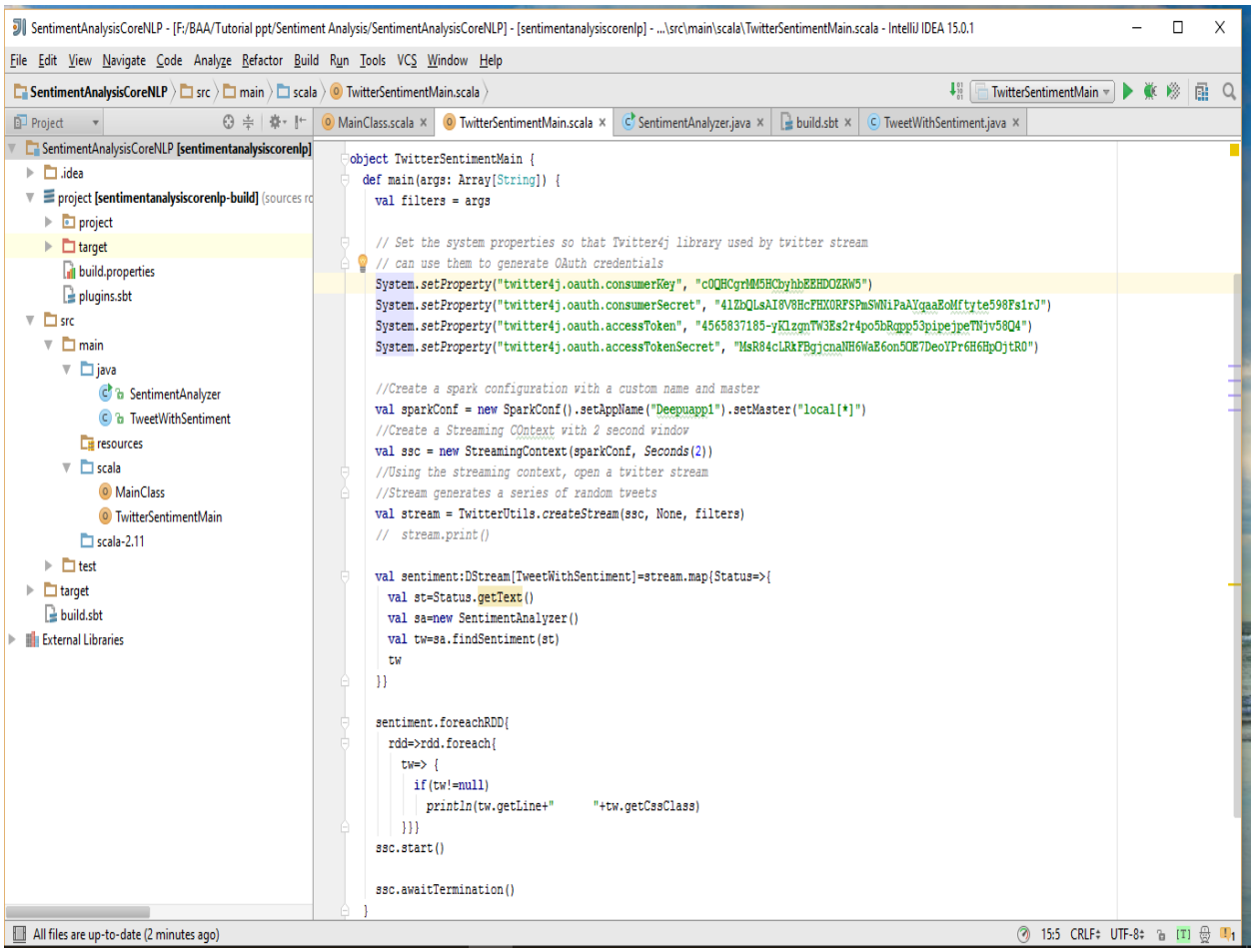
Lab Assignment 7

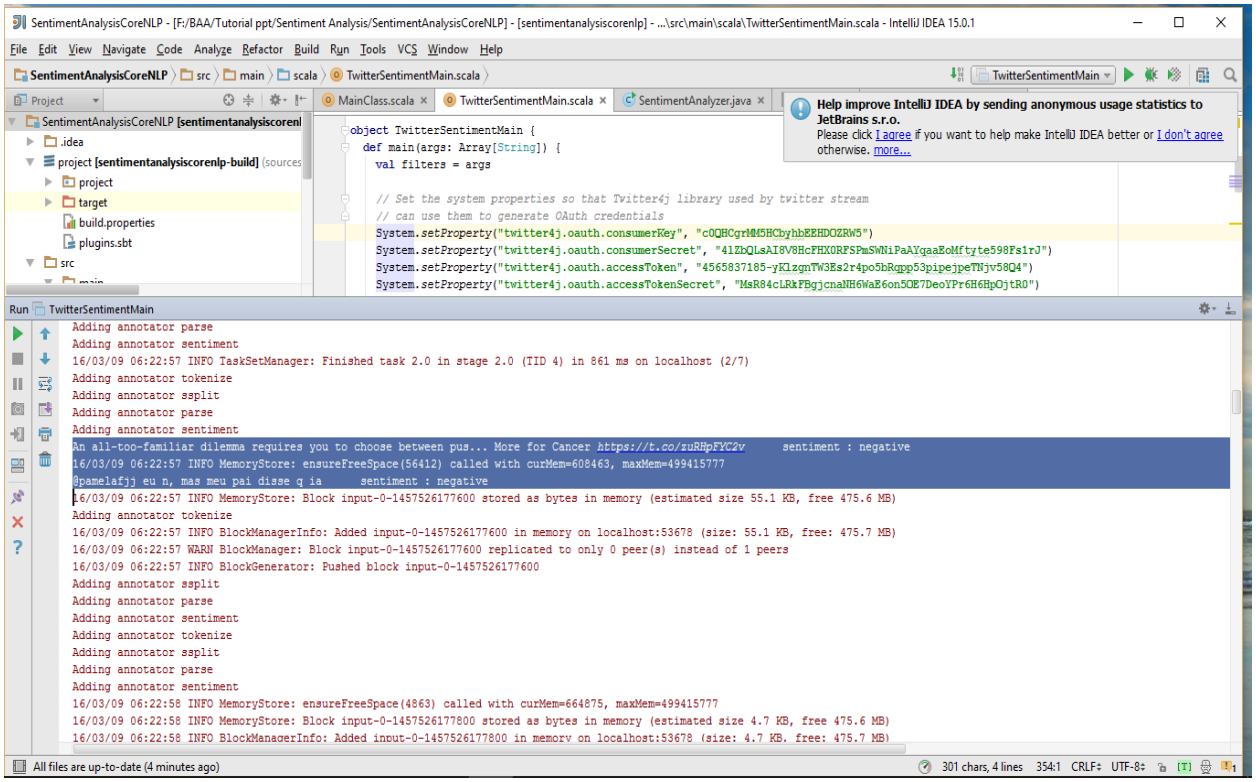
