Chapter 7

RESULTS

We have implemented this system based on java and Hadoop platform for Map Reduce framework

Dataset

- The dataset for the system is the election related downloaded tweets.
- For Sentiment analysis positive and negative thesaurus will be provided.

Results

As discussed above the system will generate volume analysis, trend analysis and sentiment analysis. The tweets downloaded from twitter database are processed with help of porter stemmer algorithm and users define functions. These filtered tweets will be used as input for various analysis modules generating the trend, volume or sentiment analysis

Centralized system: Experiment performed on single machine with core i3 processor*2.0* GHz x 2 and 4GB of main memory. The operating system is Ubuntu 16.04 and all algorithms are implemented in JDK 1.8.

The data was download from twitter and the further analysis was conducted on it as shown in results

```
SELECT text, word FROM twitter.tweets
LATERAL VIEW explode(split(text,' ')) text_ex as word;
```

The explode() is a Hive built-in User Defined Table-Generating Function (UDTF) that breaks down a array into its elements. In this case the tweet gets broken into words. The LATERAL VIEW joins the output of explode() to the input row (tweet) creating a result set that contains n rows (words) for each tweet.

I collected data for Narender Modi , BJP and election using tweets in tweets table. I want to access 12 most common hashtags on the data. For this I fired the query:

```
SELECT LOWER(hashtags.text),
COUNT(*) AS total_count
FROM tweets
LATERAL VIEW EXPLODE(entities.hashtags) t1
AS hashtags
GROUP BY LOWER(hashtags.text)
ORDER BY total_count
DESC LIMIT 14;
```

Results in:

