International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 06 | June - 2021 ISSN: 2582-3930

Fake News Detection Using Machine Learning

Manoj Lade[1] , Priyanka Kamble[2] , Shruti Barmate[3] , Ruby Khobragade[4] , Shubhangi Wasnik[5]

[1] Assistant Professor,[2345] UG Students
Department of Information Technology , Nagpur Institute Of Technology, Nagpur, India

ABSTRACT

The average news consumption via non-traditional sources such as social media, blogs and instant messaging groups has certainly gained a lot of interest in the recent time. One of the main sources of news content that people rely nowadays upon is via social media. Upon the release of a new data source, there are a set of issues associated with it. The rise of social media as a news source has the issue of generation and widespread of falsified information. This poses a threat not only specific to a platform in which the news content is published but also to the Internet as a source of medium for communicating news as a whole. In this project, we are trying to implement a proactive methodology to predict fake news that is prevalent in social media. The essential idea behind the system is to exploit the social relationship in social media to build a semi-supervised detection system.

Keywords: *Internet Social Media, Type of fake news, machine learning algorithm, natural language classification technique;*

INTRODUCTION

The usage of the web as a medium for perceiving information is increasing daily. The amount of information loaded in social media at any point is enormous, posing a challenge to the validation of the truthfulness of the information. Online media for news utilization is a twofold edged sword. From one perspective, its ease, simple access, and quick spread of data lead individuals to search out and devour news from online media. Then again, it empowers the wide spread of fake news, i.e., bad quality news with deliberately false information. The broad spread of fake news adversely affects people and society. Thus, fake news detection via online media has as of late become an arising research that is attracting tremendous attention.

The main reason that drives this framework is that on an average 62% of US adults rely on social media as their main source of news. The quality of news that is being generated in social media has substantially reduced over the years. For example, during the 2016 US presidential election, a lot of fake news about presidential candidates is spread on various social platforms as well [6], e.g., 115 pro-Trump fake stories that were shared on Facebook a total of 30 million times, and 41 pro-Clinton fake stories shared a total of 7.6 million times are observed in [7]. Such a huge amount of widely spread fake news have greatly destroyed the public persona of candidates and misled the judgment of voters. It has become very critical to detect fake news on social media in time to block the spread and refute them.

To begin with, fake news is deliberately composed to deceive readers to accept false information. The generation of fake news is intentional by the unknown sources which are trivial and there are existing methodologies to individually validate the users' trustworthiness, which makes it troublesome and nontrivial to distinguish dependent on news content; In order to detect fake news more effectively, it's necessary to mine meaningful information from different views instead of focusing on the news contents solely. In fact, fake news does not exist independently in the form of an article, like news creators and news subjects relating to news articles also exist in online social media. Thus, we need to incorporate auxiliary information, for example such as user social engagements on social media, to help make an assurance. Second, exploiting this auxiliary information is challenging in and of itself as users' social engagements with fake news produce data that is big, incomplete, unstructured, and noisy. Because the issue of fake news detection on social media is both challenging and relevant.

But, analysing these features individually doesn't consider the holistic factors of measuring the news credibility. Hence, combining the auxiliary information together with the news content to measure the news credibility is a possible route to focus. There have been techniques to validate the writing style of the users to classify the news content but these methods also have their outliers and error rates.

Why solving this problem is not an easy thing to do: as Chen et al. [9] pointed out that automatic detection of fake news is not an easy problem to solve since these days a news article generally comprises images and videos (as compared to only text), which is easy to fake. Moreover, with the social media on the rise, fake news stories are very reachable and have a very high impact factor. Also, fake news detection is difficult mainly because there is no governance in-place to control over what citizens can read and what carrier they are using to get that particular news nor who is behind that particular news story. It is safe to say that traditional printed media is slowly dying and every social media account has power to be the news writer/journalist. The challenge we have is, how do we as researchers produce a tool that can help readers of any type of content (i.e. news story) to detect if what they are viewing is fake or real. Before we start coming up with new solutions, it is necessary to survey state of the art techniques for learning purposes.

. LITERATURE REVIEW

The problem addressed is very relevant in this information age, several previous works have been carried out from different perspectives, focused on different ways and using various techniques, but ultimately all seek to combat misinformation; some of these studies will be presented below. Traditional approaches based on verification by humans and expert journalists do not scale the volume of the news content that is generated online . Text classification is the fundamental task in Natural Language Processing (NLP) and researchers have addressed this problem quite extensively . Researchers proposed a model that can check the real-time credibility within 35 seconds after combining user-based, propagation-based, and content-based text .The basic idea of Naïve Bayes is that all features are independent of each other . Naïve Bayes needs a smaller data set and less storage space. Facebook post prediction through real or fake labeling can be done through naïve Bayes and it performs well . A proposed method can separate fake contents in three categories: serious fabrication, large scale hoaxes and humorous fake . It can also provide a way to filter, vet and verify the news. PHEME was a three-year research project funded by the European Commission from 2014-2017, studying NLP techniques for dealing rumour detection, stance detection and , contradiction detection and analysis of social media rumours.

