

Laxman Dutt Degala | CS560 | February 13, 2014

Lab 3

MASHUP aPP/USE HTML5 INTERNAL STORAGE/iNSTALL CLOUDERA/RUN WORDCOUNT/INSTALL MAHOUT/RUN Naive Bayes classifier

Lab3 assignment has 3 tasks

1. Mashup Application using HTML5 internal storage
2. Install Cloudera image and run WordCount Mapreduce
3. Configure Cloudera with Mahout and Run Naïve Bayes classifier

Task 1:

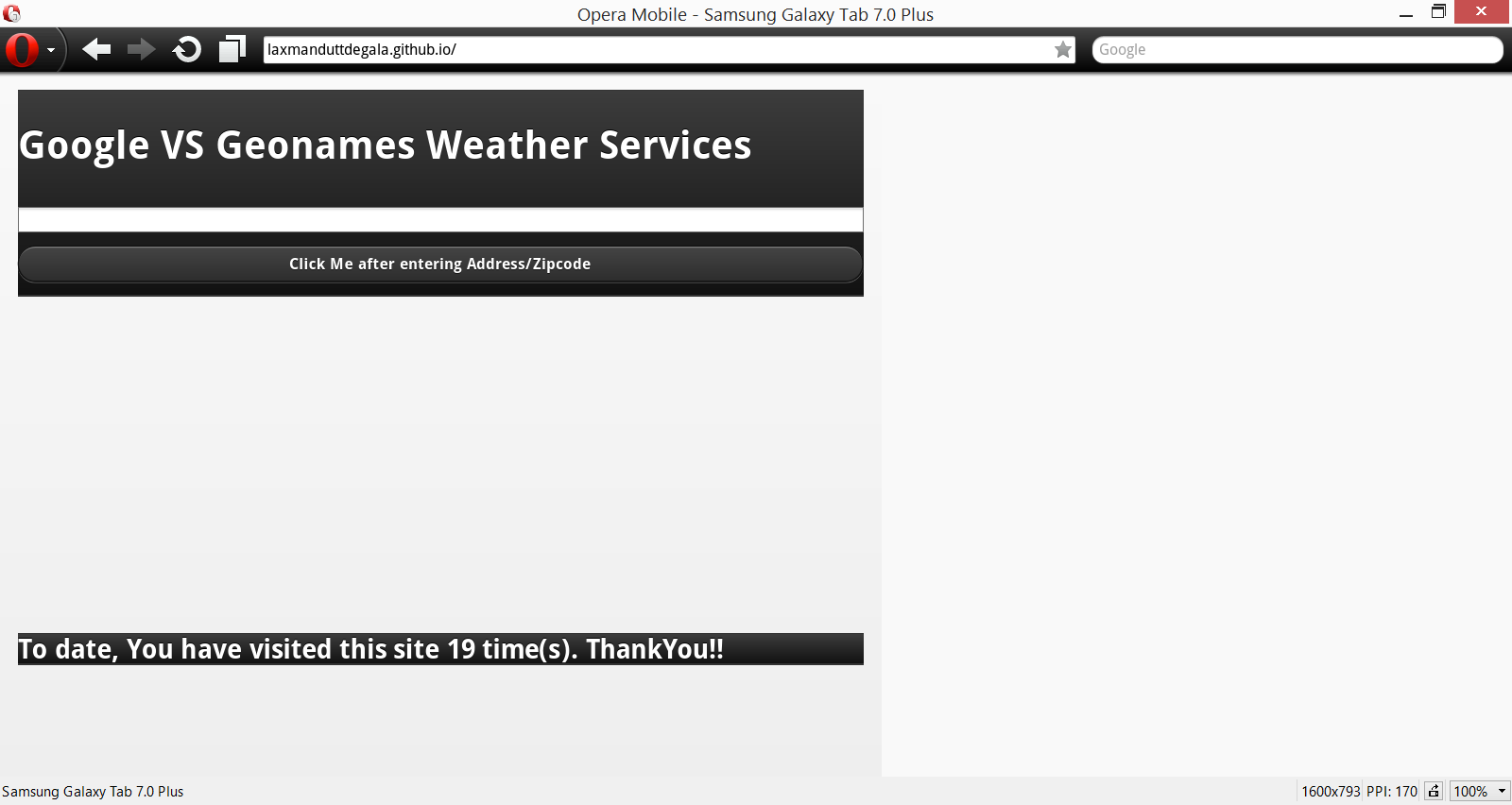
Mashup Application:

1. This application is a mash-up of Google weather services/Google Maps/Geonames weather services.
2. This application also features HTML5 Local Storage, where user is informed about the number of hits he did to the site till date.
3. Once a user enters address
4. It will retrieve weather information from Google weather services/maps.
5. It will also retrieve weather information from Geonames

Its kind of a weather compare between two weather services – Google n Geo names in this case.

