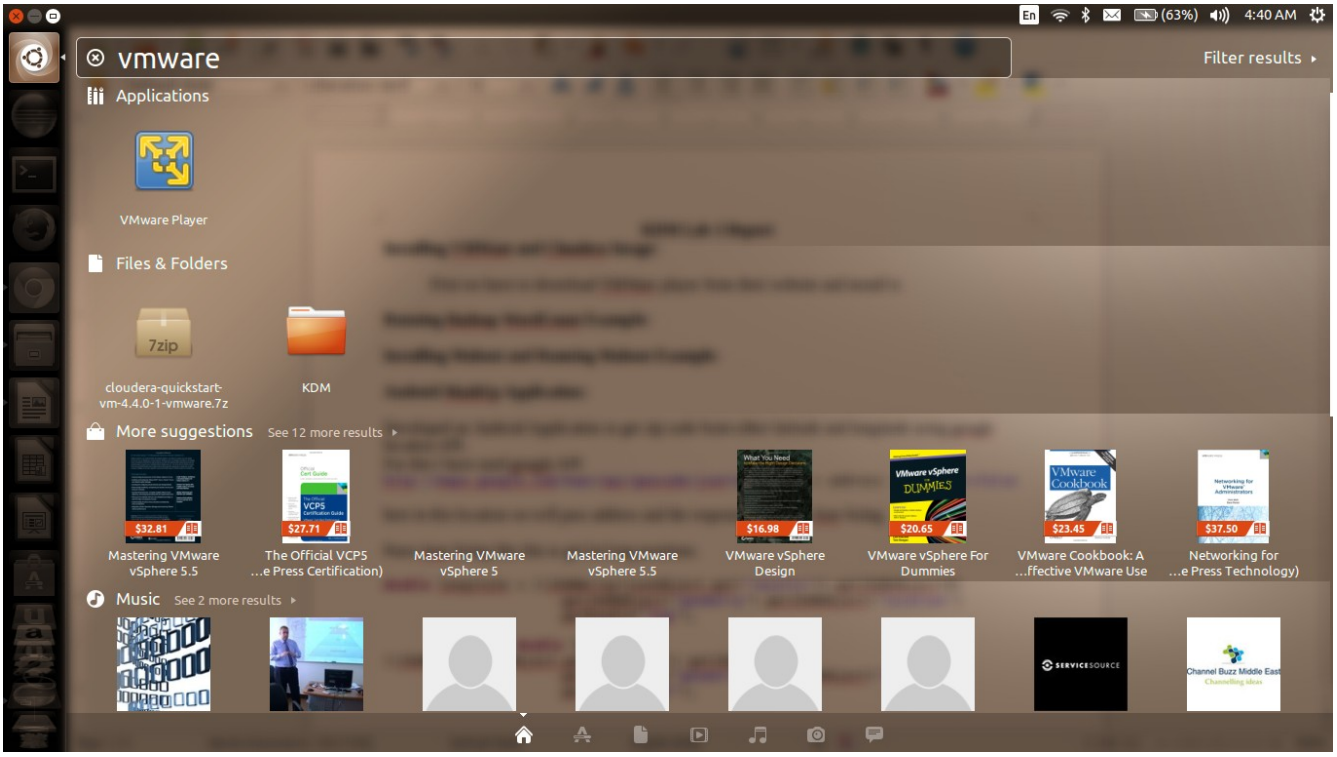
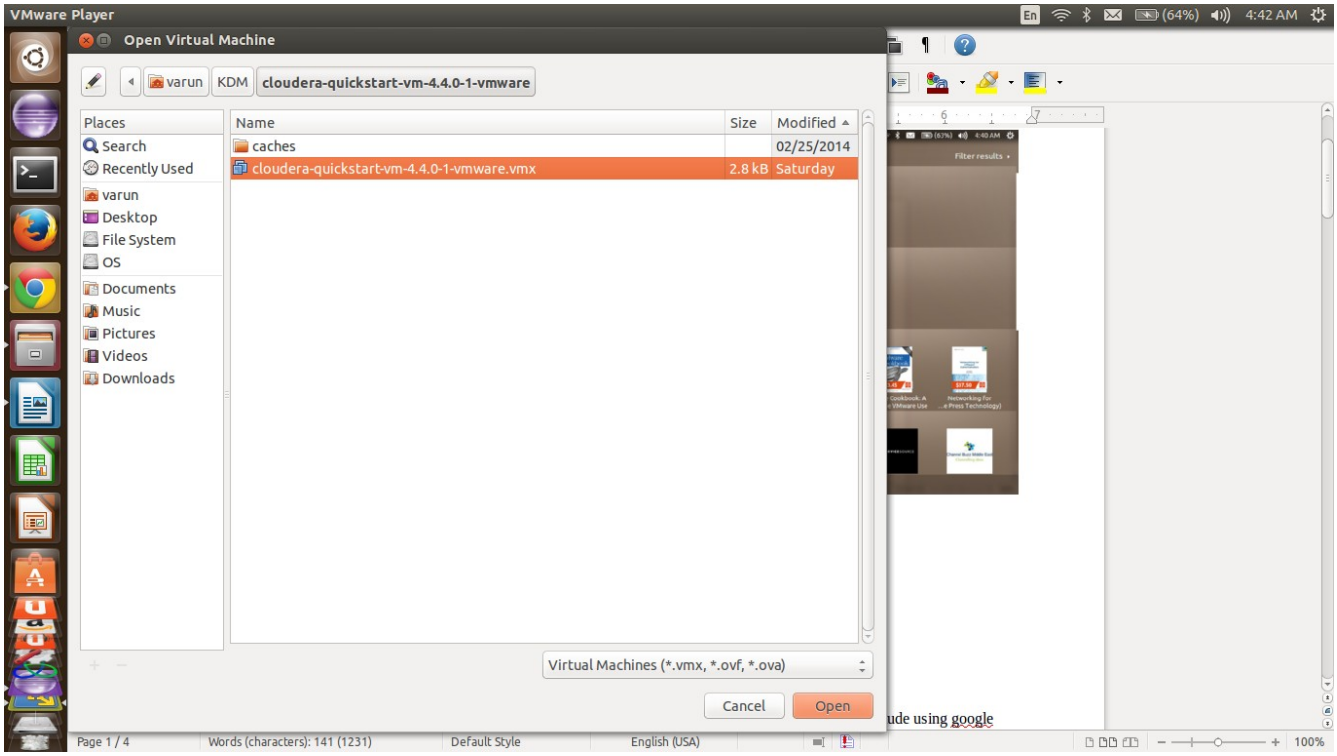
KDM Lab-3 Report

