91.542 - Natural Language Processing
Homework 4 - Write Up

Chang Liu
chang_liu@student.uml.edu

October 31, 2015

1   Summary

This write up concludes the necessary result for all the genres for my experiment. Firstly, I tried some features that I think will be enough, then I used other features and calculated the new result(precision, recall, F-measure etc...), for some technical details, please check the code.

2   Initial Features

My initial feature use the **"first word", "first letter", "last word" and "last letter"** to represent the features, it's quite simple, but for texts, we can see that usually the word has many useful information that will helps to identify the genre. If these documents are in a specific format, for example, the first word or last word contains specific category, then it's quite obvious and effective.

However, as far as I can see, maybe it's not so closely with the genres. We have to get another features that can have stronger representation ability for classifying the documents. Increasing the feature dimension should work(each dimension has new property), or counting the frequency of some words and find the most frequent one. Moreover, we can actually do some more complex work, for example, using Convolutional Neural Network to learn a good feature that can represent it better, because here we just use hand-crafted feature that contains our human domain knowledge, but maybe it's not so powerful to represent all the data.

3   Initial result On Dev Set

For category: Action
Positive, precision = 0.0, recall = 0.0, f-score = None Negative, precision = 0.857142857143, recall = 0.461538461538, f-score = 0.6 Total Accuracy = 30.00%, Error rate = 70.00%

For category: Adventure
Positive, precision = 0.142857142857, recall = 0.333333333333, f-score = 0.2
Negative, precision = 0.714285714286, recall = 0.454545454545, f-score = 0.555555555556
Total Accuracy = 33.33%, Error rate = 66.67%
For category: Comedy
Positive, precision = 0.142857142857, recall = 0.333333333333, f-score = 0.2
Negative, precision = 0.714285714286, recall = 0.454545454545, f-score = 0.555555555556
Total Accuracy = 33.33%, Error rate = 66.67%
For category: Drama
Positive, precision = 0.285714285714, recall = 0.5, f-score = 0.363636363636
Negative, precision = 0.714285714286, recall = 0.5, f-score = 0.588235294118
Total Accuracy = 41.18%, Error rate = 58.82%
For category: Fantasy
Positive, precision = 0.142857142857, recall = 0.333333333333, f-score = 0.2
Negative, precision = 0.714285714286, recall = 0.454545454545, f-score = 0.555555555556
Total Accuracy = 33.33%, Error rate = 66.67%
For category: Mystery Positive, precision = 0.285714285714, recall = 0.5,
f-score = 0.363636363636 Negative, precision = 0.714285714286, recall = 0.5,
f-score = 0.588235294118 Total Accuracy = 41.18%, Error rate = 58.82%
For category: Romance
Positive, precision = 0.571428571429, recall = 0.4, f-score = 0.470588235294
Negative, precision = 0.142857142857, recall = 0.25, f-score = 0.181818181818
Total Accuracy = 45.45%, Error rate = 54.55%
For category: Sci-Fi
Positive, precision = 0.857142857143, recall = 0.461538461538, f-score = 0.6
Negative, precision = 0.0, recall = 0.0, f-score = None Total Accuracy = 75.00%,
Error rate = 25.00%
For category: Thriller
Positive, precision = 0.428571428571, recall = 0.428571428571, f-score = 0.428571428571
Negative, precision = 0.428571428571, recall = 0.428571428571, f-score = 0.428571428571
Total Accuracy = 42.86%, Error rate = 57.14%
For category: Western
Positive, precision = 0.857142857143, recall = 0.75, f-score = 0.8 Negative,
precision = 0.714285714286, recall = 0.833333333333, f-score = 0.769230769231
Total Accuracy = 84.62%, Error rate = 15.38%

4 Final list of feature

After some experiment, I want to improve the dimension of the initial features,
and the other dimension should contain information that be more close to genres,
so I choose the **"most frequent word", "most frequent word's count",
"middle word"**, I think the most frequent word should has some information
that contains about genre. Maybe some information that about violence is
about the "Action" genre, it looks reasonable. Other features I can think of is

that the **length of whole text**, for some genre maybe it is short, like some romantic genre, so this information also should helps the representation.