Question 1-

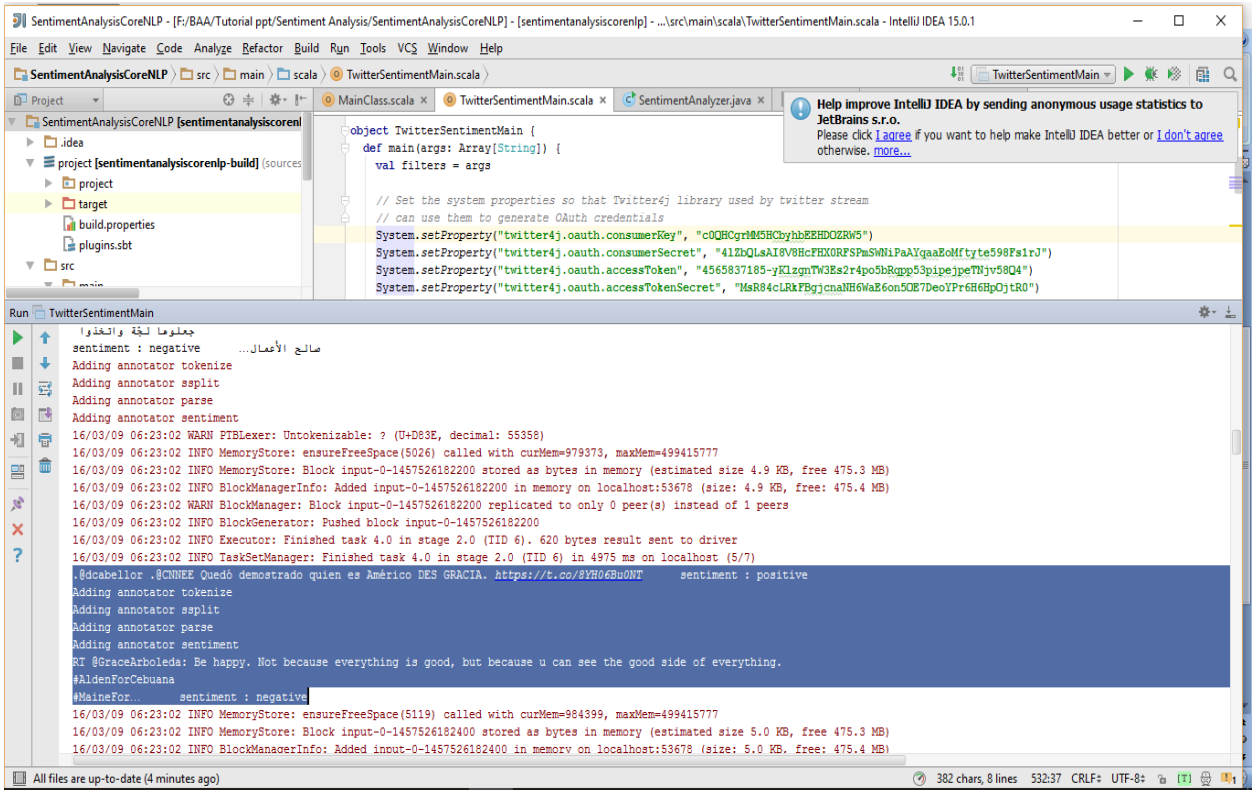
Sentimental analysis using twitter streaming