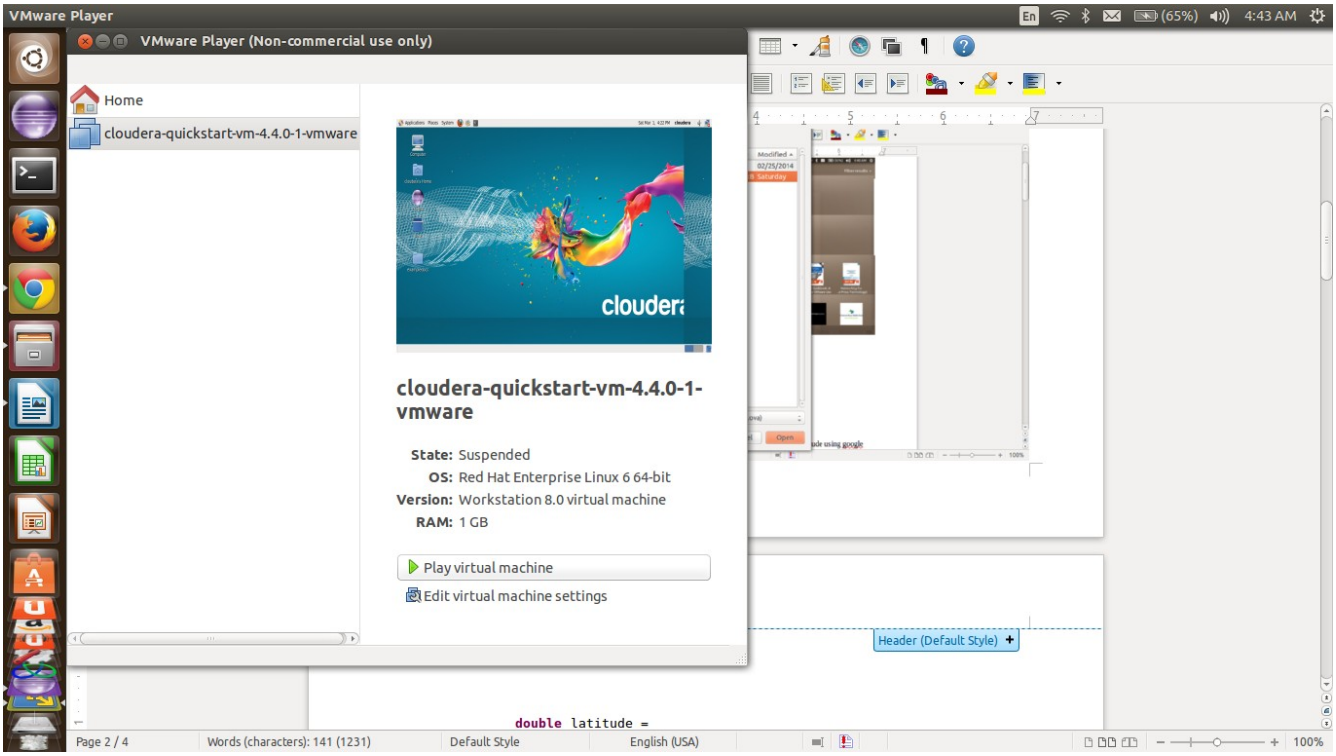
Installing VMWare and Cloudera Image:

First we have to download VMWare player from their website and install it.



