2021 International Conference on Technological Advancements and Innovations (ICTAI)

Fake News Detection using Machine Learning and Natural Language Processing

Vinit Kumar
Galgotias College of Engineering and Technology
Greater Noida,
Uttar Pradesh, India
dr.vinitkumar76@gmail.com

Arvind Kumar
Accentures Solutions Pvt. Ltd.
Noida, Uttar Pradesh
India
arvindkp504@gmail.com

Abhinav Kumar Singh
Pure Software Pvt. Ltd.
Noida, Uttar Pradesh
India
abhinavrajpoot09@gmail.com

Ansh Pachauri
Apeejay International School
Greater Noida, Uttar Pradesh
India
anshpachauri2005@gmail.com

2021 International Conference on Technological Advancements and Innovations (ICTAI) | 978-1-6654-2087-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICTAI53825.2021.9673378

***Abstract* — everyone depends on numerous online resources for news in today's world, where the internet is used as a medium. As the use of social media sites such as Facebook, Twitter, and others has grown, news has spread quickly between a large number of users in a tiny snap of time. The dissemination of false news has the implications, from swaying voting results in favour of some parties to creating skewed opinions. Furthermore, spammers use enticing news stories to generate revenue from click-bait ads. Our goal in this paper is to use AI, and ML principles to implement binary classification of different news contents available on social media sites. Our goal is to give the user the ability to identify news as fake or actual, as well as to verify the legitimacy of the website that published the news.**

***Keywords*— Fake-News, Social Media, classification, Machine Learning.**

I. INTRODUCTION

If we spend more time engaging online via social media sites, more people are using and gaining news from social-media other than traditional news sources. (i) Consuming news on social media is often more timely and less costly than conventional journalism, such as newspapers or television; and (ii) sharing, discussing, and discussing the news with friends or other readers is simpler on social media. For example, in 2016, 62 percent of adults in the United States received news from social media, compared to just 49 percent in 2012 [1].

This was also revealed that social media currently exceeds tv as a key news source. As per projections, almost one million posts have been related to fake news. The phrase "pizzagate" is often used to analyze a state in which there is a shortage of pizza "at the conclusion of the first round of the presidential campaign "False information" is a fair phrase to employ despite the abundance of this different concept." The Palgrave vocabulary was published in 2016.Fake news' widespread dissemination has the potential to damage both individuals and society.

For starters, fake news has the potential to disrupt the news ecosystem's credibility equilibrium; for example, during the 2016 presidential election in the United States, the most common fake news was much more widely shared on social-media. Second, fake news is designed to convince customers to believe misleading or inaccurate information. False propaganda is used by ideologues to promote political agendas or persuasion. as per some claims, Russia has created phoney websites and social chatbots to disseminate misleading information. Third, false news affects people's perceptions of and reactions to actual news.

Accessing news and information has become more easier and convenient. Thanks to the World Wide Web [2]. Sometimes, Internet users can follow up on events that affect them in an online format, however, with this kind of potential comes with great responsibility. To achieve those objectives, the media can manipulate information in a variety of ways. As a consequence, news articles that aren't entirely true or fully false are created. There are also blogs dedicated almost entirely to the dissemination of fake news.

People post contents on social media which cause hoaxes, fale-truths, false information, and representing false information as true news, the most common aim of websites with fake contents is to influence public opinion on specific topics.

As a result, fake news can be both a global problem and a global threat. Many scientists agree that machine learning and artificial intelligence (AI) may be used to combat false news [5]. There's an explanation for this: since hardware is cheaper and larger datasets are usable, there are a number of well-respected publications on the subject of automated deception detection. The authors of [6] provide a vivid overview of the different techniques which are used rapidly in this context. The authors explain their method for detecting fake news in [7], which is backed via feedback for specific news inside microblogs. The writers of [8] actually create two deception schemes.

978-1-6654-2087-7/21/$31.00 ©2021 IEEE

Gathering of data by requesting people to provide truthful or wrong information on a variety of subjects, including abortion, execution, and friendship. The detection accuracy achieved by the device is approximately 70%.

S-VM and the Naive Bayes classifier were used in deception detection systems (this approach is also used in the framework mentioned in this paper). They gather data by asking people to provide truthful or false information on a variety of subjects, including abortion, execution, and friendship.

The detection accuracy achieved by the device is about– Random Forest, Logistic Regression, and the nave Bayes classifier the aim of the study is to see how these methods perform in this situation, using the labelled data of datasets, and to help to implement the technique of using AI-tools for fake news identification. The implemented system was evaluated on a relatively new data-set, which allowed for a comparison of its output on recent data.

