Fake News Detection with Naïve Bayes Classifier

Vishnuji Mishra[1], Manish Verma[2], Arpit Sharma[3]
[1,2,3]*Member, Galgotias University*

Abstract **- Fake news have a huge impact on society. Fake news is being disseminated through an online-based lifestyle to reach an open audience. People use their web-based social networks to represent the only reason to spread fake news and spread the flames of falsehoods. Now people every day use social networks too much to update news. These networks aim to make social life better. Today, everyone knows and uses social media that contains an unconfirmed article, post, message, and news. Today's false stories create a variety of issues from humorous stories to finding created stories and organizing government information in certain areas. Fake News and a lack of trust in the media exacerbate issues that have far-reaching implications for our society as a whole. It is necessary to consider how techniques in the field of computer science using computer-assisted reading techniques help us to find inaccurate information. Fake stories are now regarded as one of the greatest threats to freedom of speech, journalism, and the country's democracy. In this study, a comprehensive method of detecting false information was introduced using a machine learning model trained by Fake News data is based on the latest Indian Fake news. This problem is resolved by the Machine Learning technique. In this research we are analysing the false stories (fake news) related problem that is divided into two-part training and testing data set. After fitting Naïve Bayes Algorithm, we have find the 98% accuracy.**

Index Terms **- Naïve Bayes, Fake News, Machine Learning, Accuracy, Prediction.**

I.INTRODUCTION

The term "Fake News" was less common in the last few decades but in this age of digital communication, it has become a huge beast. Fake news, information bubbles, news fraud and media dishonesty are growing issues in our society. However, in order to begin dealing with this problem, a deeper understanding of false stories and their origins is needed. Only then can one look at the various techniques and fields of machine learning (ML) and artificial intelligence (AI) that can help us combat this situation. For example, Indian non-political news data that can be used for machine learning. Since the post-presidential election and India's upcoming Lok Sabha elections, a proliferation of false news can be seen throughout the Media and newsletter. Various corporate and political groups have begun to develop their own IT cell that tries to reach as many people as possible using fake pages and fake accounts on Facebook, twitter, Instagram WhatsApp and other apps. Now even the commercial media has started to come up with discriminatory stories and begin the practice of accepting the most profitable news or the people in charge of the newsroom for some political or financial gain. Due to the unavailability of systematic data on Fake stories related to the Indian context and very little analysis was done. So one of the purposes was to compile a collection of fiction and non-fiction after going through articles, blogs, research and other fake websites. The study will therefore test it with the old data set. Some insights can be made on how to catch false stories effectively using a machine that is trained in a database of false stories.

II. RELATED WORK

Several research topics are closely correlated with this paper, including fake news analysis, spam detection which will be briefly introduced as follows.

A. Fake News Preliminary Works

Due the increasingly realized impacts of fake news since the 2016 election, some preliminary research works have been done on fake news detection. The first work on online social network fake news analysis for the election comes from Allcott et al. [4]. The other published preliminary works mainly focus on fake news detection instead [4], [5], [6], [6]. Rubin et al. [53] provides a conceptual overview to illustrate the unique features of fake news, which tends to mimic the format and style of journalistic reporting. Singh et al. [5] propose a novel text analysis based computational

approach to automatically detect fake news articles, and they also release a public dataset of valid new articles. Tacchini et al. [6] present a technical report on fake news detection with various classification models, and a comprehensive review of detecting spam and rumor is presented by Shu et al. in [3]. In this paper, we are the first to provide the systematic formulation of fake news detection problems, illustrate the fake news presentation and factual defects, and introduce unified frameworks for fake news article and creator detection tasks based on deep learning models and heterogeneous network analysis techniques.

