

## Lab 7 Assignment

Submitted by

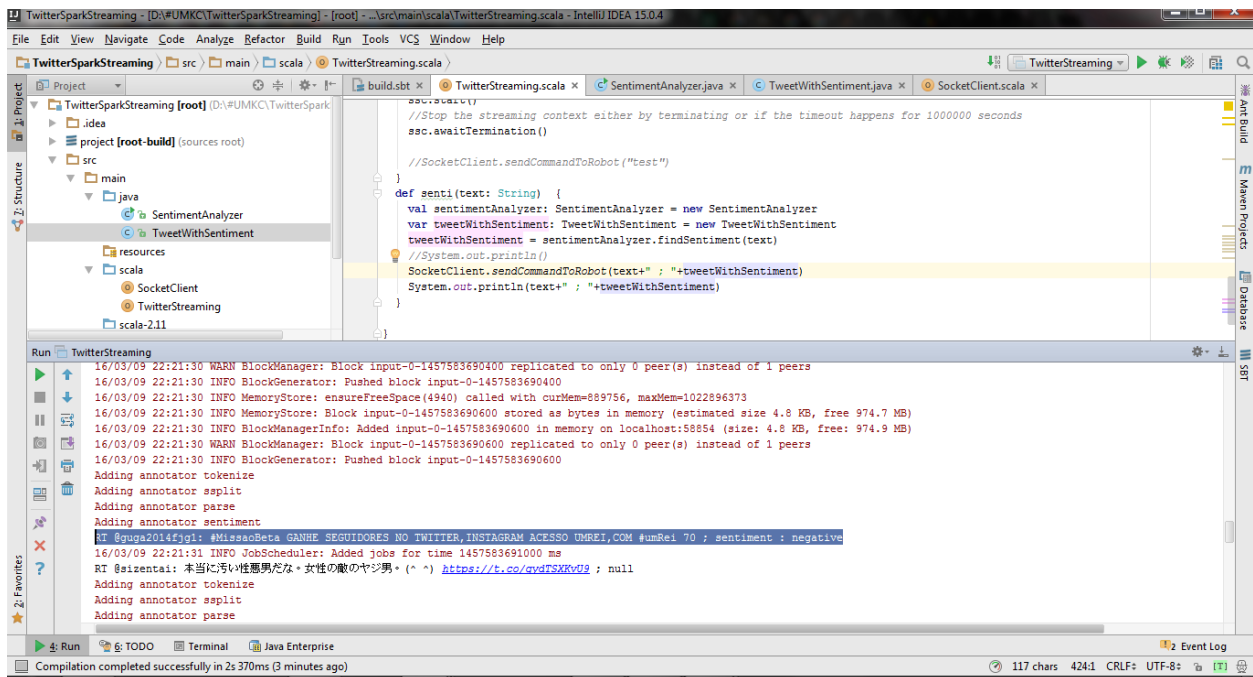
Venkata Vamsi Krishna Bhuvanam -3

Vikas Kondapalli -11

### 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.



```
TwitterSparkStreaming - [D:\UMKC\TwitterSparkStreaming] - [root] - ...src\main\scala\TwitterStreaming.scala - IntelliJ IDEA 15.0.4
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help

TwitterSparkStreaming [root] (D:\UMKC\TwitterSparkStreaming)
  project [root-build] (sources root)
  src
    main
      java
        SentimentAnalyzer
        TweetWithSentiment
      resources
      scala
        SocketClient
        TwitterStreaming
      scala-2.11

TwitterStreaming.scala x SentimentAnalyzer.java x TweetWithSentiment.java x SocketClient.scala x

//Stop the streaming context either by terminating or if the timeout happens for 1000000 seconds
ssc.awaitTermination()

//SocketClient.sendCommandToRobot("test")
}

def senti(text: String) {
  val sentimentAnalyzer: SentimentAnalyzer = new SentimentAnalyzer
  var tweetWithSentiment: TweetWithSentiment = new TweetWithSentiment
  tweetWithSentiment = sentimentAnalyzer.findSentiment(text)
  //System.out.println()
  SocketClient.sendCommandToRobot(text+" ; "+tweetWithSentiment)
  System.out.println(text+" ; "+tweetWithSentiment)
}

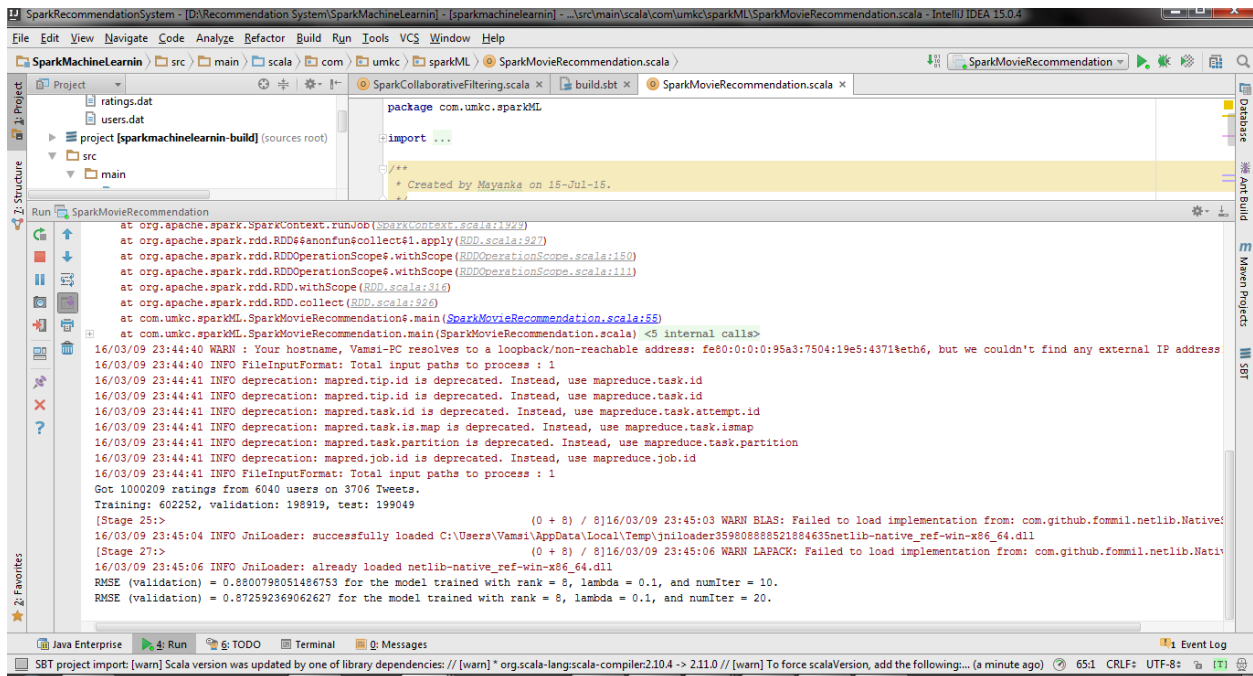
