Fake News Detection on Facebook

Md. Iftakhar Alam
American International University-Bangladesh
Dhaka, Bangladesh
mdiftakhar.alam722@gmail.com

Mushfika Jannat Bristy
American International University-Bangladesh
Dhaka, Bangladesh
mushfika022@gmail.com

Mosharaf Hossain
American International University-Bangladesh
Dhaka, Bangladesh
mosharaf111hossain@gmail.com

Md. Samiul Hauque Chowdhury
American International University-Bangladesh
Dhaka, Bangladesh
samiul.hauque98@gmail.com

Md. Asiful Islam
American International University-Bangladesh
Dhaka, Bangladesh
asiful.islam@aiub.edu

ABSTRACT

With the widespread use of social media, an additional plague has emerged: the spread of fake news. There have been numerous researches conducted on Twitter and other social media sites. Despite being one of the most popular social media sites, Facebook has undergone microscopic studies to detect fake news. The limited research that has been done on Facebook has not been able to come up with a viable answer to this problem. We collected user interactions data of Facebook posts and utilized them on machine learning algorithms to obtain reliable and accurate results. We used three classifiers algorithms and achieved a 96.414% accuracy by using the Decision Tree.

CCS CONCEPTS

• **Computing methodologies** → *Artificial intelligence*; **Supervised learning by classification**.

KEYWORDS

Fake News, Facebook, Social Media, Classification, Fact-checking Websites, Politifact, Lead Stories

ACM Reference Format:
Md. Iftakhar Alam, Mushfika Jannat Bristy, Mosharaf Hossain, Md. Samiul Hauque Chowdhury, and Md. Asiful Islam. 2018. Fake News Detection on Facebook. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The idea or concept behind fake news is not novel. It has been in existence before the inception of the internet but not at this massive scale. With an ever-increasing user base, social media platforms have become the go-to site for a quick injection of news or headlines from a variety of sources in a short time. According to 2017 research, nearly two-thirds of adults in the United States get their news from social media [20].

"Articles or contents that publishers purposefully produce to deceive or mislead the readers" is a generally accepted definition of fake news [5]. We can divide fake news into three categories [19]. Those are as follows: 1. Serious fabrications, i.e., deliberate fabrication or creation of an event that never occurred. 2. Hoaxes define the news that contains misinformation about an event with the intent to be picked up by trustworthy publishers. 3. Satires are created solely for the sake of amusement.

The widespread dissemination of fake news can have severe consequences for people and society. Fake news undermines the credibility of news organizations. It harms politics and other fields like health, science, sports, and financial markets [3, 11]. In the 2016 US presidential election, the widespread fake news mostly supported Donald Trump over Hilary Clinton. It had a large-scale implication on the election outcome [22]. Fake news can disrupt financial markets [9] and on medical science. It can be detrimental to public health, affecting health care and non-compliance with health requirements [10, 16]. For example, on October 14, 2021, a Facebook post stated that the "Corona PCR' test' is implanting a microchip," [1] which was fake, but the masses widely accepted the news. "During this coronavirus pandemic, 'fake news is putting lives at risk": UNESCO. So, it is essential to recognize fake news on social media because it negatively impacts our personal and social lives.

Researchers have used multiple approaches to detect fake news in their studies. We can divide the studies conducted in this domain into four categories. Two methods (Text and sentiment analysis) heavily rely on language to distinguish fake news. Thus, failing to perform when the news contains media files. Feature-based and Propagation approaches require a large quantity of data and are computationally expensive. Addiontionally, Propagation path needs an adequate number of user characteristics, and hackers can hack a user account objectively to disseminate fake news. For example, "In 2016 Russian hackers hacked 10,000 Twitter account of the defense department of USA and also created fake accounts and social bots to spread false stories" [7].