B. Spam Detection Research and Applications

Spams usually denote unsolicited messages or emails with unconfirmed information sent to a large number of recipients on the Internet. The concept web spam was first introduced by Convey in [11] and soon became recognized by the industry as a key challenge [11]. Spam on the Internet can be categorized into content spam [19] link spam [2] cloaking and redirection [9] and click spam [12]. Existing detection algorithms for these spams can be roughly divided into three main groups. The first group involves the techniques using content-based features, like word/language model [18] and duplicated content analysis [16], [17]. The second group of techniques mainly rely on the graph connectivity information [7] like link-based trust/distrust propagation [14] pruning of connections [6]. The last group of techniques use data like click stream [12] user behaviour [11], and HTTP session information [15] for spam detection. The differences between fake news and conventional spams have been clearly illustrated in Section I, which also make these existing spam detection techniques inapplicable to detect fake news articles.

III. DATASET COLLECTION

The News data set from the Kaggle is used for this study. The News dataset is divided into Training Data and Testing Data. Training Data Consist of 5456 records and 4 attributes. Testing Data consists of 2338 records and 4 attributes. The news contains attributes (URl, Headline, Body, Label).

A. Data pre-processing phase

In this dataset have some null values that can provide inaccurate calculation So first we will reduces the Null values row By Using this Statement: print (df.isnull().sum())

df.dropna(inplace=True

```
URLs          0
Headline      0
Body         21
Label         0
dtype: int64
```

```
: df.isnull().sum()

: URLs        0
  Headline    0
  Body        0
  Label       0
  dtype: int64
```

IV. BACKGROUND WORK

Finding false news on social media creates new problems and research challenges. While non-fiction is not a new problem - nations or groups have been using the media to propagate or influence work for centuries - the proliferation of web-based stories on social media makes false stories more powerful against traditional journalistic practices. There are several aspects of this problem that make it a different challenge to automatic detection. First, false stories are deliberately written to mislead readers, making them less likely to be easily found based on the content of the news. The content of false stories is very different in terms of topics, styles and forums, and non-fiction stories of trying to distort the truth in different language styles while at the same time making fun of true stories. For example, false stories can tell evidence of truth within an indirect context to support a false claim. Therefore, existing handmade and data-specific features are often not sufficient to detect false news. Some helpful information should also be used to improve access, such as the knowledge base and user engagement of the user. Second, using this helpful information actually leads to another critical challenge: the quality of the data itself. Fake stories are often related to recent, critical time events, which may not have been properly substantiated by the available information sources due to a lack of relevant evidence or claims. In addition, user engagement in social and non-productive issues produces large, incomplete, irregular, and noisy data. Effective methods of segregating trusted users, extracting useful post features and exploiting network connections are an

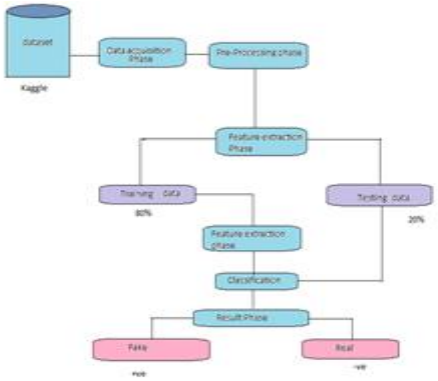
open area of research and require further investigation. Naive Bayes is also well known is used on text organisation. This article introduces a naïve Bayes Classifier for classifying news.

V. PROPOSED METHOD

The problem is not just with hackers, hacking accounts, and sending false information. The big problem here is why we call "Fake Stories". Where in the case of false stories: the story itself is made up of, without verified facts, sources, or quotations. When someone (or something like a bot) pretends to be someone or a trusted source in spreading false information, that can be considered false. In many cases, people who make up this false information have an agenda, which may be political, conservative, or morally degrading. There are countless sources of fake news these days, mostly from organized bots, which cannot get tired (they are machines) and continue to spread false information 24/7. The problem is real and it's hard to solve because the better bots deceive us. It is not always easy to find out if the information is true or not, so we need better systems that help us understand the patterns of non-communication issues to improve communication, communication and prevent world confusion [20].

