###### SENTIMENT ANALYSIS

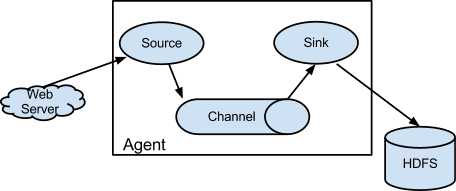
Twitter trend analysis for getting a movie review using HADOOP as the BIG DATA analysis technology.  
  
We will be performing 2 tasks:  
  
1. Download the tweets(data set) from twitter relating to a particular event using Apache Flume.  
  
2. Extract the opinion based on those tweets(opinion mining) using HDFS.

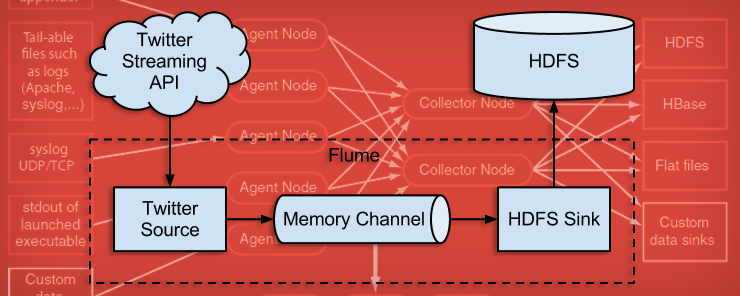