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

Despite having numerous studies on this area, very few researchers worked on the Facebook dataset rather, and they used Twitter, News Websites (Liar Dataset) dataset in their research. [26] used Facebook-like counts to analyze fake news on Facebook. Recently, Facebook has updated its reaction feature and added six new ways to interact or express users' feelings on a post [24]. A question of credibility remains as the authors of the papers who worked on the Facebook dataset did not publish their dataset.

Most researchers haven't looked at visual content on Facebook, e.g., Videos, images, and audio in their studies to create a reliable and robust dataset. Furthermore, they ignored post reactions and page engagements data in their studies. Since researchers had not focused on Facebook, we concentrated entirely on it. Because there had been no rich dataset on Facebook, we created a novel dataset (19 unique attributes) [13] containing post reactions and page engagements information. Our primary goal was to develop a content-independent model that could detect serious fabrications and hoaxes of any kind on Facebook as it is more efficient than a content-dependent model [15]. We achieved an accuracy of over 96% on our dataset.

2 RELATED WORKS

2.1 Text Analysis

One of the predominant methods is analyzing the text in the headlines or article body to find linguistic features such as n-gram matching lemmatized input using CoreNLP Lemmatizer [6]. Another method involves counting the number of pronouns while also looking at the writing style and checking the author's previous history, such as spreading false information [23]. To recognize fake news, [18] examines the text in the article body for qualities (n-grams, punctuation, psycholinguistic traits, readability, and syntax).

Text analysis is very dependent on the type of content; therefore, it cannot detect content devoid of textual elements or comprises audio, video, or photos, among other things. The perpetrators' can readily fabricate photos, videos, and audio files to advance their desire to deceive readers to achieve a specific aim. The nature of language further hinders it, which allows it to detect only one type of language. The inherent nature of this strategy necessitates a far more efficient and powerful system capable of accurately detecting any fake news.

2.2 Propagation and Hybrid Approach

A novel way is based on user characteristics analysis who is accountable for propagating the fake news, e.g., number of friends, followers, number of previous posts, location, registration date, etc. This approach can detect fake news at the early propagation stage, e.g., within the first five minutes [15]. One significant issue is authors did not investigate whether these eight user characteristics are adequate for machine learning algorithms. Furthermore, they did not elucidate whether the same approach is appropriate for Facebook since Facebook repudiates sharing users' data due to data security issues.

Hybrid approaches can detect fake news, e.g., by analyzing text content, users' characteristics, propagation path, and element-based [8]. [29] studied both content and propagation information using a supervised classifier and got outperformed. These approaches are time-consuming; instead, machine learning algorithms may come up with malfunction results if one or more features fail in the initial stage.

2.3 Features-based Approach

Features-based fake news research aims to identify valuable features for detecting fake news [21]. In a paper, researchers classified features into four classes: User ((age, account verification, number of followers, number of friends, date of account creation, etc.), Topic (Count tweets, average length, fraction re-tweets, etc.), Message (message provided in the news, words length, etc.), Propagation (depth of the re-tweet tree, max subtree, initial tweets, etc.) [8, 17]. In another paper, researchers used news content features (linguistic-based, visual-based, etc.) and social context features (user-based, post-based, network-based, etc.) [21]. [8, 21] extracted these features from Twitter and various news sources. [8] performed 3-fold cross-validation and learning schemes for feature normalization. Researchers trained datasets using SVM, Binary Classification Problem, and Deep Network Models (CNN, RNN, Geometric Deep Learning) [8, 17, 21].

Researchers have used Deep network models (CNN, Geometric Deep Learning, RNN) to detect fake news in recent years. They have certain downsides, such as requiring a large quantity of data to get higher performance than other methods and being very computationally expensive due to large and complicated data models.

