Journal of Computer Applications and Research, Volume 1, No 1, 2020

Comparative Analysis of Fake News Detection System

Theint Theint Shwe
University of Computer Studies (Mandalay)
theint2shwe@gmail.com

Abstract

Enormous amount of information is published daily via online and print media, but it is not easy to tell whether the information is a true or false. The extensive spread of fake news has the potential for extremely negative impacts on individual and society. Therefore, fake news detection has become an emerging research that is attracting tremendous attention. The purpose of the proposed system is to detect fake news with the help of text analysis using n-gram features and machine learning classification techniques. We investigate the feature extraction techniques of term frequency, term frequency –inverse document frequency. Classification of fake or real news is performed using Passive Aggressive Classifier (PAC), Naïve Bayes (NB) and Support Vector Machine (SVM) classifiers. The proposed system is evaluated using three publicly available datasets. Performance of the different classifiers is measured with precision, recall, f-measure and accuracy score. According to the analysis upon the three different datasets with three classifiers, PAC is the strongest classifier among the other two, SVM is stronger than NB.

Keywords: Keywords: fake news, term frequency (TF), term frequency –inverse document frequency (TF-IDF), Support Vector Machine (SVM), Naïve Bayes (NB), Passive Aggressive Classifier (PAC)

1. Introduction

Fake news is the deliberate spread of misinformation via traditional news media or via social media. Social media provides for easy access, little to no cost, and the spread of information at an impressive rate. On the other hand, social media provides the ideal place for the creation and spread of fake news. Sometimes, using social media as a medium for news updates is a double-edged sword. People can download articles from sites, share the information, re-share from others and by the end of the day the false information has gone so far from its original site that it becomes indistinguishable from real news.

Fake news and hoaxes have been there since before the advent of the Internet. The widely accepted definition of Internet fake news is: fictitious articles deliberately fabricated to deceive readers". Some news outlets publish fake news to increase readership or as part of psychological warfare. In general, one of the goals is profiting through clickbaits. Clickbaits lure users and entice curiosity with flashy headlines or designs to click links to increase advertisements revenues.

The purpose of this paper is to come up with a solution that can be utilized by users to detect and filter out the sites containing false and misleading information. Therefore, we collect the three publicly available datasets from the kaggle site[1] and classify the fake and real news using the machine learning algorithms. Before classifying the data set, data preprocessing has to be performed using natural language processing techniques of feature extraction methods to get the highest accuracy as much as possible. After classifying the datasets, performance of the classification algorithms: Passive Aggressive Classifier (PAC), Support Vector Machine (SVM), and Naïve Bayes classifier (NB) are measured. Their evaluation results are displayed with precision, recall, f-measure and accuracy score.

2. Related Work

Two credibility-focused Twitter datasets: CREDBANK, a crowd sourced dataset of accuracy assessments for events in Twitter, and PHEME, a dataset of potential rumors in Twitter are applied to develop a model for fake news detection. They concluded with a discussion contrasting accuracy and credibility and why models of non-experts outperform models of journalists for fake news detection in Twitter [1].

There are five different types of fake news in accordance with the literature. Various machine learning models are applied and with different datasets. Finally, news articles are labeled as fake or real news. [2].

Fake news is categorized as three different type and machine learning techniques are applied to divide as fake and real. This is especially for the information professionals can help tackle the fake news problem not only through the promotion of information literacy, but through continued research of automated deception detection [3].

Even though the problem of fake news is not a new issue, detecting fake news is believed to be a complex task given that humans tend to believe misleading information and the lack of control of the spread of fake content Fake news has been getting more attention in the last couple of years, especially since the US election in 2016 [4].

The project created a tool for detecting the language patterns that characterize fake and real news through the use of machine learning and natural language processing techniques. The results of this project demonstrated the ability for machine learning to be useful in this task. They built a model that catches many intuitive indications of real and fake news as well as an application that aids in the visualization of the classification decision [5].

3. Theory Background

Information can be accessible from everywhere and the quality of information that is true or false is one of the big problems. So, we need to eliminate or filter the propagation of misinformation to viewer/readers is necessary. This paper proposed the system to detect fake news from the different sites from the internet. The processing steps of the system are: data acquisition, preprocessing, feature extraction (TF-IDF), classification (Passive Aggressive Classifier, Support Vector Machine, Naïve Bayes), and performance evaluation (precision, recall, f-measure and accuracy). The following is the architecture of the proposed system.