II. LITERATURE REVIEW

Mykhailo Granik-et-al show a naive Bayes classifier-based technique for identifying false propaganda in its work [5]. This approach was developed into a development platform and tested using a number of Facebook new articles. They originated from three major right-wing and left-wing Facebook pages, and also three significant domestic politics news sites. He was capable to get a classification precision of around 74.05 %. The categorization of blatant propaganda is slightly less accurate. It might be related to the regression coefficients of the dataset, which contains just 4.6 percent erroneous data.

They presented a method that employed a range of machine learning algorithms to solve concerns like accuracy and time latency. To continue, they mined the HSpam14-dataset collection for 0.4 million tweets. A total of 150K spam tweets and 250K non spam tweets are categorized. They use the Bag of Words methodology to parcelled, but also the Top-35 words with the highest knowledge gain. They are able to accomplish an accuracy of 90.64%, which was approximately 17% higher than the prior method.

By combining news community and educational environment functions, Marco L. Della Vedova et al. [12] developed first ever machine learning (ML) fake news system to identify, which improves existing techniques in the research and raises efficiency to 79.2 %.

Cody Buntain et al [13] employ a two plausible oriented Twitter datasets a crowd-sourced dataset's accuracy assessments for events on Twitter, and a collection of fake statements on Twitter and journalism estimations of its consistencies rate.

To create such programme, authors used Twitter dataset from Buzz Feed's hoaxes data collection. The features are described in a feature representation, only with conclusions aligning with prior studies. They concentrate on locating frequently reposted text topics and employing Twitter's features. Such strands were used to classify tales, limiting the focus of the study towards the most popular posts.

III. METHODOLOGY

This explains the technique, which would also be divided into three stages. That the very first portion is stable and relies on a machine learning classifier. They employed four classifiers for explore and validate the prototype while choosing the best model for the final implementation. Now next portion is more complicated, it collects the user's search terms and searches on the web for content that is likely to be true.

Throughout this study, we employed Python and also Sci-kit learn modules. Python has a vast component and extension library which may be employed in machine learning. The SciKit-Learn ML collection is a yet another resource for machine learning techniques, because it gives us access to nearly all sorts of Ml techniques and allows you to evaluate them quickly. We choose Django for the web-based environment that allows for front-end development using HTML, CSS, and JavaScript.

A. System Design

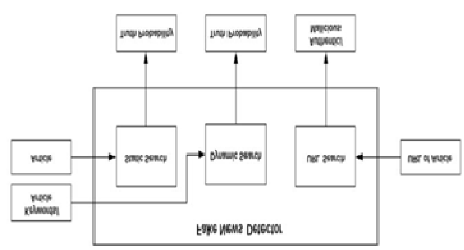

Figure 1: System Design

B. System Architecture

i. ***Static Search***: The static component of the fake news detection system's architecture is straightforward and follows the basic machine learning algorithm flow. The machine architecture is self-explanatory and seen below. The design's primary methods.

ii. ***Dynamic Search***: The site's second search area requests unique words needs to search over internet, and it returns a accurate percentage chance of the phrase being found in an article containing such words.

IV. IMPLEMENTATION

A. Gathering of Data and it's analysis

We may find news over internet from a variety of places, including social media platforms, search engines, news agency homepages, and fact-checking websites. There are different freely accessible databases for false news categorization on the Internet, such as Buzz feed News, LIAR [15], BS Detector, and others. This databases have been extensively used to determine the veracity of news in various academic papers. The origins of the dataset used in this study are briefly discussed in the following pages.

News can be found online from a variety of places, on the other hand, is a difficult challenge that normally necessitates domain experts doing close review of statements, additional facts, background, and information from credible sources. Expert writers, fact-checking blogs, industry scanners, and crowd-sourced staff are all popular sources of news data with annotations. However, no benchmark datasets for the identification of false news have been settled upon. Until it can be included in the training phase. Elections in the United States in 2016: This data was gathered using the PolitiFact API [15], a fact-checking database. It contains 13,836 labeled short statements drawn from a variety of sources, including press conferences. The USA 2016 Election Dataset dataset was used for this research, and it includes .csv files for test, practice, and validation. The data files used in this project are described in the following sections.

