1. Naive-Bayes Classifier

2. Maximum Entropy

   Principle of Maximum Entropy – Consider the testable information about a probability distribution function. Examine all the probability distribution that would satisfy the testable information. Of those examined distributions, the one with maximum entropy is the proper distribution.

   Entropy – Entropy is the as the negative of the logarithm of the probability distribution.

   The Maximum Entropy is based on the principle of maximum entropy and of all the models that fits our training data it selects the one with the largest entropy.
   Maximum Entropy model does not assume that the features are conditionally independent of each other and is normally used to solve many text classification problems such as language detection during translation.

3. Support Vector Machines

   Given labelled training data (*supervised learning*), the algorithm outputs an optimal hyperplane which categorizes new examples.
   The operation of the SVM algorithm is based on finding the hyperplane that gives the largest minimum distance to the training examples.

Challenges

1. **Sarcasm/ Implicit**

We take author, audience, environment into consideration. Use a special training set that contain all the commonly used words with sarcasm.

2. **Context or Domain Dependence**

Train the model with lot of novels and articles. Use NLP to find the different meaning and usage of those words and Use the correct meaning.

3. **Thwarted Expectations**

Give more importance to the last sentence whenever an opposite direction transition takes place.

4. **Pragmatics**

Regex use Lexical analysis [lex and yacc]

5. **Media**

Beyond the scope of the project

6. **Chat corpus**

Use a chat database to train a separate classifier

7. **Word Knowledge**

Train the model based on the domain, very common words

8. **Comparative**

Tokenize the words, depending on who is more important remove all the tokens towards left/right side post a conjunction to that word

9. ***External Link***

Web Crawling