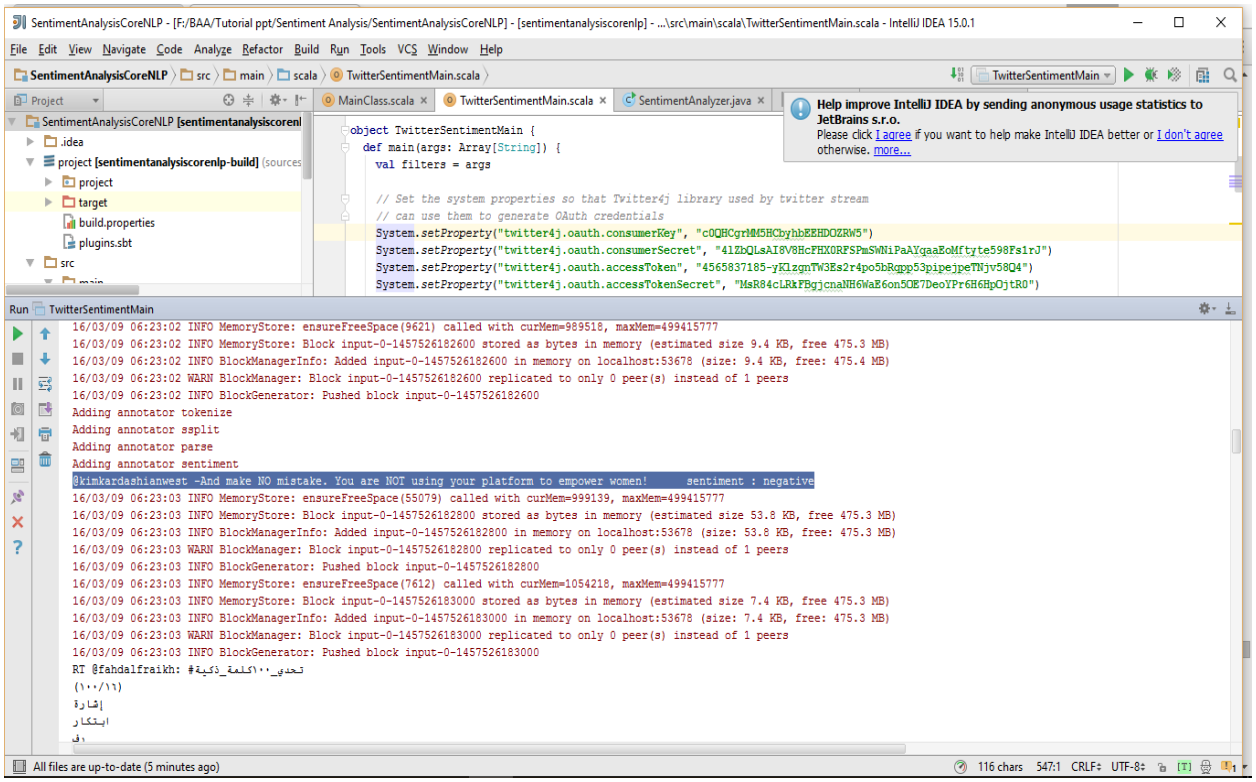
Description:

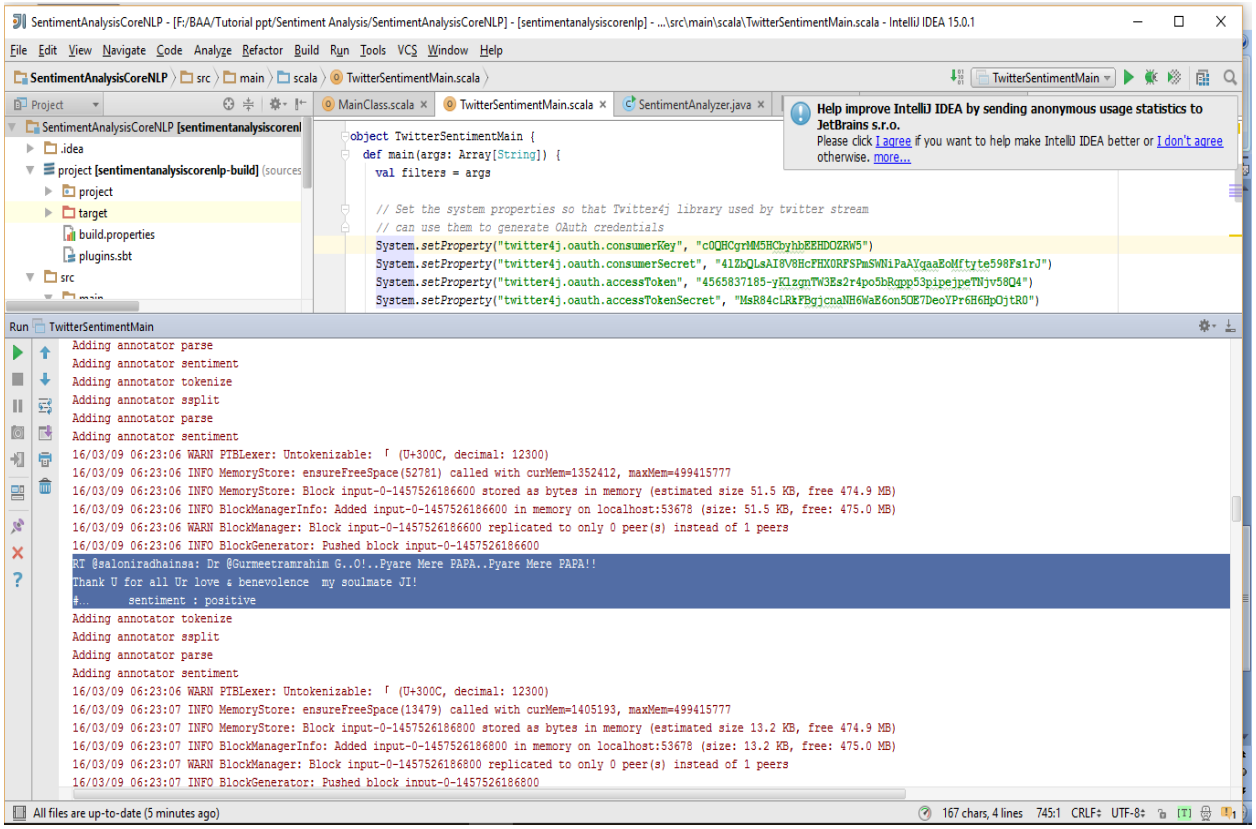
I added TwitterSentimentMain.scala to the existing project SentimentAnalysisCoreNLP given in tutorial 7. This will stream data from twitter and classify sentiment into either positive or negative.




SentimentAnalysisCoreNLP - [F:/BAA/Tutorial ppt/Sentiment Analysis/SentimentAnalysisCoreNLP] - [sentimentanalysiscorenlp] - ...\src\main\scala\TwitterSentimentMain.scala - IntelliJ IDEA 15.0.1
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
Help improve IntelliJ IDEA by sending anonymous usage statistics to JetBrains s.r.o.
Please click I agree if you want to help make IntelliJ IDEA better or I don't agree otherwise. more...
object TwitterSentimentMain {
def main(args: Array[String]) {
val filters = args
// Set the system properties so that Twitter4j library used by twitter stream
// can use them to generate OAuth credentials
System.setProperty("twitter4j.oauth.consumerKey", "c0QHCgrMM5HCbyhbEEHDOZRW5")
System.setProperty("twitter4j.oauth.consumerSecret", "4lZbQLsAI8V8HcFHX0RFSPmSWNiPaAYqaaEoMftyte598Fs1rJ")
System.setProperty("twitter4j.oauth.accessToken", "4565837185-yKlzgnTW3Es2r4po5bRqpp53pipejpeTNjv58Q4")
System.setProperty("twitter4j.oauth.accessTokenSecret", "MsR84cLRkFBgjcnaNH6WaE6on5OE7DeoYPr6H6HpOjtR0")
Run TwitterSentimentMain
Adding annotator parse
Adding annotator sentiment
16/03/09 06:22:57 INFO TaskSetManager: Finished task 2.0 in stage 2.0 (TID 4) in 861 ms on localhost (2/7)
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
An all-too-familiar dilemma requires you to choose between pus... More for Cancer https://t.co/zuRHpFYC2v sentiment : negative
16/03/09 06:22:57 INFO MemoryStore: ensureFreeSpace(56412) called with curMem=608463, maxMem=499415777
@pamelafjj eu n, mas meu pai disse q ia sentiment : negative
16/03/09 06:22:57 INFO MemoryStore: Block input-0-1457526177600 stored as bytes in memory (estimated size 55.1 KB, free 475.6 MB)
Adding annotator tokenize
16/03/09 06:22:57 INFO BlockManagerInfo: Added input-0-1457526177600 in memory on localhost:53678 (size: 55.1 KB, free: 475.7 MB)
16/03/09 06:22:57 WARN BlockManager: Block input-0-1457526177600 replicated to only 0 peer(s) instead of 1 peers
16/03/09 06:22:57 INFO BlockGenerator: Pushed block input-0-1457526177600
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
16/03/09 06:22:58 INFO MemoryStore: ensureFreeSpace(4863) called with curMem=664875, maxMem=499415777
16/03/09 06:22:58 INFO MemoryStore: Block input-0-1457526177800 stored as bytes in memory (estimated size 4.7 KB, free 475.6 MB)
All files are up-to-date (4 minutes ago)
301 chars, 4 lines 354:1 CRLF UTF-8


