ANALYSIS OF MULTINOMIAL AND MULTIVARIATE NAIVE BAYES:

EXPERIMENTAL RESULTS

- **ACCURACY :**
  We have found from TABLE 3 that the accuracy for multinomial naive bayes is at least 0.3% more than multivariate naive bayes. For each iteration in TABLE 3 , the data was well shuffled to eliminate bias.

- **PRECISION :**
  We have found from TABLE 1 and TABLE 2 that multinomial naive bayes have better precision than multivariate naive bayes.

- **RECALL :**
  We have found from TABLE 1 and TABLE 2 that multinomial naive bayes and multivariate naive bayes have equal recall.

CONCLUSION

We conclude that Multinomial naive bayes has more accuracy and misclassifies fewer spam messages as ham as compared to multivariate naive bayes.

TABLES

TABLE 1: RESULTS FOR 5 FOLD CROSS VALIDATION ON MULTIVARIATE NAIVE BAYES

MULTIVARIATE NAIVE BAYES				
FOLD NUMBER	**CONFUSION MATRIX**	**ACCURACY(in %)**	**PRECISION**	**RECALL**
1	974 1 58 81	94.704	0.94	0.99
2	955 0 72 88	93.543	0.93	1
3	980 0 55 80	95.067	0.94	1
4	956 0 61 98	94.529	0.94	1
5	961 0 58 96	94.798	0.94	1

TABLE 2: RESULTS FOR 5 FOLD CROSS VALIDATION ON MULTINOMIAL NAIVE BAYES

MULTINOMIAL NAIVE BAYES				
FOLD NUMBER	CONFUSION MATRIX	ACCURACY(in %)	PRECISION	RECALL
1	975 0 60 79	94.614	0.94	1
2	954 1 63 97	94.261	0.93	0.99
3	980 0 49 86	95.605	0.95	1
4	956 0 60 99	94.619	0.94	1
5	961 0 51 103	95.426	0.95	1

TABLE 3: AVERAGE ACCURACY OF 5 FOLD CROSS VALIDATION FOR 10 ITERATIONS

ITERATION	MULTINOMIAL NAIVE BAYES ACCURACY(IN %)	MULTIVARIATE NAIVE BAYES ACCURACY(IN %)
1	95.131	94.528
2	94.815	94.403
3	94.779	94.438
4	94.851	94.510
5	94.761	94.438
6	94.815	94.331
7	94.869	94.546
8	94.672	94.295
9	94.923	94.439
10	94.744	94.492