Run TwitterStreaming
16/03/09 22:21:30 WARN BlockManager: Block input-0-1457583690400 replicated to only 0 peer(s) instead of 1 peers
16/03/09 22:21:30 INFO BlockGenerator: Pushed block input-0-1457583690400
16/03/09 22:21:30 INFO MemoryStore: ensureFreeSpace(4940) called with curMem=889756, maxMem=1022896373
16/03/09 22:21:30 INFO MemoryStore: Block input-0-1457583690600 stored as bytes in memory (estimated size 4.8 KB, free 974.7 MB)
16/03/09 22:21:30 INFO BlockManagerInfo: Added input-0-1457583690600 in memory on localhost:88854 (size: 4.8 KB, free: 974.9 MB)
16/03/09 22:21:30 WARN BlockManager: Block input-0-1457583690600 replicated to only 0 peer(s) instead of 1 peers
16/03/09 22:21:30 INFO BlockGenerator: Pushed block input-0-1457583690600
Adding annotator tokenize
Adding annotator esplit
Adding annotator sentiment
16/03/09 22:21:31 INFO JobScheduler: Added jobs for time 1457583691000 ms
RT @isental: 本当に汚い性悪男だな。女性の敵のヤジ男。( ^ ^ ) https://t.co/qydtSxKvU9 ; null
Adding annotator tokenize
Adding annotator esplit
Adding annotator parse
Compilation completed successfully in 2s 370ms (3 minutes ago)
117 chars 424:1 CRLF UTF-8
```

## Lab 7 Assignment

### Task 2

This task was to suggest recommendations to the user based on his ratings for music genres. The following steps were performed to complete the task.

