IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

FAKE NEWS DETECTION USING NAIVE BAYES CLASSIFIER

[1]I.M.V.Krishna, [2] Dr.S.Sai Kumar

[1,2] Assistant Professor

[1,2]Departement of Information Technology

[1,2]PVP Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India

Abstract: The Paper is an application of NLP (Natural Language Processing) techniques for detecting the 'fake news', that is, misleading news stories that comes from the non-reputable sources. By building a model based on a count vectorizer (using word tallies)or (Term Frequency Inverse Document Frequency)tf-idf matrix, (word tallies relative to how often they're used in other articles in your dataset) can only get you so far. But these models do not consider the important qualities like word ordering and context. It is very possible that two articles that are similar in their word count will be completely different in their meaning. The data science community has responded by taking actions against the problem. There is a Kaggle competition called as the "Fake News Challenge" and Facebook is employing AI to filter fake news stories out of users' feeds. Combating the fake news is a classic text classification project with a straight forward proposition. Is it possible for you to build a model that can differentiate between "Real "news and "Fake" news? So a proposed work on assembling a dataset of both fake and real news and employ a Naive Bayes classifier in order to create a model to classify an article into fake or real based on its words and phrases.

Index Terms **– Natural Language Processing, Facebook, Kaggle.**

I. INTRODUCTION

Internet and social media made the access to the news information much easier and comfortable. Often Internet users can follow the events of their interest in online mode, and spread of the mobile devices makes this process even easier. But with great possibilities come great challenges. Mass media have a huge influence on the society, and as it often happens, there is someone who wants to take advantage of this fact. Sometimes to achieve some goals mass-media may manipulate the information in different ways. This leads to producing of the news articles that are not completely true or even completely false. There even exist lots of websites that produce fake news almost exclusively.

These days' fake news is creating different issues from sarcastic articles to a fabricated news and plan government propaganda in some outlets. Fake news and lack of trust in the media are growing problems with huge ramifications in our society. Obviously, a purposely misleading story is "fake news" but lately blathering social media's discourse is changing its definition. Some of them now use the term to dismiss the facts counter to their preferred viewpoints. The term 'fake news' became common parlance for the issue, particularly to describe factually incorrect and misleading articles published mostly for the purpose of making money through page views. Facebook has been at the epicenter of much critique following media attention. They have already implemented a feature to flag fake news on the site when a user sees' it they have also said publicly they are working on to distinguish these articles in an automated way. Certainly, it is not an easy task.

A given algorithm must be politically unbiased – since fake news exists on both ends of the spectrum – and also give equal balance to legitimate news sources on either end of the spectrum. In addition, the question of legitimacy is a difficult one. However, in order to solve this problem, it is necessary to have an understanding on what Fake News is. Later, it is needed to look into how the techniques in the fields of machine learning, natural language processing help us to detect fake news.

There exists a large body of research on the topic of machine learning methods for deception detection, most of it has been focusing on classifying online reviews and publicly available social media posts. Particularly since late 2016 during the American Presidential election, the question of determining 'fake news' has also been the subject of particular attention within the literature. Conroy, Rubin, and Chen outlines several approaches that seem promising towards the aim of perfectly classify the misleading articles.

They note that simple content-related n-grams and shallow parts-of-speech (POS) tagging have proven insufficient for the classification task, often failing to account for important context information. Rather, these methods have been shown useful only in tandem with more complex methods of analysis. Deep Syntax analysis using Probabilistic Context Free Grammars (PCFG) have been shown to be particularly valuable in combination with n-gram methods. Feng, Banerjee, and Choi are able to achieve 85%-91% accuracy in deception related classification tasks using online review corpora.

Feng and Hirst implemented a semantic analysis looking at 'object:descriptor' pairs for contradictions with the text on top of Feng's initial deep syntax model for additional improvement. Rubin, Lukoianova and Tatiana analyze rhetorical structure

using a vector space model with similar success. Ciampaglia et al. employ language pattern similarity networks requiring a pre-existing knowledge base.

II. PROPOSED SYSTEM

In this paper a model is build based on the count vectorizer or a tf-idf matrix (i.e) word tallies relatives to how often they are used in other articles in your dataset) can help . Since this problem is a kind of text classification, Implementing a Naive Bayes classifier will be best as this is standard for text-based processing. The actual goal is in developing a model which was the text transformation (count vectorizer vs tfidf vectorizer) and choosing which type of text to use (headlines vs full text). Now the next step is to extract the most optimal features for count vectorizer or tfidf-vectorizer, this is done by using a n-number of the most used words, and/or phrases, lower casing or not, mainly removing the stop words which are common words such as “the”, “when”, and “there” and only using those words that appear at least a given number of times in a given text dataset.

III. NAÏVE BAYES CLASSIFIER AND ITS USAGE FOR FAKE NEWS DETECTION

In machine learning, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes theorem with strong (naive) independence assumptions between the features. Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. 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. 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.

