International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.5 Sept – 2021 DOI: 10.46647/ijetms.2021.v05i05.004
ISSN: 2581-4621

FAKE NEWS DETECTION USING LINEAR REGRESSION MACHINE LEARNING ALGORITHIM

Mr. Harjender Singh
Asstt. Professor, Maharaja Surajmal Institute, New Delhi
Email :harjendersingh@msijanakpuri.com

Abstract :

Fake news has become a significant part of social media and the political realm in recent years. Fake news detection is an important research project, but it comes with its own set of obstacles. Some difficulties may arise as a result of a scarcity of resources, such as a dataset and published literature.

The majority of smartphone users prefer to read news on social media rather than on the internet. The news is published on news websites, which also serve as a source of authentication. The challenge is how to verify news and articles shared through social media platforms such as WhatsApp groups, Facebook Pages, Twitter, and other microblogging and social networking sites. Believing in rumours and pretending to be news is damaging to society. Stopping rumours, especially in emerging nations like India, and focusing on accurate, verified news items is the need of the hour.

Keyword : WhatsApp groups, Facebook Pages, society, news etc

Introduction :

What is FAKE NEWS?

Fake news is a sort of yellow journalism that encompasses pieces of news that may be hoaxes and is widely disseminated via social media and other internet platforms. This is frequently done to promote or enforce particular beliefs, and it is frequently accomplished through political agendas. Such news articles may contain incorrect or exaggerated claims, and they may get viralized by algorithms, putting consumers in a filter bubble.

Fake news is raging throughout social media platforms in large quantities. The categorization of any article, post, story, or journal as false or true has become a significant issue in this case, and it has also piqued the interest of scholars all over the world. According to a number of research studies conducted to determine the influence of any false or fictitious news on us when we return via such fake news information, Individuals begin a fundamental mental process in one problem that may not be accurate because of falsified news or news that is exploited in this way.
The pandemic scenario that is affecting the entire planet is the finest illustration of fake news. Until now, there have been a variety of news items that have been fabricated and utilized solely to confuse and disrupt people's minds, leading them to believe false news. Can anyone tell if it's real or fake?

False information on Indian social media led to form voters drinking cow wee or eating dung to prevent illness, whilst in Country, artiodactyl weewee with lime was praised as a coronavirus protection. As a cure for the potentially deadly infection, the scientists combined several stories, such as ingesting garlic, wearing heat socks, and putting goose fat on one's chest. Conspiracy theories, such as the idea that it's a bioweapon financed by a corporation to boost antigen sales, have

Website: ijetms.in Issue: 5 Volume No.5 Sept – 2021 DOI: 10.46647/ijetms.2021.v05i05.004
ISSN: 2581-4621

been assembled and watched.

Anyone in today's world can publish information on the internet. Unfortunately, fake news attracts a lot of attention on the internet, especially through social media sites. People are misled, and they don't think twice about distributing such erroneous information to the farthest reaches of the system. Such acts are harmful to society since rumours or ambiguous news might dispel unfavorable thoughts among people or a certain group of people. Preventive actions are necessary to cope with such activities at the same rate as technology advances. Broad communications has a huge role in influencing the public, and as is customary, a few persons try to profit from it. There are several websites that provide inaccurate information. Under the guise of providing real news, they purposefully try to bring forth intended publicity, deceptions, and untruth. Their primary responsibility is to maintain control over the data that allows others to have faith in it. There are several examples of similar websites all around the world. As a result, false information has an impact on people's minds. According to a study, experts believe that a variety of man-made brain power calculations can aid in the discovery of fake news.
Fake news detection is used to prohibit rumours from spreading across multiple platforms, such as social media and messaging platforms. This is done to prevent the transmission of false information, which can lead to mob lynching.

★ Knowledge Styles in Social Media Posts

As previously said, social media networking sites can scan a location in three different ways:

Computational linguistics examines **Text (multilingual)**, focusing on the semantic and consistent origin of text. Because many of the postings are written in the form of texts, a great deal of effort has gone into analysing them.

Multimedia: During a single post, many types of media are combined. Audio, video, pictures, and graphics might all be included. This is highly appealing and draws the viewers' attention without requiring them to read the content.

Hyperlinks allow the post's creator to refer to a variety of sources, therefore gaining visitors' confidence by confirming the post's origin. Cross-reference to other social media networking sites and photo embedding is also noticed.

★ Facebook Works to Stop Misinformation and False News

According to a recent report, Facebook is working on two important areas to combat the spread of fake news.
The first step is to sabotage economic incentives, as most fake news is driven by financial gain.
The second is to create new goods in order to combat the spread of incorrect information. Here are some of Facebook's safeguards:

- Improved News Feed Rankings: The incidence of fake news material is reduced as a result of improved News Feed rankings.
- Easier Reporting: Figure out what is and isn't valuable. Our community labelled stories as fake, therefore they may appear lower in the user stream.