1. Create training data through Twitter stream filtering and classification of data.
2. Manually formatting the collected data into the desired format.
3. Performing collaborative filtering on the training data to train the model.
4. Get sample data from Twitter to serve as test data.
5. Perform recommendation for the test data



The screenshot shows an IDE window titled "SparkRecommendationSystem - [D:\RecommendationSystem\SparkMachineLearning] - [sparkmachinelearning] - ...src\main\scala\com\umkc\sparkML\SparkMovieRecommendation.scala - IntelliJ IDEA 15.0.4". The editor displays the file "SparkMovieRecommendation.scala" with the following code:

```
package com.umkc.sparkML

import ...

/**
 * Created by Mayanka on 15-Jul-15.
 */
```

The Run console shows the following output:

```
at org.apache.spark.SparkContext.runJob(SparkContext.scala:172)
at org.apache.spark.rdd.RDD.$anonfun$collect$1.apply(RDD.scala:927)
at org.apache.spark.rdd.RDDOperationScope$.withScope(RDDOperationScope.scala:150)
at org.apache.spark.rdd.RDDOperationScope$.withScope(RDDOperationScope.scala:111)
at org.apache.spark.rdd.RDD.withScope(RDD.scala:316)
at org.apache.spark.rdd.RDD.collect(RDD.scala:926)
at com.umkc.sparkML.SparkMovieRecommendation$.main(SparkMovieRecommendation.scala:55)
at com.umkc.sparkML.SparkMovieRecommendation.main(SparkMovieRecommendation.scala) <5 internal calls>
16/03/09 23:44:40 WARN : Your hostname, Vamsi-PC resolves to a loopback/non-reachable address: fe80:0:0:0:95a3:7504:19e5:4371%eth6, but we couldn't find any external IP address
16/03/09 23:44:40 INFO FileInputFormat: Total input paths to process : 1
16/03/09 23:44:41 INFO deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
16/03/09 23:44:41 INFO deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
16/03/09 23:44:41 INFO deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
16/03/09 23:44:41 INFO deprecation: mapred.task.is.map is deprecated. Instead, use mapreduce.task.ismap
16/03/09 23:44:41 INFO deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
16/03/09 23:44:41 INFO deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
16/03/09 23:44:41 INFO FileInputFormat: Total input paths to process : 1
Got 1000209 ratings from 6040 users on 3706 Tweets.
Training: 602252, validation: 198919, test: 199049
[Stage 25:> (0 + 8) / 8]16/03/09 23:45:03 WARN BLAS: Failed to load implementation from: com.github.fommil.netlib.Native
16/03/09 23:45:04 INFO JniLoader: successfully loaded C:\Users\Vamsi\AppData\Local\Temp\jni-loader35980888521884635netlib-native_ref-win-x86_64.dll
[Stage 27:> (0 + 8) / 8]16/03/09 23:45:06 WARN LAPACK: Failed to load implementation from: com.github.fommil.netlib.Native
16/03/09 23:45:06 INFO JniLoader: already loaded netlib-native_ref-win-x86_64.dll
RMSE (validation) = 0.8800798051486753 for the model trained with rank = 8, lambda = 0.1, and numIter = 10.
RMSE (validation) = 0.872592369062627 for the model trained with rank = 8, lambda = 0.1, and numIter = 20.
```

The bottom status bar shows "SBT project import: [warn] Scala version was updated by one of library dependencies: // [warn] \* org.scala-lang:scala-compiler:2.10.4 -> 2.11.0 // [warn] To force scalaVersion, add the following... (a minute ago) 65:1 CRLF: UTF-8: [icon]