After that download cloudera image and open it using vmware player





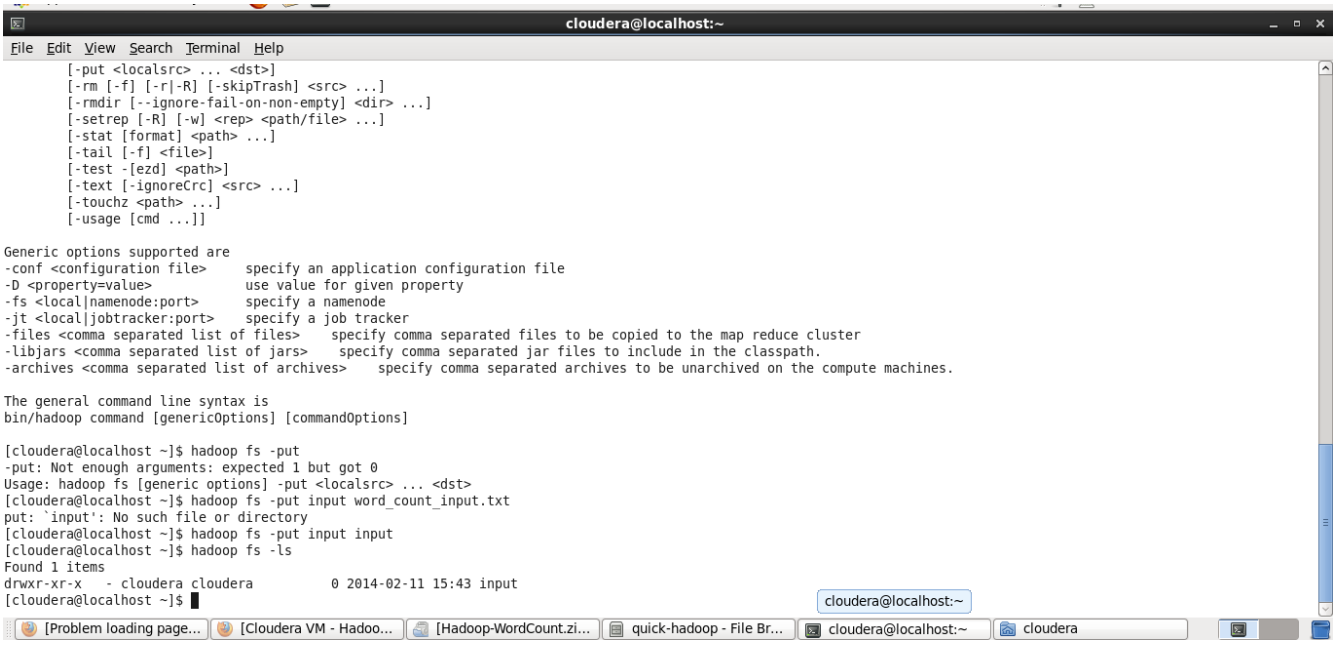
We can login to cloudera manager with
username - “cloudera”
password - “cloudera”

Running Hadoop WordCount Example:

First download HadoopWordcount example jar file and copy it to some location on our virtual machine.

Now go to that location and create a input folder and drop the input data to that folder and now we need to push this input to the hadoop distributed file system this can be done using command

