CS529 - Assignment1 report
Roll: 190101103

Key observations and Analysis:

1. When I applied feature selection (chisquare and mutual information) over the tfidf features, the accuracy of the classifiers increased slightly or remained the same.
2. When I applied dimensionality reduction using PCA or LSI/LSA (i.e. SVD) or word embedding, the accuracies are even higher than those achieved by feature selection. The reason for this is that - in feature selection, we are taking the top 500 features and eliminating the remaining ones while in dimension reduction we are projecting the feature vectors to a lesser dimension without any elimination. In the word embedding, we are even using better features since the Glove feature vectors also encode the semantic meaning of the words in the word vector representation.
3. The low accuracies of the Naive Bayes classifier in almost all the cases can be attributed to the fact that the naive bayes algorithm internally makes an assumption that the features are conditionally independent and makes the classification according to this assumption.

The accuracies (in percentage) are summarised in the 3 tables shown below -

FNC-multiclass

Classifier ↓ / Features →	only tfidf	tfidf+chisqr	tfidf+mutual_info	tfidf+PCA	tfidf+LSI/LSA	Word Embedding
SVM	64.7	71.4	72.9	73.3	73.0	69.2
Naive Bayes	47.9	17.3	17.5	16.0	16.0	47.7
Decision Tree	73.3	73.3	72.3	73.3	73.1	73.3
Random Forest	73.3	73.3	73.3	73.3	73.3	73.3

FNC-binary

Classifier ↓ / Features →	only tfidf	tfidf+chisqr	tfidf+mutual_info	tfidf+PCA	tfidf+LSI/LSA	Word Embedding
SVM	61.1	65.9	64.5	69.7	68.6	64.7
Naive Bayes	42.1	55.2	52.8	30.1	30.1	56.5
Decision Tree	70.3	70.3	70.8	70.5	70.5	68.8
Random Forest	70.5	70.5	70.5	70.5	70.5	70.5

NELA-binary

Classifier ↓ / Features →	only tfidf	tfidf+chisqr	tfidf+mutual_info	tfidf+PCA	tfidf+LSI/LSA	Word Embedding
SVM	54.6	56.9	57.3	57.0	58.2	48.2
Naive Bayes	49.5	48.7	48.6	57.3	56.5	62.0
Decision Tree	60.7	60.7	60.7	58.1	60.2	54.9
Random Forest	59.9	59.2	58.4	57.5	58.5	60.4