#SoniaGandhi 211

#Indian 142

#RahulGandhiInMumbai 139

#BJP 112

#Modi 106

#RahulGandhi 73

#Congress 60

#Kejriwal 40

#NarendraModi 38

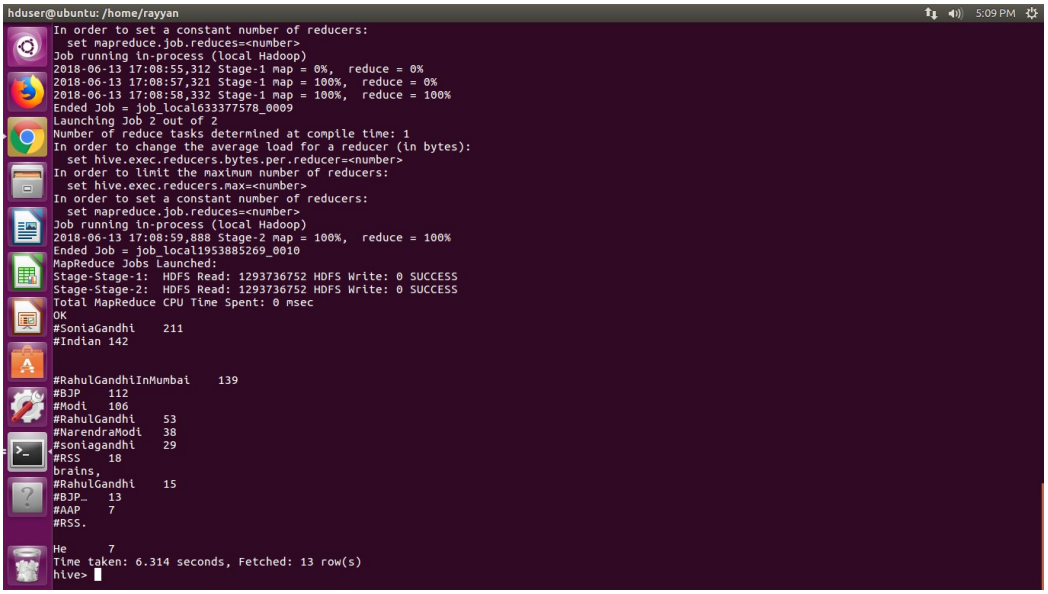
#soniagandhi 29

#RSS 18

#Congress 17

#AAP 17

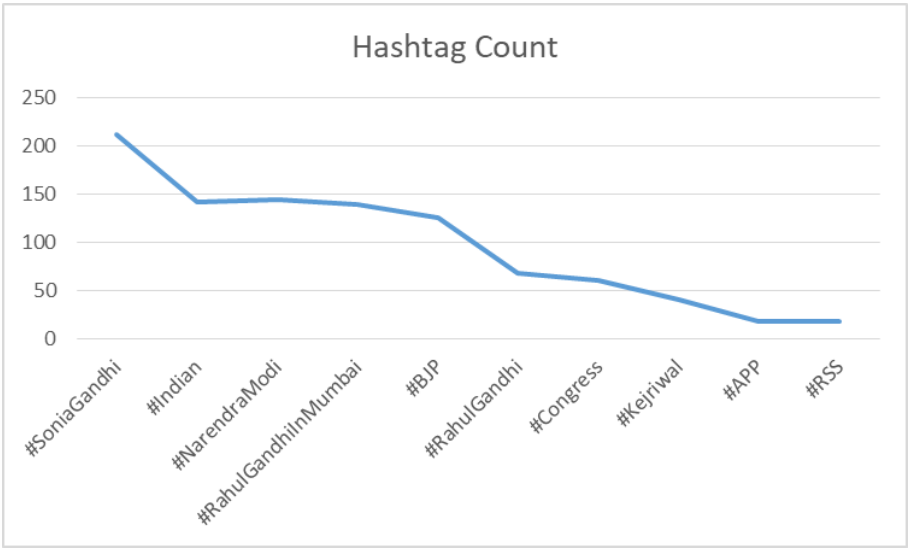
#RSS 17



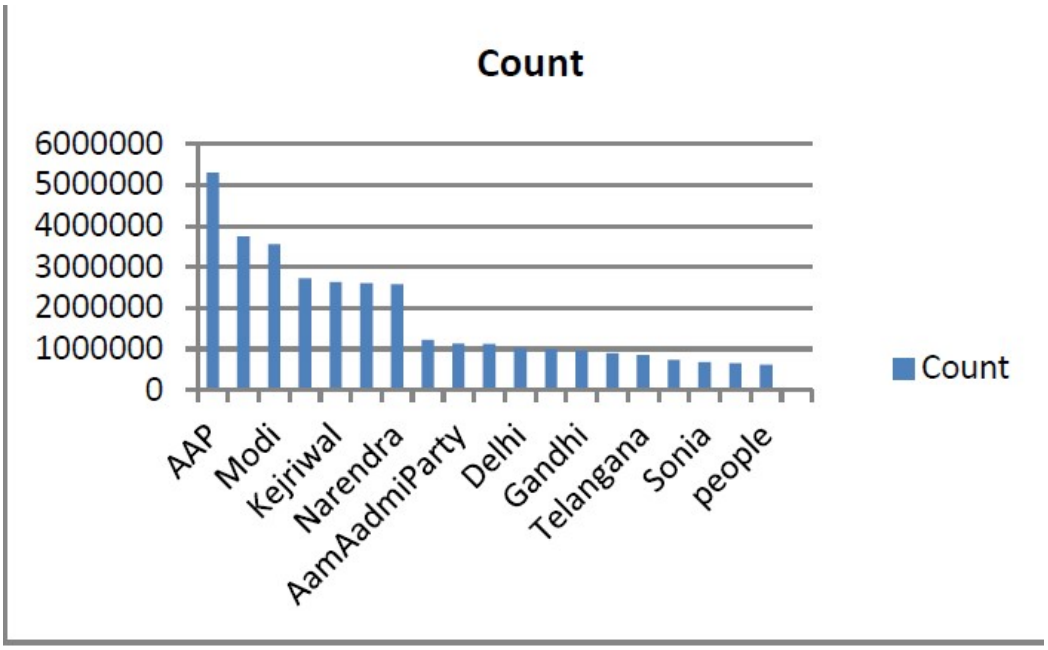
7.1 Hash Tag Wise Tweet

Hashtag's	Count	Hashtag's	Count
#SoniaGandhi	211	#RahulGandhi	68
#Indian	142	#Congress	60
#NarendraModi	144	#Kejriwal	40
#RahulGandhiInMumbai	139	#APP	17
#BJP	125	#RSS	17

7.2 Hashtag Wise Tweet Count



7.3 Trending Topic of Elections



7.4 Sentiment Analysis of Tweet Dataset