B. USA 2016 Election dataset for Fake News Detection

Features or column of dataset-

- Column 1: Title
- Column 2: Author
- Column 3 : Statement
- Column 4: Label (0,1)

C. Definition and Description

i. ***Data Processing**:* The social media data is in un-structured format in a bulk amount contact with typos, slang, and poor grammar, among other things [17]. In order to improve efficiency and reliability, strategies for using tools to make educated decisions must be developed [18]. Until predictive modelling may be used, the data must be cleaned in order to get deeper insights. Basic pre-processing of the News training data was performed for this reason. The following items were included in this step:-Data Cleaning: We get data in either a formatted or un-formatted format when we read it. Unstructured data lacks order, whereas structured data has a well-defined pattern. We have a semi-structured configuration in between the two structures, which is comparable to an unstructured manner in terms of structure.

To demonstrate properties that we want in ML frame-works to continue, dataset need to clean up the textual info. Cleaning or data preprocessing the data usually involves several steps are:

a. ***Remove Punctuation**:* Punctuation can help one interpret a sentence by providing grammatical meaning

b. ***Tokenization**:* Tokenization is the process of breaking down large amounts of text into smaller bits, such as sentences or phrases. It provides structure to previously unstructured text. For example, "Hello is world" means "Hello," "is," and "world."

c. ***Remove Stopwords**:* Stopwords are common words that may be found in nearly any writing. We remove them since they contain no information about our files. I'm ok with green or red, for example -> green, red, ok

d. ***Stemming**:* Stemming is a technique for reducing a word to its stem shape. Treating similar terms in the same manner makes a lot of sense. It uses a basic rule-based technique to delete suffices such as "ing," "ly," "s," and so on. It decreases the number of words in the corpus, but the individual words are often overlooked. Entitling, Entitled -> Entitle, for example. It's worth noting that some search engines consider terms of the same stem to be synonyms [18].

e. ***Vectorizing of Data: Using TF IDF**:* It calculates a word's "relative frequency" in a text which is compared to it's occurrence in all documents. The relative value of a word in the text and throughout the corpus is represented by TF-IDF weight [17].

Term Frequency (TF) is a calculation that determines how often a term occurs in a text. Since paper sizes vary

$$\text{TF (t,d)} = \frac{\text{(Occurrence of t in document d)}}{\text{(total word frequency in document d)}}$$

IDF stands for Inverse Text Frequency, which states that a term is of no value if it appears in any document. Certain words, such as "a," "an," "the," "in," "of," and others, occur often in a record but have no meaning. The value of these terms is reduced by the IDF, while the importance of unusual terms is increased. The higher the IDF meaning, the more special the term [17].

$$\text{IDF (t,d)} = \frac{\text{(Count of all documents)}}{\text{(Count of all documents with term-t in d)}}$$

$$\text{TF-IDF (t, d)} = \text{TF (t , d)} * \text{IDF (t , d)}$$

D. Algorithms used for classification

The classifier is trained in this part. To predict the text's class, various classifiers were investigated.-

i. Naïve Bayes Classifier: The Bayes theorem says that the presence of one function in a class does not preclude the inclusion of any other characteristic. The posterior probability calculated with

$$P(x) = P(c) * P(c) / p(x)$$

$P(c \mid x)$ = Posterior-probability
$P(c)$ = Prior-probability
$P(x \mid c)$ = likelihood
$P(x)$ = predictor prior-probability

ii. *Logistic Regression:* In basic terms, it fits data to a logit function to estimate the likelihood of an event occurring. As a result, it's sometimes called logit regression. Since it forecasts chance, its production values fall somewhere in the middle.

E. *Steps of Implementation*

i. *Static Search Implementation*

Phase 1: In the very first stage, we retrieved properties from either the pre-processed data from dataset. Illustrations of such attributes are bag-of-words, Tf-Idf Character traits, and N-grams.

Phase 2: Here is where we developed all of the classifiers for predicting the detection of fake news. The collected characteristics are given to several classifiers. The Naive-bayes, Logistic Regression, and Random Forest classifiers from sklearn were utilized. Each of the generated functions was utilized by all of the classifiers.

Phase 3: While estimating the model, we evaluated the accuracy rate and assessed the ambiguity matrix.

Phase 4: The two best-performing-models were designated as base classifiers for identifying false news for training all of the classifiers-models.

Phase 5: To feature tune these alternative prototypes and pick the best performing parameters for these classifiers, we employed Grid-Search-CV approaches.

Phase 6: Finally, based on the likelihood of actuality, the proposed model was utilized to recognize falsehoods.