3.1 Mathematical model of Naïve Bayes Classifier for Fake News Detection

The main idea is to treat each word of the news article independently. As were already mentioned, fake news articles often use the same set of words, which may indicate, that the specific article is indeed a fake news article. Of course, it is impossible to claim that the article is a fake news just because of the fact, that some words appear in it, but these words affect the probability of this fact. The formula for calculating the conditional probability of the fact, that news article is fake given that it contains some specific word looks as following:

$$Pr(F|W) = Pr(W|F)\cdot Pr(F)/(Pr(W|F)\cdot Pr(F)+Pr(W|T)\cdot Pr(T))$$

where:

$Pr(F|W)$ – conditional probability, that a news article is fake given that word W appears in it; $Pr(W|F)$ – conditional probability of finding word W in fake news articles;

$Pr(F)$ – overall probability that given news article is fake news article;

$Pr(W|T)$ – conditional probability of finding word W in true news articles;

$Pr(T)$ – overall probability that given news article is true news article.

This formula is derived from Bayes’ theorem.

3.2 Evaluation Metrics

Confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known.

- **True Positive** (TP): when predicted fake news pieces are actually annotated as fake news;
- **True Negative** (TN): when predicted true news pieces are actually annotated as true news
- **False Negative** (FN): when predicted true news pieces are actually annotated as fake news
- **False Positive** (FP): when predicted fake news pieces are actually annotated as true news.

Recall - Out of all the positive classes, how much we predicted correctly.

Precision -Out of all the classes, how much we predicted correctly

F-measure - F-score helps to measure Recall and Precision at the same time.

IV.RESULTS

```
In [36]: # Set index
         df = df.set_index("Unnamed: 0")

         # Print first lines of `df`
         df.head()
```

Out[36]:

Unnamed: 0	title	text	label
8476	You Can Smell Hillary's Fear	Daniel Greenfield, a Shillman Journalism Fello...	FAKE
10294	Watch The Exact Moment Paul Ryan Committed Pol...	Google Pinterest Digg Linkedin Reddit Stumbleu...	FAKE
3608	Kerry to go to Paris in gesture of sympathy	U.S. Secretary of State John F. Kerry said Mon...	REAL
10142	Bernie supporters on Twitter erupt in anger ag...	— Kaydee King (@KaydeeKing) November 9, 2016 T...	FAKE
875	The Battle of New York: Why This Primary Matters	It's primary day in New York and front-runners...	REAL

Figure No. 1: After Dataset Cleaning

```
In [49]: clf.fit(tfidf_train, y_train)
         pred = clf.predict(tfidf_test)
         score = metrics.accuracy_score(y_test, pred)
         print("accuracy:   %0.3f" % score)
         cm = metrics.confusion_matrix(y_test, pred, labels=['FAKE', 'REAL'])
         plot_confusion_matrix(cm, classes=['FAKE', 'REAL'])

         accuracy:   0.857
         Confusion matrix, without normalization
```

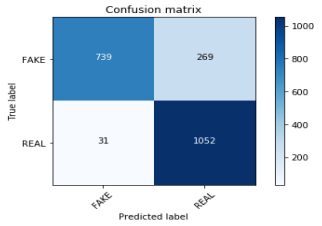

Figure No. 2: Confusion matrix with Multinomial Naïve Bayes using tf-idf vectorizer

```
clf.fit(count_train, y_train)
pred = clf.predict(count_test)
score = metrics.accuracy_score(y_test, pred)
print("accuracy:   %0.3f" % score)
cm = metrics.confusion_matrix(y_test, pred, labels=['FAKE', 'REAL'])
plot_confusion_matrix(cm, classes=['FAKE', 'REAL'])

accuracy:   0.893
Confusion matrix, without normalization
```

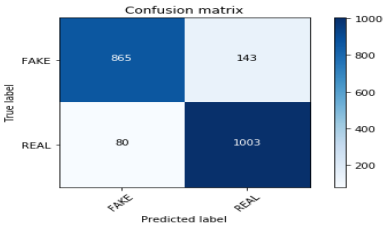

Figure No. 3: Confusion matrix with Multinomial Naïve Bayes using count vectorizer

```
clf.fit(hash_train, y_train)
pred = clf.predict(hash_test)
score = metrics.accuracy_score(y_test, pred)
print("accuracy:   %0.3f" % score)
cm = metrics.confusion_matrix(y_test, pred, labels=['FAKE', 'REAL'])
plot_confusion_matrix(cm, classes=['FAKE', 'REAL'])

accuracy:   0.902
Confusion matrix, without normalization
```

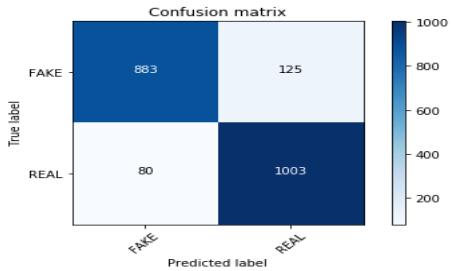

Figure No. 4: Confusion matrix with Multinomial Naïve Bayes using hashing vectorizer