**TASK –I**

**Download the tweets(data set) from twitter relating to a particular event using Apache Flume.**

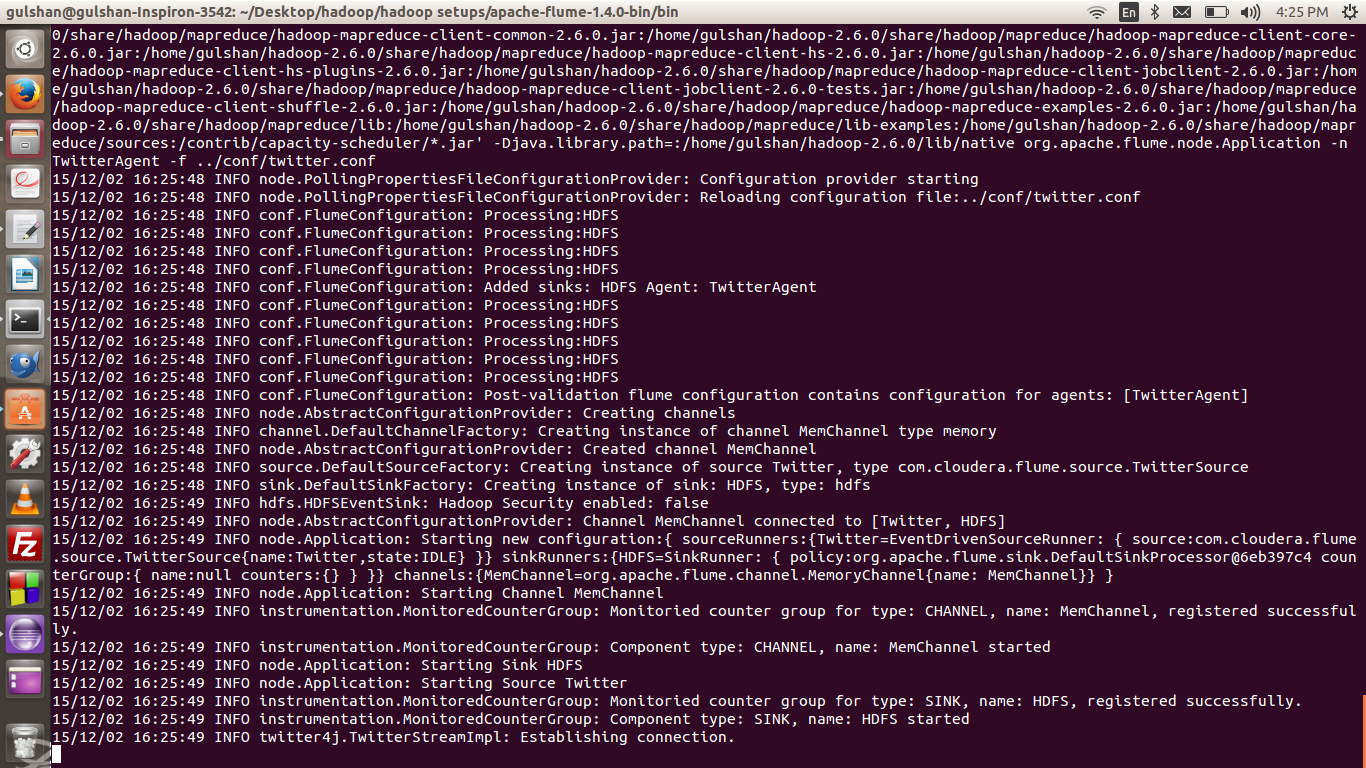
Apache Flume

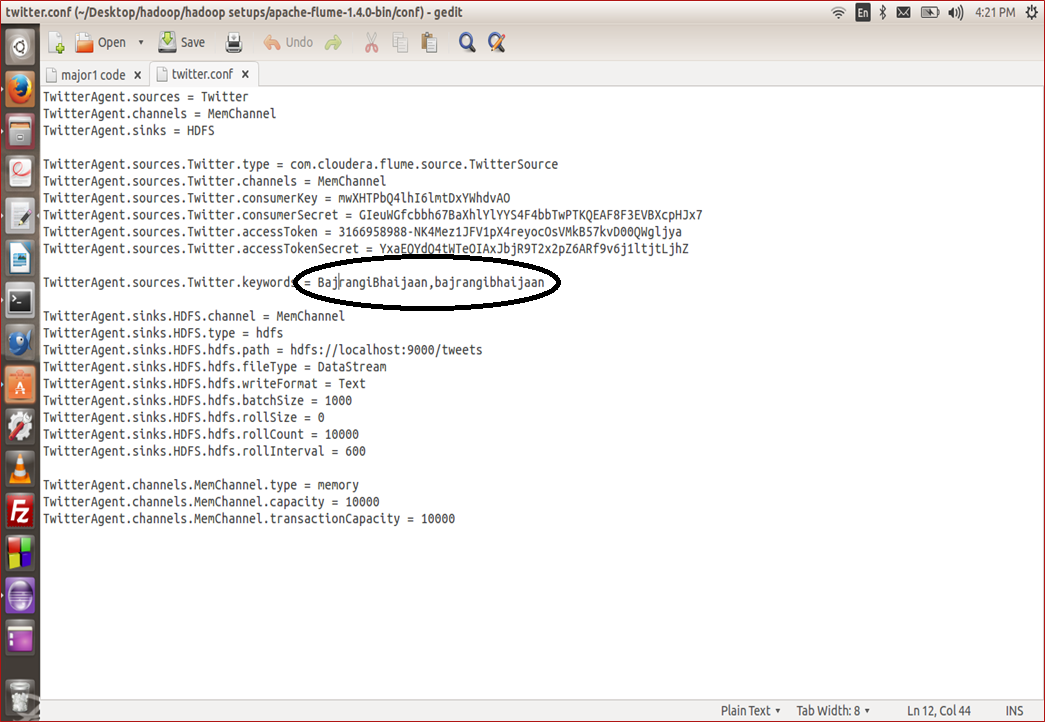
* Flume is a distributed, reliable, and available service for efficiently collecting, aggregating, and moving large amounts of log data.
* It has a simple and flexible architecture based on streaming data flows.
* It uses a simple extensible data model that allows for online analytic application.