2.4 Sentiment Analysis

Twitter [4] and other blogs such as Weibo [12] have been the subject of research to extract emotions or sentiments. Emotions were categorized and assessed based on several characteristics such as emotional intensity, category, and expression [12] by using the "Emotion-based Fake News Detection framework" to detect fake news on Weibo. [4] used the "PHEME dataset for rumor detection and veracity classification" for the analysis and LIWC (Linguistic Inquiry and Word Count) corpus to score a sentiment on Twitter. [14] trained DNN (Deep Neural Networks) with the Flair Library to detect fake news using sentiment analysis. Most recent research has relied on pre-built public datasets gathered from Twitter, although a few others have attempted to detect fake news on websites, blogs, and other platforms. Text and sentiment analysis depends on a specific language and fails to deliver when the news contains media files. The large number of characteristics required to attain increased accuracy is difficult to determine in the feature-based approach, making it computationally demanding. In the propagation method, whether a user is a fake news carrier requires an adequate number of characteristics to identify.

Despite being the most popular social media network regarding user count [2], few academics used Facebook to analyze fake news. According to an article on Forbes, Facebook spreads fake news faster than any other social media platform [27]. As a result, a new horizon in detecting fake news on Facebook has opened up.

3 PROBLEM DEFINITION

We assume that P is the number of Facebook posts sharing news or information on Facebook. Here, T = t1, t2, t3. . . ., tn and F = f1, f2, f3…., fn is a subset of N that represents true news and fake news respectively. Our model M will assess P to decide the subset.

4 DATA COLLECTION METHOD

We used several fact-checking websites to collect fake news ($politifact^1, republicworld^2, thequint^3, leadstories^4$). These sites provided the news or posts that they independently checked on their websites. Here, figure-1 depicts the data collection method.

Later on, we retrieved fake news from Facebook posts published on public pages. As for the true news, we collected it from prominent news pages on Facebook. Additionally, we collected the posts that had a significant amount of engagement. We gathered them between June 19, 2021, and August 30, 2016, using Facepager (Version 4.3) [30] and imported them as a CSV (Comma Separated Value) file. We split the dataset into two CSV files. The first contains general information about the Facebook post or page, such as the title, number of shares, total comments, etc. The latter only includes comments linked to the postings, each identified by a unique identifier. We checked every piece of data manually thus, taking a considerable amount of time to build this one-of-a-kind dataset.