Mobile version of the application can be seen using the url : **laxmanduttdegala.github.io**

Below snapshot shows the initial screen, when we enter the above stated URL in *Opera mobile*.



Entered ‘64110’, which is the zip code of Kansas city



**Code for HTML5 Internal storage**:

if(typeof(Storage)!=="undefined")

{

if (localStorage.clickcount)

{

localStorage.clickcount=Number(localStorage.clickcount)+1;

}

else

{

localStorage.clickcount=1;

}

if (localStorage.clickcount==1)

{

document.getElementById("sitecount").innerHTML="Welcome!! It's the first time you are visiting this web page...";

}

else

{

document.getElementById("sitecount").innerHTML="To date, You have visited this site " + localStorage.clickcount + " time(s). ThankYou!!";

}

}

else

{

document.getElementById("sitecount").innerHTML="Sorry, your browser does not support web storage...";

}

Task 2: Install Cloudera image and execute WordCount Mapreduce

Step 1: Download Oracle VM VirtualBox

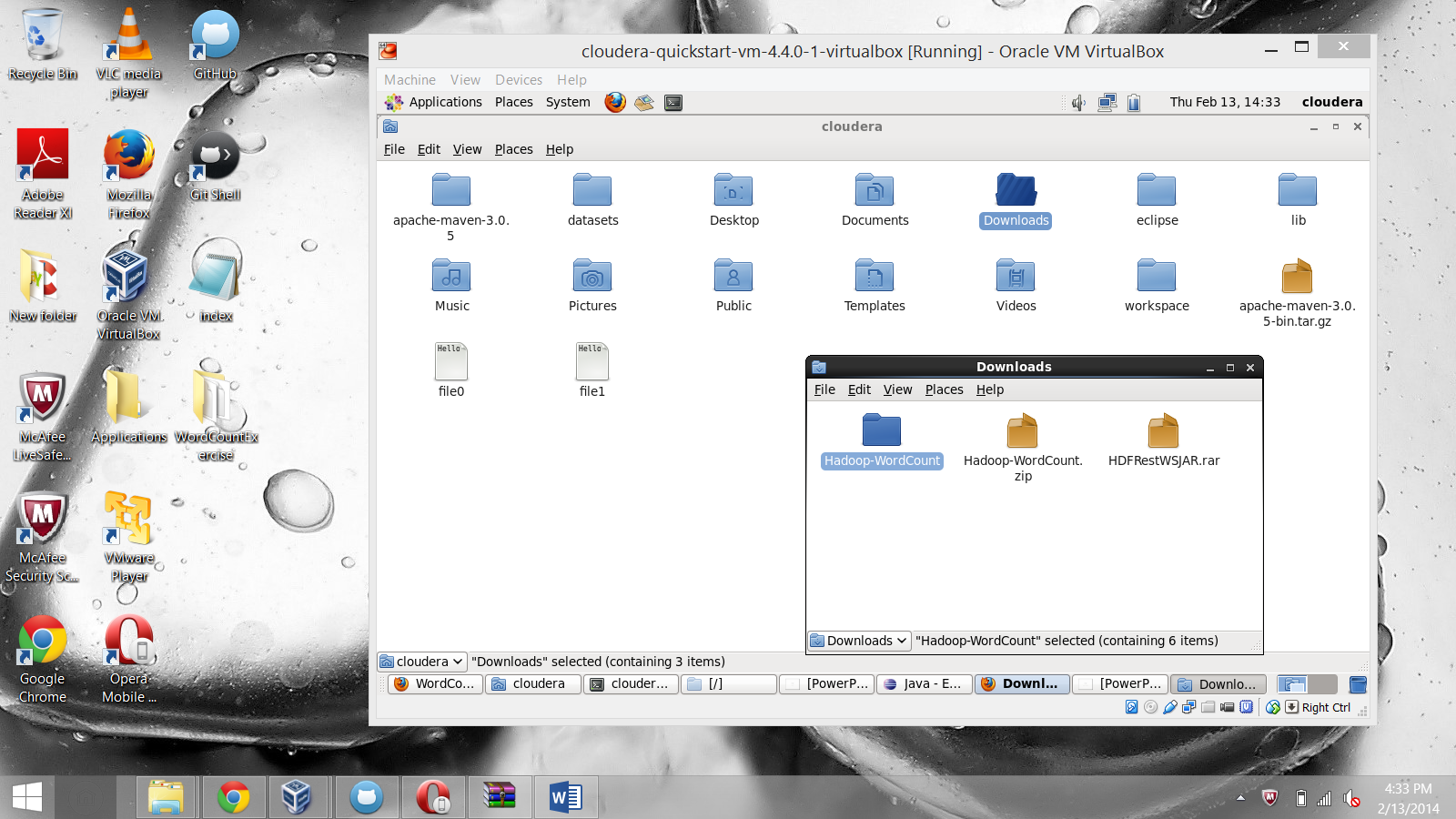
Step 2: Installed cloudera on CentOS system