5   Final result on Dev Set

For category: Action
Positive, precision = 0.285714285714, recall = 0.285714285714, f-score = 0.285714285714
Negative, precision = 0.285714285714, recall = 0.285714285714, f-score = 0.285714285714
Total Accuracy = 28.57%, Error rate = 71.43%
    For category: Adventure
Positive, precision = 0.285714285714, recall = 0.333333333333, f-score = 0.307692307692
Negative, precision = 0.428571428571, recall = 0.375, f-score = 0.4 Total Accuracy = 33.33%, Error rate = 66.67%
    For category: Comedy
Positive, precision = 0.714285714286, recall = 0.555555555556, f-score = 0.625
Negative, precision = 0.428571428571, recall = 0.6, f-score = 0.5 Total Accuracy = 66.67%, Error rate = 33.33%
    For category: Drama
Positive, precision = 0.571428571429, recall = 0.666666666667, f-score = 0.615384615385
Negative, precision = 0.714285714286, recall = 0.625, f-score = 0.666666666667
Total Accuracy = 60.00%, Error rate = 40.00%
    For category: Fantasy
Positive, precision = 0.571428571429, recall = 0.444444444444, f-score = 0.5
Negative, precision = 0.285714285714, recall = 0.4, f-score = 0.333333333333
Total Accuracy = 50.00%, Error rate = 50.00%
    For category: Mystery
Positive, precision = 0.857142857143, recall = 0.666666666667, f-score = 0.75
Negative, precision = 0.571428571429, recall = 0.8, f-score = 0.666666666667
Total Accuracy = 83.33%, Error rate = 16.67%
    For category: Romance
Positive, precision = 0.571428571429, recall = 0.5, f-score = 0.533333333333
Negative, precision = 0.428571428571, recall = 0.5, f-score = 0.461538461538
Total Accuracy = 53.85%, Error rate = 46.15%
    For category: Sci-Fi
Positive, precision = 0.571428571429, recall = 0.5, f-score = 0.533333333333
Negative, precision = 0.428571428571, recall = 0.5, f-score = 0.461538461538
Total Accuracy = 53.85%, Error rate = 46.15%
    For category: Thriller
Positive, precision = 0.285714285714, recall = 0.333333333333, f-score = 0.307692307692
Negative, precision = 0.428571428571, recall = 0.375, f-score = 0.4 Total Accuracy = 33.33%, Error rate = 66.67%
    For category: Western
Positive, precision = 0.571428571429, recall = 0.571428571429, f-score = 0.571428571429
Negative, precision = 0.571428571429, recall = 0.571428571429, f-score = 0.571428571429
Total Accuracy = 57.14%, Error rate = 42.86%

6   Final result on Test Set

For category: Action Positive, precision = 0.8, recall = 0.615384615385, f-score = 0.695652173913 Negative, precision = 0.5, recall = 0.714285714286, f-score = 0.588235294118 Total Accuracy = 76.47%, Error rate = 23.53%

For category: Adventure Positive, precision = 0.3, recall = 0.375, f-score = 0.333333333333 Negative, precision = 0.5, recall = 0.416666666667, f-score = 0.454545454545 Total Accuracy = 36.36%, Error rate = 63.64%

For category: Comedy Positive, precision = 0.5, recall = 0.5, f-score = 0.5 Negative, precision = 0.5, recall = 0.5, f-score = 0.5 Total Accuracy = 50.00%, Error rate = 50.00%

For category: Drama Positive, precision = 0.2, recall = 0.285714285714, f-score = 0.235294117647 Negative, precision = 0.5, recall = 0.384615384615, f-score = 0.434782608696 Total Accuracy = 30.43%, Error rate = 69.57%

For category: Fantasy Positive, precision = 0.4, recall = 0.5, f-score = 0.444444444444 Negative, precision = 0.6, recall = 0.5, f-score = 0.545454545455 Total Accuracy = 45.45%, Error rate = 54.55%

For category: Mystery Positive, precision = 0.5, recall = 0.454545454545, f-score = 0.47619047619 Negative, precision = 0.4, recall = 0.444444444444, f-score = 0.421052631579 Total Accuracy = 47.37%, Error rate = 52.63%