★ WhatsApp Works for Fake News Detection

WhatsApp has introduced certain security measures as well as false news identification to combat the spread of disinformation, however these are still in the alpha stage and are not yet available to

Website: ijetms.in Issue: 5 Volume No.5 Sept – 2021 DOI: 10.46647/ijetms.2021.v05i05.004
ISSN: 2581-4621

beta users. WhatsApp is experimenting with a new function called "Suspicious Link Detection." This feature will notify users by placing a red label on URLs that it recognizes as leading to a false or alternative website/news. Additionally, if a message has been forwarded more than 25 times from a device, it may be blocked.

★ **Outcome(of above social media FNDs)**

As stated in the preceding sections, all of the world's most powerful companies are attempting to keep themselves hidden from rumours, and the focus should instead be on real news and verified publications. Machine learning and natural language processing are, for the most part, the techniques used in the extraction. For the authentication of news stories, classifiers, models, and analytical algorithms must all operate together.

★ **Various Types of Fake News**

Below is a breakdown of the different types of fake news:

1. Visual-based: These fake news postings make extensive use of graphics, which might include altered images, doctored video, or a mix of the two.

2. User-generated news: This type of false news is created by phoney accounts and aimed at certain audiences, including age groups, gender, culture, and political affiliations.

3. Knowledge-based: these sorts of posts provide scientific (so-called) rationalisations for unsolved issues, leading people to assume they are genuine. Natural treatments for elevated sugar levels in the physical body, for example.

4. Style-based postings are authored by photojournalists from the UN agency, who are false and replica journalists.

5. Stance-based: It is a depiction of true statements in such a way that their meaning and intent are altered.

AIM/Objective

This paper intends to:

(1) Identifies a fake and true news detection using machine learning;

(2) Previous research have used machine learning to distinguish between false and real news; and

(3) Attempt to direct future study on the issue in this section, using past research works in the paper.

LITERATURE REVIEW

Their square assesses certain techniques that have been created to detect false news by evaluating lexical choices in headlines and various intensive language patterns (Chen, Conroy, and Rubin 2015b). Another technology (Atodiresei, Tnăselea, and Iftene 2018) created to detect false news on Twitter includes a component called the Twitter Crawler, which gathers and saves tweets in a form. When a Twitter user wants to check the accuracy of the news they've found, they'll copy a link into this app, where it will be analyzed for fake news identification. This method is made on associate degree rule is a known as the NER (Named Entity Recognition) (Atodiresei, Tănăselea, and Iftene 2018).
Their square measures several on the market approaches to assist the public to spot fake news and this paper aims to reinforce understanding of those by categorizing these approaches as found in existing literature.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.5 Sept – 2021 DOI: 10.46647/ijetms.2021.v05i05.004
ISSN: 2581-4621

METHODOLOGY

The aim of this research is to create a fake news detection model utilising three machine learning techniques. Because the focus is on model building in machine learning using jupyter notebook, this isn't continually developing new usual package systems. Machine learning normally necessitates a significant amount of time for model training and testing, as well as a large volume of high-quality dataset. In other words, if the model yields predictable results, such as the prediction of fake and actual news, we might consider it to be fairly accurate.

- Data Administration

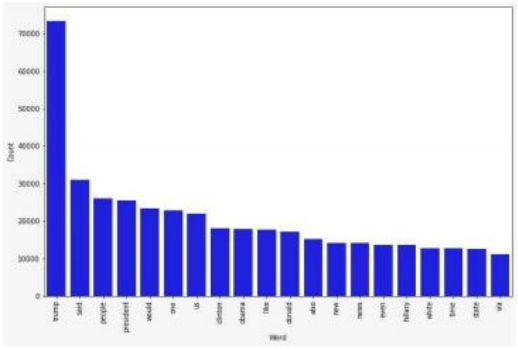
This section collects a body of knowledge (dataset), which could be a collection of report articles, stories, news, or blog postings. Once the dataset has been collected, nltk is used to identify a collection of written or spoken material stored on a computer and used to discover how language is utilised: the data is investigated to obtain a better understanding of its structure, which means stopwords are deleted.

- Data Exploration

The charting of graphs according to the fake and true news anticipated by the machine learning algorithm is the major focus of the information exploration section. Word clouds are created, which are essentially visual image ways for conveying text information in which the magnitude of each word represents its frequency or importance.
A word cloud is used to emphasise important information points about the subject. Tokenization is completed during this method.

- Model Training

The machine learning model can then be trained after the data has been adequately analysed and controlled. During the Model Training phase, many methodologies are considered, and a learning problem that is a prediction task is determined. Whatever possibilities are available within the training data set are then investigated. After that, the model is trained using an approved algorithm. We utilize three methods in this study: Logistic Regression,



Decision Tree Classifier, and Random Forest Classifier. The dataset is then matched into the algorithm's rule for training purposes, completing the testing.