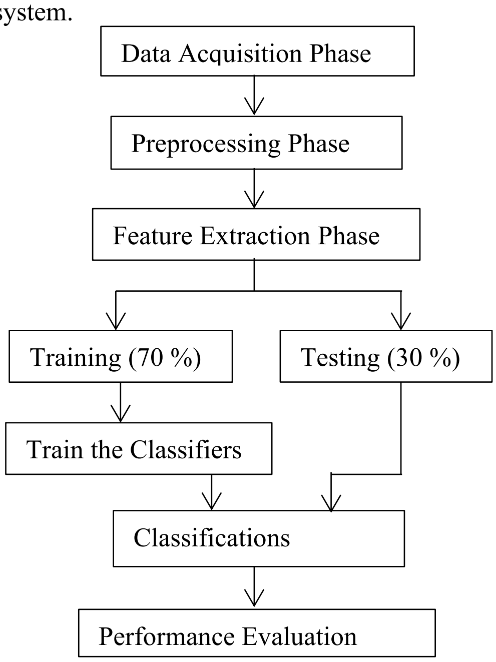

Figure1. Architecture of fake news detection system

The details of the system architecture are clearly explained in the following sections.

3.1. Data Acquisition Phase

It is the first phase of the fake news detection. It collects datasets from the different sources that are publicly available [6] [7] [8]. Three different datasets are used for classification, two datasets are from the kaggle.com and the other dataset is from the data-flair.training web sites. Real-world data is often incomplete, inconsistent, and/or lacking in certain behaviors or trends, and is likely to contain many errors. Data pre-processing is a proven method of resolving such issues. This will help in getting better results through the classification algorithms.

3.2. Preprocessing Phase

Preprocessing steps of the proposed system includes: remove blank rows, change all text to lower case, word tokenization, and remove stop words and non-alpha text, and word lemmatization [9].

3.2.1. Word Tokenization

Word tokenization is a process of breaking a stream of text up into words, phrases, symbols, or other meaningful elements called tokens. The list of tokens becomes input for further processing.

3.2.2. Stop Word Removal

Stop words are insignificant words in a language that will create noise when used as features in text classification. These are words frequently used in sentences to help connect thought or to assist in the sentence structure. Articles, prepositions, and conjunctions and some pronouns are considered stop words. We removed common words such as a, about, an, are, as, at, be, by, for, from, how, in, is, of, on, or, that, the, these, this, too, was, what, when, where, who, will, and so on. Those words were removed from each document, and the processed documents were stored and passed on to the next step.

3.2.3. Lemmatization

Lemmatization is the process of reducing the inflectional forms of each word into a common base or root. Example of lemmatized form is:

Table1. Lemmatization Form

Form	Lemmatization
studies	study
studying	study
beautiful	beautiful
beautifully	beautifully

3.3. Feature Extraction Phase

There are a large number of terms, words, and phrases in documents that lead to high computational burden for the learning process of fake news detection. Furthermore, irrelevant and redundant features can hurt the accuracy and performance of the classifiers. In order to train a classifier, it is necessary to select and extract a set of features that can be converted into numeric values. In the proposed system, Term Frequency (TF) and Term Frequency- Inverse Document Frequency (TF-IDF) are applied [3].

Let D be a dataset, or set of documents. Let d denote a document, $d \in D$; we define a document as a set of words w. Let $n_w(d)$ denote the number of times the word w appears in document d. Hence, the size of document d is $|d| = \Sigma_{w \in d}\, n_w(d)$ [10].

The normalized TF for word w with respect to document d is defined as follows:

$$TF(w)_d = \frac{n_{w(d)}}{d} \quad (1)$$

The inverse document frequency (IDF) for a term w with respect to dataset D, denoted $IDF(w)_D$, is the logarithm of the total number of documents in the corpus divided by the number of documents where this particular term appears, and is computed as follows:

$$IDF(w)_d = 1 + \log\left(\frac{|D|}{|\{d: D|w \in d\}|}\right) \quad (2)$$

One of the main characteristics of IDF is that it weights down the TF while scaling up the rare ones. TF will dominate the frequency count; however, using IDF scales lessens the impact of these terms. [2] So, TF-IDF for the word w with respect to document d and corpus D is calculated as follows:

$$TF - iDF(w)_{d,D} = TF(w)_d * IDF(w)_d \quad (3)$$

3.4. Classification

Classification is a very important area of supervised machine learning. A large number of important machine learning problems fall within the fake news classification area. There are many classification methods, and we have chosen the Passive Aggressive Classifier (PAC), Support Vector Machine (SVM) and Naïve Bayes classifier (NB).