Tried using VMware player for Virtual box but it didn’t run due to some issue. Then, used Oracle VM Virtual Box.

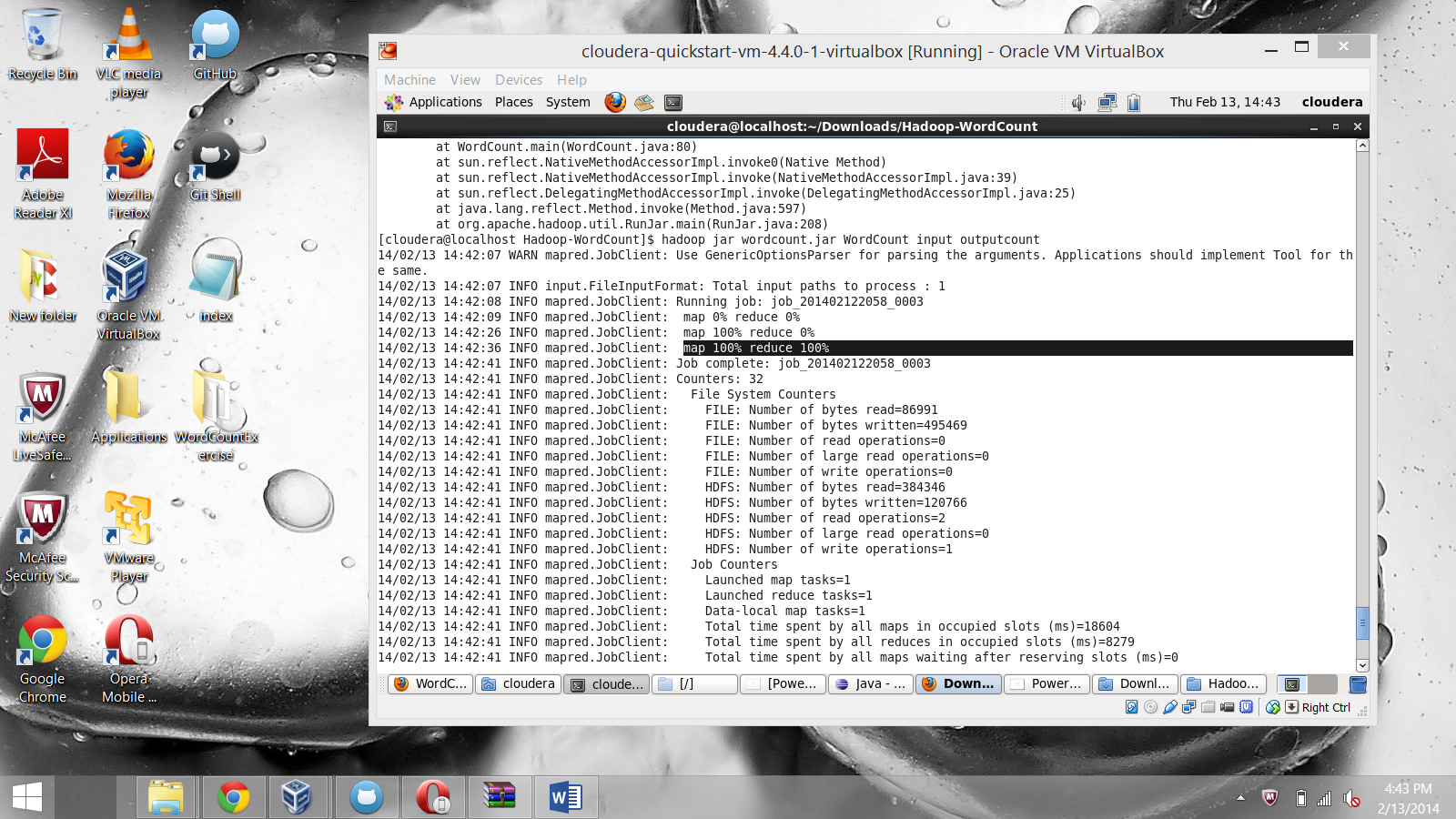
Step 3: Downloaded WordCount.zip from web and unzipped it to run MapReduce

