

# CMPS143\_ass4

Carlos del Rey Garcia 1710268

February 2019

Feature set	# of features selected	Model	Accuracy	Dataset
word_features (not binned)	All	Naïve Bayes	0.8642857142857143	dev
word_features (not binned)	All	Naïve Bayes	0.8308457711442786	test
word_features (not binned)	32	Naïve Bayes	0.6857142857142857	dev
word_features (not binned)	32	Naïve Bayes	0.681592039800995	test
word_features (not binned)	256	Naïve Bayes	0.8178571428571428	dev
word_features (not binned)	256	Naïve Bayes	0.8134328358208955	test
word_features (not binned)	8192	Naïve Bayes	0.8214285714285714	dev
word_features (not binned)	8192	Naïve Bayes	0.8159203980099502	test
word_features (not binned)	32768	Naïve Bayes	0.875	dev
word_features (not binned)	32768	Naïve Bayes	0.835820895522388	test

  

Feature set	Model	Dataset	Accuracy
word_features (not binned)	Naïve Bayes NLTK	dev	0.8308457711442786
word_features (not binned)	Naïve Bayes NLTK	test	0.8308457711442786
word_features (not binned)	Naïve Bayes Scikit Learn	dev	0.7321428571428571
word_features (not binned)	Naïve Bayes Scikit Learn	test	0.7412935323383084
word_features (not binned)	Decision Tree Scikit Learn	dev	0.75
word_features (not binned)	Decision Tree Scikit Learn	test	0.6940298507462687
word_features (not binned)	SVM	dev	0.5285714285714286
word_features (not binned)	SVM	test	0.5497512437810945
w2vec	SVM	dev	0.7285714285714285
w2vec	SVM	test	0.7189054726368159