For category: Romance Positive, precision = 0.6, recall = 0.5, f-score = 0.545454545455 Negative, precision = 0.4, recall = 0.5, f-score = 0.444444444444 Total Accuracy = 55.56%, Error rate = 44.44%

For category: Sci-Fi Positive, precision = 0.5, recall = 0.555555555556, f-score = 0.526315789474 Negative, precision = 0.6, recall = 0.545454545455, f-score = 0.571428571429 Total Accuracy = 52.38%, Error rate = 47.62%

For category: Thriller Positive, precision = 0.6, recall = 0.461538461538, f-score = 0.521739130435 Negative, precision = 0.3, recall = 0.428571428571, f-score = 0.352941176471 Total Accuracy = 52.94%, Error rate = 47.06%

For category: Western Positive, precision = 0.5, recall = 0.555555555556, f-score = 0.526315789474 Negative, precision = 0.6, recall = 0.545454545455, f-score = 0.571428571429 Total Accuracy = 52.38%, Error rate = 47.62%

7   Appendix I: Initial Feature Output

```
Now printing all the result:
NOTE: When F−score is None, then precision and recall both are zero for positive

predict:Not_Action, actual:Action
predict:Not_Action, actual:Action
predict:Not_Action, actual:Action
predict:Not_Action, actual:Action
predict:Not_Action, actual:Action
predict:Not_Action, actual:Action
```

```
predict:Not_Action,  actual:Action
predict:Not_Action,  actual:Not_Action
predict:Not_Action,  actual:Not_Action
predict:Not_Action,  actual:Not_Action
predict:Not_Action,  actual:Not_Action
predict:Action,  actual:Not_Action
predict:Not_Action,  actual:Not_Action
predict:Not_Action,  actual:Not_Action
true_positive = 0,  true_negative = 6
false_positive = 7,  false_negative = 1
For category:  Action
Positive,  precision = 0.0,  recall = 0.0,  f-score = None
Negative,  precision = 0.857142857143,  recall = 0.461538461538,  f-score = 0.6
Total Accuracy = 30.00%,  Error rate = 70.00%

predict:Adventure,  actual:Adventure
predict:Not_Adventure,  actual:Adventure
predict:Not_Adventure,  actual:Adventure
predict:Not_Adventure,  actual:Adventure
predict:Not_Adventure,  actual:Adventure
predict:Not_Adventure,  actual:Adventure
predict:Not_Adventure,  actual:Adventure
predict:Adventure,  actual:Not_Adventure
predict:Not_Adventure,  actual:Not_Adventure
predict:Not_Adventure,  actual:Not_Adventure
predict:Not_Adventure,  actual:Not_Adventure
predict:Adventure,  actual:Not_Adventure
predict:Not_Adventure,  actual:Not_Adventure
predict:Not_Adventure,  actual:Not_Adventure
true_positive = 1,  true_negative = 5
false_positive = 6,  false_negative = 2
For category:  Adventure
Positive,  precision = 0.142857142857,  recall = 0.333333333333,  f-score = 0.2
Negative,  precision = 0.714285714286,  recall = 0.454545454545,  f-score = 0.55555
Total Accuracy = 33.33%,  Error rate = 66.67%

predict:Not_Comedy,  actual:Comedy
predict:Not_Comedy,  actual:Comedy
predict:Comedy,  actual:Comedy
predict:Not_Comedy,  actual:Comedy
predict:Not_Comedy,  actual:Comedy
predict:Not_Comedy,  actual:Comedy
predict:Not_Comedy,  actual:Comedy
predict:Not_Comedy,  actual:Not_Comedy
predict:Not_Comedy,  actual:Not_Comedy
predict:Not_Comedy,  actual:Not_Comedy
```

```
predict:Not_Comedy, actual:Not_Comedy
predict:Not_Comedy, actual:Not_Comedy
predict:Comedy, actual:Not_Comedy
predict:Comedy, actual:Not_Comedy
true_positive = 1, true_negative = 5
false_positive = 6, false_negative = 2
For category: Comedy
Positive, precision = 0.142857142857, recall = 0.333333333333, f-score = 0.2
Negative, precision = 0.714285714286, recall = 0.454545454545, f-score = 0.55555
Total Accuracy = 33.33%, Error rate = 66.67%

predict:Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Drama, actual:Not_Drama
true_positive = 2, true_negative = 5
false_positive = 5, false_negative = 2
For category: Drama
Positive, precision = 0.285714285714, recall = 0.5, f-score = 0.363636363636
Negative, precision = 0.714285714286, recall = 0.5, f-score = 0.588235294118
Total Accuracy = 41.18%, Error rate = 58.82%

predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
```