Motivation and Justifications – Naive Bayes is an efficient eager learning classifier, which classification techniques is particularly used for real time applications since it is capable of producing instantaneous results in a modest amount of time, it does not requires large set of data for analogy. The estimation of test data can be carried out with a minimal amount of training data.

So here we can design a basic model for the detection of false information.



Fake News Discovery Model

This discovery model has designed for fake news detection. This model has four phases that are: Data acquisition phase, pre- processing phase, feature extraction phase and result phase.

A. Data set

This data set is cited from Kaggle and all details explained in above Dataset section [2].

B. Data acquisition phase

This phase is mainly used for gathering, filtering, cleaning the data before pre-processing phase in this phase have large amount of data is most commonly governed by four of the V's like: Velocity, volume, variety and value.

C. Pre-processing phase

This phase is a Process the data that coming from data acquisition phase. In this phase we have transfer the data into understandable format as: remove the noisy data, remove the null values etc.

D. Feature extraction phase

E. feature extraction phase is processed the data that coming from pre-processing phase. In this phase we are converting the data into features. Feature is a dimensional reduction that provide the simplest way to handling the dataset in Machine learning technique.

F. Training data

In this component we have trained the data and performing the some specific operation. Data set is divided in two format training and testing data that percentage ratio may be 80%:20%, 80 % is trained data set.

G. Testing data

In this component, we take the input and test how well the algorithm was trained that percentage ratio may be 20% out of total dataset.

H. Classification

In this phase we have characterised the model means here we describes the which type of model we are using, which is suitable for our problem So here we are using naïve bayes classifier model for fake news detection. Then after we will applying the NBModel() on the train data set.

I. Result phase

After applying the NBModel() we find the output of the problem that gives the output in label format : True/False.

V. METHODOLOGY

A machine learning method that is ready to check the reliability of news articles from Facebook using emotional analysis, The natural language method of transforming the native language into a specific format [20]. Using the convolution Neural Network, a method of dividing the Naive bayes is a method based on the use of non-news acquisition instead of using content-based methods [14]. The author introduced a set of features to measure prediction performance and non-automated news acquisition using the Lexical Features method, Syntax Features, Semantic Features, Language Features [18]. There are two main types of supervised learning where one is for supervised learning and the other is for non-supervised learning. Using the naïve Bayes separator we can use the fake news detection system [20]. Algorithm [8].

A. Algorithm

Naive Bayes is a machine learning model used for large data volumes (Big Data), even if you are working with data with millions of data values the method recommended by Naive Bayes. It provides very effective results when referring to NLP activities such as mood analysis. A quick and easy class algorithm. Bayes theory. So first we talk about Bayes Theorem.

Bayes Theorem:

The Bayes theorem works with conditional opportunities. So a conditional opportunity is a chance that something is going to happen, as something else has already happened. The Terms of Use may provide us with an opportunity for the event using its prior information.

Conditional opportunities:

$$P\left(\frac{H}{E}\right) = \frac{P(E|H) * P(H)}{P(E)}$$

Where,

P (H): Chances are H, if H is true. Known as previous opportunities.

P (E): Chances of E if E is true.

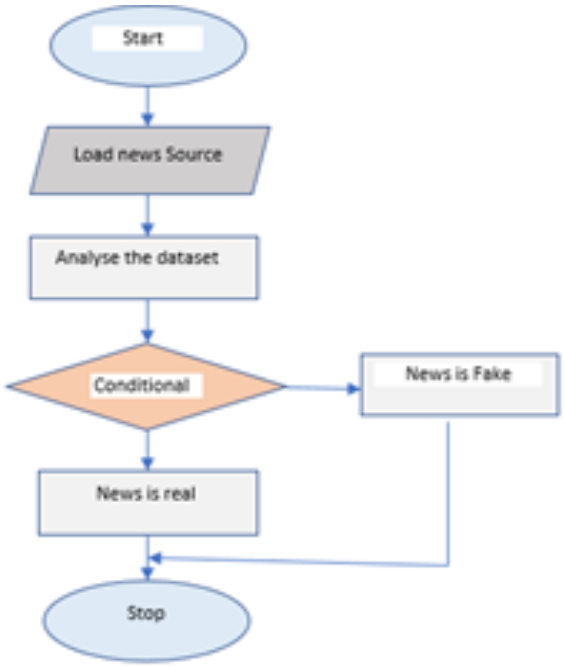
P (E | H): The probability of E when given that H is true.

