

## Fake News Detection

Submitted by:

Amit Kumar Jena

**ACKNOWLEDGMENT**

My project was inspired by N. J. Conroy, V. L. Rubin, and Y. Chen, “Automatic deception detection: Methods for finding fake news,”

And

Shlok Gilda,Department of Computer Engineering, Evaluating Machine Learning Algorithms for Fake News Detection

**INTRODUCTION**

The problem here is what we call "Fake News". A fake are those news stories that are false: the story itself is fabricated, with no verifiable facts, sources, or quotes.

When someone (or something like a bot) impersonates someone or a reliable source to false spread information, that can also be considered as fake news. In most cases, the people creating this false information have an agenda, that can be political, economical or to change the behavior or thought about a topic.

There are countless sources of fake news nowadays, mostly coming from programmed bots, that can't get tired (they're machines hehe) and continue. to spread false information 24/7.

Serious studies in the past 5 years, have demonstrated big correlations between the spread of false information and elections, the popular opinion or feelings about different topics.

The problem is real and hard to solve because the bots are getting better are tricking us. Is not simple to detect when the information is true or not all the time, so we need better systems that help us understand the patterns of fake news to improve our social media, communication and to prevent confusion in the world.

Conceptual Background of the Domain Problem:

Mostly we will be working with NLP as we will be doing text analysis and the

**REQUIREMENTS**

* Python
* numpy
* pandas
* itertools
* matplotlib
* sklearn

**Analytical Problem Framing**

The used approach differs based on the problem type. When the proposed problem is about determining to which label a given observation belongs to, then we can use classification methods to find the solution. Classification methods are based on classifying the observations in different categories depending on different factors. A dataset with a number of observations is provided for the classifier to help it finding the category of a new given observation. The classifier uses the dataset to train and create patterns. to recognize every observation to which category it suits. This method belongs to the supervised learning approach. Any algorithm that performs the classification method and classifies a given feature into a category is called a classifier. In the following, we present classification approaches useful for text classification problems. a) Naive Bayes Classifier: Naive Bayes classifiers are probabilistic classifiers3 which apply Bayes theorem in their decision rule assuming strong independence between features. Naive Bayes classifiers are useful for text classification problems, because they can be trained precisely in the supervised learning case. Given a class variable (label) y and a vector of features x = (x1, x2, ..., xn), where n is the number of features, the probability of y depending on the vector of features x can be calculated using the formula: P(y|x1, ..., xn) = P(y) Qn i=1 P(xi |y) P(x1, ..., xn) Since the value of P(x1, ..., xn) is the same for all class variables (labels), we can say that the result of a prediction probability r is the maximum probability of the calculated probabilities for all class variables: r = max(P(y) Yn i=1 P(xi |y))

**Model/s Development and Evaluation :**

The accuracy is calculated based on the test data we had when we split the dataset to train data and test data. The items labels in the test data were removed and stored in another separate variable. We pass the test data to the classifier to see what predictions it results. Then we compare the classifier predictions with the labels we have removed from the test data. The accuracy score is the percentage of the true predictions. We have the following abbreviations to represent the results of the classifiers in the tables of accuracy and ROC curves: 1) Bag-of-Words: BoW 2) Bi-gram: bigram 3) Term Frequency-Inverse Document Frequency: TF-IDF 4) Multinomial Naive Bayes: MN 5) Linear Support Vector Classifier: LSVC A. Content Detection Results We used a dataset with 6335 news articles and performed random split on this dataset into two parts: training data and testing data. Training data consists of 66.6% of the data and testing data consists of 33.3% of the data , we observe that the LSVC classifier performed well with the TF-IDF model; at the same time the MN gave the worst result with TF-IDF. This is because the MN usually requires integer feature counts to predict with a higher accuracy. Also, the results show that representing documents through weighted terms vectors (TF-IDF vectors) is more efficient in the case of linear classification. This is because the linear classifiers do not necessary require integer feature counts. Both classifiers gave the same accuracy score with the BoW model. On the other hand, the MN classifier gave a better score accuracy in case of Bi-gram model. This depends on the Bi-gram term-document matrix, which has pairs of tokens frequencies as features. This means that using the frequencies of pairs to calculate the probabilities gives better results than mapping these frequencies in a linear space; the reason is that having a pair of tokens with a high frequency can indicate a bigger probability for one of the labels than the other.

**CONCLUSION**

The problems of fake news and disinformation play an important role on nowadays life. This is because the advanced level of technology and communication methods we have enabled information spreading among people without any verification. This is a reason why researchers started searching for solutions to stop fake news and disinformation from spreading easily. However, it is well known that controlling the flow of information online is impossible. In this paper, we performed an attempt to verify the news articles credibility depending on their characteristics. At this aim, we implemented an algorithm combining several classification methods with text models. It performed well, and the accuracy results were relatively satisfying. As future work, we plan to better study the combination between the feature extraction methods and the classifiers as we will be able to choose the text representation model that performs best with the classifier. Moreover, to achieve a higher accuracy, we will have to implement a more sophisticated algorithm which may use data mining technologies with big data, because creating a big dataset including more types of news articles with more class variables (labels) will help raising the accuracy score.