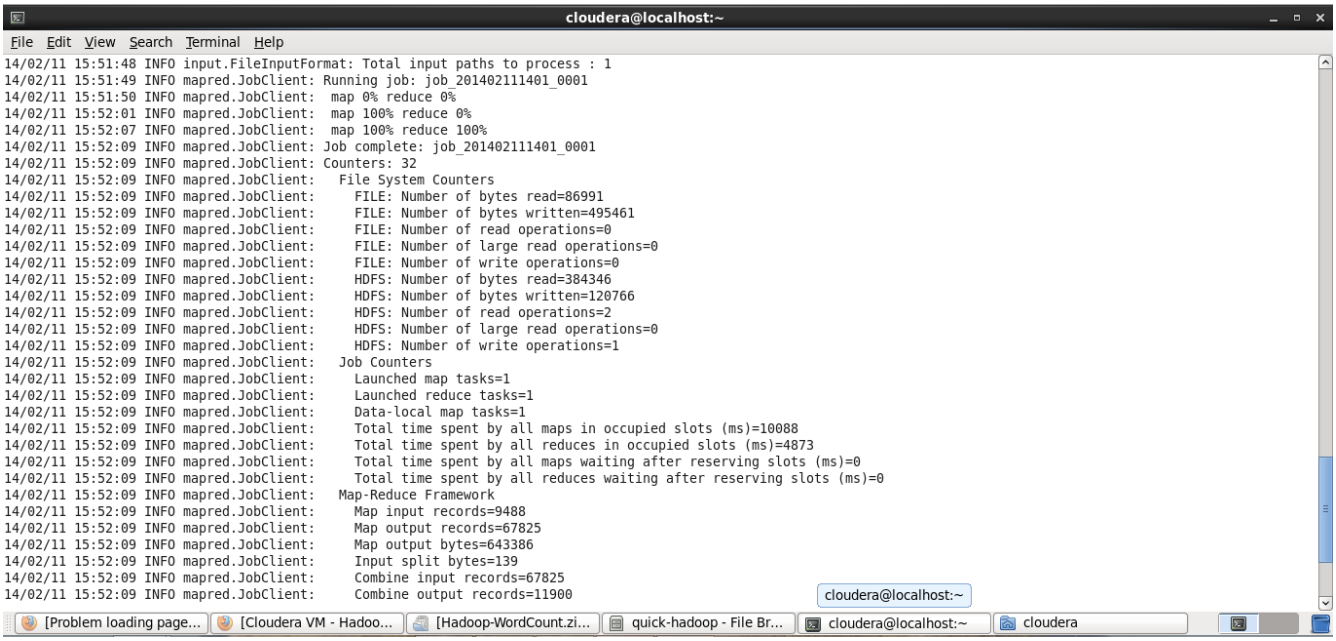
hadoop fs -put input input



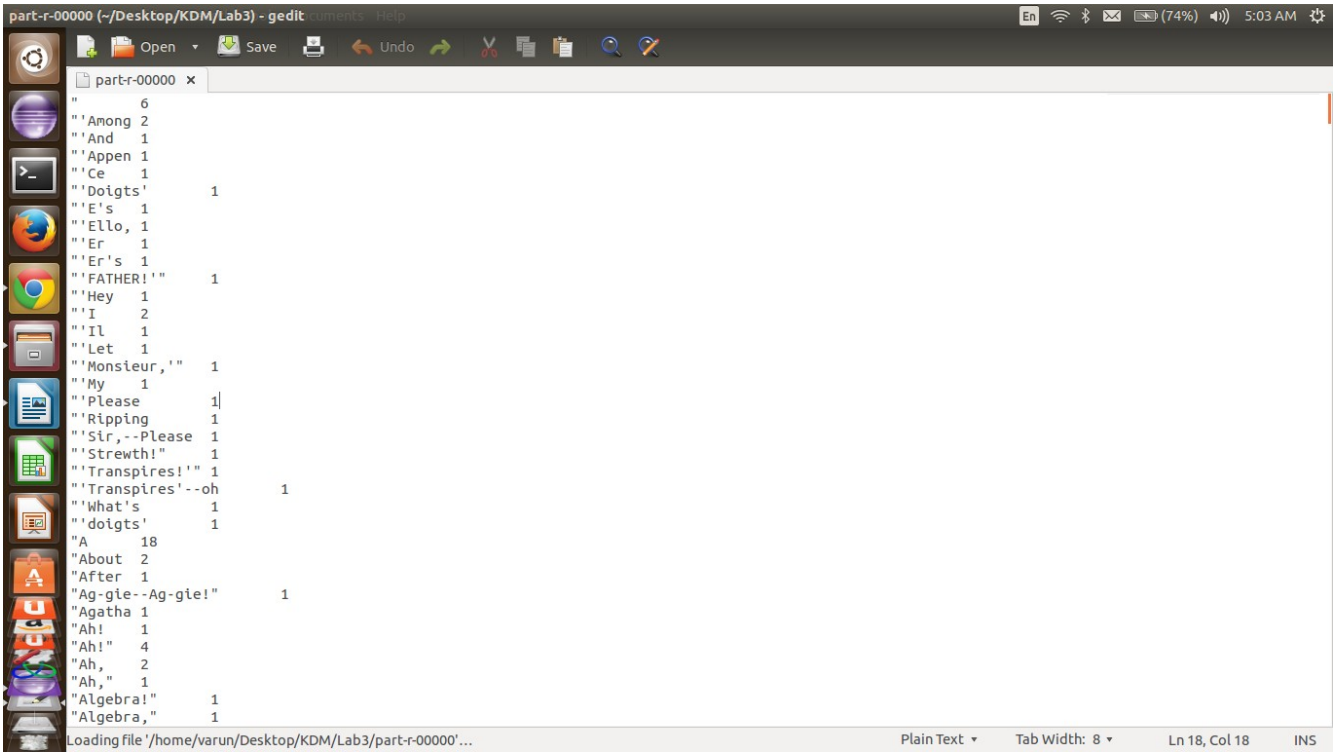
After pushing the content to hadoop HDFS we can run wordcount example on that input data.

This can be done using command

hadoop jar wordcount.jar WordCount input output



After completion of map reduce function we will get output in output folder we need to get that output to our virtual machine filesystem and can check the output.



