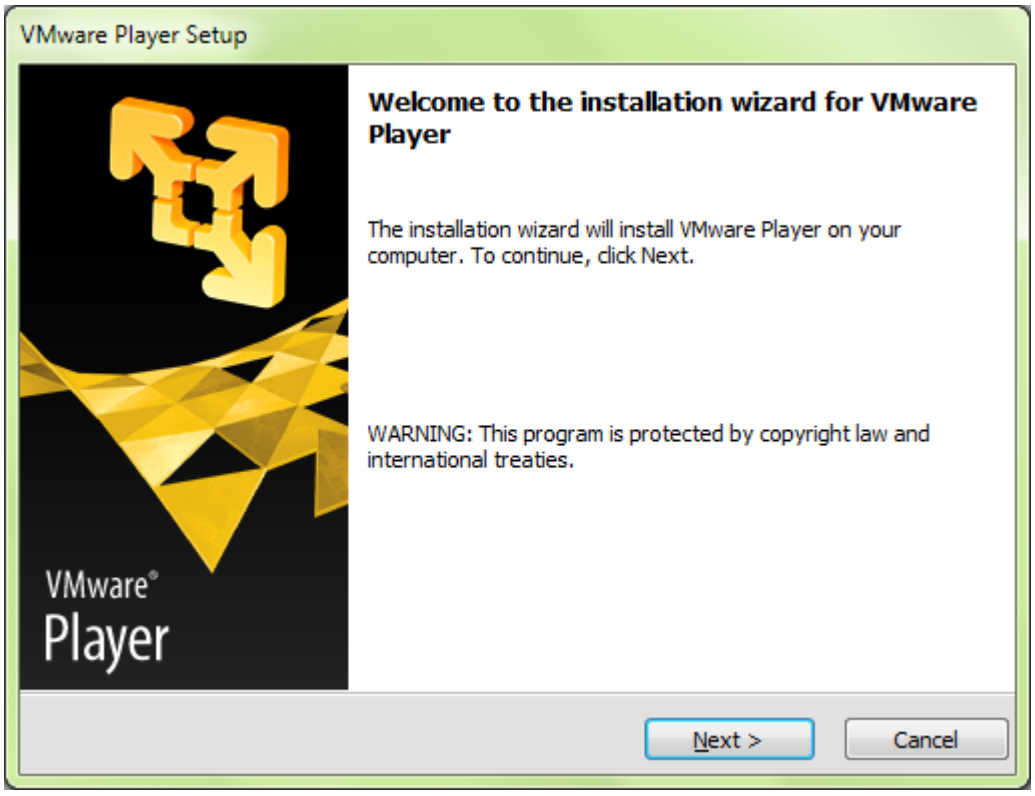
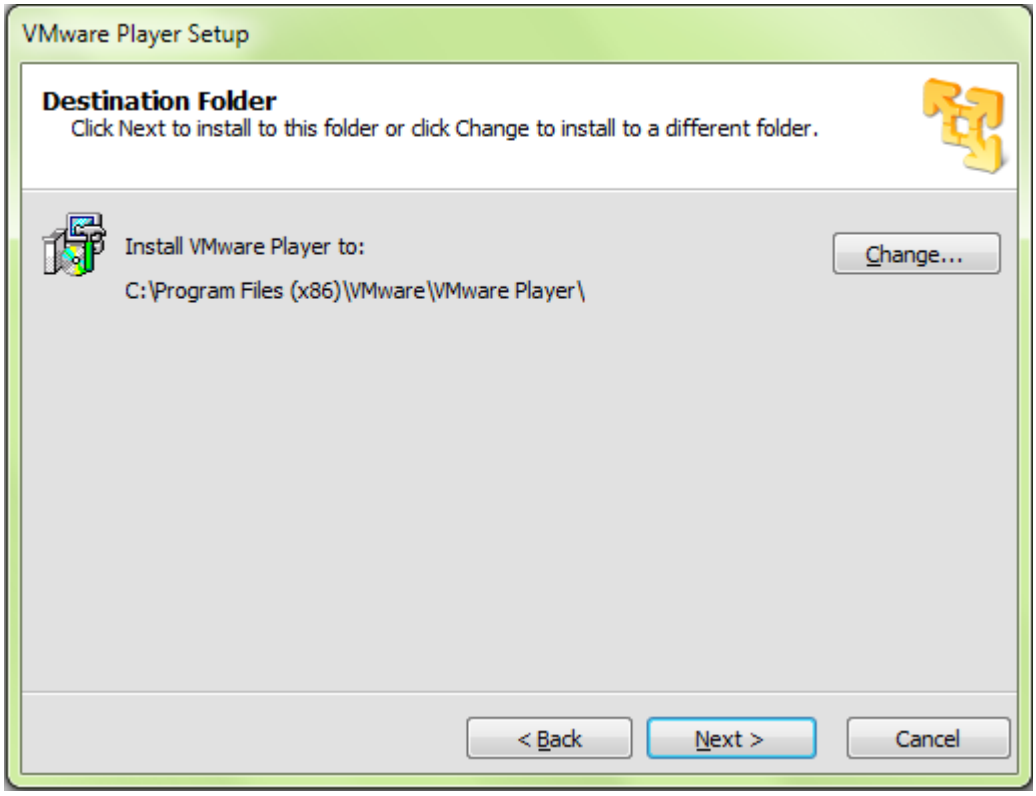
Mahout Installation Steps

Step 1: Download and install VMware Player

Download the “VMware player” from the link shown above and install it.

