SPAM FILTERING

NAÏVE BAYES

Pre- Filtering:

Ham Accuracy – 93.39 % ----- 325 Correct Predictions out of 348

Spam Accuracy – 97.69 % ----- 127 Correct Predictions out of 130

Total Accuracy – 94.56 % ----- 452 Correct Predictions out of 478

Post- Filtering:

Ham Accuracy – 93.39 % ----- 325 Correct Predictions out of 348

Spam Accuracy – 97.69 % ----- 127 Correct Predictions out of 130

Total Accuracy – 94.56 % ----- 452 Correct Predictions out of 478

Note that the accuracy before filtering out stop words and after filtering stop words are same because stop words are not the words that would affect the outcome. They are just the words like 'a', 'the', 'of' etc. which would not make much of difference if we keep or discard them.

LOGISTIC REGRESSION

Pre- Filtering

Steps	Learning Rate	Regularization	Accuracy
10	**0.003**	**0**	87.44
10	**0.003**	**0.02**	86.82
10	**0.03**	**0.02**	78.03
10	**0.03**	**0.002**	91.21
10	**0.003**	**0.002**	87.44
20	**0.003**	**0**	89.74
20	**0.003**	**0.02**	89.12
20	**0.03**	**0.02**	78.03
20	**0.03**	**0.002**	92.46
20	**0.003**	**0.002**	89.53
30	**0.003**	**0**	91
30	**0.003**	**0.02**	90.58
30	**0.03**	**0.02**	78.03
30	**0.03**	**0.002**	92.67
30	**0.003**	**0.002**	91
40	**0.003**	**0**	91.63
40	**0.003**	**0.02**	91
40	**0.03**	**0.02**	78.03
40	**0.03**	**0.002**	92.88
40	**0.003**	**0.002**	91.63
50	**0.003**	**0**	92.05
50	**0.003**	**0.02**	91
50	**0.03**	**0.02**	78.03
50	**0.03**	**0.002**	**93.96**
50	**0.003**	**0.002**	92.25

Post- Filtering

Steps	Learning Rate	Regularization	Accuracy
10	**0.003**	**0**	86.82
10	**0.003**	**0.02**	86.61
10	**0.03**	**0.02**	93.09
10	**0.03**	**0.002**	**94.56**
10	**0.003**	**0.002**	86.82
20	**0.003**	**0**	88.49
20	**0.003**	**0.02**	88.07
20	**0.03**	**0.02**	93.09
20	**0.03**	**0.002**	94.35
20	**0.003**	**0.002**	88.28
30	**0.003**	**0**	90.16
30	**0.003**	**0.02**	88.70
30	**0.03**	**0.02**	93.09
30	**0.03**	**0.002**	**94.56**
30	**0.003**	**0.002**	89.95
40	**0.003**	**0**	91.42
40	**0.003**	**0.02**	89.74
40	**0.03**	**0.02**	93.09
40	**0.03**	**0.002**	94.35
40	**0.003**	**0.002**	91
50	**0.003**	**0**	93.09
50	**0.003**	**0.02**	89.95
50	**0.03**	**0.02**	93.09
50	**0.03**	**0.002**	94.14
50	**0.003**	**0.002**	92.88