P (H | E): The probability of H when given that E is true.

Naïve Bayes Division:

- It is the types of divisions that apply to the Bayes theorem.
- Prediction of membership opportunities is made in all classes as opportunities for data points related to a particular category.
- The high probability category is considered the most appropriate category.
- This is also called the Max A Posteriori (MAP).
- To make the result more common. The effect will not be affected by subtraction (E).
- NB separators conclude that all variables or unrelated features.
- The presence or absence of a variable does not affect the presence or absence of another variable.

B. Flow chart



VI. RESULT AND EVALUATION

This model provides a total accuracy of between 98-99%.

Here, we have to calculate the accuracy of the angry separator and get the 99.0% accuracy of the uncounted data.

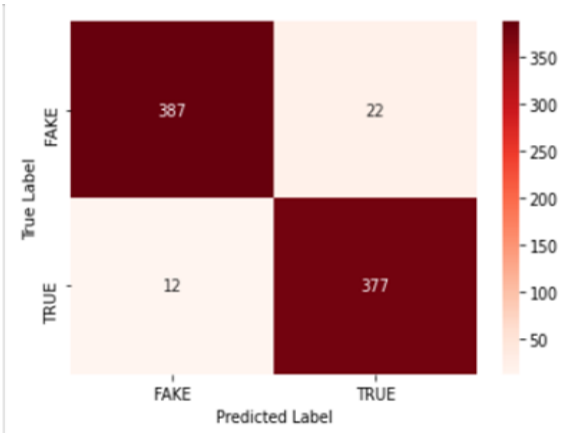
Also, here we use multinomialNB() to get the output as follows:

	precision	recall	f1-score	support
0	0.97	0.95	0.96	409
1	0.94	0.97	0.96	389
accuracy			0.96	798
macro avg	0.96	0.96	0.96	798
weighted avg	0.96	0.96	0.96	798

After calculating the result of the confusion matrix of finding the percentage of predictions means by using this method. So that, we get the accuracy of the prediction and how much data will be predicted by percent.

Output:

```
[[387  22]
 [ 12 377]]
```



This Matrix represents the confusion matrix that represents the model in 2 dimensional X and Y where X is a True Label and Y is a Prediction Label that are contains some specific values: True Positive, False Positive, False Negative, True Negative.

True Positive (TP) - The value 377 is represent the True positive values that means Prediction Label is Yes and True Label value is also Yes.

False Positive (FP) - The value 12 is represent the False Positive value that means Prediction Label is Yes but True Label is No.

False Negative (FN) - The value 22 is represent the False Negative value that means Prediction Label is No but True Label is Yes.

True Negative (TN) - The value 387 is represent the True Negative value that means Prediction Label is No and True Label is also No.

Then after we calculate the Accuracy By using this formula:

Total= TN+TP+FN+FP

Accuracy= (TN+TP)/Total

Where, A confusion matrix is a table that is used to describe the performance of a classification model on a set of test data.

VII. Interpretation of result

This model is designed to receive news, Whether the news is fake or not. If, the news is false it will show the output "False news". Otherwise, it will show the result “Real News".

First, we will enter the data into this model and then compare the data with the data already uploaded to the model using multinomialNB(). Where, After obtaining the accuracy of the entered data (input data). Create a confusion matrix that input values come from the NB separator.

