**Fake News Detection Using Natural Language Processing**

**Harun-Or-Rashid (2966), Sanad Bhowmik (2952) , Robiul Islam (2994)**

Department of CSE

Daffodil International University

Ashulia, Dhaka

***Abstract —*** *This paper shows a simple approach for fake news detection using naive Bayes classifier. This approach was implemented as a software system and tested against a data set of newspaper. We achieved classification accuracy of approximately 95% on the test set which is a decent result considering the relative simplicity of the model. This results may be improved in several ways, that are described in the article as well. Received results suggest, that fake news detection problem can be addressed with artificial intelligence methods.*

**I. INTRODUCTION**

Newspapers and the Internet make news information considerably more accessible and pleasant. Users of the Internet may frequently follow events of interest in an online manner, and the widespread usage of mobile devices makes this process more easier.

However, with immense opportunities come huge responsibilities. The media has a significant effect on society, and as is frequently the case, someone wants to profit from this. To achieve certain purposes, the media may distort information in many ways. As a result, news pieces that are not entirely genuine or even wholly fake are produced. There are also a number of websites dedicated nearly entirely to the dissemination of false information. They purposefully produce hoaxes, misinformation, and disinformation masquerading as actual news, and they frequently use social media to generate traffic and enhance their impact. The primary purpose of fake news websites is to influence public opinion on certain topics (mostly political). Ukraine, the United States of America, Germany, China, and a slew of other nations have examples of similar websites [1]. As a result, false news is both a worldwide issue and a global task.

Many experts believe that machine learning and artificial intelligence can help to combat fake news [2]. Because technology is cheaper and more datasets are accessible, artificial intelligence algorithms have lately started to operate significantly better on a variety of categorization tasks (picture recognition, speech detection, and so on).

There are a number of well-respected articles on the subject of automated deception detection. The authors of [3] give a broad review of the existing approaches in this area. The authors present their strategy for detecting false news based on comments for specific items in microblogs in [4]. The authors of [5] build two methods for deception detection, one based on support vector machines and the other on a naive Bayes classifier (which is also utilized in the system presented in this study). They gather information by asking participants to offer accurate or false information on a variety of issues, including abortion, the death penalty, and friendship. The detection accuracy attained by the system is around 95%.

II. Materials and Methods

In the following, We discuss our suggested framework, followed by descriptions of algorithms, datasets, and performance assessment criteria.

II.1. As shown in Figure 1, our proposed approach builds on existing research by using ensemble techniques with a variety of linguistic feature sets to identify news items from several categories as true or false. The new aspect of our suggested strategy is the employment of ensemble techniques in conjunction with the Linguistic Inquiry and Word Count (LIWC) feature set. There are various reputable websites that offer credible news information, as well as fact-checking websites such as PolitiFact and Snopes. Furthermore, there are open repositories [6] that are maintained by academics to retain an up-to-date list of currently accessible datasets and connections to possible fact-checking sites that may aid in the fight against the spread of fake news. However, we selected three datasets for our experiments which contain news from multiple domains (such as politics, entertainment, technology, and sports) and contain a mix of both truthful and fake articles. The datasets are available online and are extracted from the World Wide Web. The dataset are publicly available at Kaggle.

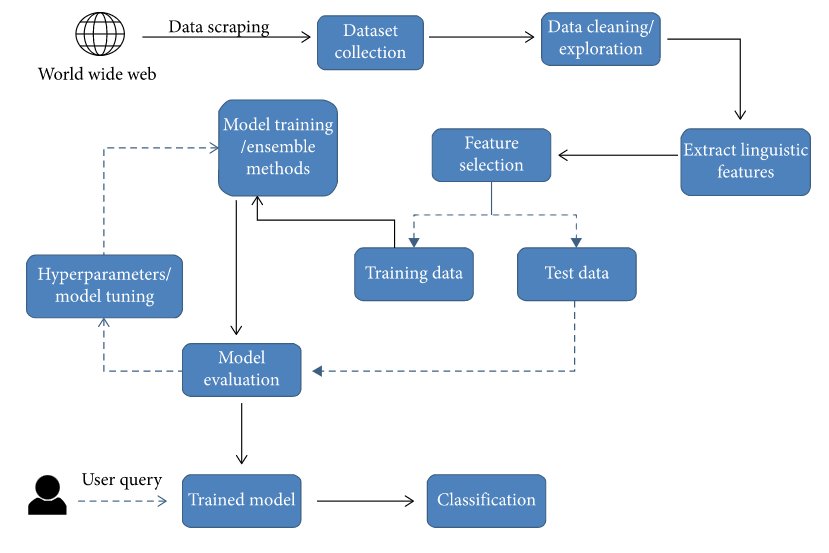


Figure 1

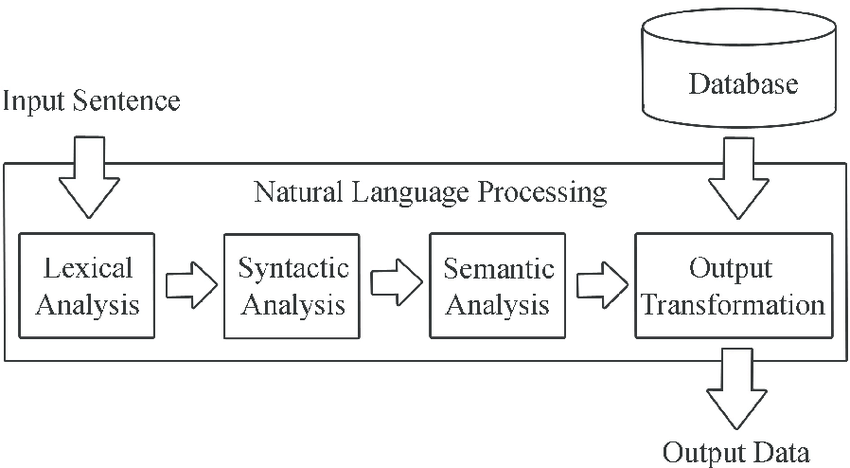


Figure 2

##### II.2. Algorithms. We used the following learning algorithms in conjunction with our proposed methodology to evaluate the performance of fake news detection classifiers.