VMware player: https://www.vmware.com/tryvmware/?p=player





Step 2: Downloading the Cloudera CDH3 Setup file

Download the Cloudera VM file from **any** of the below link.

https://docs.google.com/file/d/0B-_P02gj6T2mTGZKS3JzUTM3bjA/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mSUllcmhIdlZiM0E/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mVktsWHdXN3ktRWc/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mSXF2SUEyMS1wVDA/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mS1FtNG9sUzhzcWc/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mOGFfaE52RjNuenc/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mOTVhaDlFVUlQaDQ/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mQkZJNS1fTUVtMEE/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mZ1JaZEp6T2xYYTg/edit?usp=sharing
https://docs.google.com/file/d/0B-_P02gj6T2mcHhBeE5WV2JielE/edit?usp=sharing

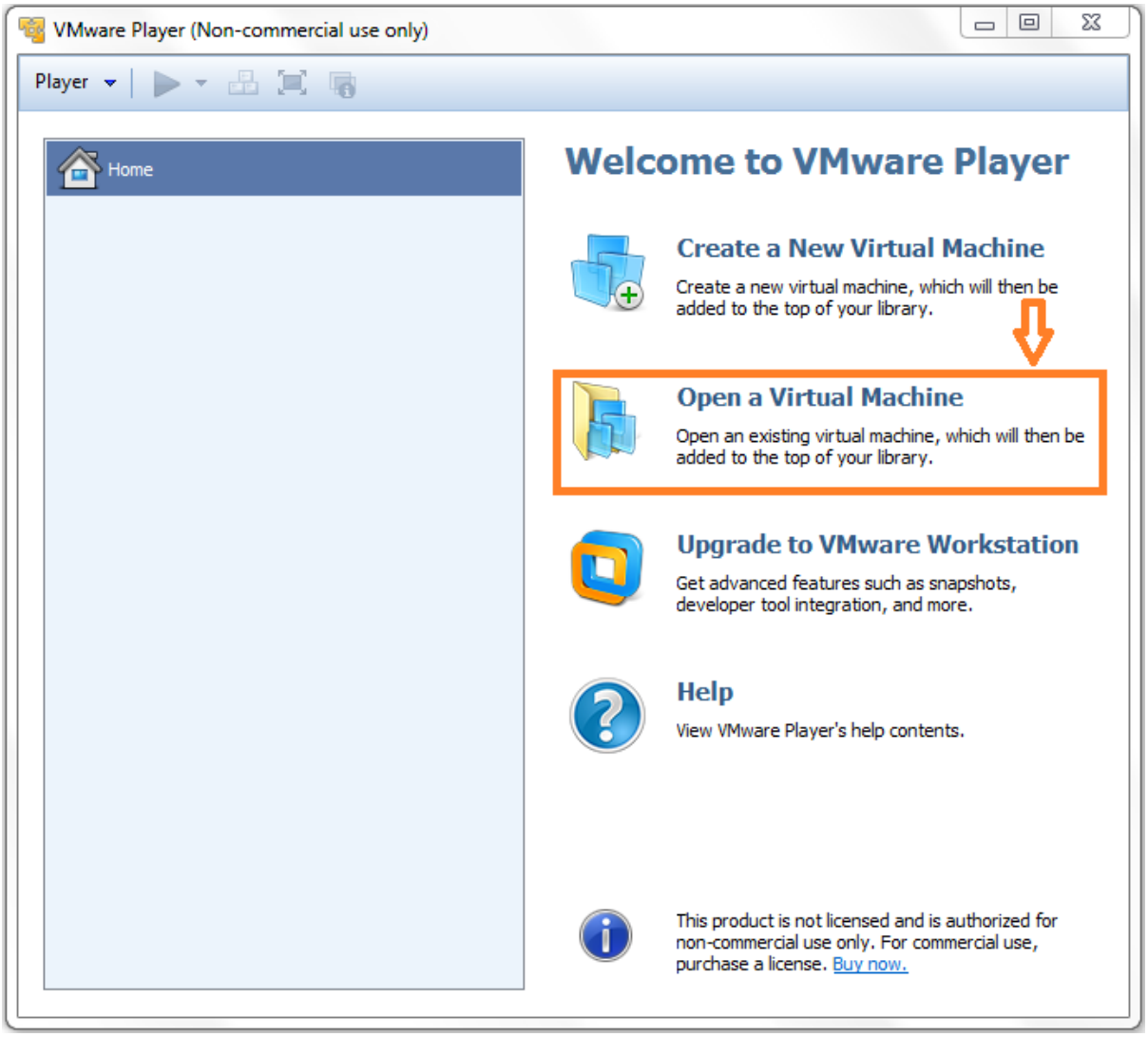
Step 3: Extracting the Cloudera Downloaded.