After finding the result of the confusion matrix of finding the percentage of predictions means by using this method we can get the accuracy of the prediction, how much data will be predicted by percent.

Output:

```
[[387  22]
 [ 12 377]]
```

VIII.CONCLUSION

In this paper in Fake News Discovery model we can conclude that any News from large or small dataset can be classified as False News or by comparing it with previous database values in a short period of time which helps users to believe certain stories.

REFERENCES

[1] Marco L, E. Tacchini, S. Moret, G. Ballarin, “Automatic Online Fake News Detection Combining Content and Social Signals,” [Online]. Available: https://en.wikipedia.org/wiki/Fake_news_website. Accessed Feb. 6, 2017.

[2] Conroy, N. Rubin and Chen. Y, “CIMT Detect: A Community Infused Matrix-Tensor Coupled Factorization,” 52(1), pp.1-4, 2018

[3] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu. Fake news detection on social media: A data mining perspective. SIGKDD Explor. Newsl., 2017.
[4] V. Rubin, N. Conroy, Y. Chen, and S. Cornwell. Fake news or truth? using satirical cues to detect potentially misleading news. In NAACLCADD, 2016.
[5] V. Singh, R. Dasgupta, D. Sonagra, K. Raman, and I. Ghosh. Automated fake news detection using linguistic analy- sis and machine learning. In SBP-BRiMS, 2017.
[6] E. Tacchini, G. Ballarin, M. Della Vedova, S. Moret, and L. de Alfaro. Some like it hoax: Automated fake news detection in social networks. CoRR, abs/1704.07506, 2017.
[7] Pal, S., Kumar, T. S., & Pal, S. (2019). Applying Machine Learning to Detect Fake News. Indian Journal of Computer Science, 4(1), 7-12
[8] Sadia Afroz, Michael Brennan, and Rachel Greenstadt. Detecting hoaxes, frauds,and deception in writing style online. In 2012 IEEE Symposium on Security and Privacy, pages 461{475. IEEE, 2012.
[9] Satwinder Singh, R.Singla, "Classification of defective modules using object-oriented metrics, International Journal of Intelligent Systems Technologies and Application" Vol.16, No. 1, pp 1-13,2017.
[10] Y. Liu, M. Zhang, S. Ma, and L. Ru. User behavior-oriented web spam detection. In WWW, 2008.
[11] M. Henzinger, R. Motwani, and C. Silverstein. Challenges in web search engines. In SIGIR Forum. 2002.
[12] Z. Dou, R. Song, X. Yuan, and J. Wen. Are click-through data adequate for learning web search rankings? In CIKM, 2008.
[13] Fake news detection in social media Kelly Stahl * B.S. Candidate, Department of Computer Sciences, California State University Stanislaus, 1 University Circle, Turlock, CA 95382 Received 20 April, 2018; accepted 15 May 2018
[14] Z. Gyongyi, H. Garcia-Molina, and J. Pedersen. Combating web spam ¨ with trustrank. In VLDB, 2004.
[15] S. Webb, J. Caverlee, and C. Pu. Predicting web spam with http session information. In CIKM, 2008. [16]. evaluating machine learning algorithms for fake news detection, In IEEE 2017.
[16] FakeNewsTracker: A Tool for Fake News Collection, Detection, and Visualization, In IEEE 2017.
[17] S.Z. Mishu, S. Ra_uddin, in Computer and Information. In WebDB 2004.
[18] R. E. Schapire, "A brief introduction to boosting," IJCAI, vol. 99, pp. 1401–1406, 1999
[19] I. Drost and T. Scheffer. Thwarting the nigritude ultramarine: Learning to identify link spam. In ECML, 2005.
[20] D. Fetterly, M. Manasse, and M. Najork. Spam, damn spam, and statistics: Using statistical analysis to locate spam