****

****

**DOWNLOADING DATA USING APACHE FLUME**

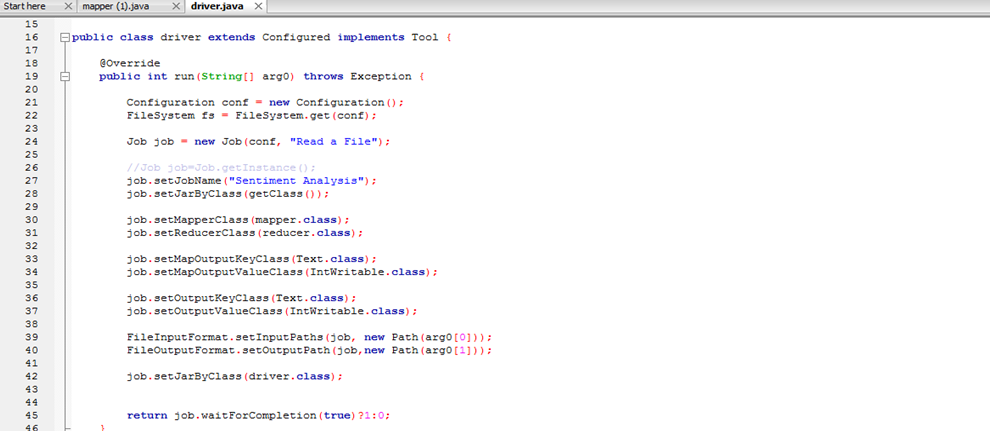
****

****

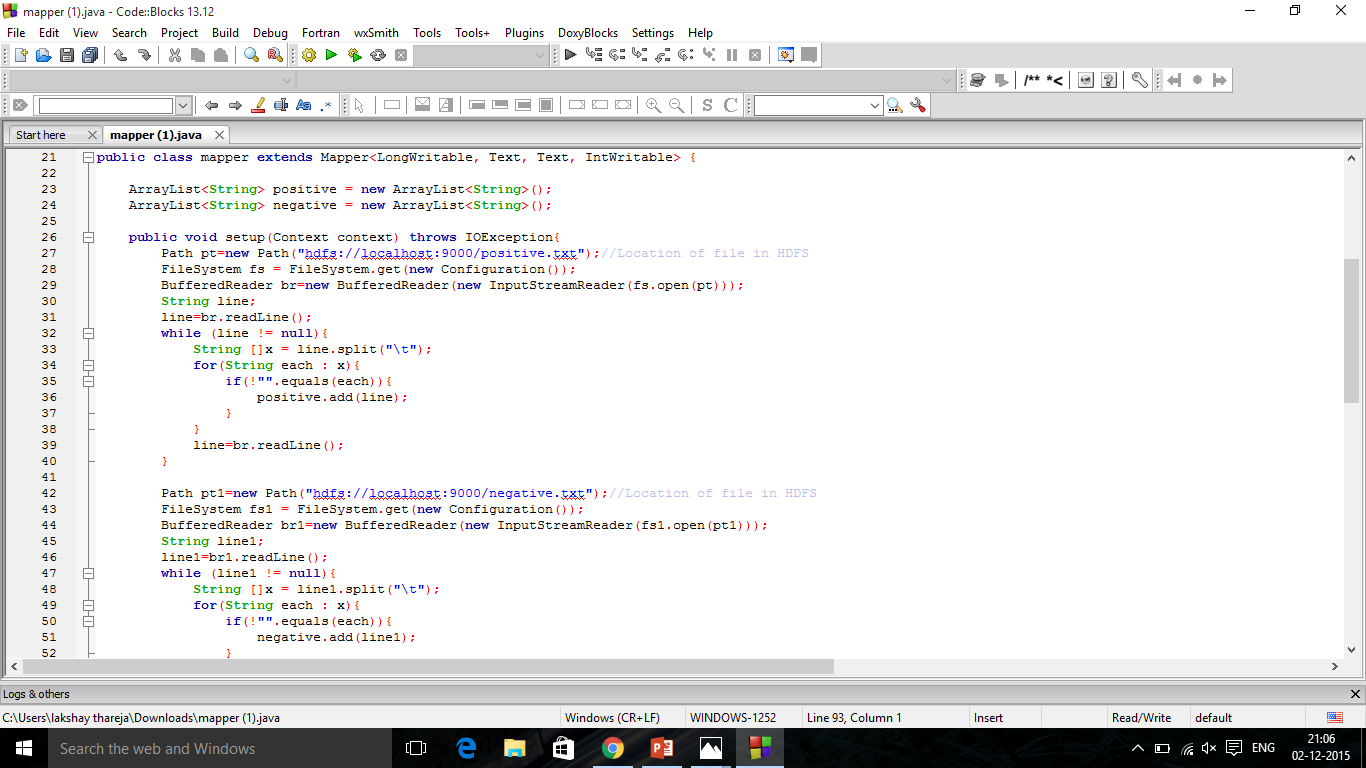
**TASK –II**

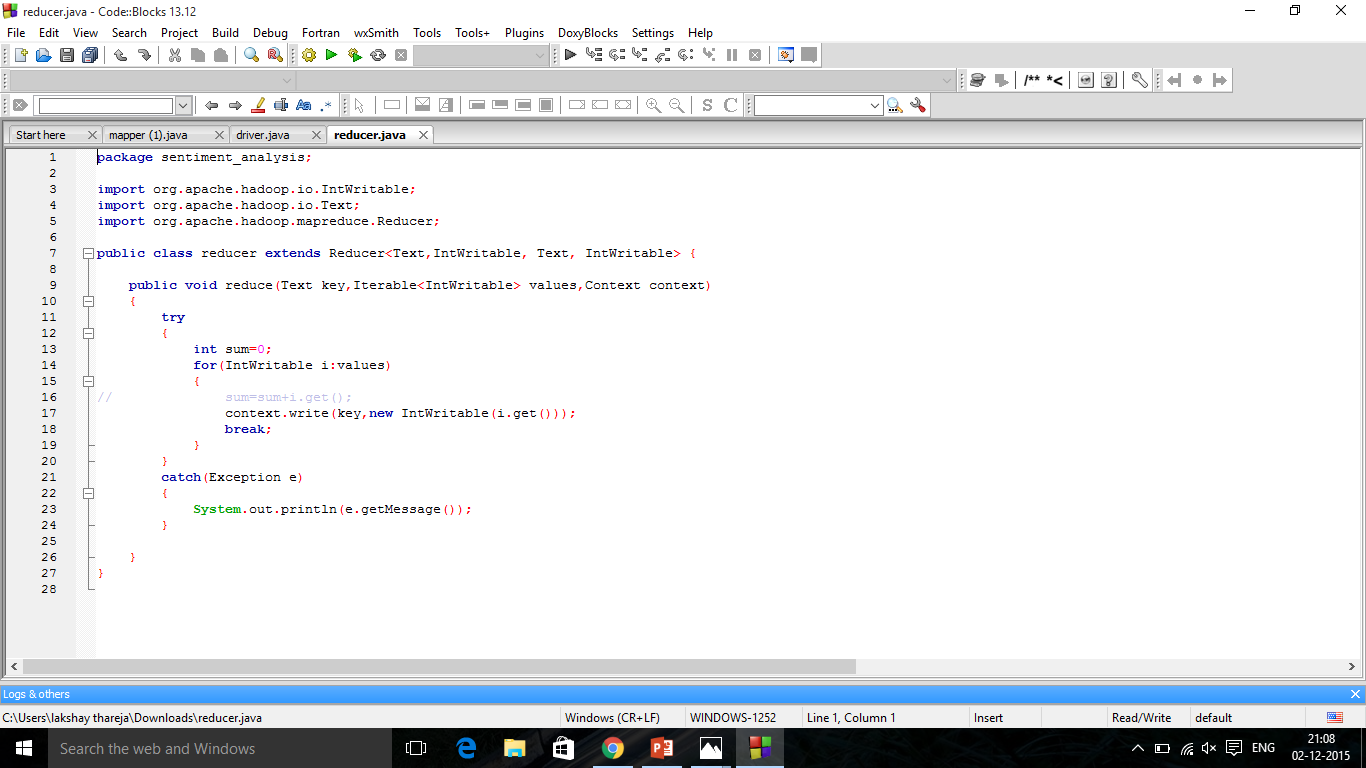
Extract the opinion based on those tweets(opinion mining) using HDFS.

DRIVER CODE SCREENSHOT

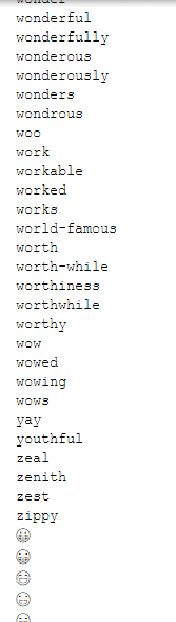


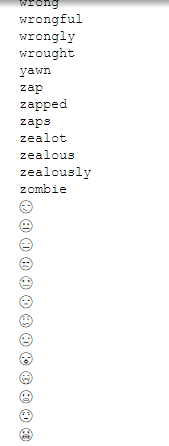
MAPPER CODE SCREENSHOT



REDUCER CODE SCREENSHOT

KNOWLEDGE BASE OF POSITIVE AND NEGATIVE WORDS





RESULT

**BENEFITS AND CONCLUSION**

* Scalable to millions and billions of data records using the power of hadoop.
* Parallel processing improves efficiency in distributed environment.
* Other uses of sentiment analysis:
  1. Opinon polls for elections.
  2. Product review by companies and businesses.
  3. Understanding customer needs and creating business solutions using the analysis result.

REFRENCES:

1. APACHE HADOOP WEBSITE - [https://**hadoop**.**apache**.org/](https://hadoop.apache.org/)
2. Hadoop: The Definitive Guide, 4th Edition - O'Reilly Media.
3. www.ubuntu.com
4. Twitter.com for data set – tweets.
5. Apache flume setup guide - https://flume.apache.org