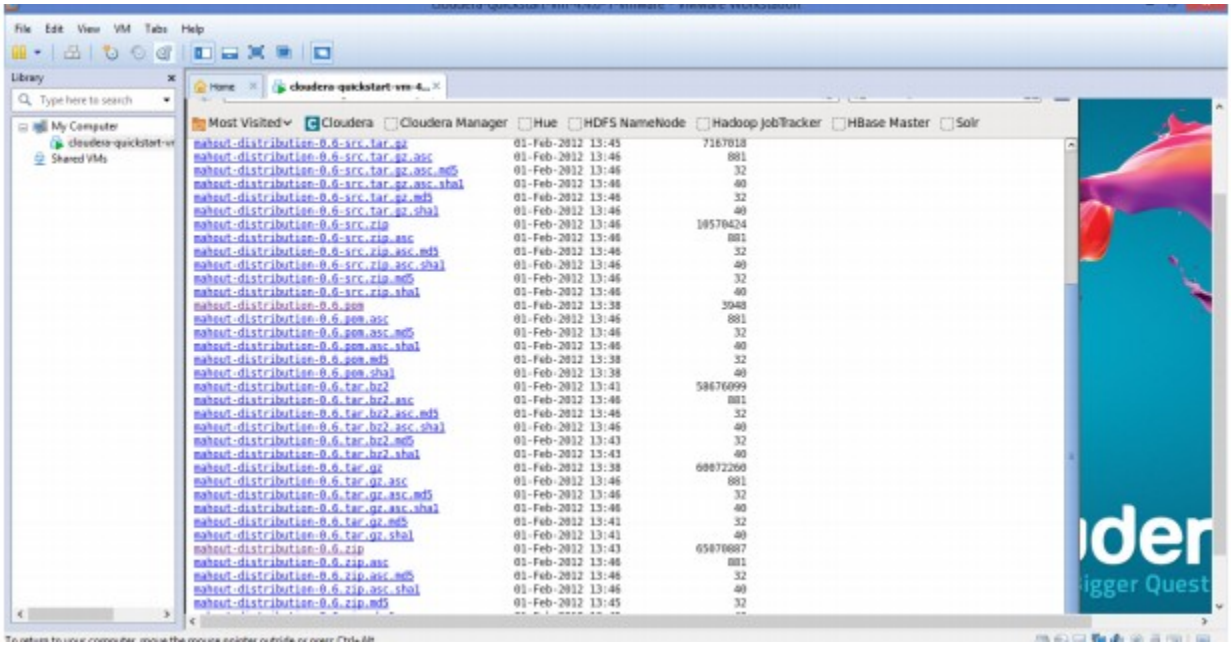
Installing Maven ,Mahout and Running Mahout Example:

we can download Maven from command line by typing
wget http://www.apache.org/dist//maven/binaries/apache-maven-3.0.5-bin.tar.gz

This will download Maven to our virtual machine.

Now we need to extract and export HOME and PATH of Maven folder.

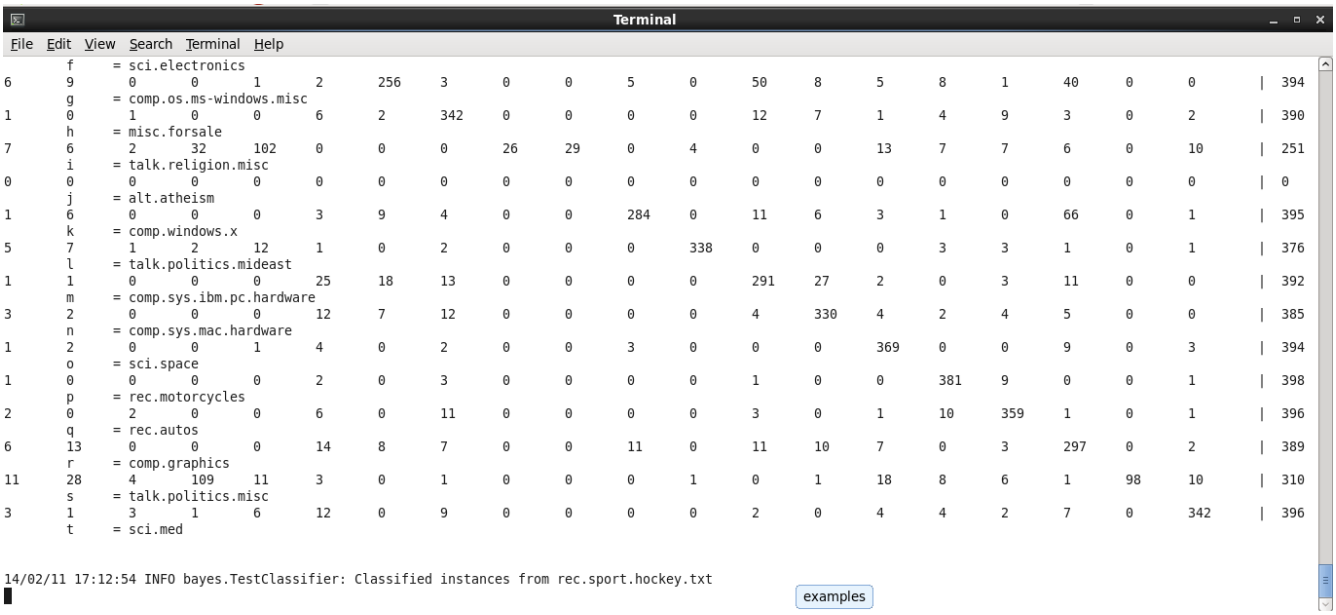
Now Download Mahout form the link mentioned in the tutorial document