Phase 7: The most recent and longest-running classifier, Logistic Regression, was stored to disc. It should be used to recognize false information.

F. *Evaluation Matrices*

The success of algorithms for the false news detection challenge was assessed using a variety of stable solution. We'll go through some of the most widely used criteria for spotting fake news in this part. The bulk of existing approaches address the false information problem as a classification problem, estimating whether such a news storey is true or not.

- When anticipated false news articles are actually labelled as fake news, this is called a True Positive (TP).
- When true news articles that were forecasted are truly categorized as accurate news, this is called true negative.
- When expected real news articles are labelled as misinformation is called False Negative.
- When falsified news is recognized as real news despite the fact that it was forecasted, this is called False Positive.

G. *Brief Description about Confusion Matrix*

A classifying model performance on a test dataset in which the real values are known is shown in a confusion matrix.

It allow the results about a procedure to be simulated. A classifying conundrum predicted conclusion is summarized in a confusion matrix.

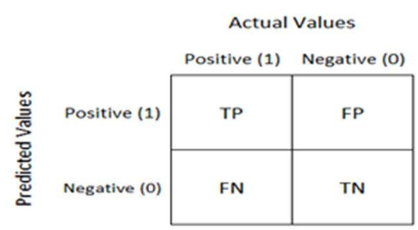

Figure-1: Confusion Matrix

We can describe the following metrics by recasting this as a classification problem:

Precision = True Positive / (Tue Positive + False Positive)
Recall = True Positive / (True Positive + False Negative)

i. *Static System*

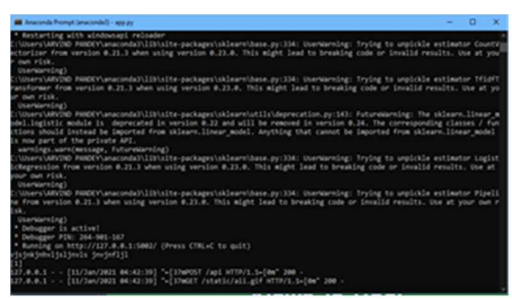

Figure 2: System View

ii. *Dynamic System*

Figure 3: System View

If News is fake

Figure 4: output (False)

If News is real

V. RESULT

At the Word and Ngram levels, the above algorithms were used for Vector features such as Count Vectors and Tf-Idf vectors. Both models were found to be accurate. To increase the models' efficacy, we used the K-fold cross validation method.

A. ***Dataset split into Training and Test:*** Using this cross-validation approach, the dataset was mainly split into k-folds. The model was built using (k-1) folds, and the model's efficacy was tested using the kth fold. This was done before each of the k-folds could be used as a test collection. For this experiment, I applied 3-fold cross validation, using 68 percent of the data to train the model and 32 percent is test data.

B. ***Confusion Matrices for Static System:*** Following the application of various derived features (Bag-of-words, Tf-Idf. N-grams) to three separate classifiers (Nave Bayes, Logistic Regression, and Random Forest), the uncertainty matrix displaying the real and expected sets is shown below:

TABLE II. CONFUSION MATRIX FOR NAÏVE BAYES CLASSIFIER

Total = 10240	Naïve Bayes Classifier	
	Fake (Predicted)	True (Predicted)
Fake (Actual)	941	4647
True (Actual)	527	6325

TABLE III. CONFUSION MATRIX FOR LOGISTIC REGRESSION USING TF-IDF FEATURES

Total = 10240	Logistic Regression	
	Fake (Predicted)	True (Predicted)
Fake (Actual)	2617	3871
True (Actual)	2097	5655

Our best model, as seen above, was Logistic Regression, which had a 72 percent accuracy. As a result, we might suppose if a consumer feeds a certain news story or head-line or title into our model, it has an 86% probability of being graded of its true existence.

TABLE IV. CONFUSION MATRIX FOR RANDOM FOREST CLASSIFIER USING TF-IDF FEATURES

Total = 10240	Random Forest	
	Fake (Predicted)	True (Predicted)
Fake (Actual)	2979	3509
True (Actual)	2630	5122

TABLE V. COMPARISON OF PRECISION, RECALL, F1 SCORE AND ACCURACY FOR ALL THREE CLASSIFIERS

Classifier	Precision	Recall	F1-Score	Accuracy
Naïve Bayes	0.69	0.79	0.81	0.63
Random Forest	0.72	0.81	0.78	0.68
Logistic Regression	0.90	0.74	0.86	0.72