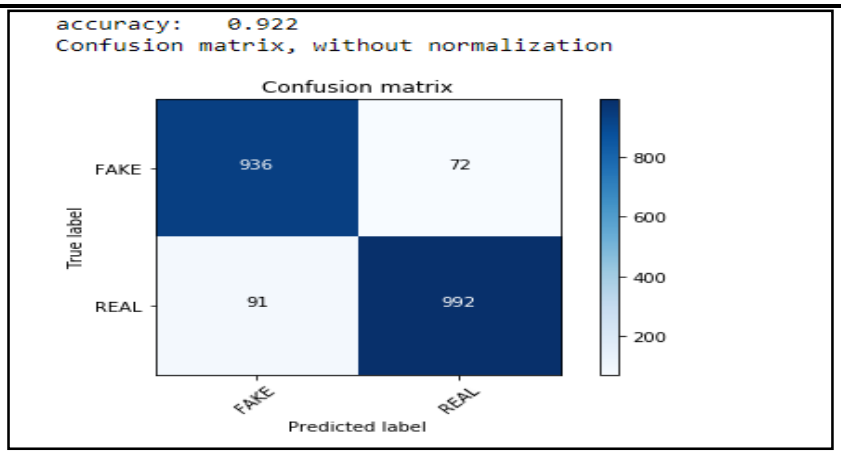

Figure No. 5: Confusion matrix with Passive Aggressive Classifier using hashing vectorizer

Figure No. 6: Final Output screen for Prediction

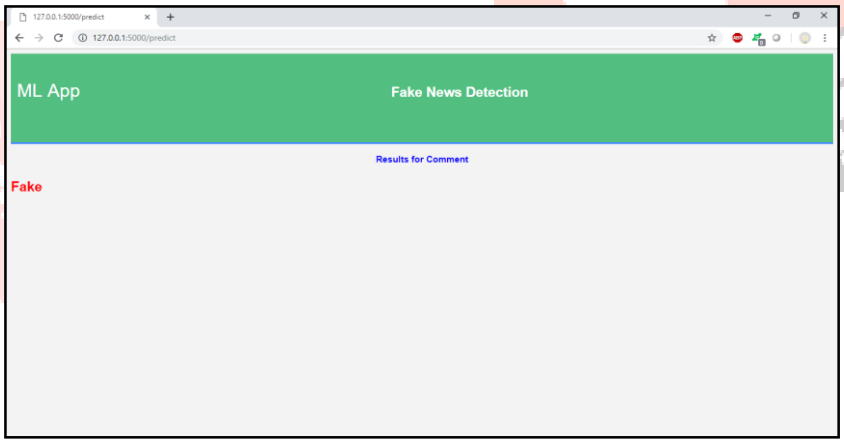

Figure No. 7: After prediction of entered news

V. CONCLUSION

The paper 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 suggest even more, that artificial intelligence techniques may be successfully used to tackle this important problem. Get more data and use it for training. In machine learning problems it is often the case when getting more data significantly improves the performance of a learning algorithm. Use stemming. In linguistic morphology and information retrieval, stemming is the process of reducing inflected words to their word stem. Such technique helps to treat similar words (like “write” and “writing”) as the same words and may improve classifier’s performance as well. Use group of words instead of separate words for calculating probabilities. This will help to use more meaningful syntax constructions for Bayes classifier.

REFERENCES

[1] "**Fake News Detection Using Naive Bayes Classifier**"- 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering.

[2] Shu, Kai, et al. "**Fake news detection on social media: A data mining perspective.**"ACM SIGKDD Explorations Newsletter19.1 (2017): 22-36. 2

[3] Ruchansky, Natali, Sungyong Seo, and Yan Liu."**CSI: A Hybrid Deep Model for Fake News Detection.**" Proceedings of the 2017 ACM on Conference on Information Knowledge Management.ACM,2017.

[4] Le, Quoc, and Tomas Mikolov. "Distributed representations of sentences and documents" International Conference on Machine Learning. 2014.

[5**]** Shu, Kai, H. Russell Bernard, and Huan Liu."**Studying Fake News via Network Analysis: Detection and Mitigation.**" arXiv preprint arXiv:1804.10233 (2018).

[6] **Beginning Python**-Peter Norton, Alex Samuel, David Aitel, Eric Foster-Johnson, Leonard Richardson, Jason Diamond, Aleatha Parker, Michael Roberts

[7] **Building Machine Learning Systems with Python**-Willi Richert Luis Pedro Coelho

[8] https://www.datacamp.com/community/tutorials/scikit-learn-fake-news

[9] **Data Mining-Practical Machine Learning Tools and Techniques,** Second Edition-Ian H.Witten , Eibe Frank.

[10] **Machine Learning**-Tom M. Mitchell

[11] **Data Mining** Third Edition, Jiawei Han ,Micheline Kamber ,Jian Pei .

[12] https://stackoverflow.com/

[13] **Software Engineering**, A Precise Approach, Pankaj Jalote, Wiley.

[14] **Software Engineering**, 7/E , Roger S. Pressman , TMH.

[15] **Design Patterns** By Erich Gamma, Pearson Education.