

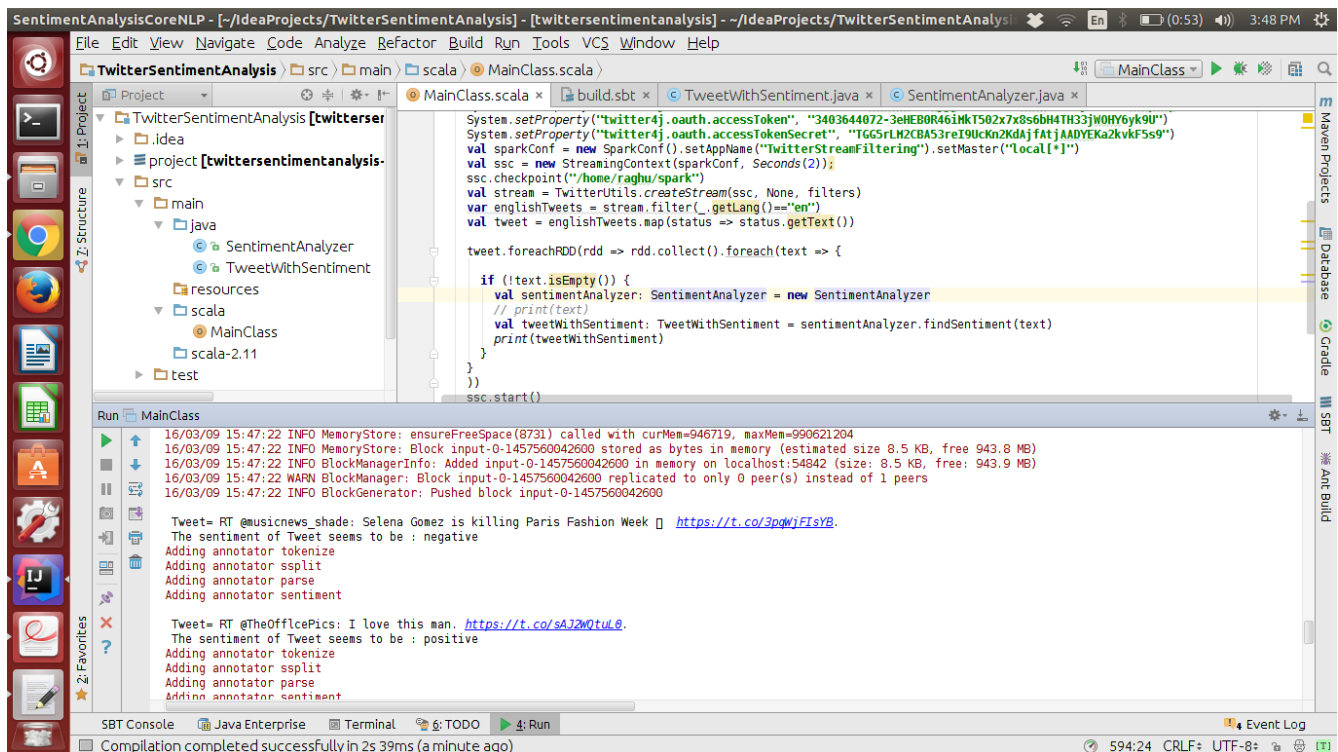
## Lab 8 Assignment

Submitted by  
**Ravi Kiran Yadavalli -13**  
**Ragunandan Rao Malangully -31**

### Task 1:

The task here was to perform sentiment analysis on Twitter streaming data. I performed the following steps to complete the task.

1. Create Twitter stream using Twitter4j package.
2. Filter the Tweets where the language is English.
3. Extract only the Tweet text from the Tweet.
4. For every Tweet text perform sentiment analysis using the Stanford NLP class.
5. Output the Tweet Text along with the sentiment of the Tweet text.



```
System.setProperty("twitter4j.oauth.accessToken", "3403644072-3eHEB0R46iMkT502x7x8s6bH4TH33jW0HY6yk9U")
System.setProperty("twitter4j.oauth.accessTokenSecret", "1G65rLH2CBAS3reI9Uckn2KdJfATjAADYEKa2kvkF5s9")
val sparkConf = new SparkConf().setAppName("TwitterStreamFiltering").setMaster("local[*]")
val ssc = new StreamingContext(sparkConf, Seconds(2))
ssc.checkpoint("/home/raghu/spark")
val stream = TwitterUtils.createStream(ssc, None, filters)
val englishTweets = stream.filter(_.getLang() == "en")
val tweet = englishTweets.map(status => status.getText())

tweet.foreachRDD(rdd => rdd.collect().foreach(text => {
  if (!text.isEmpty()) {
    val sentimentAnalyzer: SentimentAnalyzer = new SentimentAnalyzer
    // print(text)
    val tweetWithSentiment: TweetWithSentiment = sentimentAnalyzer.findSentiment(text)
    print(tweetWithSentiment)
  }
}))
ssc.start()
```

Run MainClass

```
16/03/09 15:47:22 INFO MemoryStore: ensureFreeSpace(8731) called with curMem=946719, maxMem=990621204
16/03/09 15:47:22 INFO MemoryStore: Block input-0-1457560042600 stored as bytes in memory (estimated size 8.5 KB, free 943.8 MB)
16/03/09 15:47:22 INFO BlockManagerInfo: Added input-0-1457560042600 in memory on localhost:54842 (size: 8.5 KB, free: 943.9 MB)
16/03/09 15:47:22 WARN BlockManager: Block input-0-1457560042600 replicated to only 0 peer(s) instead of 1 peers
16/03/09 15:47:22 INFO BlockGenerator: Pushed block input-0-1457560042600

Tweet= RT @musicnews_shade: Selena Gomez is killing Paris Fashion Week | https://t.co/3pQWjFISyB.
The sentiment of Tweet seems to be : negative
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment

Tweet= RT @TheOfficePics: I love this man. https://t.co/SA72W0tUu0.
The sentiment of Tweet seems to be : positive
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
```

SBT Console Java Enterprise Terminal 6: TODO 4: Run Event Log

Compilation completed successfully in 2s 39ms (a minute ago) 594/24 CRLF: UTF-8