install mahout using following command
mvn -f mahout-distribution-0.6.pom-DskipTestsinstall.

```
File  Edit  View  Search  Terminal  Help
drwxrwxr-x 3 cloudera cloudera     4096 Feb 11 16:40 integration
drwxrwxr-x 2 cloudera cloudera     4096 Feb 11 16:40 lib
-rw-r--r-- 1 cloudera cloudera    39588 Feb  1  2012 LICENSE.txt
-rw-r--r-- 1 cloudera cloudera  1662876 Feb  1  2012 mahout-core-0.6.jar
-rw-r--r-- 1 cloudera cloudera 11190212 Feb  1  2012 mahout-core-0.6-job.jar
-rwxrw-rw- 1 cloudera cloudera     3948 Feb 11 16:34 mahout-distribution-0.6.pom
-rw-r--r-- 1 cloudera cloudera   379461 Feb  1  2012 mahout-examples-0.6.jar
-rw-r--r-- 1 cloudera cloudera 23593299 Feb  1  2012 mahout-examples-0.6-job.jar
-rw-r--r-- 1 cloudera cloudera   284781 Feb  1  2012 mahout-integration-0.6.jar
-rw-r--r-- 1 cloudera cloudera   288914 Feb  1  2012 mahout-math-0.6.jar
drwxrwxr-x 3 cloudera cloudera     4096 Feb 11 16:40 math
-rw-r--r-- 1 cloudera cloudera     1888 Feb  1  2012 NOTICE.txt
-rw-r--r-- 1 cloudera cloudera     1200 Feb  1  2012 README.txt
[cloudera@localhost mahout-distribution-0.6]$ mvn -f mahout-distribution-0.6.pom -DskipTests install
[INFO] Scanning for projects...
Downloading: http://repo.maven.apache.org/maven2/org/apache/mahout/mahout/0.6/mahout-0.6.pom
Downloaded: http://repo.maven.apache.org/maven2/org/apache/mahout/mahout/0.6/mahout-0.6.pom (32 KB at 48.6 KB/sec)
Downloading: http://repo.maven.apache.org/maven2/org/apache/apache/9/apache-9.pom
Downloaded: http://repo.maven.apache.org/maven2/org/apache/apache/9/apache-9.pom (15 KB at 45.6 KB/sec)
[INFO]
[INFO] ------------------------------------------------------------------------
[INFO] Building Mahout Release Package 0.6
[INFO] ------------------------------------------------------------------------
Downloading: http://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-assembly-plugin/2.2/maven-assembly-plugin-2.2.pom
Downloaded: http://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-assembly-plugin/2.2/maven-assembly-plugin-2.2.pom (17 KB at 66.8 KB/sec)
Downloading: http://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-plugins/18/maven-plugins-18.pom
Downloaded: http://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-plugins/18/maven-plugins-18.pom (13 KB at 73.8 KB/sec)
Downloading: http://repo.maven.apache.org/maven2/org/apache/maven/maven-parent/16/maven-parent-16.pom
Downloaded: http://repo.maven.apache.org/maven2/org/apache/maven/maven-parent/16/maven-parent-16.pom (23 KB at 70.6 KB/sec)
Downloading: http://repo.maven.apache.org/maven2/org/apache/apache/7/apache-7.pom
Downloaded: http://repo.maven.apache.org/maven2/org/apache/apache/7/apache-7.pom (15 KB at 79.6 KB/sec)
Downloading: http://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-assembly-plugin/2.2/maven-assembly-plugin-2.2.jar
Downloaded: http://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-assembly-plugin/2.2/maven mahout-distribution-0.6 jar (214 KB at 202.2 KB/sec)
```

After completing installation go to examples and run the navi-base classification



Android MashUp Application:

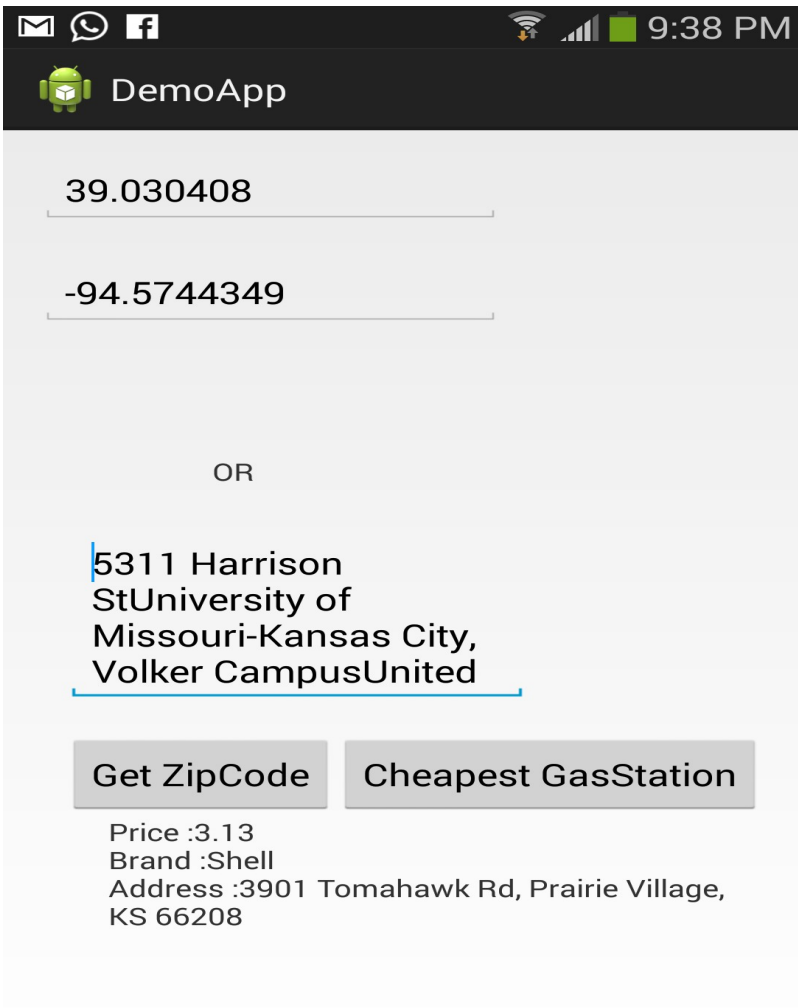
Developed an Android Application to get zip code from either latitude and longitude using google location API.
For this I have used google API
`http://maps.google.com/maps/api/geocode/json?address=" + address + "&sensor=false`