true_positive = 1, true_negative = 5
false_positive = 6, false_negative = 2
For category: Fantasy
Positive, precision = 0.142857142857, recall = 0.333333333333, f−score = 0.2
Negative, precision = 0.714285714286, recall = 0.454545454545, f−score = 0.55555
Total Accuracy = 33.33%, Error rate = 66.67%

predict:Not_Mystery, actual:Mystery
predict:Mystery, actual:Mystery
predict:Not_Mystery, actual:Mystery
predict:Not_Mystery, actual:Mystery
predict:Mystery, actual:Mystery
predict:Not_Mystery, actual:Mystery
predict:Not_Mystery, actual:Mystery
predict:Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
predict:Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
true_positive = 2, true_negative = 5
false_positive = 5, false_negative = 2
For category: Mystery
Positive, precision = 0.285714285714, recall = 0.5, f−score = 0.363636363636
Negative, precision = 0.714285714286, recall = 0.5, f−score = 0.588235294118
Total Accuracy = 41.18%, Error rate = 58.82%

predict:Romance, actual:Romance
predict:Not_Romance, actual:Romance
predict:Romance, actual:Romance
predict:Romance, actual:Romance
predict:Not_Romance, actual:Romance
predict:Not_Romance, actual:Romance
predict:Romance, actual:Romance
predict:Not_Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
true_positive = 4, true_negative = 1
false_positive = 3, false_negative = 6
For category: Romance
Positive, precision = 0.571428571429, recall = 0.4, f−score = 0.470588235294

Negative, precision = 0.142857142857, recall = 0.25, f−score = 0.181818181818
Total Accuracy = 45.45%, Error rate = 54.55%

predict:Sci−Fi, actual:Sci−Fi
predict:Sci−Fi, actual:Sci−Fi
predict:Not_Sci−Fi, actual:Sci−Fi
predict:Sci−Fi, actual:Sci−Fi
predict:Sci−Fi, actual:Sci−Fi
predict:Sci−Fi, actual:Sci−Fi
predict:Sci−Fi, actual:Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
true_positive = 6, true_negative = 0
false_positive = 1, false_negative = 7
For category: Sci−Fi
Positive, precision = 0.857142857143, recall = 0.461538461538, f−score = 0.6
Negative, precision = 0.0, recall = 0.0, f−score = None
Total Accuracy = 75.00%, Error rate = 25.00%

predict:Not_Thriller, actual:Thriller
predict:Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Thriller, actual:Thriller
predict:Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
true_positive = 3, true_negative = 3
false_positive = 4, false_negative = 4
For category: Thriller
Positive, precision = 0.428571428571, recall = 0.428571428571, f−score = 0.42857
Negative, precision = 0.428571428571, recall = 0.428571428571, f−score = 0.42857
Total Accuracy = 42.86%, Error rate = 57.14%

predict:Western, actual:Western

```
predict : Western ,  actual : Western
predict : Western ,  actual : Western
predict : Western ,  actual : Western
predict : Western ,  actual : Western
predict : Western ,  actual : Western
predict : Not_Western ,  actual : Western
predict : Not_Western ,  actual : Not_Western
predict : Not_Western ,  actual : Not_Western
predict : Western ,  actual : Not_Western
predict : Not_Western ,  actual : Not_Western
predict : Western ,  actual : Not_Western
predict : Not_Western ,  actual : Not_Western
predict : Not_Western ,  actual : Not_Western
true_positive = 6,  true_negative = 5
false_positive = 1,  false_negative = 2
For category : Western
Positive ,  precision = 0.857142857143 ,  recall = 0.75 ,  f−score = 0.8
Negative ,  precision = 0.714285714286 ,  recall = 0.833333333333 ,  f−score = 0.76923
Total Accuracy = 84.62% ,  Error rate = 15.38%
```