3.4.1. Passive Aggressive Classifier (PAC)

It is a family of machine learning algorithm for both classification and regression. The idea is very simple and the performance has been proofed to many other alternative methods [11]. The procedural steps of passive aggressive classifier are as follow: [12]

Initialize weight factor w= (0,…..,0)
monitor a stream:
    receive new doc d=(d_1 , …, d_v)
    apply tf-idf, normalize ||d|| = 1
    predict positive if $d^T w>0$
    obswerve true class: y=+1
    want to have:
        $d^T w >= +1$ if positive (y= +1)
        $d^T w <= -1$ if negative (y= -1)
same as: $y(d^T w) >=1$
loss: :=max(0, 1-$y(d^T w)$)

3.4.2. Support Vector Machine (SVM)

A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After giving SVM model sets of labeled training data for each category, they're able to categorize new text. We apply this algorithm for text classification, fake news detection [3] [13] [14]. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall [15].

3.4.3. Naïve Bayes Classifier (NB)

Naive Bayes assumes the independence of features which is not true for news articles as words usually come together. In a Naive Bayes' classifier, probabilities of individual features are multiplied to provide an overestimation of the probability which decides the class of the input (Eqn. 4 represents the Bayes' Theorem). Thus, it seems that Naive Bayes won't perform well for text classification. However, it may perform well even with strongly dependent features since the dependencies tend to cancel each other out [16].

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (4)$$

3.4.3. Performance Evaluation

After classifying the news from different datasets are as fake or real, performance of each classifier is measured with precision, recall, f-measure and accuracy score. Their respective formulations are described in the following:

$$precision = \frac{TP}{TP+FP} \quad (5)$$

$$recall = \frac{TP}{TP+FN} \quad (6)$$

$$f - measure = \frac{2*precision*recall}{Precision+recall} \quad (7)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

where,

True Positive (TP) – predicted and actual class both positive (e.g. fake news classified as fake)
True Negative (TN) – predicted and actual class both negative (e.g. real news classified as real)
False Negative (FN) – incorrect prediction of negative class (e.g. fake news classified as real)
False Positive (FP) – incorrect prediction of positive class (e.g. real news classified as fake)

4. Experimental Results

Three different data sets are tested with different machine learning techniques. The datasets are split into two, training and testing. The training data set will be used to fit the model and the predictions will be performed on the test data set. The training data will have 70% of the dataset and test data will have the remaining 30%. The performance measure of three different machine learning techniques are displayed in the following tables.

The dataset-1 is collected from the kaggle.com web site [6]. It is publicly available for fake news detection with size 11.99 MB. It contains 4 columns. The first column represents the links of the news web site (URLs), second column is the headline, the third represents the message body, and the fourth column denotes the label 1 for real news and 0 for fake news. According to the experiments on dataset-1 with three different algorithms, accuracy score of the PAC is highest. Precision of SVM is slightly higher than PAC.

Table2. Experimental results of dataset-1

Classifier	Precision	Recall	F-measure	Accuracy
PAC	0.929	0.948	0.938	93.78 %
SVM	0.918	0.937	0.927	92.77 %
NB	0.865	0.888	0.876	87.67 %

The dataset-2 is downloaded from [7] for fake news detection with the size of 29.2 MB. The dataset contains news, title, text, and label that denotes whether the news is real of fake. In dataset-2, precision is highest in PAC but accuracy score is highest in SVM.

Table3. Experimental results of dataset-2

Classifier	Precision	Recall	F-measure	Accuracy
PAC	0.986	0.987	0.986	98.58 %
SVM	0.987	0.980	0.983	98.24 %
NB	0.960	0.901	0.930	92.73 %

The dataset-3 is the political news dataset of size 29.27 MB concerned with 2016 U.S. election [8]. The evaluation results show that all of the evaluation features of PAC are higher than the other two classifiers.

Table4. Experimental results of dataset-3

Classifier	Precision	Recall	F-measure	Accuracy
PAC	0.920	0.927	0.923	92.11 %
SVM	0.909	0.941	0.925	92.48 %
NB	0.882	0.897	0.847	89.06 %

The following figure represents the accuracy score of classification techniques upon fake news datasets. According to the study, PAC algorithm has the highest accuracy of two datasets out of three. SVM has the highest accuracy in one of the three datasets and NB classifier has lower performance than PAC and NB.

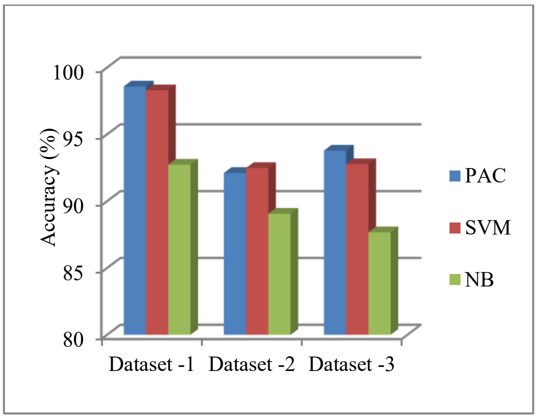

Figure2. Accuracy score

5. Conclusion

The accessibility of information from many different sources may be real or fake. The spread of information (news) is voluminous and veracity. So, fake news may affect the society and people need to distinguish the truth, real news. This system is proposed intended to help the readers/ reviewers to discriminate fake and real news. Machine Learning techniques are applied with the combination of natural language processing's feature extraction method. Performance of each techniques are also analyzed and can also be seen their accuracy.

References

[1]. C. Buntain, J.Golbeck, "Automatically Identifying Fake News in Popular Twitter Threads", arXiv:1705.01613v2 [cs.s1], 30 May 2018.

[2]. S. Kaur, P. Kumar, P. Kumaraguru, "Automating Fake News Detection System Using Multi-level Voting Model", *Soft Comput* (2019). https://doi.org/10.1007/s00500-019-04436-y.

[3]. J. Tompkins, "Disinformation Detection: A review of linguistic feature selection and classification models in news veracity assessments", State University of New York at Albany, College of Emergency Preparedness, Homeland Security and Cybersecurity Department of Information Science, IIST 669, August 2018.

[4]. N. Lemann, "Solving the Problem of Fake News", The New Yorker (2017). http://www. newyorker.com/news/news-desk/solving-the-problem-of-fake-news.

[5]. N.O. Brien, T. Poggio, "Machine Learning for Detection of Fake News", © Massachusetts Institute of Technology 2018.

[6]. https://www.kaggle.com/jruvika/fake-news-detection

[7]. https://data-flair.training/blogs/advanced-python-project-detecting -fake-news

[8]. https://www.kaggle.com/hassanamin/textdb3

[9]. https://medium.com/@bedigunjit/ simple-guide-to-text-classification-nlp-using-svm-and-naive-bayes-with-python-421db3a72d34

[10]. S. Aphiwongsophon, P. Chongstitvatana, "Detecting Fake News with Machine Learning Method", 15th international conference on Electrical Engineering/Electronics, Computer, Telecommunication and Information Technology (ECTI- CON), july 2008, DOI:10.1109/ ECTI- CON. 2018.8620051.

[11]. https://www.bonaccorso.ed/2017/10/06/ml-algorithms-addendum-passive-aggessive-algorithm.
[12]. K. Crammer, O. Dekel, J.Keshet, S.S. Shwartz, Y, Singer, “Online Passive Aggressive algorithms”, journal of machine learning research, 7 (2006)551-585.
[13]. https://www.svm-tutorial.com/text-classification-prepare-data
[14]. https://www.svm-tutorial.com/2017/02/svms-overview-support-vector-machines
[15]. K. Poddar et. al, “Comparison of Various Machine Learning Models for Accurate Detection of Fake News”, 2019 Innovations in Power and Advanced Computing Technology (i-PACT), 978-1-5386-8189-3/19/$31.00 ©2019 IEEE.
[16]. Y. Shukla, N. Yadav, A. Hari, “A Unique Approach for Detection of Fake News using Machine Learning”, International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177 Volume 7 Issue VI, June 2019.