here in this location we will pass address and the response will be in Json String.

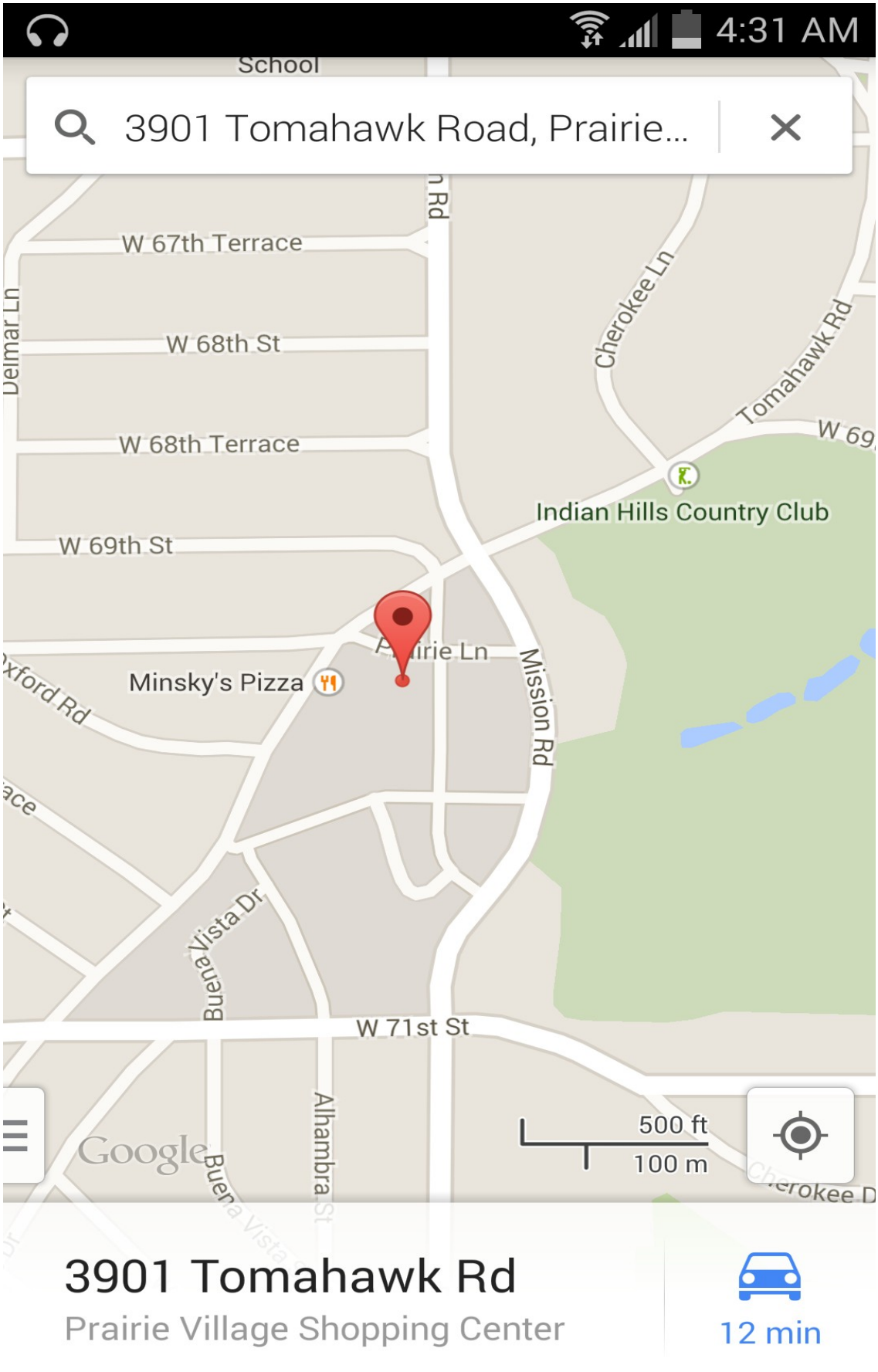
Parse Json String like this to get lat long positions.

