

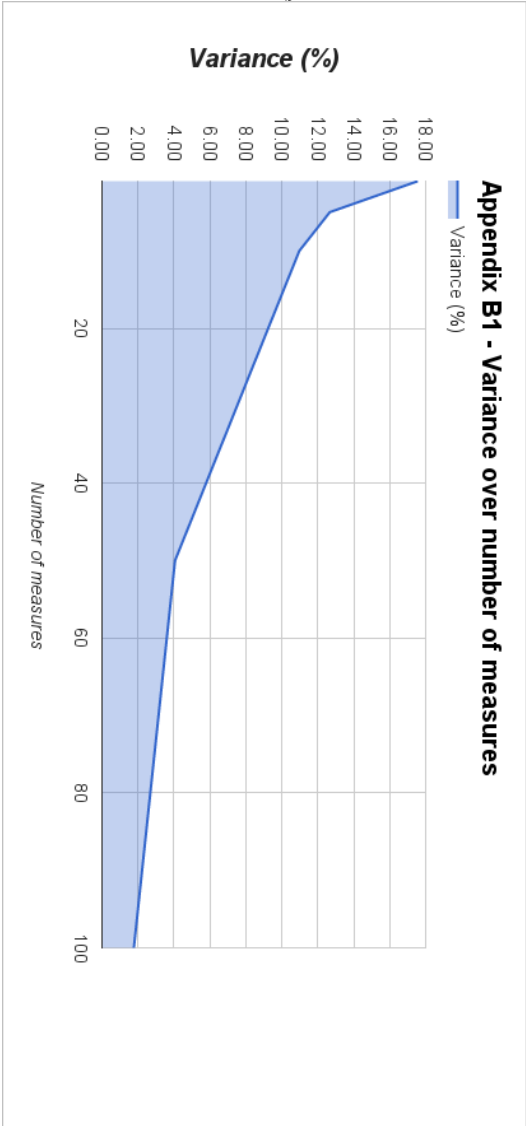
Appendix B - Charts

The chart on the right represents the imprecision (+/-) of the classifier performance.

The vertical axis represents the variance over the measure taken (in %).

The horizontal axis represents the number of times a value has been measured.

It is possible to see with in the chart, as more measures are taken of the same value the variance decreases.



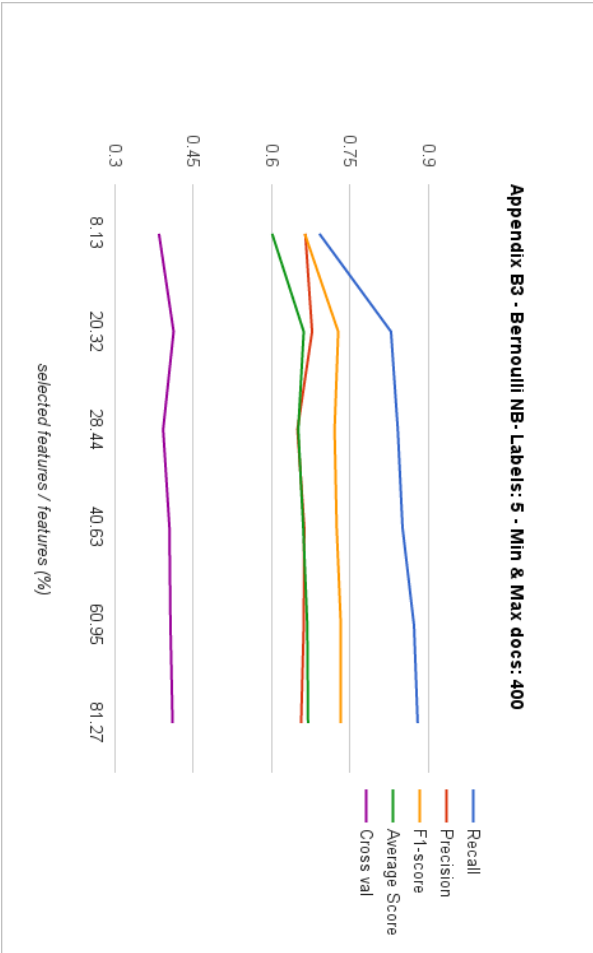
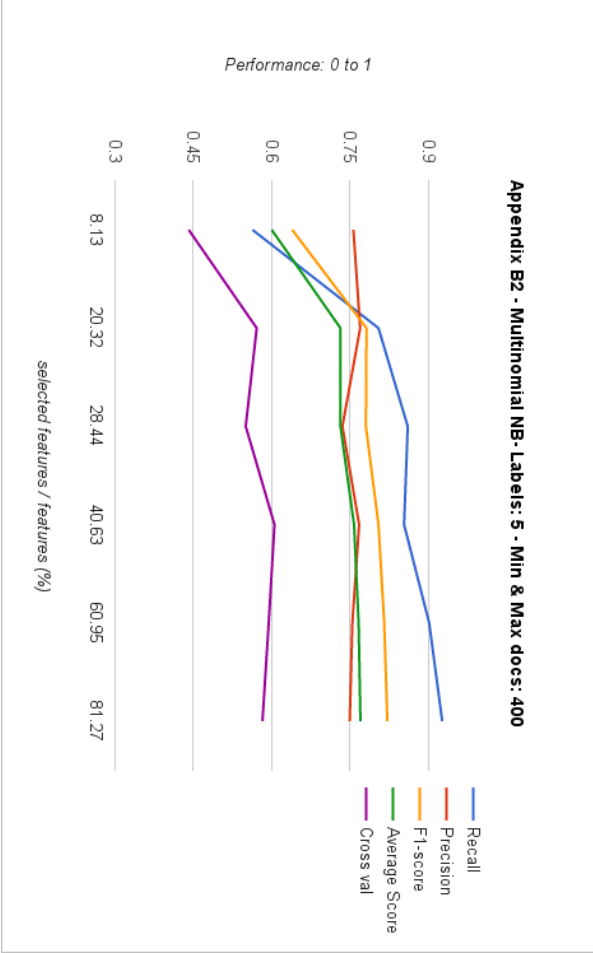
The charts below represent the performance measures of the classifiers built. The top label in bold specifies the type of the classifier and the variables used to build it.

The vertical axis represents the actual performance with values from 0 to 1. The higher the better.

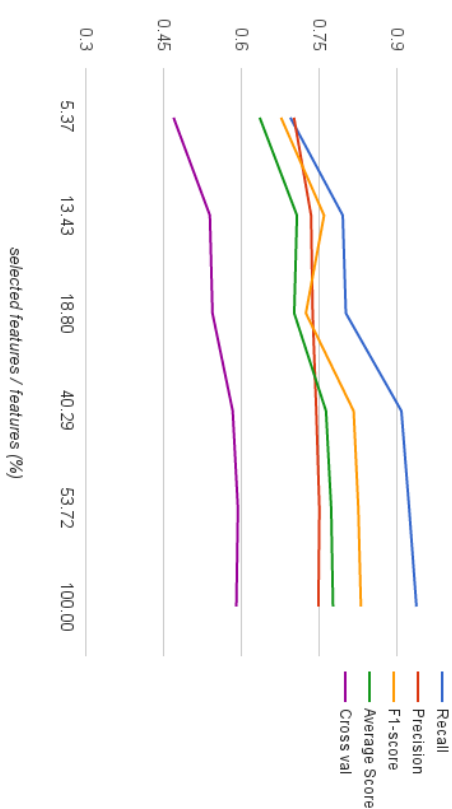
The horizontal axis is the ratio of the features used to train the classifier over the total number of features in the Vocabulary (expressed as a percentage).

Average Score, the green line, represents the average between all the measures, Recall - Precision - F1-score - Cross-validation.

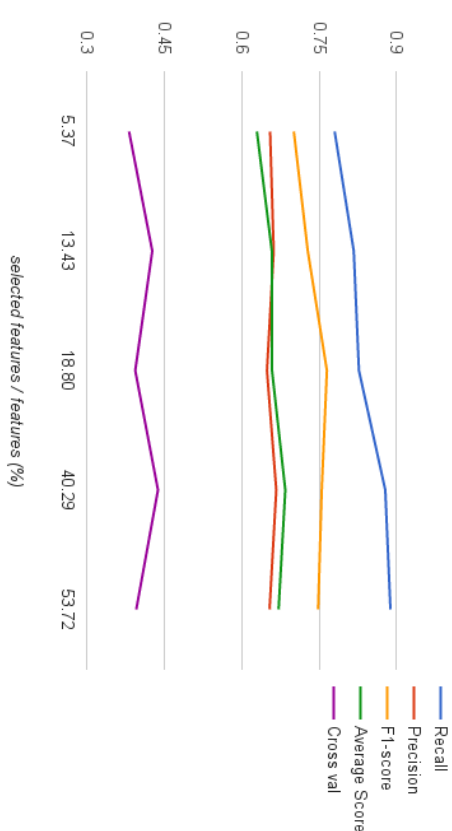
When both Min and Max documents are specified, the same amount of documents have been selected for each label. When only Min is specified, its value is the minimum amount of documents per label used to train the classifier.



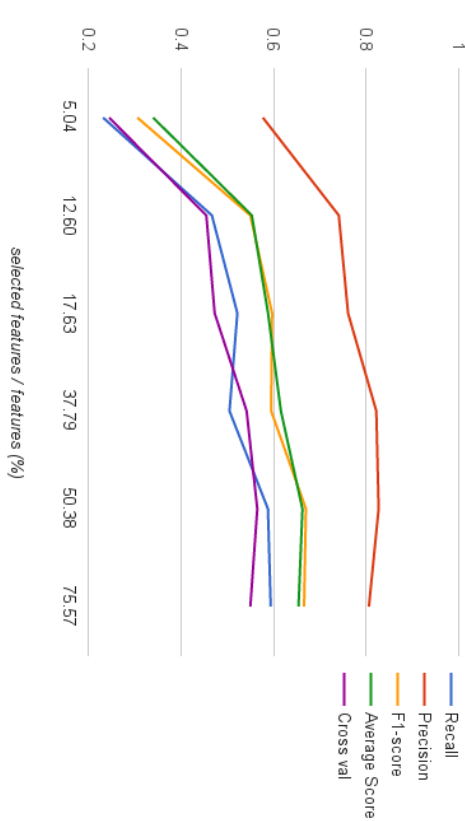
Appendix B3 - Multinomial NB- Labels: 5 - Min docs: 400



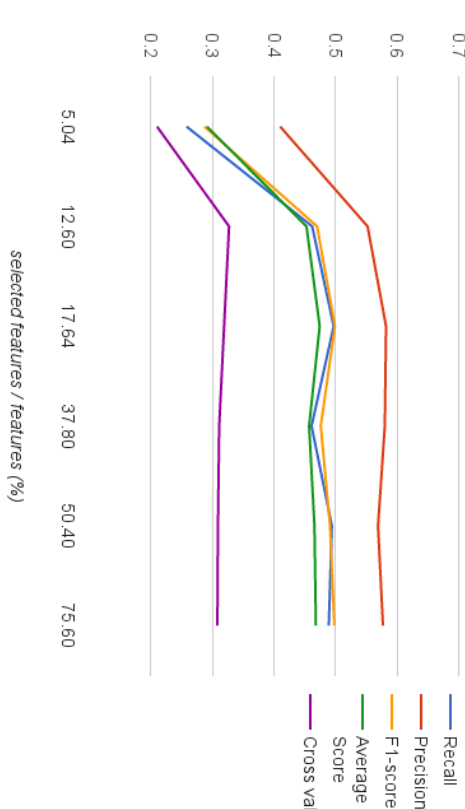
Appendix B4 - Bernoulli NB- Labels: 5 - Min docs: 400



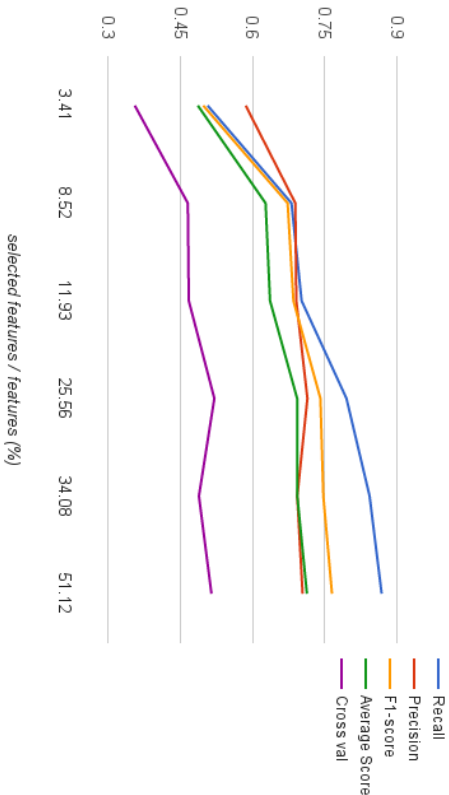
Appendix B5 - Multinomial NB- Labels: 10 - Min & Max docs: 200



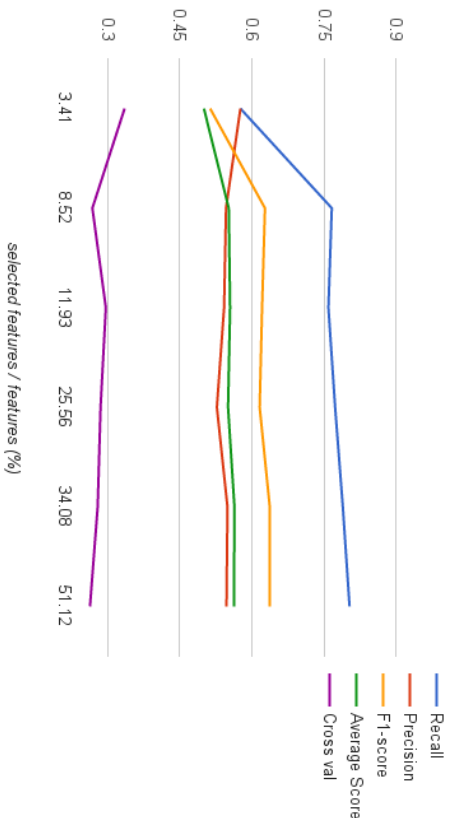
Appendix B6 - Bernoulli NB- Labels: 10 - Min & Max docs: 200



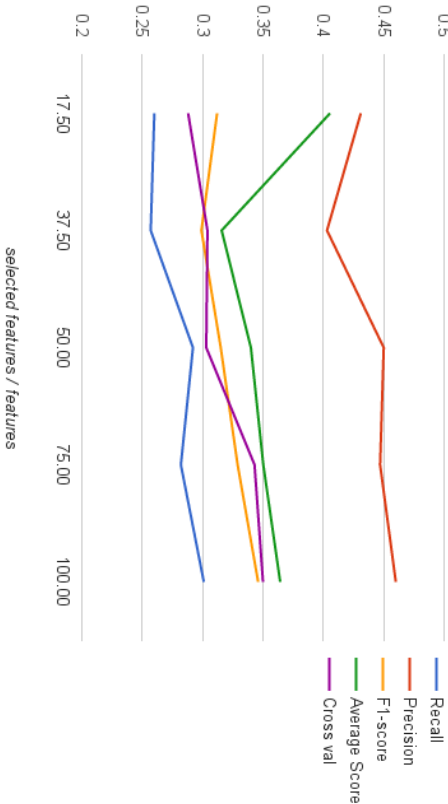
Appendix B7 - Multinomial NB- Labels: 10 - Min docs: 200



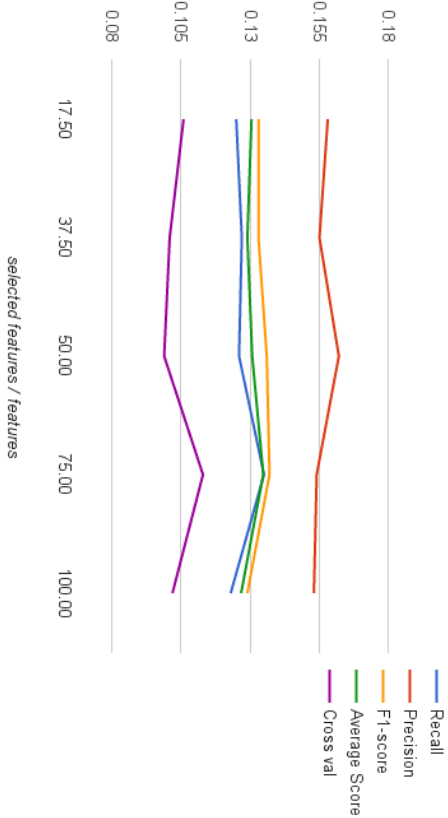
Appendix B8 - Bernoulli NB- Labels: 10 - Min docs: 200



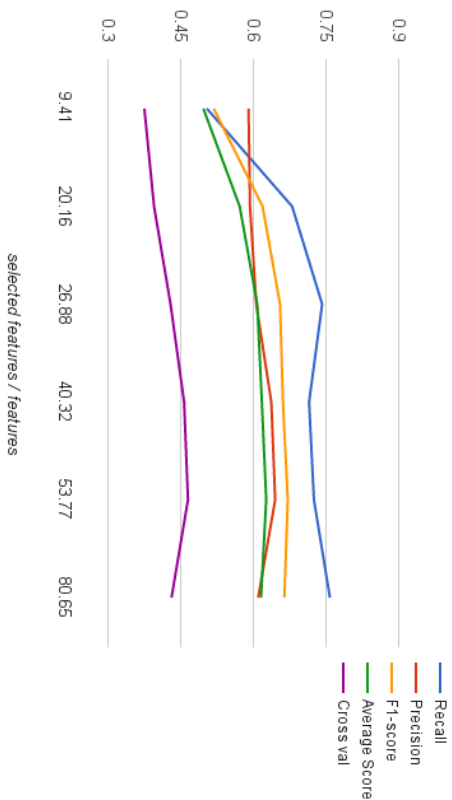
Appendix B9 - Multinomial NB- Labels: 20 - Min & Max docs: 99



Appendix B10 - Bernoulli NB- Labels: 20 - Min & Max docs: 99



Appendix B11 - Multinomial NB- Labels: 20 - Min docs: 99



Appendix B12 - Bernoulli NB- Labels: 20 - Min docs: 99