Please find below the screen shots for the same.

Download WordCount and unzip using the terminal command : unzip hadoop-WordCount.zip



Map : 100% Reduce : 100%

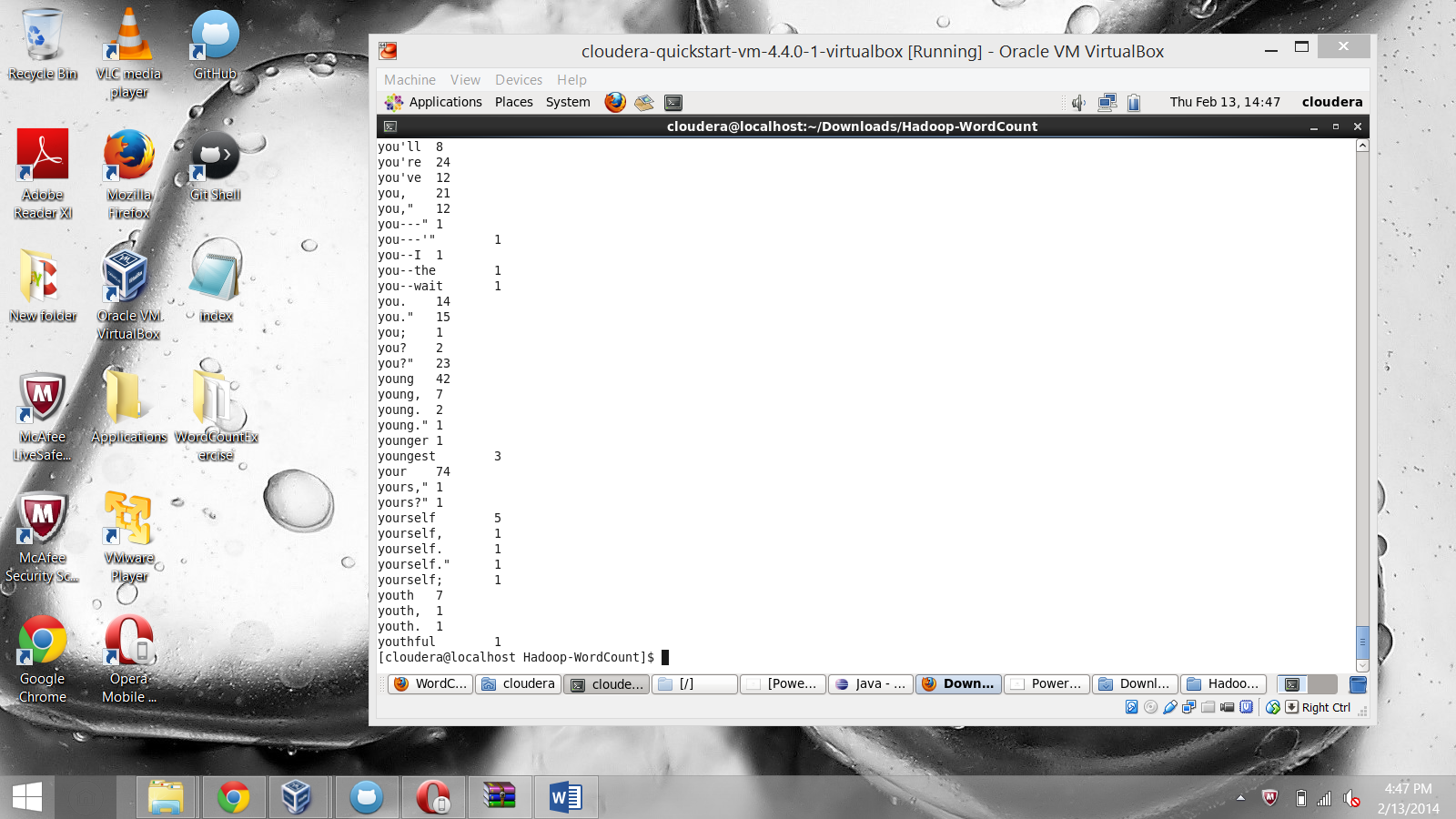


To see the contents of Output:

Hadoop fs –cat outputcount/\*



Final output of WordCount:

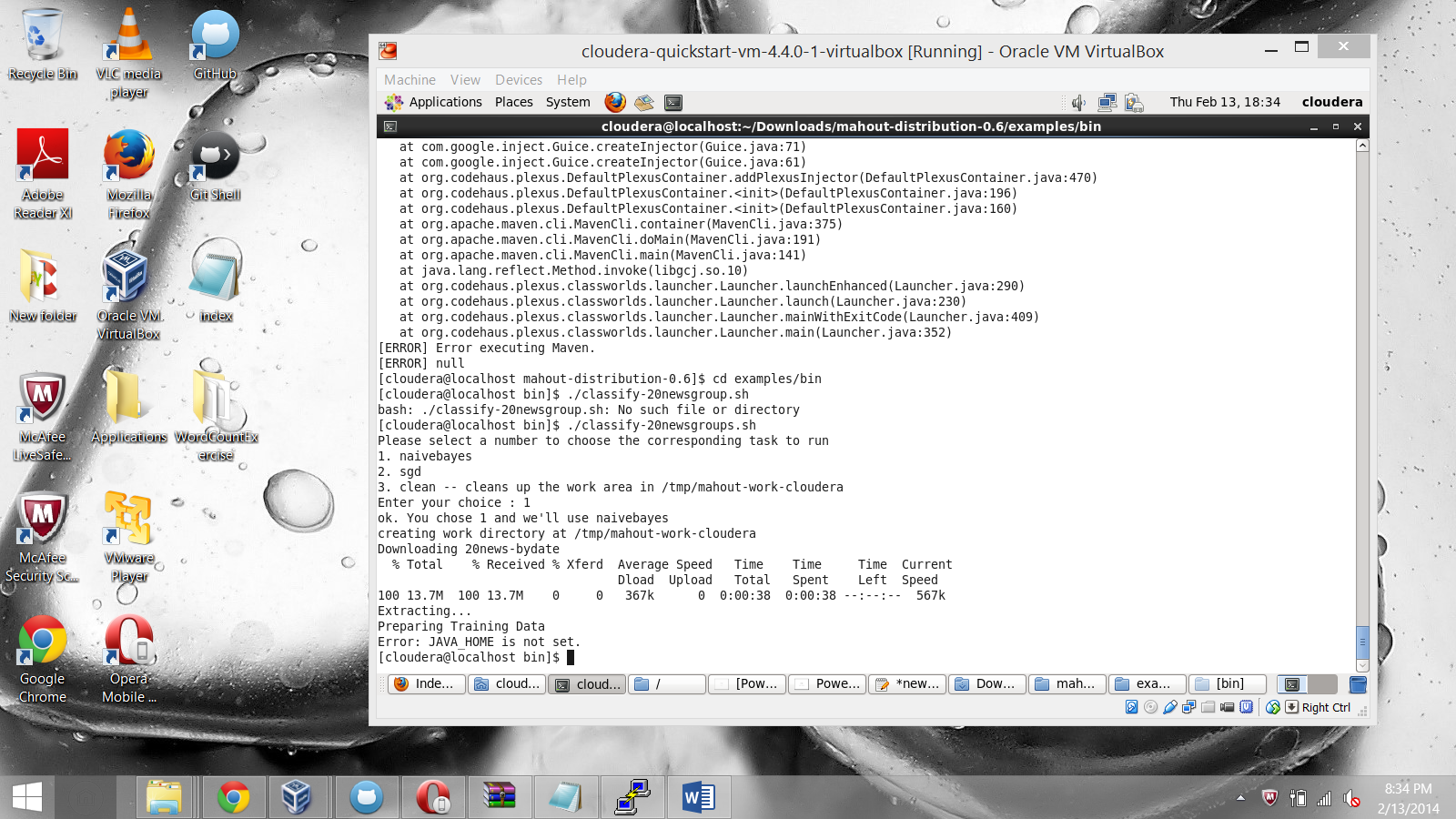


Task 3:

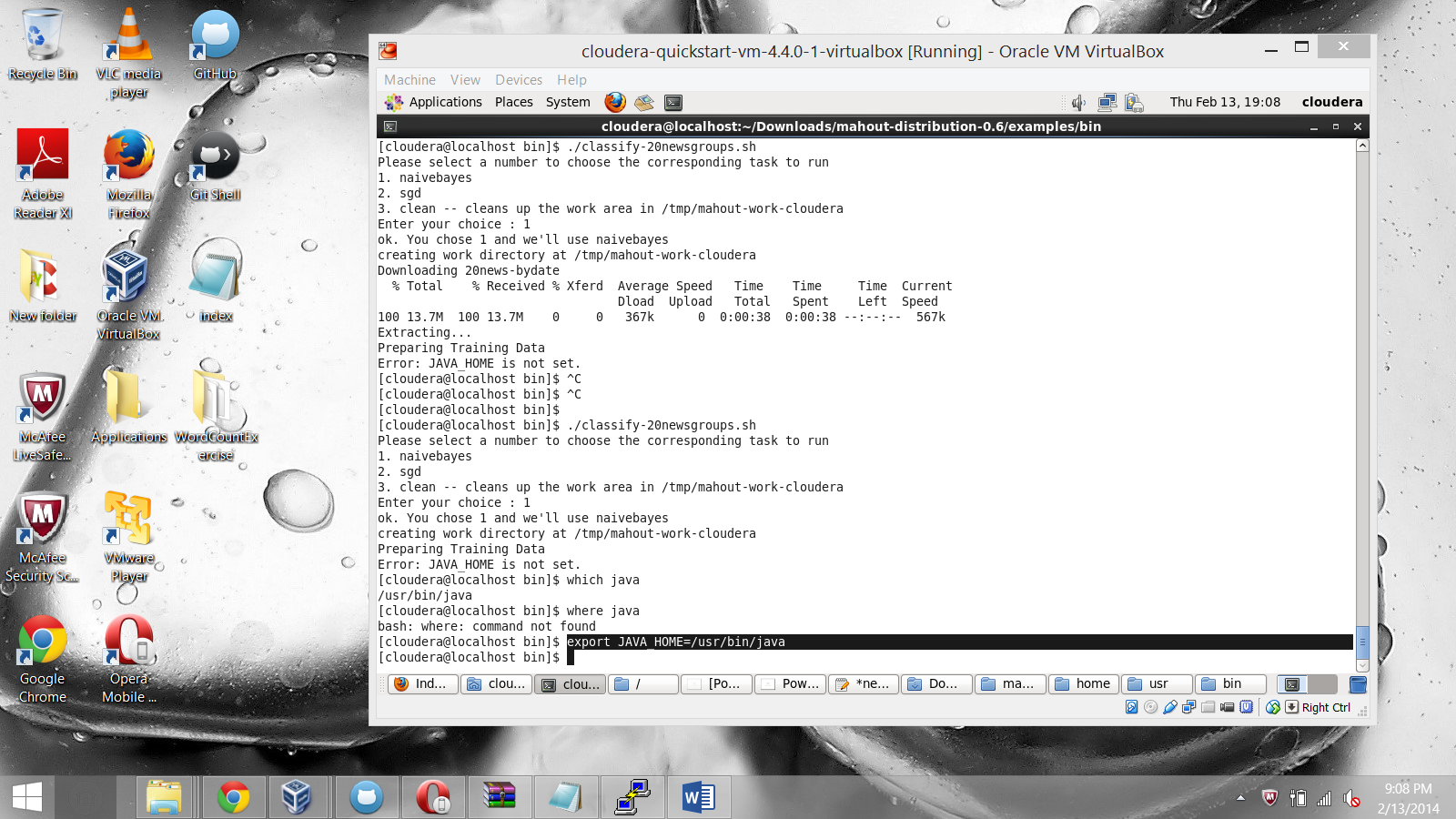
 Configure Cloudera with Mahout. Run Naive Bayes classifier.