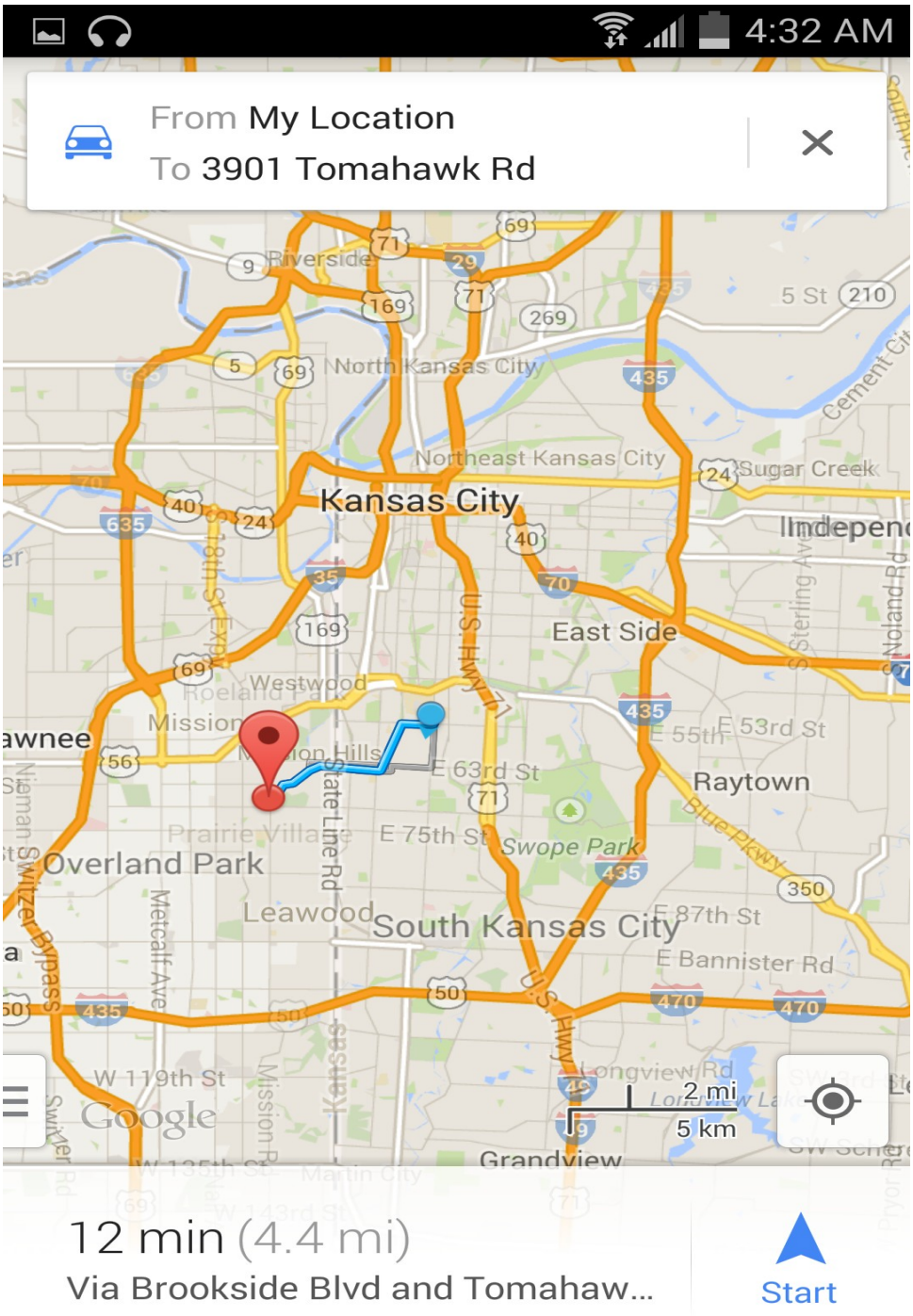
```
double longitute = ((JSONArray)jsonObject.get("results")).getJSONObject(0)
                        .getJSONObject("geometry").getJSONObject("location")
                        .getDouble("lng");

                    double latitude =
((JSONArray)jsonObject.get("results")).getJSONObject(0)
                        .getJSONObject("geometry").getJSONObject("location")
                        .getDouble("lat");
```

and to get cheapest GasStationn details for that given location for this service I have used http://www.mshd.net/api/gasprices/+ zipcode API



After clicking on the GasStation we will show directions to that GasStation from current location. Using Google maps.




4:32 AM
From My Location
To 3901 Tomahawk Rd
Riverside
North Kansas City
Northeast Kansas City
Kansas City
Sugar Creek
East Side
Westwood
Mission
Roeland Park
Prairie Village
Overland Park
Leawood
South Kansas City
Raytown
Swope Park
E 63rd St
E 75th St
E 87th St
E 53rd St
E Bannister Rd
Longview Rd
W 119th St
Grandview
Martin City
State Line Rd
Metcalf Ave
Switzer Bypass
U.S. Hwy 71
Blue Pkwy
Sterling Ave
Noland Rd
Cement City
2 mi
5 km
Google
12 min (4.4 mi)
Via Brookside Blvd and Tomahaw...
Start