SentimentAnalysisCoreNLP - [F:/BAA/Tutorial ppt/Sentiment Analysis/SentimentAnalysisCoreNLP] - [sentimentanalysiscorenlp] - ...\src\main\scala\TwitterSentimentMain.scala - IntelliJ IDEA 15.0.1
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
Help improve IntelliJ IDEA by sending anonymous usage statistics to JetBrains s.r.o.
Please click I agree if you want to help make IntelliJ IDEA better or I don't agree otherwise. more...
object TwitterSentimentMain {
def main(args: Array[String]) {
val filters = args
// Set the system properties so that Twitter4j library used by twitter stream
// can use them to generate OAuth credentials
System.setProperty("twitter4j.oauth.consumerKey", "c0QHCgrMM5HCbyhbEEHDOZRW5")
System.setProperty("twitter4j.oauth.consumerSecret", "4lZbQLsAI8V8HcFHX0RFSPmSWNiPaAYqaaEoMftyte598Fs1rJ")
System.setProperty("twitter4j.oauth.accessToken", "4565837185-yKlzgnTW3Es2r4po5bRqpp53pipejpeTNjv58Q4")
System.setProperty("twitter4j.oauth.accessTokenSecret", "MsR84cLRkFBgjcnaNH6WaE6on5OE7DeoYPr6H6HpOjtR0")
Run TwitterSentimentMain
جعلوها لجّة واتخذوا
sentiment : negative صالح الأعمال...
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
16/03/09 06:23:02 WARN PTBLexer: Untokenizable: ? (U+D83E, decimal: 55358)
16/03/09 06:23:02 INFO MemoryStore: ensureFreeSpace(5026) called with curMem=979373, maxMem=499415777
16/03/09 06:23:02 INFO MemoryStore: Block input-0-1457526182200 stored as bytes in memory (estimated size 4.9 KB, free 475.3 MB)
16/03/09 06:23:02 INFO BlockManagerInfo: Added input-0-1457526182200 in memory on localhost:53678 (size: 4.9 KB, free: 475.4 MB)
16/03/09 06:23:02 WARN BlockManager: Block input-0-1457526182200 replicated to only 0 peer(s) instead of 1 peers
16/03/09 06:23:02 INFO BlockGenerator: Pushed block input-0-1457526182200
16/03/09 06:23:02 INFO Executor: Finished task 4.0 in stage 2.0 (TID 6). 620 bytes result sent to driver
16/03/09 06:23:02 INFO TaskSetManager: Finished task 4.0 in stage 2.0 (TID 6) in 4975 ms on localhost (5/7)
.@dcabellor .@CNNEE Quedó demostrado quien es Américo DES GRACIA. https://t.co/8YH06Bu0NT sentiment : positive
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
RT @GraceArboleda: Be happy. Not because everything is good, but because u can see the good side of everything.
#AldenForCebuana
#MaineFor... sentiment : negative
16/03/09 06:23:02 INFO MemoryStore: ensureFreeSpace(5119) called with curMem=984399, maxMem=499415777
16/03/09 06:23:02 INFO MemoryStore: Block input-0-1457526182400 stored as bytes in memory (estimated size 5.0 KB, free 475.3 MB)
All files are up-to-date (4 minutes ago)
382 chars, 8 lines 532:37 CRLF UTF-8