## Lab 7 Assignment

```
SparkRecommendationSystem - [D:\Recommendation System\SparkMachineLearnin - [sparkmachinelearnin] - ...src\main\scala\com\umkc\sparkML\SparkMovieRecommendation.scala - IntelliJ IDEA 15.0.4
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
SparkMachineLearnin src main scala com umkc sparkML SparkMovieRecommendation.scala
Project [sparkmachinelearnin-build] (sources root)
  src
    main
      java
      resources
      stopwords.txt
      scala
        SparkMovieRecommendation.scala
          .take(50)
          var i = 1
          SocketClient.sendCommandToRobot("Tweets recommended for you:\n")
          println("Tweets recommended for you:")
          recommendations.foreach { r =>
            SocketClient.sendCommandToRobot("%2d".format(i) + " : " + movies(r.product))
            println("%2d".format(i) + " : " + movies(r.product))
            i += 1
          }
Run SparkMovieRecommendation
[Stage 1378]:=====
RMSE (validation) = 3.7558695311242833 for the model trained with rank = 12, lambda = 10.0, and numIter = 20.
[Stage 1685]:
The best model improves the baseline by 22.01%.
Tweets recommended for you:
1: Danny Jones who used to wear his mum's dressing gown and a Tommy Cooper-style fez to bed thanks judge as he is jailed for seven years
2: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide https://www.reddit.com/r/unitedkingdom/comments/48nlix/how\_the\_daily\_mail\_bough
3: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide https://www.reddit.com/r/unitedkingdom/comments/48nlix/how\_the\_daily\_mail\_bough
4: ""Diabetic Collins""; worst gangland nickname ever.
5: from YouTube - American exceptionalism meets Australian straight talking
6: from YouTube - American exceptionalism meets Australian straight talking
7: We've all been there- ur halfway through an online article & accidentally start writing erotic fiction.....
8: "Caller @BBCRadio2: ""I'm not listening to the experts on the EU. Experts built the Titanic.""
9: #VandyeMataram #JaiHind #TransformingIndia #WakeUpIndia Amit Singh: isiddhantsharma proudhindu9 ??? India ?? ??? https://t.co/XhswvqQXh4
10: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide https://www.reddit.com/r/unitedkingdom/comments/48nlix/how\_the\_daily\_mail\_bough
11: Best article correction ever
12: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide https://www.reddit.com/r/unitedkingdom/comments/48nlix/how\_the\_daily\_mail\_bough
13: #VandyeMataram #JaiHind #TransformingIndia #WakeUpIndia Amit Singh: isiddhantsharma proudhindu9 ??? India ?? ??? https://t.co/XhswvqQXh4
14: from YouTube - American exceptionalism meets Australian straight talking
15: from YouTube - American exceptionalism meets Australian straight talking
16: ""Diabetic Collins""; worst gangland nickname ever.
17: #VandyeMataram #JaiHind #TransformingIndia #WakeUpIndia Amit Singh: isiddhantsharma proudhindu9 ??? India ?? ??? https://t.co/XhswvqQXh4
18: Display of @realDonaldTrump wines steaks & spring water for tonight's press conference in Jupiter.
Compilation completed successfully in 3s 598ms (4 minutes ago)
```

## Lab 7 Assignment

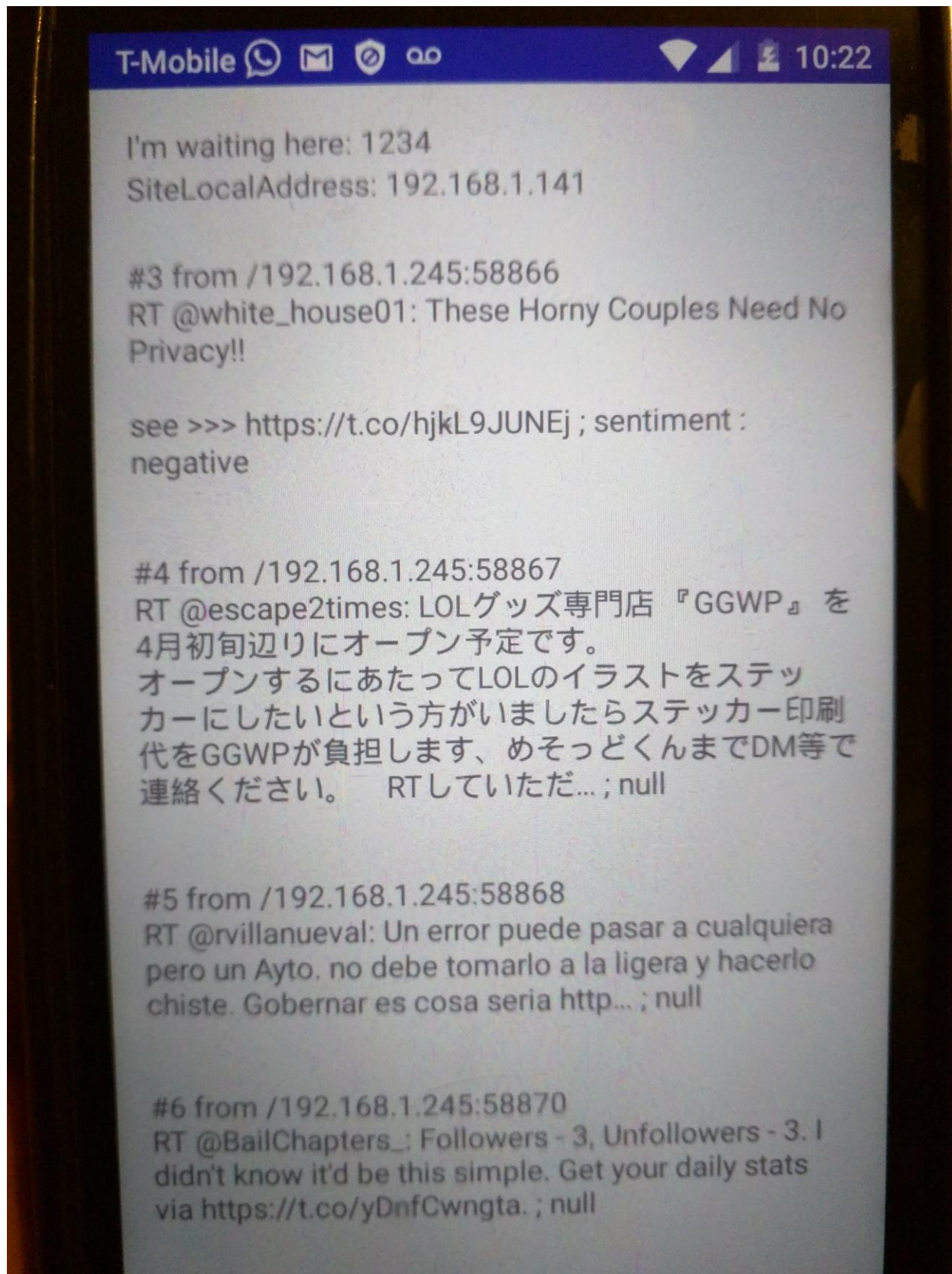
### Task 3:

The task here was to send the recommendations to phone through socket connection.

The following steps were performed to accomplish the task.

1. Create a socket server at Android phone side.
2. Analyze the tweet using sentiment analysis.
3. Send the Analyzed sentiment data and tweet data to socket client.
4. Display the received string on the UI and the Android client end.

## Lab 7 Assignment



## Lab 7 Assignment

### Task 4:

The task here was to send the recommendations to phone through socket connection.

The following steps were performed to accomplish the task.

5. Create a socket server at Android phone side.
6. Create a client socket for the recommendation process that provides recommendations.
7. Send the recommendations as string over the socket from the client to the server.
8. Display the received string on the UI and the Android client end.



## Lab 7 Assignment

T-Mobile



12:19

I'm waiting here: 1234

SiteLocalAddress: 192.168.1.141

#1 from /192.168.1.245:54043

Tweets recommended for you:

#2 from /192.168.1.245:54044


1: Danny Jones who used to wear his mum's dressing gown and a Tommy Cooper-style fez to bed thanks judge as he is jailed for seven years

#3 from /192.168.1.245:54045

2: Best article correction ever

#4 from /192.168.1.245:54046

3: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide <https://www.reddit.com/r/unitedkingdom/comments/49nlix/>

how\_the\_daily\_mail\_bought\_hatred\_to\_my\_home\_to wn/  @BestoftheMail

#5 from /192.168.1.245:54047

4: Display of @realDonaldTrump wines steaks & spring water for tonight's press conference in Jupiter.

## Lab 7 Assignment

T-Mobile

12:19

I'm waiting here: 1234

SiteLocalAddress: 192.168.1.141

#1 from /192.168.1.245:54043

Tweets recommended for you:

#2 from /192.168.1.245:54044

1: Danny Jones who used to wear his mum's dressing gown and a Tommy Cooper-style fez to bed thanks judge as he is jailed for seven years

#3 from /192.168.1.245:54045

2: Best article correction ever

#4 from /192.168.1.245:54046

3: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide <https://www.reddit.com/r/unitedkingdom/comments/49nlix/>

how\_the\_daily\_mail\_bought\_hatred\_to\_my\_home\_to wn/ @BestoftheMail

#5 from /192.168.1.245:54047

4: Display of @realDonaldTrump wines steaks & spring water for tonight's press conference in Jupiter.



