

## WEEKLY REPORT 20/11/2014

### **List of Features:**

0-spos:: number of words in one review that map to strong positive in the dictionary

1-wpos

2-sneu

3-wneu

4-sneg

5-wneg

6-pos

7-neg

**8-negation**

9-len

10-day

11-vote

**12-pos (WordNet)**

**13-neg (WordNet)**

**14-character\_count**

15-uppercase\_count

16-lowercase\_count

**17-punctuation\_count**

18-alphabetic\_count

19-numeric\_count

**20-1 star (Bag of words)**

**21-2 star (Bag of words)**

**22-3 star (Bag of words)**

**23-4 star (Bag of words)**

**24-5 star (Bag of words)**

*Put table mutual\_info and PCC score here*

| # Features | mutual information | PCC           |
|------------|--------------------|---------------|
| 0          | 0.0927815181       | 0.0903658491  |
| 1          | 0.0943339902       | -0.0570241331 |
| 2          | 0.0503507516       | -0.1082753224 |
| 3          | 0.0647941896       | -0.136978806  |
| 4          | 0.09351242         | -0.279446049  |

|    |              |               |
|----|--------------|---------------|
| 5  | 0.0727823628 | -0.240633071  |
| 6  | 0.1358688632 | -0.0567549444 |
| 7  | 0.1457278019 | -0.2697715562 |
| 8  | 0.0757117812 | -0.2575723079 |
| 9  | 0.6385285802 | -0.153142069  |
| 10 | 0.0026563788 | -0.0016561843 |
| 11 | 0.0449040492 | -0.026564253  |
| 12 | 1.4597446389 | -0.0405194947 |
| 13 | 1.4621774296 | -0.2398235775 |
| 14 | 1.1786996041 | -0.145964132  |
| 15 | 0.1797516526 | -0.1263577924 |
| 16 | 1.1031838378 | -0.1429638135 |
| 17 | 0.2191398531 | -0.1341208976 |
| 18 | 1.1266395011 | -0.1433751799 |
| 19 | 0.060229426  | -0.1403685205 |
| 20 | 0.9917703132 | -0.0133283845 |
| 21 | 0.7933696034 | 0.1204474404  |
| 22 | 0.6799529815 | 0.1197528927  |
| 23 | 0.9904469334 | -0.1388707886 |
| 24 | 0.8339643766 | 0.1487493243  |

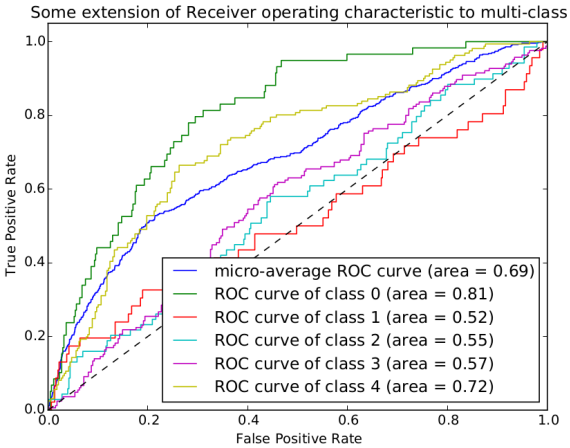
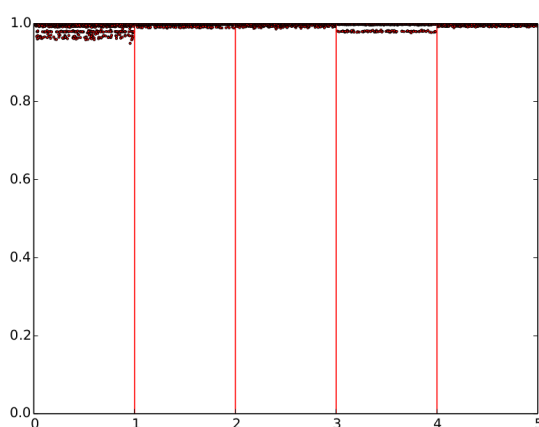
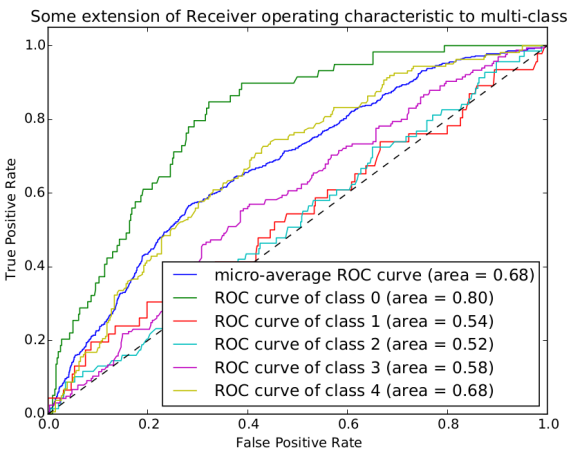
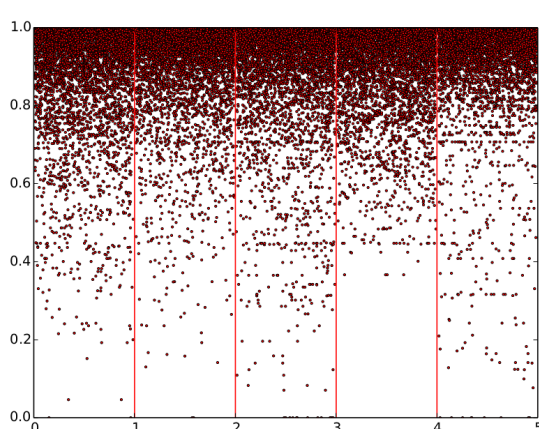
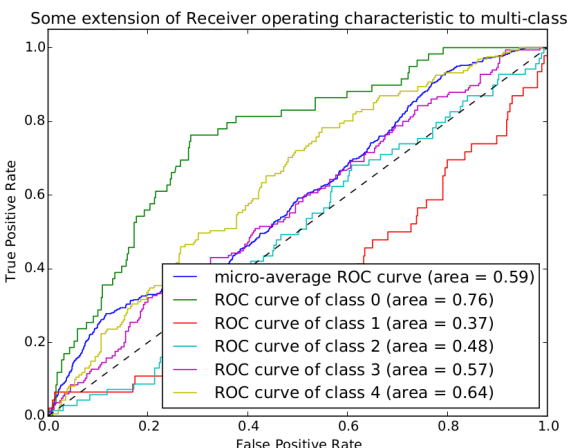
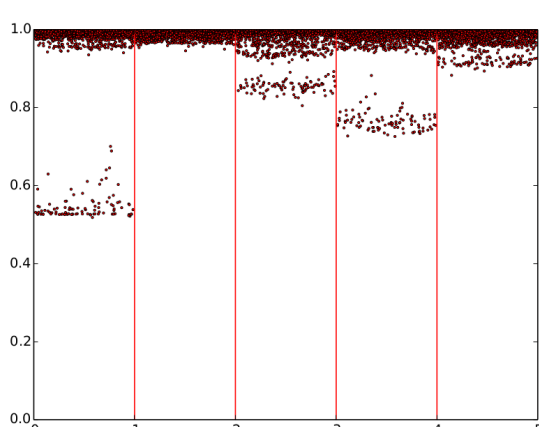
Put table  $r^2$  score here

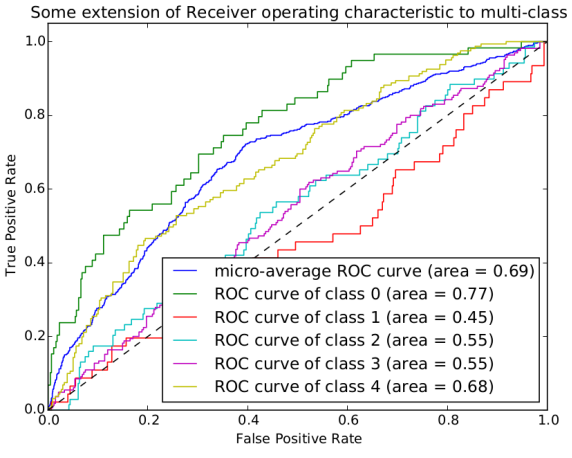
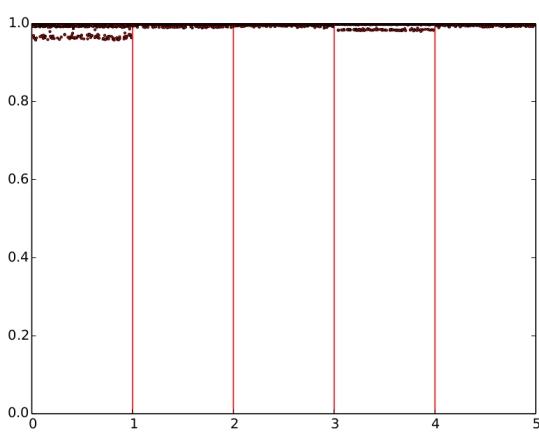
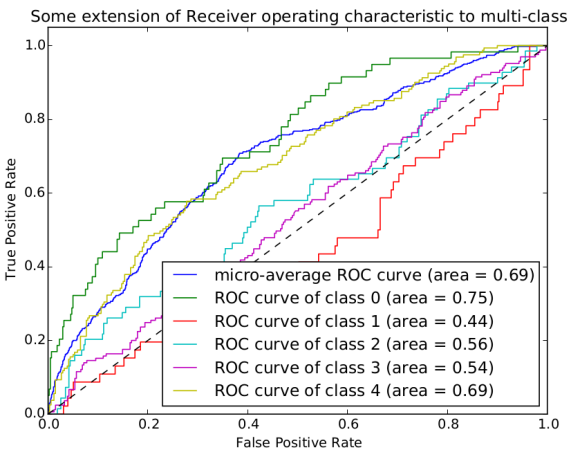
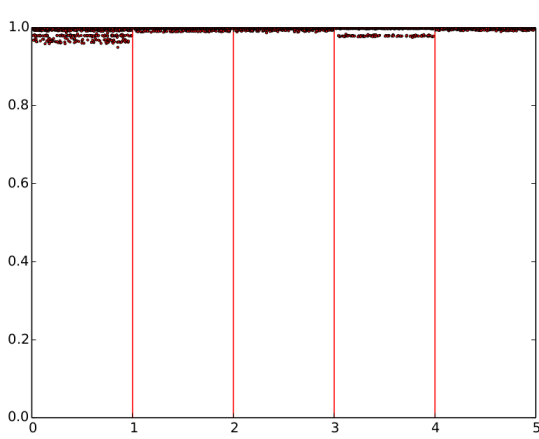
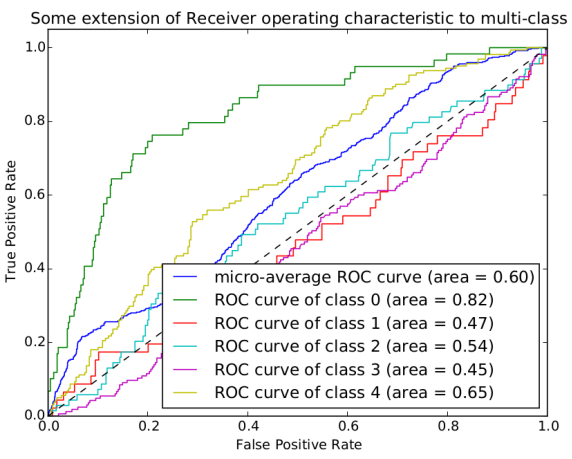
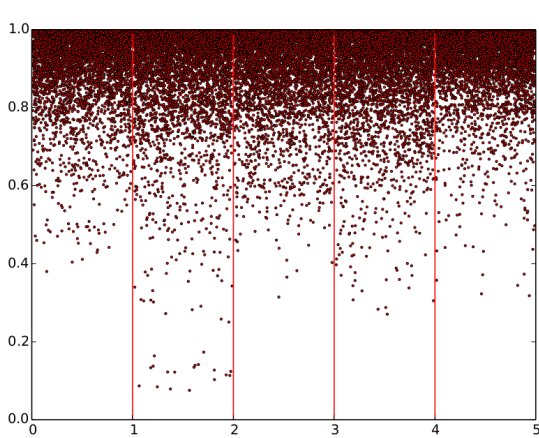
| # | Manually<br>Selected<br>Features | Accuracy<br>SVM | Wrong<br>rate<br>SVM | $R^2$ score<br>SVM | Accurac<br>y RF | Wrong<br>rate<br>RF | $R^2$ score<br>RF |
|---|----------------------------------|-----------------|----------------------|--------------------|-----------------|---------------------|-------------------|
| 1 | [0, 1, 2, 3, 4,                  | 0.48            | 0.215                | 0.000339827        | 0.35            | 0.305               | -0.28285002       |

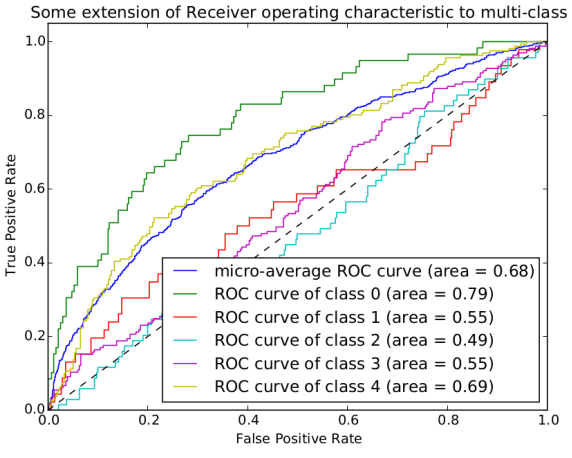
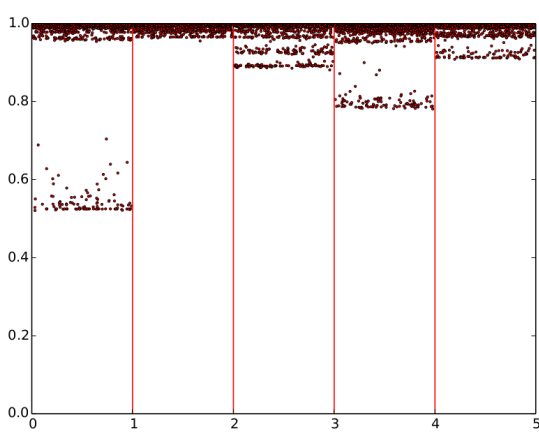
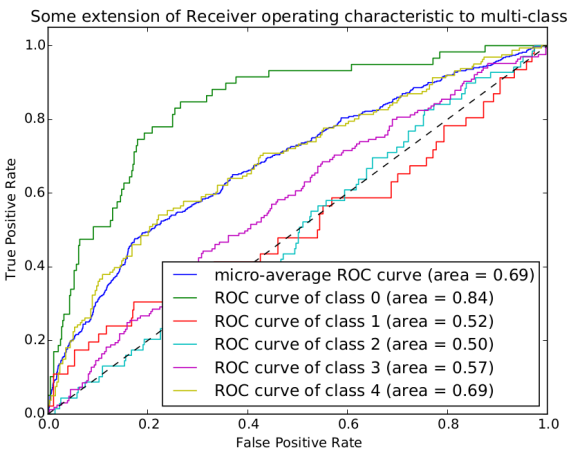
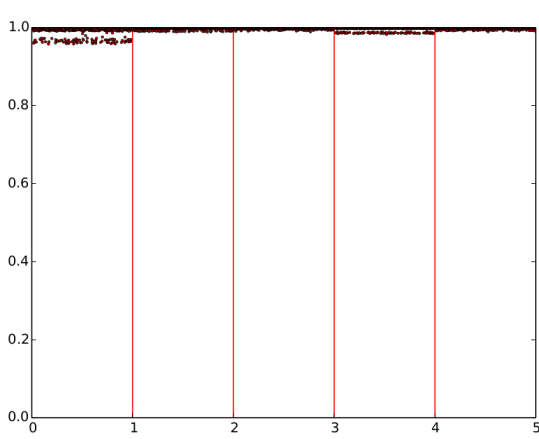
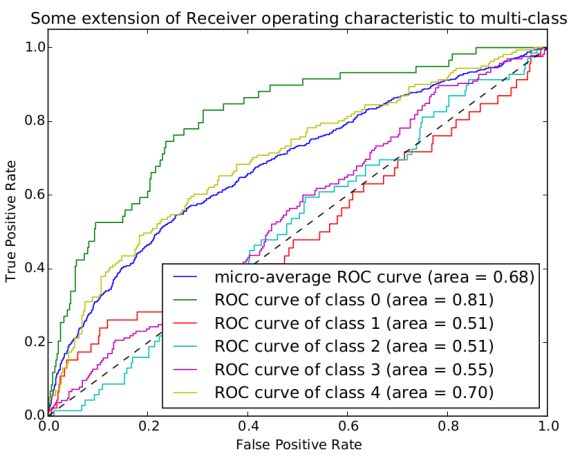
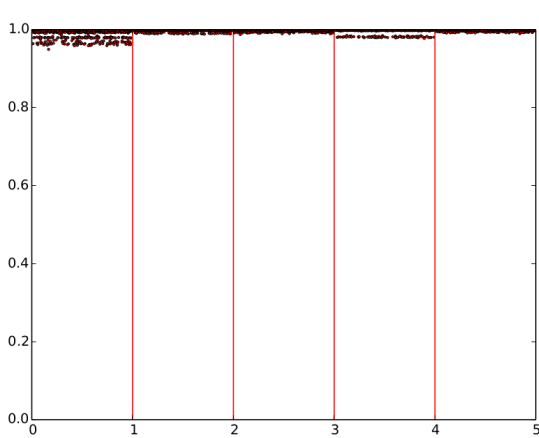
|    |  |       |       |                   |       |       |                   |
|----|--|-------|-------|-------------------|-------|-------|-------------------|
|    | 5, 8]  |       |       | 8                 |       |       | 27                |
| 2  | [0, 1, 2, 3, 4, 5, 8, 9, 10, 11]                         | 0.43  | 0.205 | -0.03930675<br>12 | 0.405 | 0.225 | -0.03930675<br>12 |
| 3  | [0, 1, 2, 3, 4, 5, 8, 14, 15, 16, 17, 18, 19]            | 0.47  | 0.225 | -0.11859990<br>94 | 0.435 | 0.23  | -0.07895333<br>03 |
| 4  | [0, 1, 2, 3, 4, 5, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19] | 0.44  | 0.205 | -0.05346624<br>38 | 0.435 | 0.23  | -0.01098776<br>62 |
| 5  | [6, 7, 8]  | 0.485 | 0.22  | -0.20922066<br>15 | 0.37  | 0.24  | -0.35081558<br>68 |
| 6  | [6, 7, 8, 9, 10, 11]                                     | 0.4   | 0.21  | -0.22621205<br>26 | 0.4   | 0.215 | -0.25736293<br>61 |
| 7  | [6, 7, 8, 14, 15, 16, 17, 18, 19]                        | 0.48  | 0.205 | -0.11859990<br>94 | 0.415 | 0.215 | -0.22904395<br>11 |
| 8  | [6, 7, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19]             | 0.45  | 0.21  | -0.10444041<br>69 | 0.43  | 0.195 | -0.02231536<br>02 |
| 9  | [12, 13, 8]  | 0.505 | 0.205 | -0.08461712<br>73 | 0.355 | 0.295 | -0.33665609<br>42 |
| 10 | [12, 13, 8, 9, 10, 11]                                   | 0.425 | 0.205 | -0.13559130<br>04 | 0.345 | 0.245 | -0.25736293<br>61 |
| 11 | [12, 13, 8, 14, 15, 16, 17, 18, 19]                      | 0.48  | 0.2   | -0.05629814<br>23 | 0.405 | 0.23  | -0.14125509<br>74 |
| 12 | [12, 13, 8, 9, 10, 11, 14, 15, 16, 17,                   | 0.43  | 0.195 | -0.05629814<br>23 | 0.385 | 0.225 | -0.22054825<br>56 |

|    |   |       |       |                   |       |       |                   |
|----|---|-------|-------|-------------------|-------|-------|-------------------|
|    | 18, 19]   |       |       |                   |       |       |                   |
| 13 | [20, 21, 22,<br>23, 24]                               | 0.535 | 0.165 | 0.150430448<br>6  | 0.49  | 0.185 | 0.031490711<br>4  |
| 14 | [20, 21, 22,<br>23, 24, 9, 10,<br>11]                 | 0.46  | 0.19  | -0.07895333<br>03 | 0.445 | 0.175 | -0.04497054<br>83 |
| 15 | [20, 21, 22,<br>23, 24, 14,<br>15, 16, 17,<br>18, 19] | 0.495 | 0.18  | 0.045650203<br>9  | 0.515 | 0.165 | 0.153262347<br>1  |

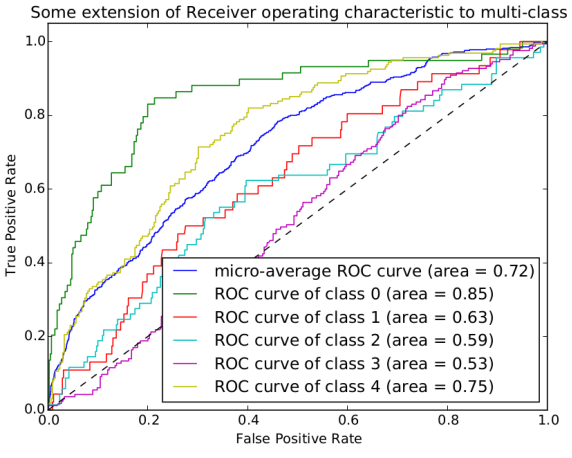
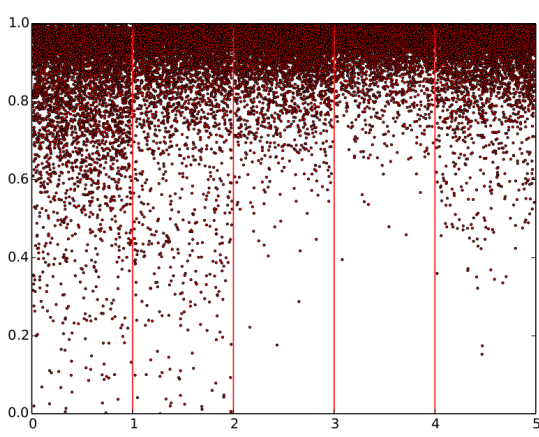
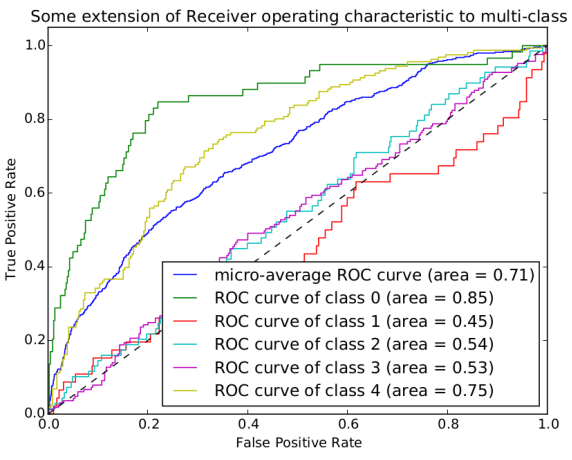
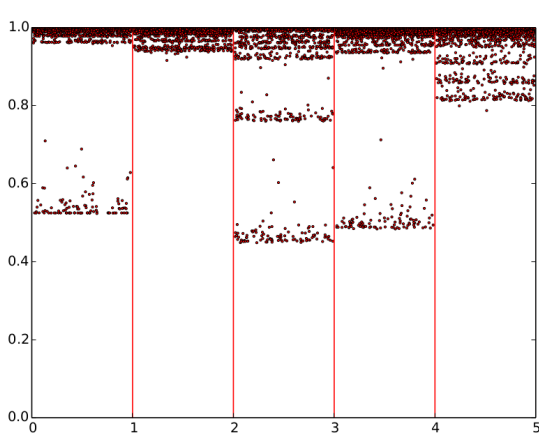
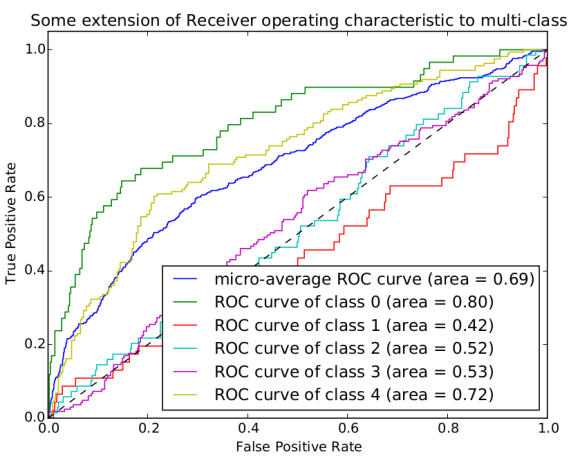
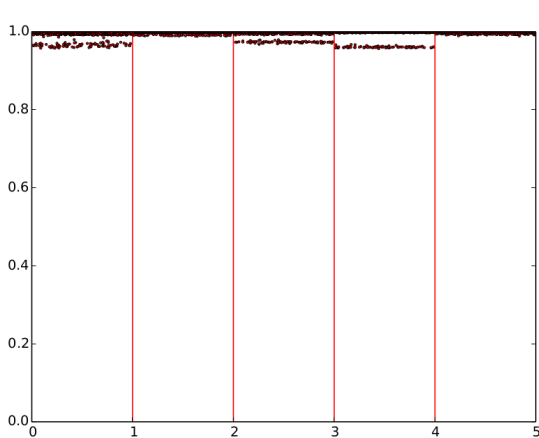
| # | Features                                      | ROC Curve  | COS Distribution |
|---|---|--|------------------|
| 1 | [0, 1, 2, 3, 4, 5, 8]                         | <p>Some extension of Receiver operating characteristic to multi-class</p> <p>             micro-average ROC curve (area = 0.68)<br/>             ROC curve of class 0 (area = 0.82)<br/>             ROC curve of class 1 (area = 0.64)<br/>             ROC curve of class 2 (area = 0.48)<br/>             ROC curve of class 3 (area = 0.60)<br/>             ROC curve of class 4 (area = 0.71)           </p> |                  |
| 2 | [0, 1, 2, 3, 4, 5, 8, 9, 10, 11]              | <p>Some extension of Receiver operating characteristic to multi-class</p> <p>             micro-average ROC curve (area = 0.64)<br/>             ROC curve of class 0 (area = 0.79)<br/>             ROC curve of class 1 (area = 0.61)<br/>             ROC curve of class 2 (area = 0.41)<br/>             ROC curve of class 3 (area = 0.56)<br/>             ROC curve of class 4 (area = 0.70)           </p> |                  |
| 3 | [0, 1, 2, 3, 4, 5, 8, 14, 15, 16, 17, 18, 19] | <p>Some extension of Receiver operating characteristic to multi-class</p> <p>             micro-average ROC curve (area = 0.71)<br/>             ROC curve of class 0 (area = 0.81)<br/>             ROC curve of class 1 (area = 0.55)<br/>             ROC curve of class 2 (area = 0.54)<br/>             ROC curve of class 3 (area = 0.56)<br/>             ROC curve of class 4 (area = 0.71)           </p> |                  |

|   |  |  |   |
|---|--|--|---|
| 4 | [0, 1, 2, 3, 4, 5, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19] | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>— micro-average ROC curve (area = 0.69)<br/> — ROC curve of class 0 (area = 0.81)<br/> — ROC curve of class 1 (area = 0.52)<br/> — ROC curve of class 2 (area = 0.55)<br/> — ROC curve of class 3 (area = 0.57)<br/> — ROC curve of class 4 (area = 0.72)</p>   |    |
| 5 | [6, 7, 8]  | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>— micro-average ROC curve (area = 0.68)<br/> — ROC curve of class 0 (area = 0.80)<br/> — ROC curve of class 1 (area = 0.54)<br/> — ROC curve of class 2 (area = 0.52)<br/> — ROC curve of class 3 (area = 0.58)<br/> — ROC curve of class 4 (area = 0.68)</p>  |   |
| 6 | [6, 7, 8, 9, 10, 11]                                     | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>— micro-average ROC curve (area = 0.59)<br/> — ROC curve of class 0 (area = 0.76)<br/> — ROC curve of class 1 (area = 0.37)<br/> — ROC curve of class 2 (area = 0.48)<br/> — ROC curve of class 3 (area = 0.57)<br/> — ROC curve of class 4 (area = 0.64)</p> |  |

|   |  |  |   |
|---|--|--|---|
| 7 | [6, 7, 8, 14, 15, 16, 17, 18, 19]            | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.69)<br/>       ROC curve of class 0 (area = 0.77)<br/>       ROC curve of class 1 (area = 0.45)<br/>       ROC curve of class 2 (area = 0.55)<br/>       ROC curve of class 3 (area = 0.55)<br/>       ROC curve of class 4 (area = 0.68)     </p>   |    |
| 8 | [6, 7, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19] | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.69)<br/>       ROC curve of class 0 (area = 0.75)<br/>       ROC curve of class 1 (area = 0.44)<br/>       ROC curve of class 2 (area = 0.56)<br/>       ROC curve of class 3 (area = 0.54)<br/>       ROC curve of class 4 (area = 0.69)     </p>  |   |
| 9 | [12, 13, 8]                                  | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.60)<br/>       ROC curve of class 0 (area = 0.82)<br/>       ROC curve of class 1 (area = 0.47)<br/>       ROC curve of class 2 (area = 0.54)<br/>       ROC curve of class 3 (area = 0.45)<br/>       ROC curve of class 4 (area = 0.65)     </p> |  |

|    |  |  |   |
|----|--|--|---|
| 10 | [12, 13, 8, 9, 10, 11]                         | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.68)<br/>       ROC curve of class 0 (area = 0.79)<br/>       ROC curve of class 1 (area = 0.55)<br/>       ROC curve of class 2 (area = 0.49)<br/>       ROC curve of class 3 (area = 0.55)<br/>       ROC curve of class 4 (area = 0.69)     </p>   |    |
| 11 | [12, 13, 8, 14, 15, 16, 17, 18, 19]            | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.69)<br/>       ROC curve of class 0 (area = 0.84)<br/>       ROC curve of class 1 (area = 0.52)<br/>       ROC curve of class 2 (area = 0.50)<br/>       ROC curve of class 3 (area = 0.57)<br/>       ROC curve of class 4 (area = 0.69)     </p>  |   |
| 12 | [12, 13, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19] | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.68)<br/>       ROC curve of class 0 (area = 0.81)<br/>       ROC curve of class 1 (area = 0.51)<br/>       ROC curve of class 2 (area = 0.51)<br/>       ROC curve of class 3 (area = 0.55)<br/>       ROC curve of class 4 (area = 0.70)     </p> |  |



|    |  |  |   |
|----|--|--|---|
| 13 | [20, 21, 22, 23, 24]                         | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.72)<br/>       ROC curve of class 0 (area = 0.85)<br/>       ROC curve of class 1 (area = 0.63)<br/>       ROC curve of class 2 (area = 0.59)<br/>       ROC curve of class 3 (area = 0.53)<br/>       ROC curve of class 4 (area = 0.75)     </p>   |    |
| 14 | [20, 21, 22, 23, 24, 9, 10, 11]              | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.71)<br/>       ROC curve of class 0 (area = 0.85)<br/>       ROC curve of class 1 (area = 0.45)<br/>       ROC curve of class 2 (area = 0.54)<br/>       ROC curve of class 3 (area = 0.53)<br/>       ROC curve of class 4 (area = 0.75)     </p>  |   |
| 15 | [20, 21, 22, 23, 24, 14, 15, 16, 17, 18, 19] | <p>Some extension of Receiver operating characteristic to multi-class</p>  <p>       micro-average ROC curve (area = 0.69)<br/>       ROC curve of class 0 (area = 0.80)<br/>       ROC curve of class 1 (area = 0.42)<br/>       ROC curve of class 2 (area = 0.52)<br/>       ROC curve of class 3 (area = 0.53)<br/>       ROC curve of class 4 (area = 0.72)     </p> |  |