The Cloudera file you have downloaded will be having the extension tar.bz2 to open it we need to extract the Cloudera. You can use the WinRAR or WinZip.

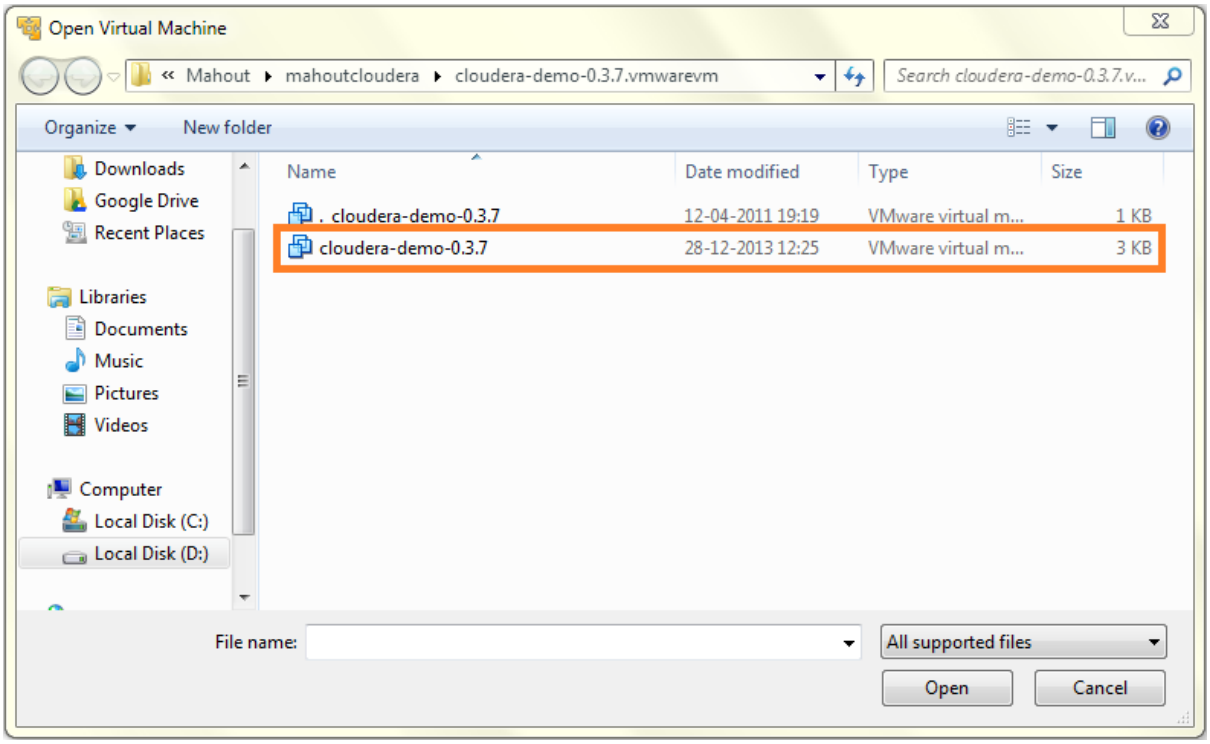
Note: If you do not have them you can download the WinRAR from the below link based on your OS and the bit-version.
http://www.rarlab.com/download.htm

Step 4: Opening the Cloudera using VMWare Player.

To open Cloudera start the VMPlayer and click on **Open a Virtual Machine** as shown in the below image.