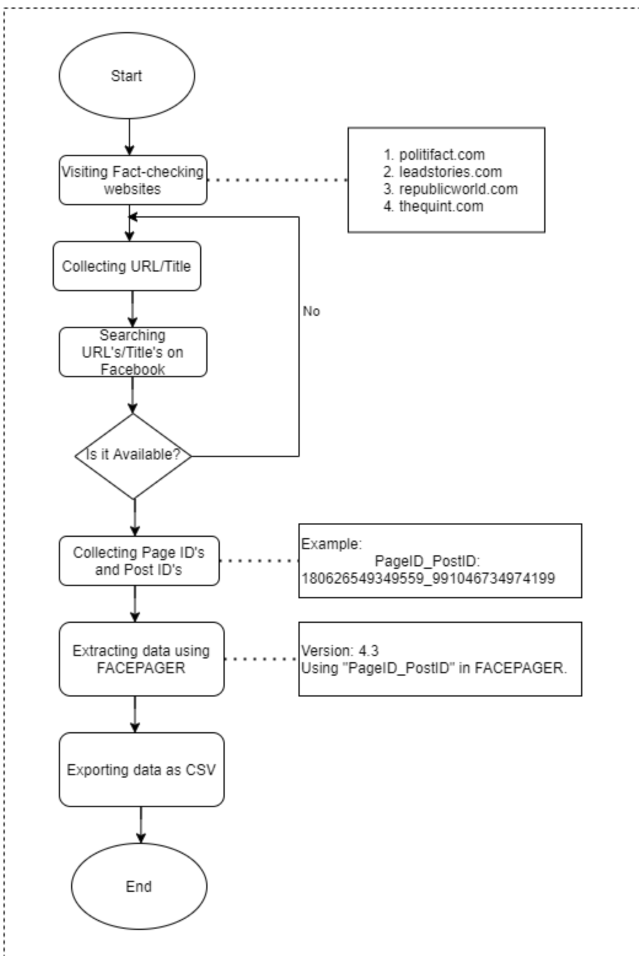

Figure 1: Data Collection Process

5 DATASET DESCRIPTION

The Dataset is named "Fake News Detection on Facebook." It encompasses 834 labeled data points, 614 "true" and 220 "false." There are 19 attributes grouped into two groups: post information and page information (Figure-2). The Post information segment contains general information about a post, i.e., type, URL, label, comment, share and reaction count, and six individual reactions. Page information includes the data regarding a page, i.e., name, about, fan count, followers. We identified the sections by a unique identifier named Object_id and id.

Group	Attribute
Page Information	Object_id
	name
	about
	category
	fan_count
	talking_about_count
Post Information	id
	type
	url
	Comment_count
	Share_count
	label
	reactions_count
	like_count
	love_count
	wow_count
	haha_count
	sad_count
	angry_count

Figure 2: Data Description

6 PROPOSED MODEL

This section consists of four parts describing the data source model to select the best model for the dataset (Figure-3).

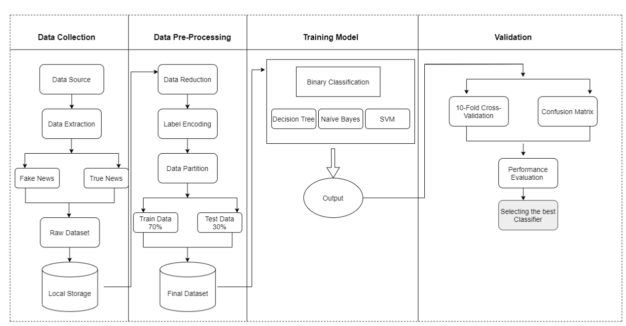

Figure 3: Proposed Model

6.1 Data Collection

This study used Fact-checking websites and prominent news pages on Facebook to retrieve fake and true news, and we extracted them using Facepager (Version 4.3) [25]. We merged fake and true news to create a single raw file for subsequent processing.

6.2 Data Pre-processing

After merging the raw dataset, there were 22 columns; some were unnecessary information, such as information unrelated to the posts or pages but associated with the application. We deleted them afterward and renamed the remaining 19 columns' names for better understanding. We utilized label encoding on the dataset to convert the columns from categorical (Boolean and Objective type) to numerical for the classifier and divided the Dataset into Training (70%) and Test (30%) data. Finally, the dataset was ready to be used to train a model.

6.3 Training Model

Classification is a method of separating two or more classes. There are three steps to it [28]. The initial step is to develop a training set containing data with known classes, and the next phase is to select the features relevant to the model. Finally, we will have to test a classifier's accuracy using test sets. To differentiate between true and fake news, we tested three classifiers to determine the best fit for our dataset.

6.4 Validation

This phase plays the most crucial role in objectively determining a model's performance and reliability. We applied two statistical validation approaches to assess a model's performance. To empirically compare these three models and select the best one, we used Confusion Matrix and K-Fold Cross Validation on the models.

7 RESULT

We have applied three classifiers to the dataset to distinguish fake news (Figure-4). Decision Tree scored the highest score among the three classifiers at 96.414%, whereas Gaussian Naïve Bayes scored the lowest, at 85.259%. Support Vector Machine (SVM) falls between Naïve Bayes and Decision tree at 87.649%.

