Fake News Classification

Shipra Kumari
Boise State University
Boise, USA
shiprakumari@u.boisestate.edu

Uwaila Ekhator
Boise State University
Boise, USA
uwailaekhator@u.boisestate.edu

Desmond Kofi Boateng
Boise State University
Boise, USA
desmondboateng@u.boisestate.edu

ABSTRACT

With more news being shared on social media, it's becoming harder to tell what's real and what's fake. This project focus on developing a classification model to predict whether a given piece of news is real or fake. We looked at three major thing i.e, the content of the news, how users share the news, and the relationships between users who follow each other.By using these factors together, we can build a model that makes more accurate predictions about fake news. This project helps in the fight against misinformation and makes it easier to trust the news we see online.

1 INTRODUCTION

Fake news, often created with the intent to manipulate public opinion or generate ad revenue, can lead to widespread misinformation, influencing public perception, political outcomes, and social behavior. Therefore, identifying and combating fake news has become a crucial objective for researchers, journalists, and technologists alike.

This project aims to explore the distinguishing characteristics of real and fake news using various textual, linguistic, and sentiment-based features. By analyzing features such as text readability, keyword frequency, named entities, and emotional tones, we seek to understand the stylistic and semantic differences that can signal a news story's authenticity. Leveraging these features, the project will apply machine learning models to classify news articles as either real or fake, aiming to develop a reliable approach to automated fake news detection.

1.1 Problem Definition

In today's world of digital, Fake news is growing problems on Social media specially, where false information can spread quickly and influence people's opinions. Its hard to know whether a news article ois real or fake just by looking at text.Traditional methods of checking the truth of news can take a long time and aren't always effective.

A better solution is needed—one that not only looks at the news content but also considers how people share news and their connections on social media. By understanding both the news itself and the way people interact with it, we can create a smarter system that accurately predicts if news is real or fake. This project aims to build such a system by using the news content, how it's shared by users, and the relationships between those users.

2 RELATED WORKS

This section describes other research that have been done on classifying fake information. Several methods have been used to detect misinformation, these include feature-based methods, graph-based methods, and modeling-based methods [6]. We would focus on reviewing feature-based methods as that is the scope of this project.

2.1 Text-Based Feature Engineerng

In the domain of fake reviews and ratings, Ott et al. [8] found that "faked" writing styles are typically more exaggerated and consists of more verbs, adverbs, pronouns and pre-determiners. They also found that fake negative review writers exaggerate too negatively, including a large number of words which communicated negative emotion far more than normal reviews. In another fake reviews paper, by Harris [4], it was found that deceptive opinion spam is generally harder to read than truthful reviews (based on the Average Readability Index). These fake reviews are also more extreme in their emotions and opinions compared to real reviews, which aligns with findings from Ott et al.

In the context fake news detection, Silverman [12] discovered that about 13% of over 1600 analyzed news articles had incoherent headline and content body, for instance, some articles used declarative headlines paired with article bodies which expressed skepticism about the truthfulness of the information. [5], studied the textual characteristics of fake news, and compared them with Satirical news, they found that fake news article titles are usually very long and tend to compress the main claim into their titles, have frequently capitalized words to create emphasis, and closely resemble titles of satirical articles. In contrast, the main body of the articles are usually shorter, repetitive, and contain fewer nouns and analytical terms, making them easier to read.

Qazninian et al. [10] conducted an study on identifying Twitter rumors by manually annotating over 10,000 tweets. The authors utilized three feature categories: text-based (unigrams, bigrams, and part-of-speech), user-specific (prior false information), and Twitter-specific (hashtags and URLs). These features were converted into log-likelihood true/false ratios based on training data distribution, and combined for classification, achieving a mean average precision of 95%. It was also found that the content based features performed better than network and Twitter specific features. Another study on false news detection by Perez-Rosas et al. [9], utilized a large dataset of text based features for classification, consisting of n-grams, punctuations, LIWC, readability, and syntax features.