##### II.2.1. Naive bayes classifier. Naive Bayes classifiers are a type of basic probabilistic classifiers based on the Bayes theorem and strong (naive) independence assumptions between features in machine learning. Naive Bayes is a straightforward method for building classifiers, which are models that give class labels to problem cases represented as vectors of feature values, with the class labels selected from a limited set. It is a family of techniques for training such classifiers, all based on the same principle: all naive Bayes classifiers assume that the value of one feature is independent of the value of any other feature, given the class variable [7].

##### Naive Bayes classifiers are a popular statistical technique of email filtering. They emerged in the middle of the 90s and were one of the first attempts to tackle spam filtering problem [8].

##### Naive Bayes typically use bag of words features to identify spam e-mail, an approach commonly used in text classification. Naive Bayes classifiers work by correlating the use of tokens (typically words, or sometimes other constructions, syntactic or not), with spam and non-spam e-mails and then using Bayes theorem to calculate a probability that an email is or is not a spam message [8].

##### III. Overview of the training dataset

##### Dataset, collected from Kaggle, was used for learning and testing the naive Bayes classifier.

##### The dataset contains information about Newspaper, each of which represent a news article. They were collected from three large site each from the right and from the left, as well as three large mainstream political news pages (Politico, CNN, ABC News). All nine pages have earned the coveted verified blue checkmark from Facebook, which gives them an additional layer of credibility on the platform. The smallest of these public pages has over 450 thousand of followers, and the largest – over 4.1 million followers [9].

##### They also noted whether the news was a link, photo, video, or text. Raters were asked to provide notes and sources to explain their rulings of “mixture of true and false” or “mostly false.” They could also indicate whether they were unsure of a given rating, which would trigger a second review of the same post in order to ensure consistency. Any discrepancies between the two ratings were resolved by a third person. That same person conducted a final review of all posts that were rated mostly false to ensure they warranted that rating [9]

##### IV. Received Results

##### Table-1 summarizes the results that were obtained.

##### True news and false news pieces have almost same categorization accuracy, although true news has somewhat lower classification accuracy. This might be due to the dataset's skewness: only 47% of it is true news.

##### Table-1

|  |  |  |
| --- | --- | --- |
| News article type | Total number of news in test dataset | Classification accuracy |
| True | 21418 |  |
| Fake | 23503 | 0.94 |
| Total | 44921 |  |

##### The precision of a classifier is calculated as follows: Precision = tp / (tp + fp).

##### where

##### tp – number of true positive examples;

##### fp – number of false positive examples.

##### The recall of a classifier is calculated as follows:

##### Recall = tp / (tp + fn).

##### V. Conclusion

The research showed, that even quite simple artificial intelligence algorithm (such as naive Bayes classifier) may show a good result on such an important problem as fake news classification. Therefore the results of this research suggest even more, that artificial intelligence techniques may be successfully used to tackle this important problem.

REFERENCE

*[1] Fake news websites. (n.d.) Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Fake\_news\_website. Accessed Feb. 6, 2017.*

*[2] Cade Metz. (2016, Dec. 16). The bittersweet sweepstakes to build an AI that destroys fake news. [Online]. Available: https://www.wired.com/ 2016/12/bittersweet-sweepstakes-build-ai-destroys-fake-news/*

*[3] Conroy, N., Rubin, V. and Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology, 52(1), pp.1-4.*

*[4] Markines, B., Cattuto, C., & Menczer, F. (2009, April). Social spam detection. In Proceedings of the 5th International Workshop**on Adversarial Information Retrieval on the Web (pp. 41-48)*

*[5] Rada Mihalcea , Carlo Strapparava, The lie detector: explorations in the automatic recognition of deceptive language, Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, August 04-04, 2009, Suntec, Singapore*

*[6] F. T. Asr and M. Taboada, “Misinfotext: a collection of news articles, with false and true labels,” 2019.*

##### *[7] Naive Bayes classifier. (n.d.) Wikipedia. [Online]. Available:*

##### *https://en.wikipedia.org/wiki/Naive\_Bayes\_classifier. Accessed Feb. 6, 2017.*

##### *[8] Naive Bayes spam filtering. (n.d.) Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Naive\_Bayes\_spam\_filtering. Accessed Feb. 6, 2017.*

##### *[9] Craig Silverman, Lauren Strapagiel, Hamza Shaban, Ellie Hall, Jeremy Singer-Vine. (2016, Oct. 20).* *Hyperpartisan Facebook pages are publishing false and misleading information at an alarming rate. [Online]. Available: https://www.buzzfeed.com/craigsilverman/partisan-fb-pagesanalysis?utm\_term=.twM44ywz1B#.cxEnnGWD6g*