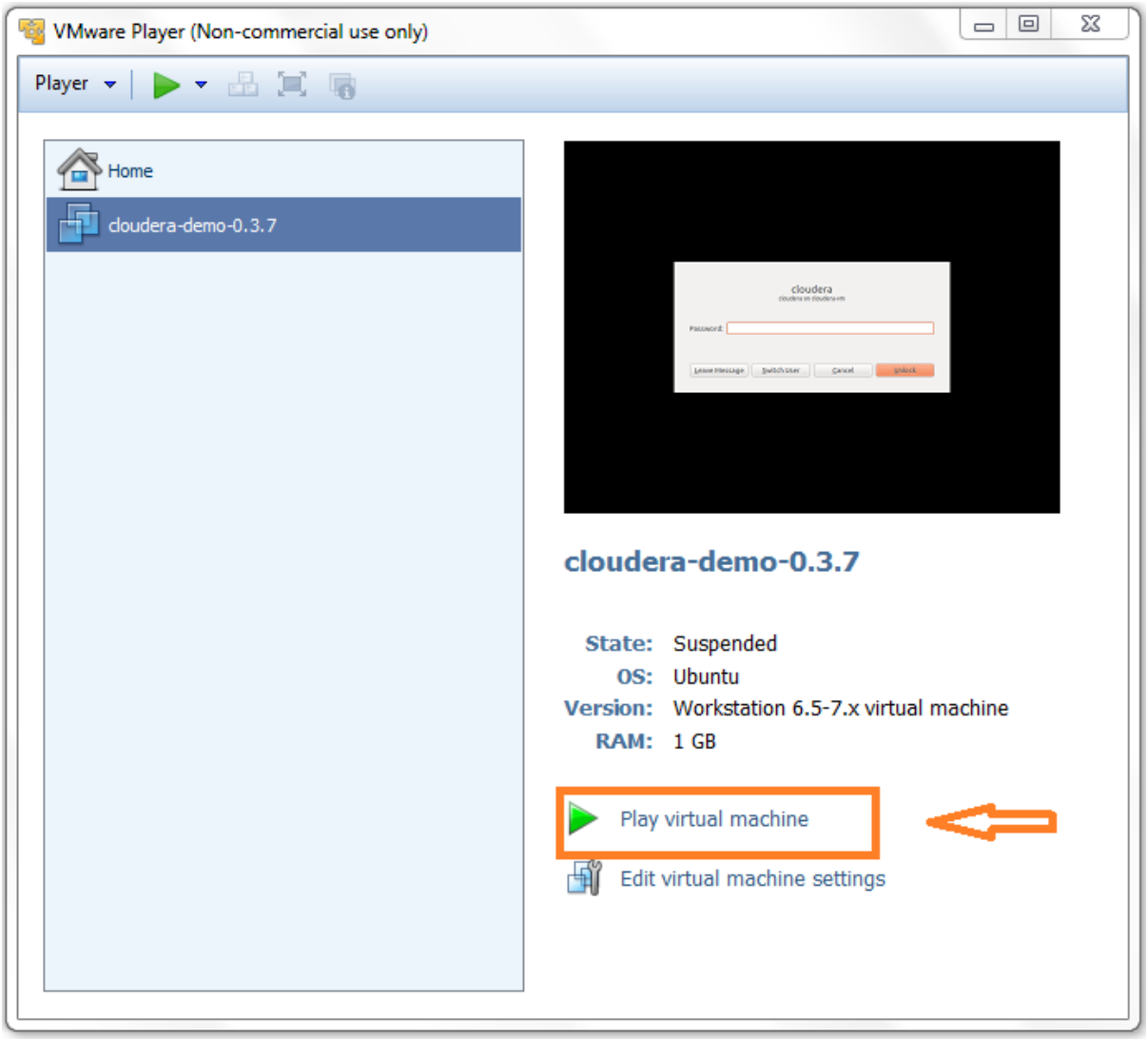


Select the location where you have extracted the Cloudera and then select the file which is of 3 KB size as shown in the below image.

