

Name:- M. Hassan Ibrar

Reg:- 2020306

So we used 2 datasets, i.e. newsgroup and Reuters and while using these data sets we filtered them we trained several models on these datasets after that these were evaluated

and we got following results.

Model	accuracy	Precision	Recall	F1-score
NaiveBayes	0.81	0.85	0.82	0.80
Logistic Regress	0.90	0.91	0.94	0.90
SVM	0.90	0.91	0.94	0.90
Random Forest	0.83	0.85	0.83	0.81

The reason for the SVM to perform better is that all the models modeled that data pretty differently and SVM performed well because due to the large dimmantiety of the dataset. It was easy for SVM to learn lines and differentiate between the samples.