8   Appendix II: Final Feature Output

```
*************************** Start of Dev Set ************************************
Now printing all the result for Dev set :
NOTE: When F−score is None , then precision and recall both are zero for positive

predict : Not_Action ,  actual : Action
predict : Action ,  actual : Action
predict : Not_Action ,  actual : Action
predict : Not_Action ,  actual : Action
predict : Not_Action ,  actual : Action
predict : Action ,  actual : Action
predict : Not_Action ,  actual : Action
predict : Action ,  actual : Not_Action
predict : Not_Action ,  actual : Not_Action
predict : Action ,  actual : Not_Action
predict : Action ,  actual : Not_Action
predict : Action ,  actual : Not_Action
predict : Action ,  actual : Not_Action
predict : Not_Action ,  actual : Not_Action
true_positive = 2,  true_negative = 2
false_positive = 5,  false_negative = 5
For category : Action
Positive ,  precision = 0.285714285714 ,  recall = 0.285714285714 ,  f−score = 0.28571
Negative ,  precision = 0.285714285714 ,  recall = 0.285714285714 ,  f−score = 0.28571
Total Accuracy = 28.57% ,  Error rate = 71.43%
```

```
predict : Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Adventure ,  actual : Not_Adventure
predict : Adventure ,  actual : Not_Adventure
predict : Not_Adventure ,  actual : Not_Adventure
predict : Adventure ,  actual : Not_Adventure
predict : Not_Adventure ,  actual : Not_Adventure
predict : Not_Adventure ,  actual : Not_Adventure
predict : Adventure ,  actual : Not_Adventure
true_positive = 2,  true_negative = 3
false_positive = 5,  false_negative = 4
For  category :  Adventure
Positive ,  precision = 0.285714285714 ,  recall = 0.333333333333 ,  f−score = 0.30769
Negative ,  precision = 0.428571428571 ,  recall = 0.375 ,  f−score = 0.4
Total  Accuracy = 33.33% ,  Error  rate = 66.67%

predict : Not_Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Not_Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Not_Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Not_Comedy
predict : Comedy ,  actual : Not_Comedy
true_positive = 5,  true_negative = 3
false_positive = 2,  false_negative = 4
For  category :  Comedy
Positive ,  precision = 0.714285714286 ,  recall = 0.555555555556 ,  f−score = 0.625
Negative ,  precision = 0.428571428571 ,  recall = 0.6 ,  f−score = 0.5
Total  Accuracy = 66.67% ,  Error  rate = 33.33%

predict : Not_Drama ,  actual : Drama
predict : Drama ,  actual : Drama
predict : Drama ,  actual : Drama
```

```
predict:Drama, actual:Drama
predict:Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
true_positive = 4, true_negative = 5
false_positive = 3, false_negative = 2
For category: Drama
Positive, precision = 0.571428571429, recall = 0.666666666667, f-score = 0.61538
Negative, precision = 0.714285714286, recall = 0.625, f-score = 0.666666666667
Total Accuracy = 60.00%, Error rate = 40.00%

predict:Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Fantasy, actual:Fantasy
predict:Fantasy, actual:Fantasy
predict:Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
true_positive = 4, true_negative = 2
false_positive = 3, false_negative = 5
For category: Fantasy
Positive, precision = 0.571428571429, recall = 0.444444444444, f-score = 0.5
Negative, precision = 0.285714285714, recall = 0.4, f-score = 0.333333333333
Total Accuracy = 50.00%, Error rate = 50.00%

predict:Mystery, actual:Mystery
predict:Mystery, actual:Mystery
predict:Mystery, actual:Mystery
predict:Mystery, actual:Mystery
predict:Mystery, actual:Mystery
predict:Mystery, actual:Mystery
predict:Not_Mystery, actual:Mystery
```

```
predict:Not_Mystery, actual:Not_Mystery
predict:Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
predict:Mystery, actual:Not_Mystery
predict:Mystery, actual:Not_Mystery
predict:Not_Mystery, actual:Not_Mystery
true_positive = 6, true_negative = 4
false_positive = 1, false_negative = 3
For category: Mystery
Positive, precision = 0.857142857143, recall = 0.666666666667, f-score = 0.75
Negative, precision = 0.571428571429, recall = 0.8, f-score = 0.666666666667
Total Accuracy = 83.33%, Error rate = 16.67%

predict:Romance, actual:Romance
predict:Not_Romance, actual:Romance
predict:Romance, actual:Romance
predict:Romance, actual:Romance
predict:Not_Romance, actual:Romance
predict:Romance, actual:Romance
predict:Not_Romance, actual:Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Not_Romance, actual:Not_Romance
predict:Not_Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Not_Romance, actual:Not_Romance
true_positive = 4, true_negative = 3
false_positive = 3, false_negative = 4
For category: Romance
Positive, precision = 0.571428571429, recall = 0.5, f-score = 0.533333333333
Negative, precision = 0.428571428571, recall = 0.5, f-score = 0.461538461538
Total Accuracy = 53.85%, Error rate = 46.15%

predict:Not_Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
predict:Sci-Fi, actual:Not_Sci-Fi
predict:Sci-Fi, actual:Not_Sci-Fi
```

```
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Sci−Fi, actual:Not_Sci−Fi
predict:Not_Sci−Fi, actual:Not_Sci−Fi
true_positive = 4, true_negative = 3
false_positive = 3, false_negative = 4
For category: Sci−Fi
Positive, precision = 0.571428571429, recall = 0.5, f−score = 0.533333333333
Negative, precision = 0.428571428571, recall = 0.5, f−score = 0.461538461538
Total Accuracy = 53.85%, Error rate = 46.15%

predict:Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Thriller, actual:Thriller
predict:Not_Thriller, actual:Thriller
predict:Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
true_positive = 2, true_negative = 3
false_positive = 5, false_negative = 4
For category: Thriller
Positive, precision = 0.285714285714, recall = 0.333333333333, f−score = 0.30769
Negative, precision = 0.428571428571, recall = 0.375, f−score = 0.4
Total Accuracy = 33.33%, Error rate = 66.67%

predict:Western, actual:Western
predict:Western, actual:Western
predict:Western, actual:Western
predict:Not_Western, actual:Western
predict:Not_Western, actual:Western
predict:Western, actual:Western
predict:Not_Western, actual:Western
predict:Not_Western, actual:Not_Western
predict:Not_Western, actual:Not_Western
predict:Western, actual:Not_Western
predict:Western, actual:Not_Western
predict:Not_Western, actual:Not_Western
predict:Western, actual:Not_Western
predict:Not_Western, actual:Not_Western
true_positive = 4, true_negative = 4
```