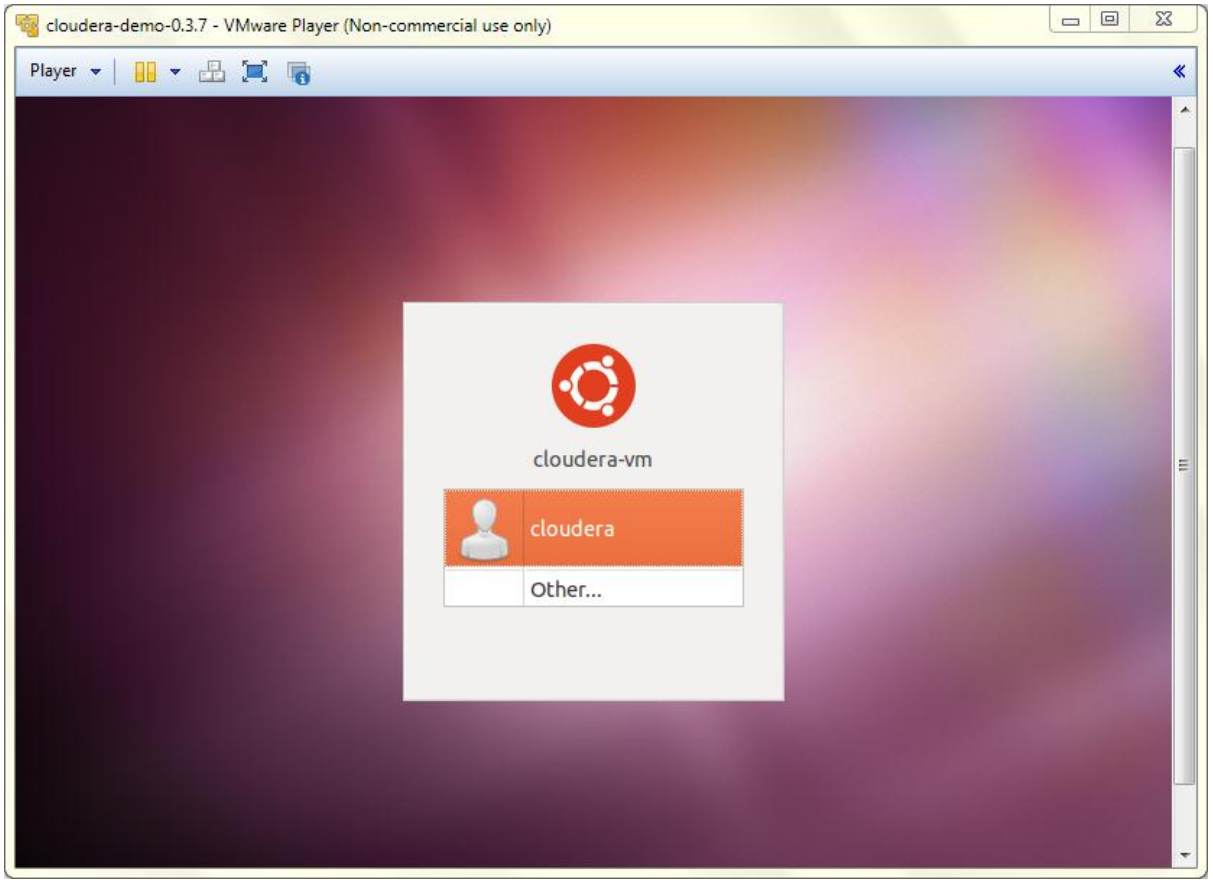


Step 5: Play virtual machine.

Now click on Play Virtual Machine as shown in the following image.

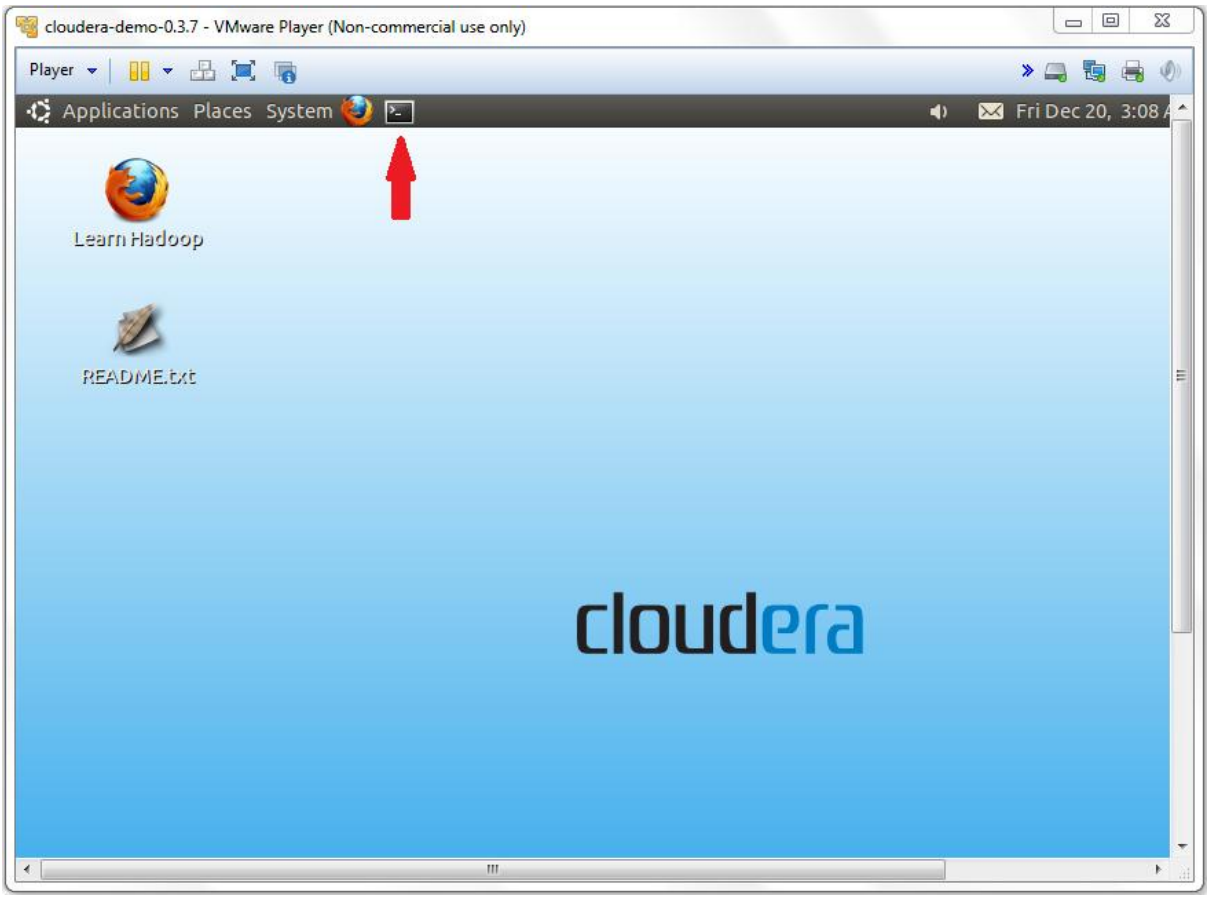


You will find the below image when the Cloudera is opened.
Select the User Cloudera and give the Password **– cloudera**


cloudera-demo-0.3.7 - VMware Player (Non-commercial use only)
Player
cloudera-vm
cloudera
Other...