The problem of this news has been a matter in today's social media. There are researchers who have observed that these news are no longer a retain for advertising. The participation of users in spreading of news is also a concern for researchers. Earlier the motive of such hoax were just to attract web traffic and some monetary gains by using advertisements but these type of activities has emerged into the count that comprises dangerous organization and their intent can have nation- wide or world- wide effects or change in opinions of users or political leaders.

The most precise way is to input this news as a binary allocation problem. Hence all the information is divided in two parts (False or true) which is tough due to the fact there are cases where the information is partially actual and partially faux. To cope with this hassle, including further prospect is a regular exercise. These are the special paths that include extra training. The first one is: setting the class for the news which is hypothetical (not sure whether it is true or fake).

The second one is: setting levels which are greater than in the case of correctness. The Writers have given these fake news reverting work, where the output is in terms of numeric for correctness. Assigning the case in this order can make it less sincere to do the checking. There are many times where checking is done by calculating the difference between the for seen score and the ground score. Since the datasets which are available are having ground truth scores, there are challenges for how to change the different labels into numeric form. Some authors have also used Rhetorical Approach also known as Rhetorical Structure Theory (RST).This theory is used to identify relation between different words in a text. It identifies meaning in a text irrespective of its size and it is helpful in analysis of new texts.

METHODOLOGY

Data Exploration

The dataset used for classification was drawn from a public domain. Fake news articles were collected from an open source Kaggle dataset [33] that was published during the 2016 election cycle. The collection is made up of 18000 news articles highlighted. These articles were collected from news organizations NYT, Guardian, and Bloomberg during the election period. Articles are separated through binary labels 0 and 1. The dataset is already sorted qualitatively with fake, non-fake and not clear labels. This division can be seen. where we have 15,115 articles from the false category and 1,846 from the true category. The remaining articles are classified as not clear due to some other reasons like unique ID missing, source not clear etc. The task itself leads to a quite imbalanced dataset, wherefrom the total articles, roughly 12% are in the true category. This imbalance is typical in this task, and also seen in previous similar works [29], [30]. The second dataset contains 5000 real news articles collected from the Signal Media News dataset [25], in which 2,541 belong to the false class and 299 to the true class. We skipped the unclear class due to the missing values.

Models Description

Different classification models can be applied in this case, but to choose the most adequate one and to tune its parameters we run several experiments on different models. We started experimenting with classification models that have proven to be effective and give good results in related sentence classification tasks. Some of the models did not give good results and were discarded, one of them was Logistics Regression, while Support Vector Machines, naïve Bayes and Passive Aggressive gave promising results and we continued to experiment on them. To check the accuracy, we compare our results with other datasets through performance metrics.

- Naïve Bayes: It is a powerful classification model that performs well when we have a small dataset and it requires less storage space. It does not produce good results if words are co related between each other [18].the Naïve Bayes formula that explains the probability of an attribute that belongs to a class independent from other classes.

Our approach is to use this algorithm which despite its simplicity worked well for this application. In this theorem we find most the most suitable hypothesis (H) from given data. One of the easiest ways of selecting most precise hypothesis of the given data that one should try to use data that is already known to be true or false. Naive Bayes classifier provides a way that we can find out the probability of a given hypothesis from our prior knowledge.

Bayes' Theorem is stated as:

P (M|N) = P (N|M)*P (M)\P (N)

Where,

• P (M|N) is the probability of hypothesis M when we have already witnessed data N, which is also called the posterior probability (probability an event will happen after all previous information has been taken into account.) • P (N|M) is the probability of data N when we know that hypothesis M is true. • P (M) is the probability of the hypothesis M being true, that is calculated before new data is collected. • P (N) is the probability of the input data.

• Passive Aggressive: These algorithms are mainly used for classification [27]. The idea is very simple and the performance has been proven with many other alternative methods like Online Perceptron and MIRA.

CLASSIFICATION MODEL

TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY:

TF-IDF stands for Term Frequency-Inverse Document Frequency. In this model, the words are assigned a weight based on the frequency of appearance. The model has 2 parameters as mentioned in the name.

