

Extra Credit: Deception Detection

- This system was implemented using Naïve Bayes.
- Tokenized the words using NLTK
- Removed stop words using NLTK
- Lower cased them and dropped all the special characters in a word.
- Formed two dictionaries of T and F reviews, smoothed them using Laplace
- Each test review is passed through these two dictionaries and it's given a tag based on maximum probability that the review gets after checking both dictionaries.
- Ignored words that are not present in dictionary.
- Created 80-20 train-test sets.
- Accuracy during test was 53.9%