

Algorithm	R Package	Accuracy											
		Kaggle Fake News Dataset 1		Kaggle Fake News Dataset 2		Victory University Dataset		Gofaaas Fake News Dataset		Final Model 1: Combination of the 3 Datasets		Final Model 2: Victory University Dataset	
		20,800 News articles		4,009 News articles		35,309 News articles		511 News articles		60,118 News articles		35,309 News articles	
		Data split ratio: Train data: 70% Test data: 30%		Data split ratio: Train data: 70% Test data: 30%		Data split ratio: Train data: 70% Test data: 30%		Data split ratio: Train data: 70% Test data: 30%		Data split ratio: Train data: 100%		Data split ratio: Train data: 100%	
		Train Accuracy	Test Accuracy	Train Accuracy	Test Accuracy	Train Accuracy	Test Accuracy	Train Accuracy	Test Accuracy	Train Accuracy	Test Accuracy	Train Accuracy	Test Accuracy (Tested over the Gofaaas Dataset)
Naive Bayes	e1071	72.03%	72.02%	80.90%	78.80%	97.93%	97.95%	92.74%	81.05%	86.06%	67.12%		
Support vector machine	We used RTextTools , which depends on e1071	99.74%	94.33%	99.57%	93.43%	99.97%	99.17%	100%	75.16%	99.57%	59.69%		
Extreme Gradient Boosting	xgboost	99.97%	97.13%	100%	98.25%	100%	99.80%	100%	83.66%	99.99%	73.97%	100%	78.86%

Train Accuracy: The accuracy of the Model over the Training data.

Test Accuracy: The accuracy of the Model over the Test data.

Final Model 1: This model was built by using 3 datasets combined: Kaggle Fake News Datasets 1 and 2 and the Victory University Dataset. With this model we got a good accuracy (73.97%) when classifying the Gofaaas Dataset. Therefore, we keep it as one of our Final Models.

Final Model 2: This model was built using the Victory University Dataset (Training data: 100%). The best accuracy that we got when classifying the Gofaaas Dataset was with this model. Therefore, we keep it as one of our Final Models.