Step 6: Open a terminal

After logging you will find the below image. Open the terminal by clicking on the terminal symbol highlighted in the below image.



Step 7: Downloading the Mahout Distribution file.

Command:

wget http://apache.spinellicreations.com/mahout/0.8/mahout-distribution-0.8-src.tar.gz

```
cloudera@cloudera-vm:~$ wget http://apache.spinellicreations.com/mahout/0.8/maho
ut-distribution-0.8-src.tar.gz
--2013-12-27 01:53:04--  http://apache.spinellicreations.com/mahout/0.8/mahout-d
istribution-0.8-src.tar.gz
Resolving apache.spinellicreations.com... 72.88.94.11
Connecting to apache.spinellicreations.com|72.88.94.11|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 7405016 (7.1M) [application/x-gzip]
Saving to: `mahout-distribution-0.8-src.tar.gz'

100%[======================================>] 7,405,016    460K/s   in 19s

2013-12-27 01:53:23 (388 KB/s) - `mahout-distribution-0.8-src.tar.gz' saved [740
5016/7405016]
```

Step 8: Check if it is downloaded.

Command: ls

```
cloudera@cloudera-vm:~$ ls
cloudera  Desktop  mahout-distribution-0.8-src.tar.gz
```

Step 9: Extracting the downloaded tar file

Command: tar -xvf mahout-distribution-0.8-src.tar.gz

```
cloudera@cloudera-vm:~$ tar -xvf mahout-distribution-0.8-src.tar.gz
```

Step 10: Check if the folder is extracted.

Command: ls

```
cloudera@cloudera-vm:~$ ls
cloudera  Desktop  mahout-distribution-0.8  mahout-distribution-0.8-src.tar.gz
```

Step 11: Make directory for Mahout

Command: sudo mkdir /usr/lib/mahout

Note: As we do not have the permissions to create a directory in the directory /usr/lib we use sudo which makes you as root user and gives us that permissions.

```
cloudera@cloudera-vm:~$ sudo mkdir /usr/lib/mahout
[sudo] password for cloudera:
```

When it asks for password give **cloudera.**

Step 12: Move the extracted folder to the Mahout directory

Command: sudo mv mahout-distribution-0.8/* /usr/lib/mahout/

```
cloudera@cloudera-vm:~$ sudo mv mahout-distribution-0.8/* /usr/lib/mahout/
```

Step 13: Change file permission of the bin folder of Mahout to rwx.

Command: sudo chmod -R +rwx /usr/lib/mahout/bin

```
cloudera@cloudera-vm:/usr/lib/mahout$ sudo chmod -R +rwx /usr/lib/mahout/bin
```

Step 14: Download Maven tar file

Command: wget http://www.trieuvan.com/apache/maven/maven-3/3.1.1/binaries/apache-maven-3.1.1-bin.tar.gz

```
cloudera@cloudera-vm:~$ wget http://www.trieuvan.com/apache/maven/maven-3/3.1.1/
binaries/apache-maven-3.1.1-bin.tar.gz
--2013-12-27 05:31:45--  http://www.trieuvan.com/apache/maven/maven-3/3.1.1/bina
ries/apache-maven-3.1.1-bin.tar.gz
Resolving www.trieuvan.com... 66.201.46.168
Connecting to www.trieuvan.com|66.201.46.168|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 5494427 (5.2M) [application/x-gzip]
Saving to: `apache-maven-3.1.1-bin.tar.gz'

100%[======================================>] 5,494,427    330K/s   in 17s

2013-12-27 05:32:02 (320 KB/s) - `apache-maven-3.1.1-bin.tar.gz' saved [5494427/
5494427]
```

Step 15: Check if it is downloaded.

Command: ls

```
cloudera@cloudera-vm:~$ ls
apache-maven-3.1.1-bin.tar.gz  mahout-distribution-0.8
cloudera                       mahout-distribution-0.8-src.tar.gz
Desktop
```

Step 16: Extract the downloaded file

Command: tar -xvf apache-maven-3.1.1-bin.tar.gz

```
cloudera@cloudera-vm:~$ tar -xvf apache-maven-3.1.1-bin.tar.gz
```

Step 17: Check the extracted folder

Command: ls

```
cloudera@cloudera-vm:~$ ls
apache-maven-3.1.1             cloudera  mahout-distribution-0.8
apache-maven-3.1.1-bin.tar.gz  Desktop   mahout-distribution-0.8-src.tar.gz
```

Step 18: Make a directory for Maven

Command: sudo mkdir /usr/lib/maven

```
cloudera@cloudera-vm:~$ sudo mkdir /usr/lib/maven
```

Step 19: Move the extracted folder to the created directory

Command: sudo mv apache-maven-3.1.1/* /usr/lib/maven

```
cloudera@cloudera-vm:~$ sudo mv apache-maven-3.1.1/* /usr/lib/maven
[sudo] password for cloudera:
```

Step 20: Export the path.

