**Classification Algorithm Analysis Table**

The table below show the results received when testing the algorithms used and their different scores in terms of Precision, Recall and F-Measure.

Please note:

* Oversampling has been applied on the dataset for this analysis.
* We are using binary feature to represent the presence of the word in a sentence

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm | Training Set or Cross Validation (10 Folds) | Positive Class (risky) | | | Negative Class (non-risky) | | | Result Buffer from WEKA |
| Precision | Recall | F-Measure | Precision | Recall | F-Measure |
| Random Forest | Training Set | 1.000 | 0.994 | 0.997 | 0.998 | 1.000 | 0.999 | <ORFTS1> |
| Random Forest | Cross Validation | 0.980 | 0.897 | 0.937 | 0.969 | 0.994 | 0.982 | <ORFCV1> |
| Neural Network | Training Set |  |  |  |  |  |  |  |
| Neural Network | Cross Validation |  |  |  |  |  |  |  |
| Naïve Bayes | Training Set | 0.500 | 0.704 | 0.585 | 0.896 | 0.782 | 0.835 | <ONBTS1> |
| Naïve Bayes | Cross Validation | 0.481 | 0.688 | 0.566 | 0.889 | 0.770 | 0.825 | <ONBCV1> |