

CS5542

Big Data Apps and Analytics

LAB ASSIGNMENT #7

-Nihar Dudam (6)
Sricharan Pamuru(20)

1. Sentimental Analysis on Twitter Streaming:

The screenshot displays the IntelliJ IDEA IDE interface for a Scala project named 'SentimentAnalysisCoreNLP'. The project structure on the left includes 'src/main/scala' with files 'MainClass.scala', 'SentimentAnalyzer.java', and 'TweetWithSentiment.java'. The main editor shows the 'MainClass.scala' file with the following code:

```
object MainClass {  
  def main(args: Array[String]) {  
    val filters = args  
  
    System.setProperty("twitter4j.oauth.consumerKey", "1oVEx4fjDXAgQoNLExOjw91C")  
    System.setProperty("twitter4j.oauth.consumerSecret", "gj07311Q8SoZMFwJfeCqW0KUCzrK1qKyEMLcr3C4aTWQ8Qa2Xn")  
    System.setProperty("twitter4j.oauth.accessToken", "702647205807534080-UZjUbaNvQaYhTW13MkD3oyBHAamF3jd")  
    System.setProperty("twitter4j.oauth.accessTokenSecret", "E129gJVngXwIUJmPQb0MokVUeN5UBGcxdBNCrUrUveSbq")  
    val sparkConf = new SparkConf().setAppName("STweetsApp").setMaster("local[*]")  
    //Create a Streaming Context with 2 second window  
    val ssc = new StreamingContext(sparkConf, Seconds(5))  
    val stream = TwitterUtils.createStream(ssc, None, filters)  
    val english_tweets = stream.filter(_._2.getLang() == "en")  
    val tweets = english_tweets.map(status => status.getText())  
  
    tweets.foreachRDD(rdd => rdd.collect().foreach(text => {  
      println(text)  
      val sentimentAnalyzer: SentimentAnalyzer = new SentimentAnalyzer  
      val tweetWithSentiment: TweetWithSentiment = sentimentAnalyzer.findSentiment(text)  
      System.out.println(tweetWithSentiment)  
    })  
  }  
}
```

The bottom panel shows the 'Run' output, indicating the successful execution of the application. The output includes the following lines:

```
Adding annotator ssplit  
Adding annotator parse  
Adding annotator sentiment  
TweetWithSentiment [line=So fuck Meghan right?, cssClass=sentiment : negative]  
RT @ItsBrandoBro: @jvanoortl what's a riddle if you solve it in 2 seconds?  
null  
RT @ohhmmami: certain shit just isn't worth fighting for anymore. friendships, relationships, like I'm not gonna fight for something that's ...  
Adding annotator tokenize  
Adding annotator ssplit  
Adding annotator parse  
Adding annotator sentiment  
16/03/09 19:31:49 INFO BlockManagerInfo: Removed broadcast_13_piece0 on localhost:53534 in memory (size: 1117.0 B, free: 1122.4 MB)
```

In this part of Assignment, we are streaming tweets from Twitter and running sentimental analysis on the tweets.

2. TV series recommendation to the user

```
RMSE (validation) = 2.0 for the model trained with rank = 12, lambda = 10.0, and numIter = 20.
The best model was trained with rank = 8 and lambda = 10.0, and numIter = 10, and its RMSE on the test set is 2.806243040080456.
The best model improves the baseline by -151.15%.
Movies recommended for you:
1: Modern Family
2: Oz
3: Agents of S.H.I.E.L.D
4: The Sopranos
5: Louie
6: The X-Files
7: 24:action|drama|adventure
8: The Vampire Diaries
9: Sons of Anarchy
10: Flashforward
11: How I Met Your Mother
12: Homeland
13: Rescue Me
14: Grey's Anatomy:drama
15: The Big Bang Theory
16: The Simpsons
17: Supernatural
18: Two and Half Men
16/03/09 23:51:24 INFO RemoteActorRefProvider$RemotingTerminator: Shutting down remote daemon.
16/03/09 23:51:24 INFO RemoteActorRefProvider$RemotingTerminator: Remote daemon shut down; proceeding with flushing remote transports.

Process finished with exit code 0
```

In this part of assignment, data regarding TV series is collected using streaming Twitter data and manually edited data to recommend the popular TV series to the user.

3. The above TV Series is displayed in mobile phone.