Command: export PATH=$PATH:/usr/lib/maven/bin/

```
cloudera@cloudera-vm:~$ export PATH=$PATH:/usr/lib/maven/bin/
```

Step 21: Change the directory.

Command: cd /usr/lib/mahout

```
cloudera@cloudera-vm:~$ cd /usr/lib/mahout
cloudera@cloudera-vm:/usr/lib/mahout$
```

Step 22: Compile and Testing Mahout.

Command: sudo /usr/lib/maven/bin/mvn -DskipTests install

```
cloudera@cloudera-vm:/usr/lib/mahout$ sudo /usr/lib/maven/bin/mvn -DskipTests in
stall
[INFO] Scanning for projects...
Downloading: http://repo.maven.apache.org/maven2/org/apache/apache/9/apache-9.po
m
Downloaded: http://repo.maven.apache.org/maven2/org/apache/apache/9/apache-9.pom
 (15 KB at 7.0 KB/sec)
[INFO] ------------------------------------------------------------------------
[INFO] Reactor Build Order:
[INFO]
[INFO] Mahout Build Tools
[INFO] Apache Mahout
[INFO] Mahout Math
[INFO] Mahout Core
```

It will take some time. Hence please wait until you find the next command line.

Step 23: change to home directory.

Command: cd ~

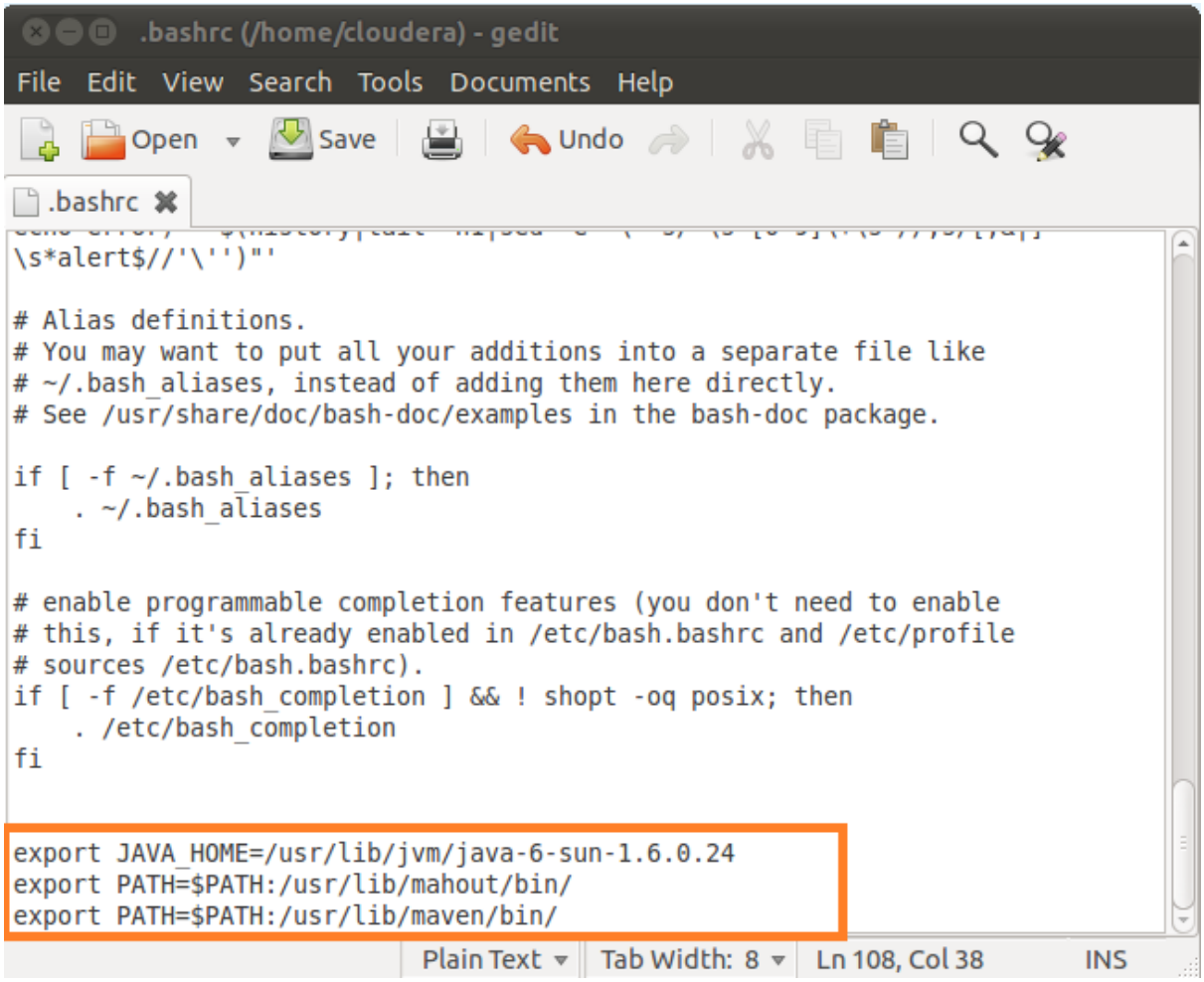
```
cloudera@cloudera-vm:/usr/lib/mahout$ cd ~
```

Step 24: Setting the environment variables in bashrc file.

Command: sudo gedit .bashrc

```
cloudera@cloudera-vm:~$ sudo gedit .bashrc
```

Go to end of the file and add the line as in the screen shot.



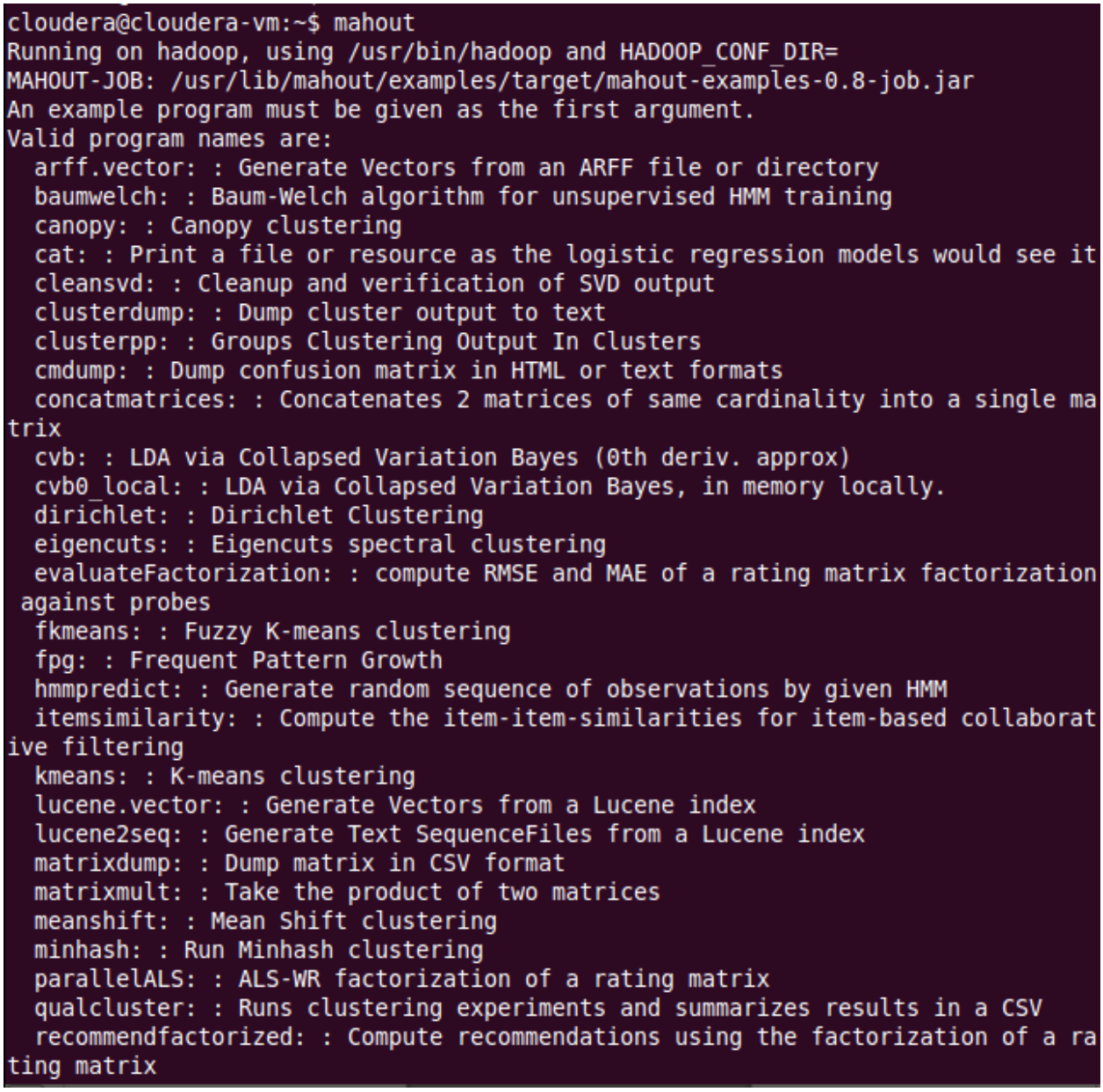
Step 25: Load the changes made.

Command: source .bashrc

```
cloudera@cloudera-vm:~$ source .bashrc
```

Step 26: Let us check if Mahout is working or not.

Command: mahout



If it displays the above image then you have successfully installed Mahout.