The term frequency component adjusts the weight proportionally with the number of times the word appears in the document with respect to the total number of words in that document. Inverse document frequency component identifies unique words in the set of documents and increases weight accordingly.

If a particular word is appearing in most of the documents, then its weight is reduced as it will not help anyway in distinguishing the documents. Though this model weights the words based on the frequency and unique factors, it is not able to capture the meaning of the word.

PROPOSED MODEL

Our proposed model starts with the extraction phase and then we have four main steps. The first step is related to the NLP models where we measure the frequency of words and build the vocabulary of known words in fake news datasets. Next, fake news is detected using NB, SVM and PA classifiers.

Finally, we test our models with several experiments and some other datasets and propose the final fake news detection model.

IMPLEMENTATION:

STATIC SEARCH IMPLEMENTATION-

In static part, we have trained and used 3 out of 4 algorithms for classification. They are Naïve Bayes, Random Forest and Logistic Regression.

Step 1: In first step, we have extracted features from the already pre-processed dataset. These features are; Bag-of-words, Tf-Idf Features and N-grams.

Step 2: Here, we have built all the classifiers for predicting the fake news detection. The extracted features are fed into different classifiers. We have used Naive-bayes, Logistic

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 06 | June - 2021
ISSN: 2582-3930

Classifier	Accuracy	Precision	Recall
Naïve Bayes	0.85%	0.89%	0.87%
Passive Aggressive	0.93%	0.92%	0.89%

Regression, and Random forest classifiers from sklearn. Each of the extracted features was used in all of the classifiers.

Step 3: Once fitting the model, we compared the f1 score and checked the confusion matrix.

Step 4: After fitting all the classifiers, 2 best performing models were selected as candidate models for fake news classification.

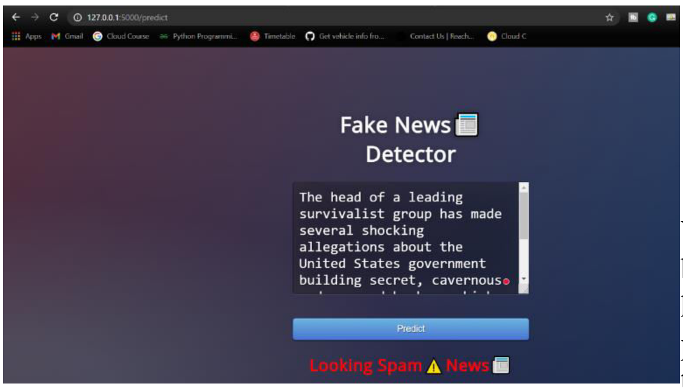
Step 5: We have performed parameter tuning by implementing GridSearchCV methods on these candidate models and chosen best performing paramters for these classifier.

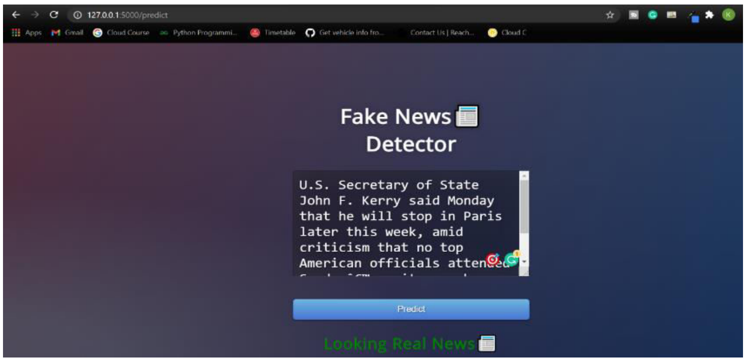
Step 6: Finally selected model was used for fake news detection with the probability of truth.

Step 7: Our finally selected and best performing classifier was Logistic Regression which was then saved on disk. It will be used to classify the fake news.

It takes a news article as input from user then model is used for final classification output that is shown to user along with probability of truth.

SCEENSHOT

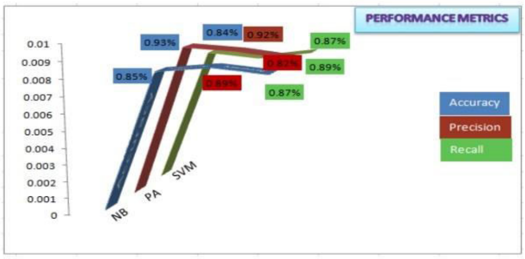




RESULT

We conducted several experiments with different feature set combinations as discussed in Section and the model selection in Section. Our proposed combination works well and obtains performance above the baseline 0.50. The best performing classifier is PA when we check the performance through accuracy and precision. However, somehow in the recall it reduced.