## Lab 7 Assignment

T-Mobile



12:19

I'm waiting here: 1234

SiteLocalAddress: 192.168.1.141

ever. "

#7 from /192.168.1.245:54049

6: Cold shower required for MailOnline journo (thx @MintRoyale)

#8 from /192.168.1.245:54050

7: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide <https://www.reddit.com/r/unitedkingdom/comments/49nlix/>

how\_the\_daily\_mail\_bought\_hatred\_to\_my\_home\_to wn/ @BestoftheMail

#9 from /192.168.1.245:54051

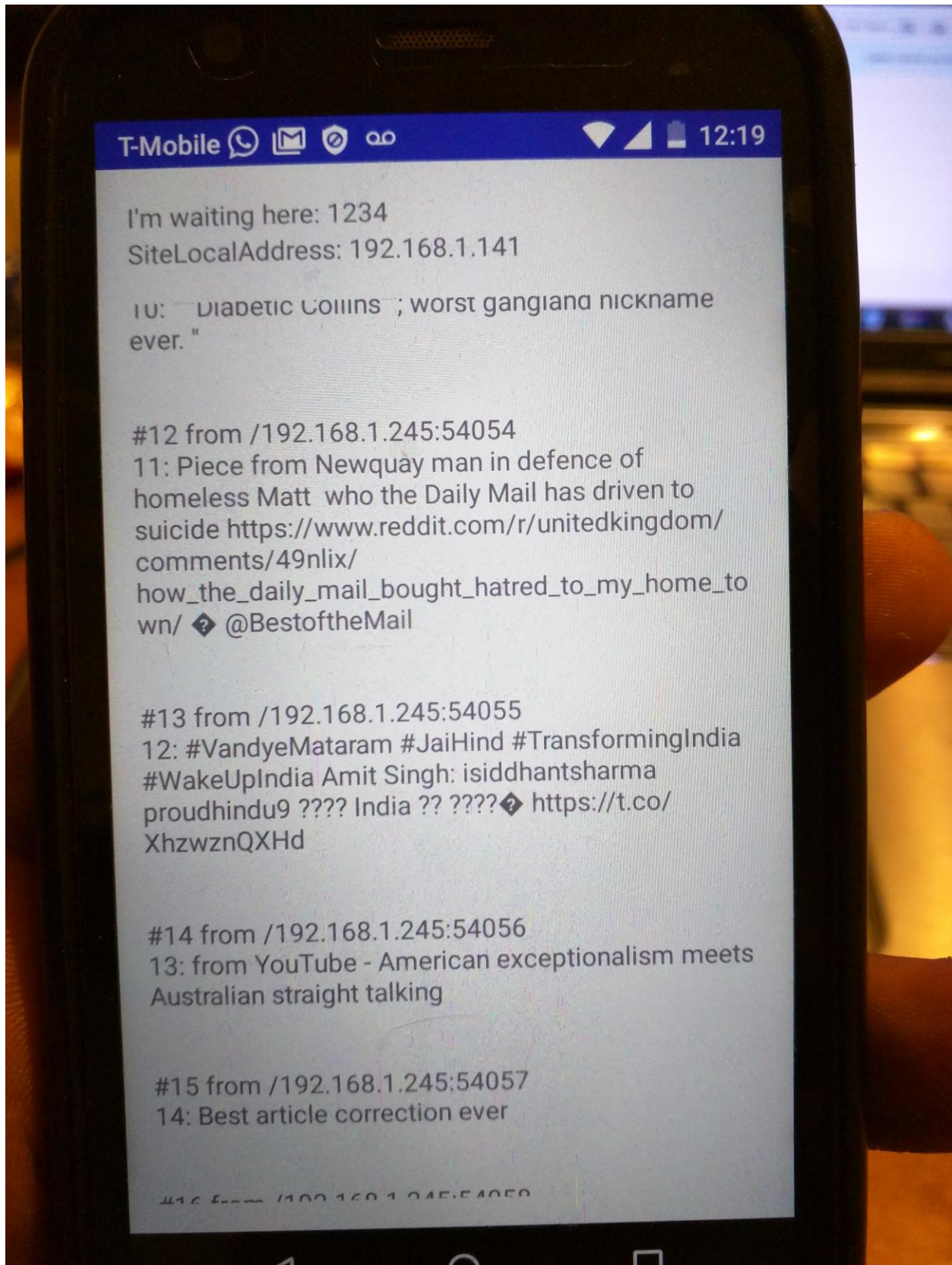
8: Piece from Newquay man in defence of homeless Matt who the Daily Mail has driven to suicide <https://www.reddit.com/r/unitedkingdom/comments/49nlix/>

how\_the\_daily\_mail\_bought\_hatred\_to\_my\_home\_to wn/ @BestoftheMail

#10 from /192.168.1.245:54052

9: "Caller @BBCRadio2: ""I'm not listening to the experts on the EU. Experts built the Titanic.""

## Lab 7 Assignment





## Lab 7 Assignment