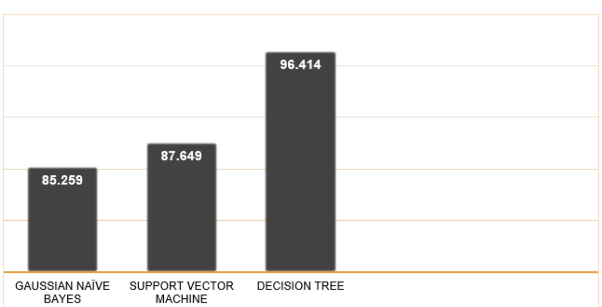

Figure 4: Classifiers Accuracy

We applied two statistical validation approaches, namely Confusion Matrix (figure-6), and K-Fold Cross-Validation (Figure-5),

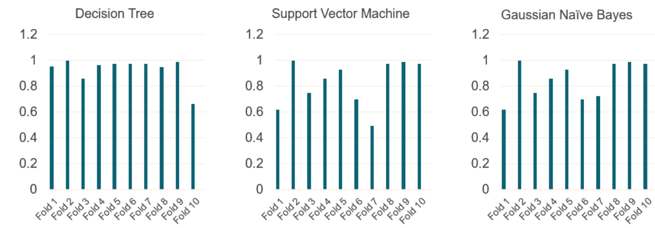

Figure 5: K-Fold Cross-Validation

Method	Class	Precision	Recall	f1-score	Support
Decision Tree	0	1	0.86	0.93	65
	1	0.95	1	0.98	186
	Accuracy			0.96	251
	Macro Avg	0.98	0.93	0.95	251
	Weighted Avg	0.97	0.96	0.96	251
Support Vector Machine	0	1.00	0.52	0.69	65
	1	0.86	1.0	0.92	186
	Accuracy			0.88	251
	Macro Avg	0.93	0.76	0.80	251
	Weighted Avg	0.89	0.88	0.86	251
Gaussian Naïve Bayes	0	1	0.43	0.6	65
	1	0.83	1	0.91	186
	Accuracy			0.85	251
	Macro Avg	0.92	0.72	0.76	251
	Weighted Avg	0.88	0.85	0.83	251

Figure 6: Confusion Matrix

to assess a model's performance and compare the three models empirically. Among the three models, the Decision Tree produced more consistent outcomes.

After examining the outcomes, we can state that the Decision Tree produces more accurate and reliable results among the three models.

8 CONCLUSION AND FUTURE WORK

To develop a content-independent model, we haven't examined the contents found in posts (Video, audio, image, etc.). We have collected the comments related to each post but weren't analyzed as we aimed to develop a content-independent model. Our model suffers from a cold start problem, i.e., it cannot detect fake news at an early stage (posts without any interactions from users). Due to Facebooks' policy regarding user data, we haven't been able to collect the data of users or groups; instead, we used the data from public pages for this study. Researchers can use a similar approach on other platforms to detect fake news.

1.https://www.politifact.com/
2.https://www.republicworld.com/fact-check
3.https://www.thequint.com
4.https://leadstories.com

REFERENCES

[1] [n.d.]. Politifact - Covid-19 tests are not part of a conspiracy to microchip people. https://www.politifact.com/factchecks/2021/oct/21/facebook-posts/covid-19-tests-are-not-part-conspiracy-microchip-p/
[2] 2021. Most used social media 2021. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

[3] Iftikhar Ahmad, Muhammad Yousaf, Suhail Yousaf, and Muhammad Ovais Ahmad. 2020. Fake news detection using machine learning ensemble methods. *Complexity* 2020 (2020).

[4] Oluwaseun Ajao, Deepayan Bhowmik, and Shahrzad Zargari. 2019. Sentiment aware fake news detection on online social networks. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2507–2511.

[5] Monther Aldwairi and Ali Alwahedi. 2018. Detecting fake news in social media networks. *Procedia Computer Science* 141 (2018), 215–222.

[6] Peter Bourgonje, Julian Moreno Schneider, and Georg Rehm. 2017. From clickbait to fake news detection: an approach based on detecting the stance of headlines to articles. In *Proceedings of the 2017 EMNLP workshop: natural language processing meets journalism*. 84–89.

[7] Massimo Calabresi. 2021. Russia's US Social Media Hacking: Inside the Information War. https://time.com/4783932/inside-russia-social-media-war-america/

[8] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web*. 675–684.

