

Analysis Report

10% of the corpus

Job	Mapper out records	Combiner out records	Map output bytes	Communications bytes
1	741548308	13531861	15230177361	144215771
2	15059684	10620313	487264450	189410883
3	10084534	No combiner	1019195977	605870185

100% of the corpus

Job	Mapper out records	Combiner out records	Map output bytes	Communications bytes
1	8224768431	160020205	171555131639	1928683093
2	2905871296	214028938	53510274112	3133862486
3	170864697	No combiner	14649227989	7685858410

Evaluation Report

10% of the corpus

Precision -> 0.955

Recall -> 0.955

F1 -> 0.954

```
Correctly Classified Instances      1643          95.5233 %
Incorrectly Classified Instances     77           4.4767 %
Kappa statistic                    0.8536
Mean absolute error                 0.1591
Root mean squared error             0.2217
Relative absolute error             49.369 %
Root relative squared error         55.2529 %
Total Number of Instances          1720
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.989	0.179	0.956	0.989	0.972	0.857	0.957	0.984	False
	0.821	0.011	0.950	0.821	0.881	0.857	0.957	0.929	True
Weighted Avg.	0.955	0.145	0.955	0.955	0.954	0.857	0.957	0.973	

100% of the corpus

Precision -> 0.988

Recall -> 0.988

F1 -> 0.988

```
Correctly Classified Instances      14171          98.7595 %
Incorrectly Classified Instances     178           1.2405 %
Kappa statistic                    0.9263
Mean absolute error                 0.0379
Root mean squared error             0.1076
Relative absolute error             22.6859 %
Root relative squared error         37.2469 %
Total Number of Instances          14349
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.992	0.059	0.994	0.992	0.993	0.926	0.991	0.998	False
	0.941	0.008	0.925	0.941	0.933	0.926	0.991	0.937	True
Weighted Avg.	0.988	0.054	0.988	0.988	0.988	0.926	0.991	0.993	