- Model Assessment

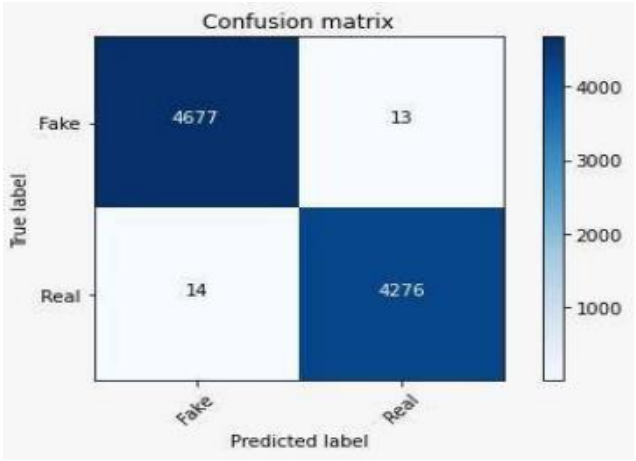
The output of the model generated is measured in many ways while evaluating it. The model's correctness is graded using performance measures such as F1 score, precision, recall, and accuracy rate, which are based on the confusion matrix report. Various adjustments are frequently made to the model till satisfaction is attained in terms of the model's creation yielding clever precision of output.

DISCUSSION & RESULTS

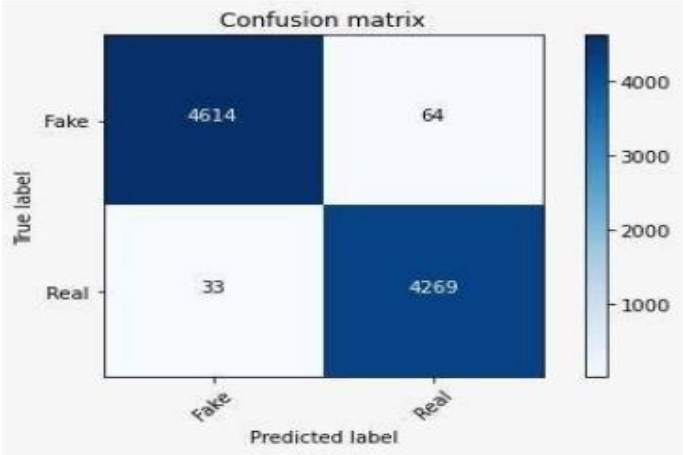
Each of the three machine learning algorithms employed in this research has its unique(own) accuracy percentage when applied to the dataset. The following are the accuracy estimates for each algorithm:

Website: ijetms.in Issue: 5 Volume No.5 Sept – 2021 DOI: 10.46647/ijetms.2021.v05i05.004
ISSN: 2581-4621

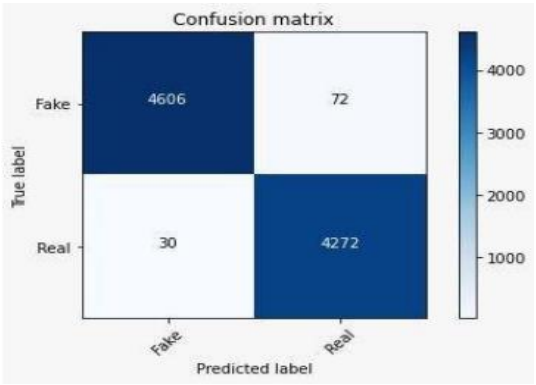
Classifier	Accuracy
Logistic Regression:	98.8%
Decision Tree Classifier	99.6%
Random Forest Classifier	98.9%



Confusion Matrix of Decision Tree Classifier



Confusion Matrix of Logistic Regression



Confusion Matrix of Random ForestClassifier

CONCLUSION & FUTURE SCOPE

Spreading false information has always had a harmful influence on society. When it comes to distinguishing between false and genuine news, there is still a lot of ambiguity in our culture. Fake news is a false alarm for anybody since it constantly misleads the audience, leaving them perplexed and unable to respond appropriately. They see their daily lives with their own eyes. So, this is when our research can be used to anticipate whether a given piece of news is false or not! People will be more conscious of fake news propagation if they consider the philosophy of our research report. This system was completed in the final year, but it will undoubtedly benefit from more enhancements in the near future, such as the use of a flask.

REFERENCES

1. https://searchenterpriseai.techtarget.com/definition/machine-learning
2. https://www.researchgate.net/figure/Pseudocode-for-KNN-classification_fig7_260397165
3. https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/

Website: ijetms.in Issue: 5 Volume No.5 Sept – 2021 DOI: 10.46647/ijetms.2021.v05i05.004
ISSN: 2581-4621

4. https://www.researchgate.net/figure/Pseudocode-of-naive-bayes-algorithm_fig2_325937073

5. Parikh, S. B., &Atrey, P. K. (2018, April). Media-Rich Fake News Detection: A Survey. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 436-441). IEEE.

6. Conroy, N. J., Rubin, V. L., & Chen, Y. (2015, November). Automatic deception detection: Methods for finding fake news. In Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community (p. 82). American Society for Information Science

7. Stahl, K. (2018). Fake News Detection in Social Media.