SentimentAnalysisCoreNLP - [F:/BAA/Tutorial ppt/Sentiment Analysis/SentimentAnalysisCoreNLP] - [sentimentanalysiscorenlp] - ...\src\main\scala\TwitterSentimentMain.scala - IntelliJ IDEA 15.0.1
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
SentimentAnalysisCoreNLP src main scala TwitterSentimentMain.scala
TwitterSentimentMain
Project
SentimentAnalysisCoreNLP [sentimentanalysiscorenl
.idea
project [sentimentanalysiscorenlp-build] (sources
project
target
build.properties
plugins.sbt
src
MainClass.scala
TwitterSentimentMain.scala
SentimentAnalyzer.java
Help improve IntelliJ IDEA by sending anonymous usage statistics to JetBrains s.r.o.
Please click I agree if you want to help make IntelliJ IDEA better or I don't agree otherwise. more...
object TwitterSentimentMain {
def main(args: Array[String]) {
val filters = args
// Set the system properties so that Twitter4j library used by twitter stream
// can use them to generate OAuth credentials
System.setProperty("twitter4j.oauth.consumerKey", "c0QHCgrMM5HCbyhbEEHDOZRW5")
System.setProperty("twitter4j.oauth.consumerSecret", "41ZbQLsAI8V8HcFHX0RFSPmSWNiPaAYqaaEoMfyte598Fs1rJ")
System.setProperty("twitter4j.oauth.accessToken", "4565837185-yKlzgnTW3Es2r4po5bRqpp53pipejpeTNjv58Q4")
System.setProperty("twitter4j.oauth.accessTokenSecret", "MsR84cLRkFBgjcnaNH6WaE6on5OE7DeoYPr6H6HpOjtR0")
Run TwitterSentimentMain
16/03/09 06:23:02 INFO MemoryStore: ensureFreeSpace(9621) called with curMem=989518, maxMem=499415777
16/03/09 06:23:02 INFO MemoryStore: Block input-0-1457526182600 stored as bytes in memory (estimated size 9.4 KB, free 475.3 MB)
16/03/09 06:23:02 INFO BlockManagerInfo: Added input-0-1457526182600 in memory on localhost:53678 (size: 9.4 KB, free: 475.4 MB)
16/03/09 06:23:02 WARN BlockManager: Block input-0-1457526182600 replicated to only 0 peer(s) instead of 1 peers
16/03/09 06:23:02 INFO BlockGenerator: Pushed block input-0-1457526182600
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
@kimkardashianwest -And make NO mistake. You are NOT using your platform to empower women! sentiment : negative
16/03/09 06:23:03 INFO MemoryStore: ensureFreeSpace(55079) called with curMem=999139, maxMem=499415777
16/03/09 06:23:03 INFO MemoryStore: Block input-0-1457526182800 stored as bytes in memory (estimated size 53.8 KB, free 475.3 MB)
16/03/09 06:23:03 INFO BlockManagerInfo: Added input-0-1457526182800 in memory on localhost:53678 (size: 53.8 KB, free: 475.3 MB)
16/03/09 06:23:03 WARN BlockManager: Block input-0-1457526182800 replicated to only 0 peer(s) instead of 1 peers
16/03/09 06:23:03 INFO BlockGenerator: Pushed block input-0-1457526182800
16/03/09 06:23:03 INFO MemoryStore: ensureFreeSpace(7612) called with curMem=1054218, maxMem=499415777
16/03/09 06:23:03 INFO MemoryStore: Block input-0-1457526183000 stored as bytes in memory (estimated size 7.4 KB, free 475.3 MB)
16/03/09 06:23:03 INFO BlockManagerInfo: Added input-0-1457526183000 in memory on localhost:53678 (size: 7.4 KB, free: 475.3 MB)
16/03/09 06:23:03 WARN BlockManager: Block input-0-1457526183000 replicated to only 0 peer(s) instead of 1 peers
16/03/09 06:23:03 INFO BlockGenerator: Pushed block input-0-1457526183000
RT @fahdalfraikh: #تحدي_''كلمة_ذكية
(١٠٠/١٦)
إشارة
ابتكار
All files are up-to-date (5 minutes ago)
116 chars 547:1 CRLF UTF-8


SentimentAnalysisCoreNLP - [F:/BAA/Tutorial ppt/Sentiment Analysis/SentimentAnalysisCoreNLP] - [sentimentanalysiscorenlp] - ...\src\main\scala\TwitterSentimentMain.scala - IntelliJ IDEA 15.0.1
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
SentimentAnalysisCoreNLP src main scala TwitterSentimentMain.scala
TwitterSentimentMain
Project
SentimentAnalysisCoreNLP [sentimentanalysiscorenl
.idea
project [sentimentanalysiscorenlp-build] (sources
project
target
build.properties
plugins.sbt
src
MainClass.scala
TwitterSentimentMain.scala
SentimentAnalyzer.java
Help improve IntelliJ IDEA by sending anonymous usage statistics to JetBrains s.r.o.
Please click I agree if you want to help make IntelliJ IDEA better or I don't agree otherwise. more...
object TwitterSentimentMain {
def main(args: Array[String]) {
val filters = args
// Set the system properties so that Twitter4j library used by twitter stream
// can use them to generate OAuth credentials
System.setProperty("twitter4j.oauth.consumerKey", "c0QHCgrMM5HCbyhbEEHDOZRW5")
System.setProperty("twitter4j.oauth.consumerSecret", "41ZbQLsAI8V8HcFHX0RFSPmSWNiPaAYqaaEoMfyte598Fs1rJ")
System.setProperty("twitter4j.oauth.accessToken", "4565837185-yKlzgnTW3Es2r4po5bRqpp53pipejpeTNjv58Q4")
System.setProperty("twitter4j.oauth.accessTokenSecret", "MsR84cLRkFBgjcnaNH6WaE6on5OE7DeoYPr6H6HpOjtR0")
Run TwitterSentimentMain
Adding annotator parse
Adding annotator sentiment
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
16/03/09 06:23:06 WARN PTBLexer: Untokenizable: 「 (U+300C, decimal: 12300)
16/03/09 06:23:06 INFO MemoryStore: ensureFreeSpace(52781) called with curMem=1352412, maxMem=499415777
16/03/09 06:23:06 INFO MemoryStore: Block input-0-1457526186600 stored as bytes in memory (estimated size 51.5 KB, free 474.9 MB)
16/03/09 06:23:06 INFO BlockManagerInfo: Added input-0-1457526186600 in memory on localhost:53678 (size: 51.5 KB, free: 475.0 MB)
16/03/09 06:23:06 WARN BlockManager: Block input-0-1457526186600 replicated to only 0 peer(s) instead of 1 peers
16/03/09 06:23:06 INFO BlockGenerator: Pushed block input-0-1457526186600
RT @saloniradhainsa: Dr @Gurmeetramrahim G..O!..Pyare Mere PAPA..Pyare Mere PAPA!!
Thank U for all Ur love & benevolence my soulmate JI!
#.. sentiment : positive
Adding annotator tokenize
Adding annotator ssplit
Adding annotator parse
Adding annotator sentiment
16/03/09 06:23:06 WARN PTBLexer: Untokenizable: 「 (U+300C, decimal: 12300)
16/03/09 06:23:07 INFO MemoryStore: ensureFreeSpace(13479) called with curMem=1405193, maxMem=499415777
16/03/09 06:23:07 INFO MemoryStore: Block input-0-1457526186800 stored as bytes in memory (estimated size 13.2 KB, free 474.9 MB)
16/03/09 06:23:07 INFO BlockManagerInfo: Added input-0-1457526186800 in memory on localhost:53678 (size: 13.2 KB, free: 475.0 MB)
16/03/09 06:23:07 WARN BlockManager: Block input-0-1457526186800 replicated to only 0 peer(s) instead of 1 peers
16/03/09 06:23:07 INFO BlockGenerator: Pushed block input-0-1457526186800
All files are up-to-date (5 minutes ago)
167 chars, 4 lines 745:1 CRLF UTF-8

Question-2

Make Recommendation(Related to your Project)

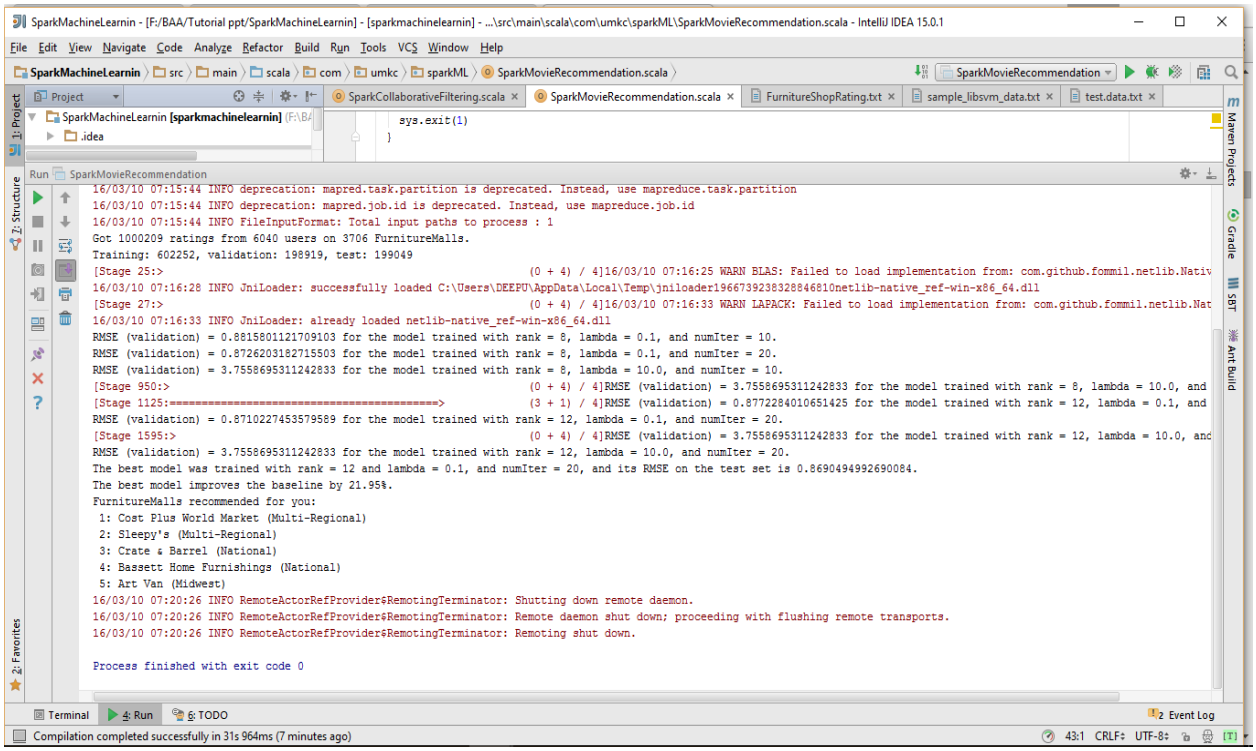
Description:

In this section, I used the SparkMovieRecommendation.scala code to recommend the best places to buy his furniture(FurnitureMalls) to the user. Using 25 handpicked FurnitureMalls, the model recommends top 5 FurnitureMalls based on user preference.

This Screen shot shows the iterations of training:

```
Got 1000209 ratings from 6040 users on 3706 FurnitureMalls.
Training: 602252, validation: 198919, test: 199049
[Stage 25:>                                                         (0 + 4) / 4]16/03/10 07:16:25 WARN BLAS: Failed to load implementation from: com.github.fommil.netlib.Nativ
16/03/10 07:16:28 INFO JniLoader: successfully loaded C:\Users\DEEPU\AppData\Local\Temp\jniloader1966739238328846810netlib-native_ref-win-x86_64.dll
[Stage 27:>                                                         (0 + 4) / 4]16/03/10 07:16:33 WARN LAPACK: Failed to load implementation from: com.github.fommil.netlib.Nat
16/03/10 07:16:33 INFO JniLoader: already loaded netlib-native_ref-win-x86_64.dll
RMSE (validation) = 0.8815801121709103 for the model trained with rank = 8, lambda = 0.1, and numIter = 10.
RMSE (validation) = 0.8726203182715503 for the model trained with rank = 8, lambda = 0.1, and numIter = 20.
RMSE (validation) = 3.7558695311242833 for the model trained with rank = 8, lambda = 10.0, and numIter = 10.
[Stage 950:>                                                        (0 + 4) / 4]RMSE (validation) = 3.7558695311242833 for the model trained with rank = 8, lambda = 10.0, and
[Stage 1125:=========================================>             (3 + 1) / 4]RMSE (validation) = 0.8772284010651425 for the model trained with rank = 12, lambda = 0.1, and
RMSE (validation) = 0.8710227453579589 for the model trained with rank = 12, lambda = 0.1, and numIter = 20.
[Stage 1595:>                                                       (0 + 4) / 4]RMSE (validation) = 3.7558695311242833 for the model trained with rank = 12, lambda = 10.0, and
RMSE (validation) = 3.7558695311242833 for the model trained with rank = 12, lambda = 10.0, and numIter = 20.
The best model was trained with rank = 12 and lambda = 0.1, and numIter = 20, and its RMSE on the test set is 0.8690494992690084.
The best model improves the baseline by 21.95%.
```

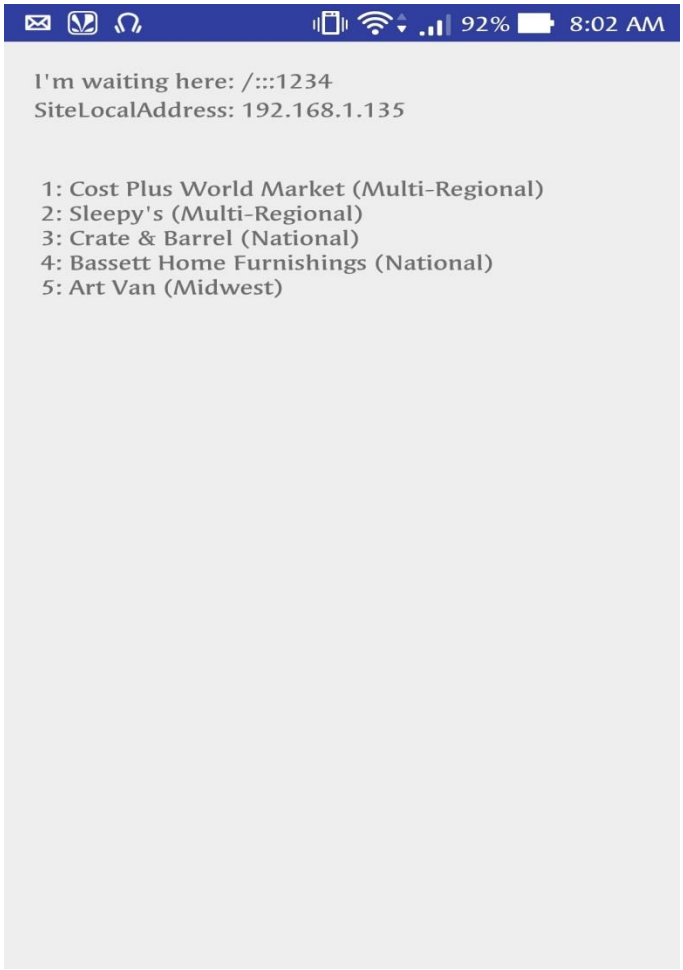
This Screenshot shows the FurnitureMalls recommended to you:



Question 3

Recommendation through smartphone:

Description:

Now the goal is to send the recommendations of the user to his smart phone. So I have taken this recommendation output above and sent it as a notification to smart phone.

Here is the Screenshot: