

Project Report—Analyze Models

Get news from New York Times

```
python news_api.py
```

Compare binary classifiers when number of Features is 200

```
bin/spark-submit tfidf_classify_spark.py
```

DATA: restaurant reviews {Positive; Negative}

When number of features is 200:

Compare the confusion matrix and test error rate of models

```
{NaiveBayes; SVMWithSGD; LogisticRegressionWithLBFGS; OnlinePerceptron(Itr=1);  
AveragePerceptron(Itr=1);OnlinePerceptron(Itr=1)+AveragePerceptron(Itr=1);AveragePerceptron(Itr=5);  
AveragePerceptron(Itr=10)}
```

```
[Yilans-MacBook-Air:spark-2.0.1-bin-hadoop2.7 Jillian$ bin/spark-submit ~/Documents/Python/large_data_pj/tfidf_classify_spark.py  
Test error rate of NaiveBayes Model: 0.236236236236  
Test error rate of SVMWithSGD Model: 0.253253253253  
Test error rate of LogisticRegressionWithLBFGS Model: 0.208208208208  
Test error rate of OnlinePerceptron: 0.29029029029  
Test error rate of AveragePerceptron(iterations=1): 0.232232232232  
Test error rate of AveragePerceptron(iterations=1) after single-pass OnlinePerceptron: 0.205205205205  
Test error rate of AveragePerceptron(iterations=5):: 0.203203203203  
Test error rate of AveragePerceptron(iterations=10): 0.212212212212
```

Confusion Matrix of models:

Model	errorRate	falseNegative	falsePositive	trueNegative	truePositive
NaiveBayes	0.23623623623624	83	153	171	592
SVMWithSGD	0.25325325325326	19	234	90	656
LogisticRegressionWithLBFGS	0.2082082082082	82	126	198	593
OnlinePerceptron	0.2902902902903	207	83	241	468
AveragePerceptronItr1	0.23223223223224	80	152	172	595
AveragePerceptronOPItr1	0.20520520520522	78	127	197	597
AveragePerceptronItr5	0.2032032032032	76	127	197	599
AveragePerceptronItr10	0.21221221221222	79	133	191	596

Compare binary classifiers when number of Features is 2000

```
bin/spark-submit tfidf_classify_spark.py
```

DATA: restaurant reviews {Positive; Negative}

When number of features is 2000:

Compare the confusion matrix and test error rate of models

```
{NaiveBayes; SVMWithSGD; LogisticRegressionWithBFGS; OnlinePerceptron(Itr=1);
AveragePerceptron(Itr=1);OnlinePerceptron(Itr=1)+AveragePerceptron(Itr=1);AveragePerceptron(Itr=5);
AveragePerceptron(Itr=10)}
```

```
(Vilans-MacBook-Air:spark-2.0.1-bin-hadoop2.7 Jillian$ bin/spark-submit ~/Documents/Python/large_data_pj/tfidf_classify_spark.py
Test error rate of NaiveBayes Model: 0.159159159159
Test error rate of SVMWithSGD Model: 0.223223223223
Test error rate of LogisticRegressionWithLBFGS Model: 0.0
Test error rate of OnlinePerceptron: 0.14014014014
Test error rate of AveragePerceptron(iterations=1): 0.17017017017
Test error rate of AveragePerceptron(iterations=1) after single-pass OnlinePerceptron: 0.106106106106
Test error rate of AveragePerceptron(iterations=5):: 0.115115115115
Test error rate of AveragePerceptron(iterations=10): 0.103103103103
```

Confusion Matrix of models:

Model	errorRate	falseNegative	falsePositive	trueNegative	truePositive
NaiveBayes	0.15915915915915915	63	96	228	612
SVMWithSGD	0.22322322322322322	18	205	119	657
LogisticRegressionWithLBFGS	0.0	0	0	324	675
OnlinePerceptron	0.14014014014014015	47	93	231	628
AveragePerceptronIter1	0.17017017017017017	58	112	212	617
AveragePerceptronOptIter1	0.1061061061061061	47	59	265	628
AveragePerceptronIter5	0.11511511511511512	38	77	247	637
AveragePerceptronIter10	0.1031031031031031	36	67	257	639

MulticlassPerceptron Training error rate and Confusion Matrices:

bin/spark-submit category_news_train_perceptron.py

Data: News from *NEW YORK TIMES* in 2016 year, {number of classes=19; number of features=2000}

Lifting perceptron binary class models to MulticlassPerceptron using One-vs-Rest

Training error rate

```
[Yilans-MacBook-Air:spark-2.0.1-bin-hadoop2.7 Jillian$ bin/spark-submit ~/Documents/Python/large_data_pj/
/category_news_train_perceptron.py
22223
+-----+-----+-----+-----+-----+-----+
| category| errorRate|falseNegative|falsePositive|trueNegative|truePositive|
+-----+-----+-----+-----+-----+-----+
| Art & Design| 0.10457633982810602| 2| 2322| 19200| 699|
| World| 0.09116680916167934| 180| 1846| 15951| 4246|
| Sports| 0.04949826756063538| 62| 1038| 16968| 4155|
| Fashion & Style| 0.1538946136885209| 2| 3418| 18055| 748|
| Books| 0.029788957386491472| 1| 661| 20360| 1201|
| Music| 0.06929757458488953| 1| 1539| 19601| 1082|
| Television| 0.04967826126085587| 1| 1103| 20588| 531|
| Movies| 0.03644872429464969| 2| 808| 20400| 1013|
| Technology| 0.11515096971605994| 0| 2559| 18871| 793|
| Science| 0.13990010349637763| 0| 3109| 18505| 609|
| Food| 0.04396346127885524| 0| 977| 20727| 519|
| Real Estate| 0.0559780407685731| 0| 1244| 20626| 353|
| Theater| 0.06754263600773973| 0| 1501| 19998| 724|
| Health| 0.11141610043648473| 2| 2474| 18819| 928|
| Travel| 0.0949916752913648| 0| 2111| 19773| 339|
| Education| 0.07397741079062232| 0| 1644| 19910| 669|
| Your Money| 0.07555235566755164| 0| 1679| 20388| 156|
| Politics| 0.05858794942177024| 19| 1283| 17929| 2992|
| Economy| 0.07109751158709446| 0| 1580| 20449| 194|
+-----+-----+-----+-----+-----+-----+
```

Compare Multiclass Models Test error rate and Confusion Matrices:

```
bin/spark-submit compare_Multiclass_Models.py
```

Data: News from *NEW YORK TIMES* in 2016 year, {number of classes=19; number of features=2000}

```
Yilans-MacBook-Air:spark-2.0.1-bin-hadoop2.7 Jillian$ bin/spark-submit ~/Documents/Python/large_data_pj/compare_Multiclass_Models.py
```

```
training data set count: 17765
```

```
test data set count: 4458
```

```
NaiveBayes Test set accuracy = 0.814266487214
```

```
DecisionTree Test set accuracy = 0.411395244504
```

```
RandomForest model: RandomForestClassificationModel (uid=rfc_56a2c3d1ed00) with 10 trees
```

```
RandomForest Test set accuracy = 0.449977568416
```

```
One-vs-Rest (LogisticRegression) Test set accuracy = 0.758187528039
```

```
MulticlassPerceptron accuracy:
```

```
+-----+-----+-----+-----+-----+-----+
|      category|      errorRate|falseNegative|falsePositive|trueNegative|truePositive|
+-----+-----+-----+-----+-----+-----+
|  Art & Design| 0.09510991475998205|      12|      412|      3902|      132|
|      World| 0.11799013010318528|     103|      423|      3141|      791|
|      Sports| 0.07514580529385374|      66|      269|      3331|      792|
|Fashion & Style| 0.16464782413638404|      36|      698|      3601|      123|
|      Books|0.036787797218483624|      18|      146|      4061|      233|
|      Music| 0.09914759982054733|      23|      419|      3828|      188|
|  Television|0.048900852400179454|       9|      209|      4150|       90|
|      Movies| 0.03297442799461642|      16|      131|      4124|      187|
|  Technology| 0.13593539703903096|      19|      587|      3713|      139|
|      Science| 0.15029161058770749|      21|      649|      3679|      109|
|      Food| 0.05361148497083894|      11|      228|      4117|      102|
|  Real Estate| 0.11238223418573351|       5|      496|      3897|       60|
|      Theater| 0.04441453566621804|       9|      189|      4126|      134|
|      Health| 0.13885150291610587|      17|      602|      3662|      177|
|      Travel| 0.09062359802602064|       6|      398|      3995|       59|
|  Education| 0.06370569762225213|       9|      275|      4051|      123|
|  Your Money| 0.06729475100942127|       3|      297|      4133|       25|
|  Politics| 0.07447285778375953|      42|      290|      3600|      526|
|  Economy| 0.08613728129205922|       4|      380|      4035|       39|
+-----+-----+-----+-----+-----+-----+
```

Compare 4 MulticlassPerceptron Models with different initial weight

```
bin/spark-submit category_tweets_perceptron.py --master local[10]
```

Training data is Feedback of correct categories of tweets (200)

Predict categories of test tweets using models after online trained every 50 tweets

- OnlinePerceptron model with weight initialized as zeros
- AveragePerceptron model with weight initialized as zeros
- AveragePerceptron model with weight initialized as trained model by news(but ignore weights history of 22223 news)
- AveragePerceptron model with weight initialized as trained model by news(and store weights history of 22223 news)

```
print "Training 4 MulticlassPerceptron models using",dataset.count(),"training data.....\n"
# train an online perceptron model use part of training data
initialModelOP.train(traindata,trainlabels,method="Online")
# train an Average perceptron model(Iteration=10) use part of training data
#initialModelAP.train(traindata,trainlabels,method="Online")
initialModelAP.train(traindata,trainlabels,method="Average",source="News")
# train an Average perceptron model(Iteration=10) use part of training data initialized using News trained Model(ignore weights)
loadModelOP.train(traindata,trainlabels,method="Average",source="Feedback")
# train an Average perceptron model(Iteration=10) use part of training data initialized using News trained Model(store weights)
loadModelAP.train(traindata,trainlabels,method="Average",source="Feedback")

print "Predicting labels of test Tweets after trained ",dataset.count()," Category Feedback Tweets:\n"
```

```
^[[AVilans-MacBook-Air:spark-2.0.1-bin-hadoop2.7 Jillian$ bin/spark-submit ~/Documents/Python/large_data_pj/category_tweets_perceptron.py
```

category	text
Others	I really hate folks w. pissy ass attitudes for nun.. shit no where near cute
Others	"... even ignoring @RuPaul in her quest to become a mindless mogul." Never! https://t.co/qr4mecWmSv
Food	Bout to eat at CB 🍕 I love their breakfast
Others	I hate my friends lmao https://t.co/Esb04AktBg
Your Money	Every tax dollar invested in family-planning saves taxpayers 7 in obvious, predictable costs.
Others	dead week, man. It gets people. https://t.co/C8zislnRa3
Others	"Where dreams are made of" is not poetic license, it's just bad grammar.
Politics	Soldiers dying in war with naksals, party ministers and people busy celebrating delhi MCD election, not good !
Others	Bad little bitch, wanna know how her lips taste
Travel	We have these in downtown Omaha, Nebraska 🍷 https://t.co/hvUg9PqGM
Others	It's weird that you can remember dreams, that make no sense. But for some reason it will stick with you.
Fashion & Style	@Phil_Lewis_Next up, underwear with artificial skid marks.
Others	@clairemackkyd Are you referring to the dinked middle? Do you have 7" "spider" adapter? Or did you mean something e... https://t.co/2bCHaynfsF
Fashion & Style	If u put actual clothes on & do your makeup everyday before class, what r u doing?!?!?
Others	@Haylie_Evald3 so jealous 🙄
Others	So tired of everybody & everything.
World	China Launches Aircraft Carrier, Boosting Military Presence
Others	motherfucking sucks.
Television	these men really are working hard for their spinoff. https://t.co/0cHaw10UEt
Health	Hospital againnnnn 🤒

only showing top 20 rows

AveragePerceptron	NewsPerceptronIgnoreWeight	NewsPerceptronStoreWeight	OnlinePerceptron
Others	Sports Fashion & Style Health Travel	Sports Fashion & Style Health Travel	Others
Others	Fashion & Style Movies Health Travel	Fashion & Style Movies Health Travel	Others
Others	Sports Fashion & Style Technology Science Health Education	Sports Fashion & Style Technology Science Health Education	Others
Others	Art & Design Sports Health Education	Art & Design Sports Health Education	Others
Others	Music Movies Real Estate Travel	Music Movies Real Estate Travel	Others
Others	World Health	World Health	Others
Others	Art & Design Fashion & Style Food Health Travel Your Money Economy	Art & Design Fashion & Style Food Health Travel Your Money Economy	Others
Others	World Fashion & Style Travel	World Fashion & Style Travel	Others
Others	Sports Fashion & Style Music	Sports Fashion & Style Music	Others
Others	World Health	World Health	Others
Others	Sports Science Health	Sports Science Health	Others
Others	Fashion & Style Music Real Estate	Fashion & Style Music Real Estate	Others
Others	Science Education	Science Education	Others
Others	Technology Health Education Your Money	Technology Health Education Your Money	Others
Others	Fashion & Style Education	Fashion & Style Education	Others
Others	Fashion & Style Music Science Health	Fashion & Style Music Science Health	Others
Others	World	World	Others
Others	Others	Others	Others
Others	World Sports Fashion & Style Health Your Money Economy	World Sports Fashion & Style Health Your Money Economy	Others
Others	World Sports Health	World Sports Health	Others

only showing top 20 rows

Training 4 MulticlassPerceptron models using 50 training data.....

Training 4 MulticlassPerceptron models using 50 training data.....

Predicting labels of test Tweets after trained 50 Category Feedback Tweets:

AveragePerceptron	NewsPerceptronIgnoreWeight	NewsPerceptronStoreWeight	OnlinePerceptron
Others	Health Travel	Fashion & Style Health Travel	Others
Others	Fashion & Style Movies Food Travel	Fashion & Style Movies Health Travel	Others
Others	Science Food	Sports Fashion & Style Technology Science	Others
Others	Sports	Sports Health Education	Others
Others	Music Movies Travel	Music Movies Travel	Others
Others	Others	World	Others
Others	Others	Art & Design Fashion & Style Food Health Economy	Others
Others	World Fashion & Style	World Fashion & Style Travel	Others
Others	Fashion & Style	Sports Fashion & Style Music	Others
Others	World	World Health	Others
Others	Sports	Sports Science Health	Others
Others	Music	Music	Others
Others	Science Education	Science Education	Others
Others	Fashion & Style	Technology Health Education	Others
Others	Others	Fashion & Style	Others
Others	Music	Music Health	Others
Others	World	World	Others
Others	Others	Others	Others
Others	World Sports Fashion & Style Economy	Sports Fashion & Style Health Your Money Economy	Others
Others	World	World Sports Health	Others

only showing top 20 rows

Training 4 MulticlassPerceptron models using 48 training data.....

Training 4 MulticlassPerceptron models using 48 training data.....

Predicting labels of test Tweets after trained 48 Category Feedback Tweets:

AveragePerceptron	NewsPerceptronIgnoreWeight	NewsPerceptronStoreWeight	OnlinePerceptron
Others	Travel	Art & Design Sports Fashion & Style Technology Health Travel	Others
Others	Fashion & Style Movies Travel	World Fashion & Style Movies Health Travel	Others
Others	Science Food	Sports Fashion & Style Technology Science Health Education	Others
Others	Others	Art & Design Sports Fashion & Style Movies Health Education	Others
Others	Others	Music Movies Travel	Others
Others	Others	World Science Health	Others
Others	Others	Art & Design Fashion & Style Science Food Health Travel Economy	Others
Others	World Fashion & Style	World Fashion & Style Travel	Others
Others	Fashion & Style	Sports Fashion & Style Music Technology	Others
Others	World	World Health	Others
Others	Sports	Sports Technology Science Health	Others
Others	Music	Sports Fashion & Style Music Science Real Estate Health	Others
Others	Education	Technology Science Education	Others
Others	Fashion & Style	Technology Health Education	Others
Others	Others	Fashion & Style Technology Health Education	Others
Others	Music	Sports Fashion & Style Music Technology Health	Others
Others	World	World Health Education	Others
Others	Others	Sports	Others
Others	World Fashion & Style Economy	Sports Fashion & Style Health Your Money Economy	Others
Others	World	World Sports Health Education	Others

only showing top 20 rows

Training 4 MulticlassPerceptron models using 48 training data.....

Training 4 MulticlassPerceptron models using 48 training data.....

Predicting labels of test Tweets after trained 48 Category Feedback Tweets:

AveragePerceptron	NewsPerceptronIgnoreWeight	NewsPerceptronStoreWeight	OnlinePerceptron
Others	Travel	Art & Design Sports Fashion & Style Technology Science Real Estate Health Travel Education	Others
Others	Fashion & Style Movies Food Travel	World Fashion & Style Movies Technology Health Travel Education	Others
Others	Science Food	Art & Design Sports Fashion & Style Technology Science Health Education	Others
Others	Sports	Art & Design Sports Fashion & Style Books Movies Technology Health Education	Others
Others	Music Movies Travel	World Music Movies Real Estate Health Travel Politics Economy	Others
Others	Others	Art & Design World Sports Fashion & Style Science Health Education	Others
Others	Others	Art & Design Sports Fashion & Style Science Food Health Travel Education Economy	Others
Others	World	World Fashion & Style Health Travel	Others
Others	Fashion & Style	Sports Fashion & Style Music Technology Health Education	Others
Others	World	Art & Design World Technology Health	Others
Others	Sports	World Sports Technology Science Health Education	Others
Others	Music	Sports Fashion & Style Music Technology Science Food Real Estate Health	Others
Others	Education	Technology Science Health Education	Others
Others	Fashion & Style	World Technology Health Education	Others
Others	Others	Art & Design Sports Fashion & Style Technology Health Education	Others
Others	Music	Sports Fashion & Style Music Technology Science Health Education	Others
Others	World	World Fashion & Style Health Education	Others
Others	Others	Sports Fashion & Style Music Health Education	Others
Others	World Sports Economy	Sports Fashion & Style Technology Health Education Economy	Others
Others	World	World Sports Fashion & Style Technology Science Health Education	Others

only showing top 20 rows

Training 4 MulticlassPerceptron models using 57 training data.....

Training 4 MulticlassPerceptron models using 57 training data.....

Predicting labels of test Tweets after trained 57 Category Feedback Tweets:

AveragePerceptron		NewsPerceptronIgnoreWeight	NewsPerceptronStoreWeight	OnlinePerceptron
Others	Health	Art & Design World Sports Fashion & Style Technology Science Real Estate Health Education	Others	
Others	Fashion & Style Travel	World Sports Fashion & Style Movies Technology Real Estate Health Education	Others	
Others	Science Food	Art & Design Sports Fashion & Style Technology Science Health Education	Others	
Others	Sports	Art & Design World Sports Fashion & Style Books Music Movies Technology Science Health Education	Others	
Others	Music Movies Travel	World Sports Music Movies Technology Real Estate Health Travel Education Politics	Others	
Others	Others	Art & Design World Sports Fashion & Style Technology Science Health Education	Others	
Others	Others	Art & Design World Sports Fashion & Style Science Food Real Estate Health Travel Education Your Money Economy	Others	
Others	World Fashion & Style Travel	World Sports Fashion & Style Technology Real Estate Health Travel Education	Others	
Others	Fashion & Style	World Sports Fashion & Style Books Music Technology Science Food Health Education	Others	
Others	World	Art & Design World Sports Fashion & Style Books Music Technology Health Education	Others	
Others	Sports Health	World Sports Fashion & Style Technology Science Health Education	Others	
Others	Music	Sports Fashion & Style Books Music Technology Science Real Estate Health Education	Others	
Others	Science Education	Sports Books Technology Science Health Education	Others	
Others	Fashion & Style	World Sports Music Technology Real Estate Health Education Your Money	Others	
Others	Others	Art & Design World Sports Fashion & Style Music Technology Science Health Education	Others	
Others	Music	Sports Fashion & Style Books Music Technology Science Health Education	Others	
Others	World	Art & Design World Sports Fashion & Style Music Technology Health Education	Others	
Others	Others	Art & Design World Sports Fashion & Style Music Technology Health Education	Others	
Others	World Sports Fashion & Style Economy	Sports Fashion & Style Music Technology Health Education Your Money Economy	Others	
Others	World Health	World Sports Fashion & Style Music Technology Science Health Education	Others	

only showing top 20 rows

Compare 4 MulticlassPerceptron Models with different iterations and memory state

```
bin/spark-submit category_tweets_perceptron.py --master local[10]
```

Training data is Feedback of correct categories of tweets (400)

Predict categories of test tweets using models after online trained every 100 tweets

- OnlinePerceptron model with weight initialized as zeros, iteration=10
- OnlinePerceptron model with weight initialized as trained model by news(ignored history of News), iteration=10
- AveragePerceptron model with weight initialized as trained model by news(ignore history of News), iteration=1
- AveragePerceptron model with weight initialized as trained model by news(ignore history of News), iteration=10

```
print "Training 4 MulticlassPerceptron models using",dataset.count(),"training data.....\n"
# train an Online Perceptron model(Iteration=10) use part of training data, initialize model with all zeros
initialModelOP.train(traindata,trainlabels,method="Online",MaxItr=10)
# train an Online Perceptron model(Iteration=10) use part of training data, initialize model with news trained model(ignore history)
initialModelAP.train(traindata,trainlabels,method="Online",MaxItr=10)
# train an Average perceptron model(Iteration=1) use part of training data, initialize using News trained Model(ignore history)
loadModelOP.train(traindata,trainlabels,method="Average",MaxItr=1)
# train an Average perceptron model(Iteration=10) use part of training data, initialize using News trained Model(ignore history)
loadModelAP.train(traindata,trainlabels,method="Average",MaxItr=10)

print "Predicting labels of test Tweets after trained ",dataset.count()," Category Feedback Tweets:\n"
```

```
Yilans-MacBook-Air:spark-2.0.1-bin-hadoop2.7 Jillian$ bin/spark-submit ~/Documents/Python/large_data_pj/category_tweets_perceptron.py
```

```
+-----+
|category|text|
+-----+
|Others  |really hate folks w. pissy ass attitudes for nun.. shit no where near cute|
|Others  |"... even ignoring @RuPaul in her quest to become a mindless mogul." Never! https://t.co/qr4mecVwSv|
|Food    |Bout to eat at CB 🍔 I love their breakfast|
|Others  |I hate my friends lmao https://t.co/Esb04AktBg|
|Your Money|Every tax dollar invested in family-planning saves taxpayers 7 in obvious, predictable costs.|
|Others  |dead week, man. It gets people. https://t.co/C8zislnRa3|
|Others  |"Where dreams are made of" is not poetic license, it's just bad grammar.|
|Politics|Soldiers dying in war with naksals, party ministers and people busy celebrating delhi MCD election, not good !|
|Others  |Bad little bitch, wanna know how her lips taste|
|Travel  |We have these in downtown Omaha, Nebraska 🍷 https://t.co/hvUg9PogGM|
|Others  |It's weird that you can remember dreams, that make no sense. But for some reason it will stick with you.|
|Fashion & Style|@Phil_Lewis_Next up, underwear with artificial skid marks.|
|Others  |@clairemacckyd Are you referring to the dinked middle? Do you have 7" "spider" adapter? Or did you mean something e.. https://t.co/2bCHaynfsF|
|Fashion & Style|If u put actual clothes on & do your makeup everyday before class, what r u doing?!?!?|
|Others  |@Haylie_Ewald3 so jealous 🙄|
|Others  |So tired of everybody & everything.|
|World   |China Launches Aircraft Carrier, Boosting Military Presence|
|Others  |motherfucking sucks.|
|Television|these men really are working hard for their spinoff. https://t.co/0cHaw10UEt|
|Health  |Hospital againnnnn 🏥|
+-----+
```

```
only showing top 20 rows
```

NewsPerceptronAPITr1 OnlinePerceptron	NewsPerceptronAPITr10	NewsPerceptronOPITr10
Sports Fashion & Style Health Travel Others	Sports Fashion & Style Health Travel	Sports Fashion & Style Health Travel
Fashion & Style Movies Health Travel Others	Fashion & Style Movies Health Travel	Fashion & Style Movies Health Travel
Sports Fashion & Style Technology Science Health Education Art & Design Sports Health Education Others	Sports Fashion & Style Technology Science Health Education Art & Design Sports Health Education	Sports Fashion & Style Technology Science Health Education Art & Design Sports Health Education
Music Movies Real Estate Travel Others	Music Movies Real Estate Travel	Music Movies Real Estate Travel
World Health Others	World Health	World Health
Art & Design Fashion & Style Food Health Travel Your Money Economy Others	Art & Design Fashion & Style Food Health Travel Your Money Economy	Art & Design Fashion & Style Food Health Travel Your Money Economy
World Fashion & Style Travel Others	World Fashion & Style Travel	World Fashion & Style Travel
Sports Fashion & Style Music Others	Sports Fashion & Style Music	Sports Fashion & Style Music
World Health Others	World Health	World Health
Sports Science Health Others	Sports Science Health	Sports Science Health
Fashion & Style Music Real Estate Others	Fashion & Style Music Real Estate	Fashion & Style Music Real Estate
Science Education Others	Science Education	Science Education
Technology Health Education Your Money Others	Technology Health Education Your Money	Technology Health Education Your Money
Fashion & Style Education Others	Fashion & Style Education	Fashion & Style Education
Fashion & Style Music Science Health Others	Fashion & Style Music Science Health	Fashion & Style Music Science Health
World Others	World	World
Others Others	Others	Others
World Sports Fashion & Style Health Your Money Economy Others	World Sports Fashion & Style Health Your Money Economy	World Sports Fashion & Style Health Your Money Economy
World Sports Health Others	World Sports Health	World Sports Health

only showing top 20 rows

Training 4 MulticlassPerceptron models using 78 training data.....

Predicting labels of test Tweets after trained 78 Category Feedback Tweets:

NewsPerceptronAPitr1	NewsPerceptronAPitr10	NewsPerceptronOPitr10	OnlinePerceptron
Others	Others	Others	Others
Fashion & Style Movies Travel	Fashion & Style Movies Food Travel	Fashion & Style Movies Food Travel	Others
Technology Science	Science Food	Science Food	Others
Sports	Others	Others	Others
Music Movies Travel	Music Movies Travel Economy	Music Movies Travel Economy	Others
World	World	World	Others
Art & Design Fashion & Style Food Travel Economy	Art & Design Fashion & Style Food Travel Economy	Art & Design Fashion & Style Food Travel Economy	Others
World Fashion & Style	World Fashion & Style Travel	World Fashion & Style	Others
Fashion & Style	Fashion & Style	Fashion & Style	Others
World	World	World	Others
Sports	Sports	Others	Others
Music	Music	Music	Others
Education	Food Education	Food Education	Others
Technology Education	Technology	Technology	Others
Fashion & Style	Fashion & Style	Fashion & Style	Others
Music	Music	Music	Others
World	World	World	Others
Others	Others	Others	Others
World Sports Fashion & Style Economy	World Fashion & Style Economy	World Fashion & Style Economy	Others
World	World	World	Others

only showing top 20 rows

Training 4 MulticlassPerceptron models using 75 training data.....

Training 4 MulticlassPerceptron models using 75 training data.....

Predicting labels of test Tweets after trained 75 Category Feedback Tweets:

NewsPerceptronAPitr1	NewsPerceptronAPitr10	NewsPerceptronOPitr10	OnlinePerceptron
Others	Others	Politics	Others
Fashion & Style Travel	Fashion & Style	Others	Others
Science	Food	Food	Others
Sports Health	Others	Others	Others
Music Movies Travel	Music Movies Travel Economy	Movies	Others
World	World	World	Others
Food Travel Economy	Economy	Others	Others
World Fashion & Style Travel	World Fashion & Style Travel	World	Others
Fashion & Style	Fashion & Style	Fashion & Style	Others
World	World	World	Others
Others	Others	Others	Others
Music	Music	Music	Others
Education	Education	Education	Others
Technology	Technology	Fashion & Style	Others
Fashion & Style	Fashion & Style	Others	Others
Music	Music	Others	Others
World	World	World	Others
Others	Others	Others	Others
World Sports Fashion & Style Economy	World Fashion & Style Economy	World Fashion & Style Economy	Others
World	Others	Health	Others

only showing top 20 rows

Training 4 MulticlassPerceptron models using 79 training data.....

Training 4 MulticlassPerceptron models using 79 training data.....

Predicting labels of test Tweets after trained 79 Category Feedback Tweets:

NewsPerceptronAPItr1	NewsPerceptronAPItr10	NewsPerceptronOPITr10	OnlinePerceptron
Health	Others	Politics	Others
Fashion & Style Movies Travel	Fashion & Style Travel	Others	Others
Sports Technology Science	Science	Technology	Others
Sports Health Education	Others	Others	Others
Music Movies Travel	Music Movies Travel Economy	Music Movies Travel	Others
World	World	World	Others
Art & Design Fashion & Style Food Health Travel Your Money Economy	Art & Design Fashion & Style Food Travel Economy	Others	Others
World Fashion & Style Travel	World Fashion & Style Travel	World Travel	Others
Fashion & Style Music	Fashion & Style	Fashion & Style	Others
World Health	World	World	Others
Sports Health	Others	Others	Others
Music	Music	Fashion & Style	Others
Science Education	Education	Others	Others
Technology Education Your Money	Technology	Others	Others
Fashion & Style	Fashion & Style	Others	Others
Music	Music	Others	Others
World	World	World	Others
Others	Others	Others	Others
World Sports Fashion & Style Health Your Money Economy	World Fashion & Style Economy	Fashion & Style Economy	Others
World Sports	Others	Health	Others

only showing top 20 rows

Training 4 MulticlassPerceptron models using 82 training data.....

Training 4 MulticlassPerceptron models using 82 training data.....

Predicting labels of test Tweets after trained 82 Category Feedback Tweets:

NewsPerceptronAPItr1	NewsPerceptronAPItr10	NewsPerceptronOPITr10	OnlinePerceptron
Sports Fashion & Style Real Estate Health Your Money	Others	Others	Others
World Fashion & Style Movies Health Travel	Fashion & Style Movies Food Travel	Others	Others
Sports Fashion & Style Technology Science	Science Food	Technology Food	Others
Sports Health Education	Sports	Others	Others
Music Movies Travel Economy	Movies Economy	Your Money	Others
World	World	Others	Others
Art & Design Sports Fashion & Style Music Food Health Travel Your Money Economy	Art & Design Fashion & Style Food Travel Economy	Others	Others
World Fashion & Style	World Fashion & Style	Politics	Others
Sports Fashion & Style Music Your Money	Fashion & Style	Others	Others
World Health	World	Travel	Others
Sports Health	Sports	Others	Others
Fashion & Style Music	Music	Fashion & Style	Others
Technology Science Education	Food Education	Others	Others
Technology Health Education Your Money	Technology	Fashion & Style	Others
Sports Fashion & Style Technology Education	Fashion & Style	Others	Others
Fashion & Style Music Technology Your Money	Music	Others	Others
World	World	World	Others
Sports	Others	Others	Others
World Sports Fashion & Style Music Health Your Money Economy	World Fashion & Style Economy	Television	Others
World Sports Health	World	Health	Others

only showing top 20 rows

Yilans-MacBook-Air:spark-2.0.1-bin-hadoop2.7 Jillian\$