VI. CONCLUSION

The bulk of activities in the twenty-first century are completed digitally. Newspapers, which were once preferred as physical copies, are now being replaced by online-only apps such as Facebook, Twitter, and news stories. Forwards from Whatsapp are another important source. The rising issue of false news further complicates matters by attempting to influence or sway people's opinions and attitudes. The bulk of activities are now completed digitally in the twenty-first century. Newspapers are being replaced by apps such as Facebook, Twitter, and web news posts, which were formerly preferred as hard copies. Forwards on Whatsapp is a significant resource of information. The growth of false news further complicates matters by attempting to influence or sway people's beliefs and attitudes. Different Natural Language Processing and ML approaches must be used to do this. The model training is done on a suitable data-set, and its accuracy is assessed a variety of performance metrics. To distinguish news headlines or posts, the best model is a model with the maximum accuracy need to be used. As a result, we might assume, if a consumer feeds a certain news storey or head-line into the model, it has a 88% probability of being listed of its true existence.

The customer will search the news storey or keywords online, as well as the website's authenticity. The dynamic system's accuracy is 93 percent, and it improves with each iteration.

We hope to create our own dataset, which will be updated as new information becomes available.

REFERENCES

[1] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, “Fake News Detection on Social Media: A Data Mining Perspective” arXiv:1708.01967v3 [cs.SI], 3 Sep 2017
[2] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, 2017, pp. 900-903.
[3] Fake news websites. (n.d.) Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Fake_news_website. Accessed Feb. 6, 2017
[4] Cade Metz. (2016, Dec. 16). The bittersweet sweepstakes to build an AI that destroys fake news.
[5] Conroy, N., Rubin, V. and Chen, Y. (2015). “Automatic deception detection: Methods for finding fake news” at Proceedings of the Association for Information Science and Technology, 52(1), pp.1-4.
[6] 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)
[7] Kushal Agarwalla, Shubham Nandan, Varun Anil Nair, D. Deva Hema, “Fake News Detection using Machine Learning and Natural Language Processing,” International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6, March 2019.
[8] H. Gupta, M. S. Jamal, S. Madisetty and M. S. Desarkar, "A framework for real-time spam detection in Twitter," 2018 10th International Conference on Communication Systems & Networks (COMSNETS), Bengaluru, 2018, pp. 380-383.
[9] M. L. Della Vedova, E. Tacchini, S. Moret, G. Ballarin, M. DiPierro and L. de Alfaro, "Automatic Online Fake News Detection Combining Content and Social Signals," 2018 22nd Conference of Open Innovations Association (FRUCT), Jyvaskyla, 2018, pp. 272-279.
[10] C. Buntain and J. Golbeck, "Automatically Identifying Fake News in Popular Twitter Threads," 2017 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, 2017, pp. 208-215.
[11] C. Buntain and J. Golbeck, "Automatically Identifying Fake News in Popular Twitter Threads," 2017 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, 2017, pp. 208-215
[12] S. B. Parikh and P. K. Atrey, "Media-Rich Fake News Detection: A Survey," 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, 2018, pp.
[13] Scikit-Learn- Machine Learning In Python
[14] Dataset- Fake News detection William Yang Wang. " liar, liar pants on _re": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648, 2017.
[15] Shankar M. Patil, Dr. Praveen Kumar, “Data mining model for effective data analysis of higher education students using MapReduce” IJERMT, April 2017 (Volume-6, Issue-4)
[16] Aayush Ranjan, “ Fake News Detection ,Engineering Delhi Technological University, July 2018.
[17] Patil S.M., Malik A.K. (2019) Correlation Based Real-Time Data Analysis of Graduate Students Behaviour. In: Santosh K., Hegadi R. (eds) Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2018. Communications in Computer and Information Science, vol 1037. Springer, Singapore.
[18] Badreesh Shetty, “Natural Language Processing (NLP) for machine learning” at towardsdatascience, Medium.
[19] NLTK 3.5b1 documentation, Nltk generate n gram
[20] Ultimate guide to deal with Text Data (using Python) – for Data Scientists and Engineers by Shubham Jain, February 27, 2018
[21] Understanding the random forest by Anirudh Palaparthi, Jan 28, at analytics vidya.
[22] Understanding the random forest by Anirudh Palaparthi, Jan 28, at analytics vidya.
[23] Dhama,“Detecting-Fake-News-with-Python”, Github, 2019
[24] What is a Confusion Matrix in Machine Learning by Jason Brownlee on November 18, 2016 in Code Algorithms From Scratch.