Install Mahout as per steps in the Presentation. Please find the screen shots below:

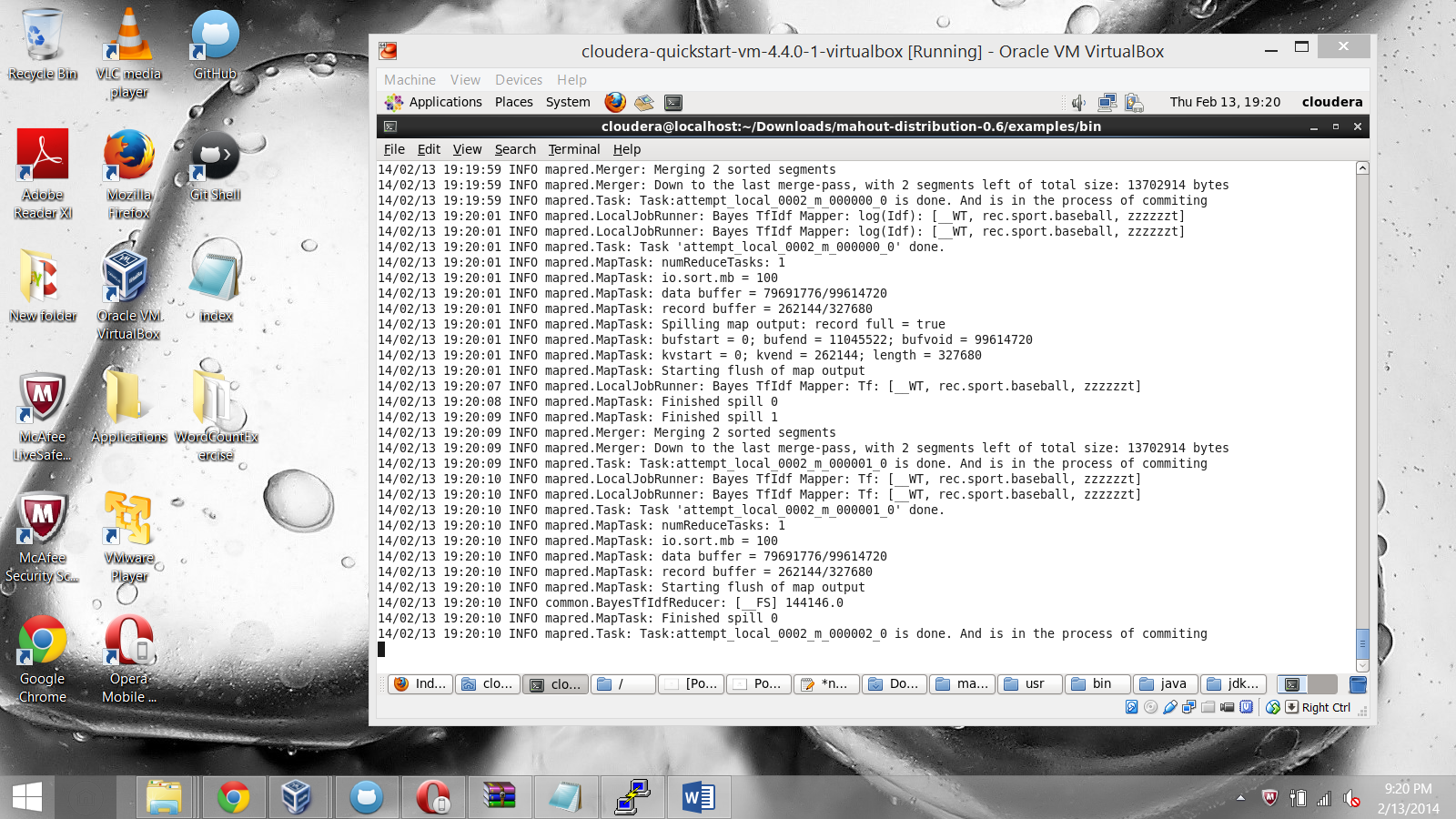
Error: JAVA HOME not found



Export JAVA\_HOME = /usr/java/java JDK



Running….



Final output of Naïve Bayes classifier:

