**Steps:**

1. **Data Processing & Cleaning**
2. Merge fake and true news data
3. Shuffle
4. Convert to lowercase
5. Removing number and special character
6. Tokenize & removing stop words
7. **Data exploration**
8. **Modeling**

* Vectorizing - Vectorize the words to numerical data (pick one)
* TF-IDF (takes word frequencies into consideration)
* Bag of Words (does not take into consideration the frequency of words)
* Split thedataset into train and test
* Classification (pick few)
* *Naïve Bayes*
* Bag of Words
* *Random Forest*
* *Decision Tree*
* *Logistic Regression*
* *K-Nearest Neighbors*
* Confusion matrix - check the accuracy of the model