false_positive = 3, false_negative = 3
For category: Western
Positive, precision = 0.571428571429, recall = 0.571428571429, f−score = 0.57142
Negative, precision = 0.571428571429, recall = 0.571428571429, f−score = 0.57142
Total Accuracy = 57.14%, Error rate = 42.86%

*************************** End of Dev Set ***********************************
*************************** Start of Test Set ********************************
Now printing **all** the result **for** Test **set**:
NOTE: When F−score **is** None, then precision **and** recall both are zero **for** positive

predict:Action, actual:Action
predict:Action, actual:Not_Action
predict:Action, actual:Not_Action
predict:Not_Action, actual:Not_Action
predict:Not_Action, actual:Not_Action
predict:Action, actual:Not_Action
predict:Not_Action, actual:Action
predict:Action, actual:Action
predict:Action, actual:Not_Action
predict:Action, actual:Action
predict:Not_Action, actual:Not_Action
predict:Action, actual:Action
predict:Action, actual:Action
predict:Action, actual:Action
predict:Action, actual:Action
predict:Action, actual:Action
predict:Action, actual:Not_Action
predict:Not_Action, actual:Not_Action
predict:Not_Action, actual:Action
predict:Not_Action, actual:Not_Action
true_positive = 8, true_negative = 5
false_positive = 2, false_negative = 5
For category: Action
Positive, precision = 0.8, recall = 0.615384615385, f−score = 0.695652173913
Negative, precision = 0.5, recall = 0.714285714286, f−score = 0.588235294118
Total Accuracy = 76.47%, Error rate = 23.53%

predict:Not_Adventure, actual:Adventure
predict:Not_Adventure, actual:Not_Adventure
predict:Not_Adventure, actual:Adventure
predict:Adventure, actual:Not_Adventure
predict:Not_Adventure, actual:Adventure
predict:Adventure, actual:Adventure
predict:Adventure, actual:Not_Adventure
predict:Adventure, actual:Not_Adventure

```
predict : Not_Adventure ,  actual : Not_Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Adventure ,  actual : Not_Adventure
predict : Not_Adventure ,  actual : Not_Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Not_Adventure
predict : Adventure ,  actual : Not_Adventure
predict : Not_Adventure ,  actual : Adventure
predict : Adventure ,  actual : Adventure
predict : Not_Adventure ,  actual : Not_Adventure
true_positive = 3 ,  true_negative = 5
false_positive = 7 ,  false_negative = 5
For category :  Adventure
Positive ,  precision = 0.3 ,  recall = 0.375 ,  f−score = 0.333333333333
Negative ,  precision = 0.5 ,  recall = 0.416666666667 ,  f−score = 0.454545454545
Total Accuracy = 36.36% ,  Error rate = 63.64%

predict : Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Comedy
predict : Not_Comedy ,  actual : Not_Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Not_Comedy
predict : Comedy ,  actual : Comedy
predict : Not_Comedy ,  actual : Comedy
predict : Not_Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Comedy
predict : Not_Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Not_Comedy ,  actual : Comedy
predict : Comedy ,  actual : Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Not_Comedy
predict : Not_Comedy ,  actual : Not_Comedy
true_positive = 5 ,  true_negative = 5
false_positive = 5 ,  false_negative = 5
For category :  Comedy
Positive ,  precision = 0.5 ,  recall = 0.5 ,  f−score = 0.5
Negative ,  precision = 0.5 ,  recall = 0.5 ,  f−score = 0.5
Total Accuracy = 50.00% ,  Error rate = 50.00%
```

predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Drama
predict:Drama, actual:Not_Drama
predict:Not_Drama, actual:Drama
predict:Drama, actual:Not_Drama
predict:Drama, actual:Not_Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Drama
predict:Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Drama, actual:Not_Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Not_Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Drama
predict:Not_Drama, actual:Not_Drama
predict:Drama, actual:Not_Drama
predict:Drama, actual:Drama
true_positive = 2, true_negative = 5
false_positive = 8, false_negative = 5
For category: Drama
Positive, precision = 0.2, recall = 0.285714285714, f−score = 0.235294117647
Negative, precision = 0.5, recall = 0.384615384615, f−score = 0.434782608696
Total Accuracy = 30.43%, Error rate = 69.57%

predict:Not_Fantasy, actual:Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Not_Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Fantasy
predict:Fantasy, actual:Not_Fantasy
predict:Not_Fantasy, actual:Fantasy

predict : Fantasy , actual : Fantasy
true_positive = 4 , true_negative = 6
false_positive = 6 , false_negative = 4
For category : Fantasy
Positive , precision = 0.4 , recall = 0.5 , f−score = 0.444444444444
Negative , precision = 0.6 , recall = 0.5 , f−score = 0.545454545455
Total Accuracy = 45.45% , Error rate = 54.55%

predict : Not_Mystery , actual : Not_Mystery
predict : Mystery , actual : Not_Mystery
predict : Not_Mystery , actual : Mystery
predict : Mystery , actual : Mystery
predict : Not_Mystery , actual : Mystery
predict : Mystery , actual : Not_Mystery
predict : Mystery , actual : Mystery
predict : Not_Mystery , actual : Mystery
predict : Mystery , actual : Mystery
predict : Not_Mystery , actual : Not_Mystery
predict : Mystery , actual : Mystery
predict : Not_Mystery , actual : Mystery
predict : Mystery , actual : Not_Mystery
predict : Not_Mystery , actual : Not_Mystery
predict : Not_Mystery , actual : Not_Mystery
predict : Mystery , actual : Not_Mystery
predict : Not_Mystery , actual : Mystery
predict : Mystery , actual : Not_Mystery
predict : Mystery , actual : Not_Mystery
predict : Mystery , actual : Mystery
true_positive = 5 , true_negative = 4
false_positive = 5 , false_negative = 6
For category : Mystery
Positive , precision = 0.5 , recall = 0.454545454545 , f−score = 0.47619047619
Negative , precision = 0.4 , recall = 0.444444444444 , f−score = 0.421052631579
Total Accuracy = 47.37% , Error rate = 52.63%

predict : Romance , actual : Not_Romance
predict : Not_Romance , actual : Romance
predict : Not_Romance , actual : Romance
predict : Not_Romance , actual : Not_Romance
predict : Romance , actual : Romance
predict : Romance , actual : Romance
predict : Romance , actual : Not_Romance
predict : Romance , actual : Not_Romance
predict : Not_Romance , actual : Romance
predict : Romance , actual : Not_Romance
predict : Not_Romance , actual : Not_Romance

```
predict:Romance, actual:Romance
predict:Not_Romance, actual:Romance
predict:Romance, actual:Romance
predict:Not_Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Not_Romance
predict:Romance, actual:Romance
predict:Romance, actual:Romance
predict:Not_Romance, actual:Not_Romance
true_positive = 6, true_negative = 4
false_positive = 4, false_negative = 6
For category: Romance
Positive, precision = 0.6, recall = 0.5, f-score = 0.545454545455
Negative, precision = 0.4, recall = 0.5, f-score = 0.444444444444
Total Accuracy = 55.56%, Error rate = 44.44%

predict:Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Not_Sci-Fi
predict:Not_Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
predict:Sci-Fi, actual:Not_Sci-Fi
predict:Not_Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Not_Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Sci-Fi, actual:Not_Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
predict:Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
predict:Not_Sci-Fi, actual:Not_Sci-Fi
true_positive = 5, true_negative = 6
false_positive = 5, false_negative = 4
For category: Sci-Fi
Positive, precision = 0.5, recall = 0.555555555556, f-score = 0.526315789474
Negative, precision = 0.6, recall = 0.545454545455, f-score = 0.571428571429
Total Accuracy = 52.38%, Error rate = 47.62%

predict:Not_Thriller, actual:Thriller
predict:Thriller, actual:Thriller
predict:Thriller, actual:Not_Thriller
```

predict:Not_Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Thriller
predict:Thriller, actual:Thriller
predict:Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Thriller
predict:Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
predict:Thriller, actual:Thriller
predict:Thriller, actual:Not_Thriller
predict:Thriller, actual:Thriller
predict:Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Not_Thriller, actual:Thriller
predict:Not_Thriller, actual:Not_Thriller
predict:Thriller, actual:Not_Thriller
predict:Thriller, actual:Thriller
predict:Thriller, actual:Thriller
true_positive = 6, true_negative = 3
false_positive = 4, false_negative = 7
For category: Thriller
Positive, precision = 0.6, recall = 0.461538461538, f-score = 0.521739130435
Negative, precision = 0.3, recall = 0.428571428571, f-score = 0.352941176471
Total Accuracy = 52.94%, Error rate = 47.06%

predict:Not_Western, actual:Not_Western
predict:Western, actual:Not_Western
predict:Not_Western, actual:Western
predict:Western, actual:Not_Western
predict:Not_Western, actual:Western
predict:Western, actual:Not_Western
predict:Western, actual:Western
predict:Not_Western, actual:Western
predict:Not_Western, actual:Western
predict:Western, actual:Western
predict:Not_Western, actual:Not_Western
predict:Not_Western, actual:Not_Western
predict:Western, actual:Western
predict:Western, actual:Western
predict:Not_Western, actual:Not_Western
predict:Not_Western, actual:Not_Western
predict:Western, actual:Not_Western
predict:Western, actual:Western
predict:Not_Western, actual:Not_Western
predict:Not_Western, actual:Western
true_positive = 5, true_negative = 6
false_positive = 5, false_negative = 4

For category: Western
Positive, precision = 0.5, recall = 0.555555555556, f-score = 0.526315789474
Negative, precision = 0.6, recall = 0.545454545455, f-score = 0.571428571429
Total Accuracy = 52.38%, Error rate = 47.62%
*************************** End of Test Set ************************************