[9] Bryan Fong. 2021. Analysing the behavioural finance impact of'fake news' phenomena on financial markets: a representative agent model and empirical validation. *Financial Innovation* 7, 1 (2021), 1–30.

[10] Roberto Grilli, Craig Ramsay, and Silvia Minozzi. 2002. Mass media interventions: effects on health services utilisation. *Cochrane database of systematic reviews* 1 (2002).

[11] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. Fake news on Twitter during the 2016 US presidential election. *Science* 363, 6425 (2019), 374–378.

[12] Chuan Guo, Juan Cao, Xueyao Zhang, Kai Shu, and Miao Yu. 2019. Exploiting emotions for fake news detection on social media. *arXiv preprint arXiv:1903.01728* (2019).

[13] Iftyy007. [n.d.]. IFTYY007/Fake-News-on-facebook: Dataset for fake news on Facebook. https://github.com/Iftyy007/Fake-News-on-Facebook

[14] Sebastian Kula, Michał Choraś, Rafał Kozik, Paweł Ksieniewicz, and Michał Woźniak. 2020. Sentiment analysis for fake news detection by means of neural networks. In *International Conference on Computational Science*. Springer, 653–666.

[15] Yang Liu and Yi-Fang Brook Wu. 2018. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In *Thirty-second AAAI conference on artificial intelligence*.

[16] Anthony Matthews, Emily Herrett, Antonio Gasparrini, Tjeerd Van Staa, Ben Goldacre, Liam Smeeth, and Krishnan Bhaskaran. 2016. Impact of statin related media coverage on use of statins: interrupted time series analysis with UK primary care data. *bmj* 353 (2016).

[17] Federico Monti, Fabrizio Frasca, Davide Eynard, Damon Mannion, and Michael M Bronstein. 2019. Fake news detection on social media using geometric deep learning. *arXiv preprint arXiv:1902.06673* (2019).

[18] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2017. Automatic detection of fake news. *arXiv preprint arXiv:1708.07104* (2017).

[19] Victoria L Rubin, Yimin Chen, and Nadia K Conroy. 2015. Deception detection for news: three types of fakes. *Proceedings of the Association for Information Science and Technology* 52, 1 (2015), 1–4.

[20] Elisa Shearer and Jeffrey Gottfried. 2020. News use across social media platforms 2017. https://www.pewresearch.org/journalism/2017/09/07/news-use-across-social-media-platforms-2017

[21] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter* 19, 1 (2017), 22–36.

[22] Craig Silverman. 2016. This analysis shows how viral fake election news stories outperformed Real News on facebook. https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

[23] Kelly Stahl. 2018. Fake news detection in social media. *California State University Stanislaus* 6 (2018), 4–15.

[24] Liz Stinson. 2016. Facebook reactions, the totally redesigned like Button, is here. https://www.wired.com/2016/02/facebook-reactions-totally-redesigned-like-button/

[25] Strohne. [n.d.]. Strohne/Facepager: Facepager was made for fetching public available data from YouTube, Twitter and other websites on the basis of apis and webscraping. https://github.com/strohne/Facepager

[26] Eugenio Tacchini, Gabriele Ballarin, Marco L Della Vedova, Stefano Moret, and Luca de Alfaro. 2017. Some like it hoax: Automated fake news detection in social networks. *arXiv preprint arXiv:1704.07506* (2017).

[27] Mark Travers. 2021. Facebook spreads fake news faster than any other social website, according to New Research. https://bit.ly/3yopM6N;fbclid=IwAR2Bf_ubp42cFtdjtX1eOSaYSJdc8BJlReFcgwZok_qwWge805bWMIesRzk

[28] Richard L. White. [n.d.]. Steps in developing a classifier. http://sundog.stsci.edu/rick/SCMA/node3.html

[29] Xinyi Zhou, Atishay Jain, Vir V Phoha, and Reza Zafarani. 2020. Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice* 1, 2 (2020), 1–25.