Fig. displays the performance metrics of all the classifiers. The next section will describe the results when we compare the proposed combination with other datasets and different classifiers but in the same domain.



We compare our results with the same model but different datasets and different features, as highlighted in Table. It is observed that the proposed models perform well and achieved the highest accuracy up to 93% with Passive Aggressive, 85% with naïve Bayes.

when they changed the Stylometric features, it achieved 84% accuracy [17]. On the other side, Horrne and Adali achieved 71% accuracy when they applied text-based features [31]. The results show that the proposed combination improves the existing performance in some categories. For further analysis, we applied different combinations to check the accuracy of the proposed model with other models.

CONCLUSION

Within the vast domain of social media and its issues, there are critical problems which are threat to social media as a platform. Among those, fake news is not platform specific and more critical because of its effect over the drop in platform usage. This is a framework, we are trying to build a proactive methodology which could leverage the logistical information for better classification. This method is not only specific to any platform or social media, but the same knowledge can be applied in the domain for Q&A forums and blogs such as StackOverflow, where the auxiliary users information is available. The future work for this framework involves integration with multiple platforms to know more about the user profile via a combinatorial credibility score. Also, making this score common across all platforms in the interest will have a more positive impact on the interest.

FUTURE WORK

The spread of misinformation has extreme harmful effects on users and social environment. Fake news is designed to mislead the user which makes it difficult to detect it in the first place. There are many sources from which fake news are spread which causes chaos among the people and society.

A Future enhancement would be to identify the source of the fake news and to stop the spreading of fake news in online platforms and in social media platforms. It would also have the capability to track and find the sources of these fake news so that we can stop the people who are trying to address the public with these malicious intent. They would also identify the social account of the people spreading fake news and roumers so that they can stop them before its too late. These all thing can engulf the society with positivity and a healthy life.

REFERENCES

[1] Beyond News Contents: The Role of Social Context for Fake News Detection; Kai Shu, Suhang Wang, Huan Liu; 2019 Association for Computing Machinery.

[2] Graph Attention Networks; Petar Velickovi, Guillem Cucurul, Arantxa Casanova, Adriana Romero, Pietro Lio, Yoshua Bengio; ICLR 2018.

[3] N. J. Conroy, V. L. Rubin, and Y. Chen, “Automatic deception detection: Methods for finding fake news,” Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, 2015

[4] S. Feng, R. Banerjee, and Y. Choi, “Syntactic stylometry for deception detection,” in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, Association for Computational Linguistics, 2012, pp. 171–175.

[5] Fake News Detection on Social Media: A Data Mining Perspective; Shu, Kai and Sliva, Amy and Wang, Suhang and Tang, Jiliang and Liu, Huan; ACM SIGKDD Explorations Newsletter 2017.

[6] Jain, A., & Kasbe, A. (2018, February). Fake news detection. In 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) (pp. 1-5). IEEE.

[7] Z. Jin, J. Cao, H. Guo, Y. Zhang, Y.

Wang, and J. Luo. Detection and analysis of 2016 us presidential election related rumors on twitter. In SBP-BRiMS, 2017.

[8] Michal Lukasik, Trevor Cohn, and Kalina Bontcheva. 2015a. Classifying tweet level judgements of rumours in social media. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP"15). 2590–2595. Michal

[9] H. Allcott and M. Gentzkow. Social media and fake news in the 2016 election. Journal of Economic Perspectives, 2017.

[10] Y. Chen, N. J. Conroy, and V. L. Rubin, "News in an online world: The need for an automatic crap detector," Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, 2015.

[11]] Kostakos, P., Nykanen, M., Martinviita, M., Pandya, A., & Oussalah, M. (2018, August). Meta-terrorism: identifying linguistic patterns in public discourse after an attack. In 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) (pp. 1079-1083). IEEE.

[12] V. L. Rubin, N. J. Conroy, and Y. Chen, "Towards news verification: Deception detection methods for news discourse," in Hawaii International Conference on System Sciences, 2015.

[13] Y. Chen, N. J. Conroy, and V. L. Rubin, "Misleading online content: Recognizing clickbait as false news," in Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection, pp. 15–19.

[14] B. Markines, C. Cattuto, and F. Menczer, "Social spam detection," in Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web. ACM, 2009, pp. 41–48.

[15] N. J. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, 2015.

[16] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.

[17] V. Perez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihal- ´ cea, "Automatic detection of fake news," arXiv preprintarXiv:1708.07104, 2017.