2.2 User-Based Feature Engineerng

According to [6], it is necessary to look beyond the appearance of fake texts, and look into who authored the text, in order to detect false information. Shrestha and Spezzano [11] demonstrated that user personality traits, emotions, and writing style are key indicators of fake news authors. In the task of classifying fake news spreaders, they achieved an average precision ranging from 0.80 on the PAN 2020 dataset to 0.99 on the FakeNewsNet-Politifact dataset, when they added demographic, behavioral, and network features, to the features mentioned earlier. Authors, Guess et al. [3], found political orientation, age, and social media usage to be the most relevant user features in their analysis of user demographics as predictors of fake news sharing on Facebook. . Kumar et al. [7] found that the creators of Wikipedis hoaxes typically have newer registered accounts and less editing experience.

2.3 Network-Based Feature Engineering

Starbird [13] found that false news domains form tightly connected clusters, which shows that users frequently mention these domains together in their false news tweets. Subrahmanian et al. [14] discovered that some bot accounts that spread misinformation cluster near one another, appear as groups in Twitter's follower-followee network, and have significant overlap between their followers and followees. In addition, Bessi et al. [1], applied the k-core analysis to twitter follower-followee network, and realized that the proportion of bots increases in higher network cores, suggesting that bots become increasingly central in the spreading of false information.

Zhou et al. [16] identified the following network-based fake news patterns: More-Spreader Pattern iė. more users spread fake news than true news, Father-Distance Pattern iė. fake news spreads farther than true news, Stronger-Engagement Pattern iė. spreaders engage more strongly with fake news than with true news, and Denser-Network Pattern which means fake news spreaders form denser networks compared to truth spreaders.

3 DATASET DESCRIPTION

In this project, we utilize comprehensive fake news detection datasets from FakeNewsNet to build the classification model. The dataset is collected from the fact-checking platform Politifact and contains both news content and social context information. This allows extracting features from the social network and the news items that will improve the classification of fake news. The news items are collected from X, formerly known as Twitter and the relationships between users who shared the news items are also shown in the data giving us insights into how fake news is propagated through social networks.

- News.txt: This file contains a list of 240 news IDs where each ID represents a particular news article and it is classed as whether it is fake or real. It uses the format "PolitiFact_Real_X' or 'Politifact_Fake_X' where X is the article number.
- User.txt: The file contains the unique user IDs of people who have interacted with the news content from Twitter. They are 23,865 in total
- FakeNewsContent: This folder contains a detailed metadata for each fake news article. The file is saved as a JSON file and each of them is named after the respective news ID in this format 'FakeNewsContent/news_id_Webpagejson'. The metadata contains attributes such as the news sources, headline, image, body_text, and publication date of the news item.
- RealNewsContent: This folder is similar to the FakeNews-Content folder. This folder contains metadata for real news articles that have been saved as JSON files. The metadata structure includes the same attributes as the fake news files.
- PolitiFactNewsUser.txt: This file contains the relationship between the users and the news articles that they have posted or shared. Each line shows the news_id refers to the news articles, 'user_id' refers to the user who posted the articles and 'post-count' indicates the number of times the user has shared that particular news item.
  The graph has 23,865 nodes and 32,791 edges. In total, there are 32,791 unique news-user relationships.
- PolitiFactUserUser.txt: This file records the relationship between the users as to which users are following which users. Each line is structured as 'follower_id followed_id' is the user's ID following another user, and 'followed_ID' is the user being followed.
  It is a directed graph with 23,865 nodes (users), and 574,744 edges (connections) between the nodes.

4 PROPOSAL OF FEATURES

4.1 Methodology and Feature Engineering

4.1.1 *News Features.*

- **Punctuation count-** Punctuation count can indicate the reliability of weather news. Excessive exclamation marks or question marks may suggest sensationalism or uncertainty, while proper punctuation enhances clarity. Analyzing punctuation can help develop features for fake news detection models, highlighting writing styles typical of credible versus dubious sources.
- **Number of nouns -** The number of nouns can affect fake news detection by indicating specificity and detail. Articles with many specific nouns may suggest credible reporting, while excessive vague or generic nouns can imply a lack of substance or sensationalism, potentially signaling less trustworthy information.
- **Number of pronouns-** The number of pronouns can indicate subjectivity and personal opinion. A high count of pronouns, especially first-person ones, may suggest a biased or emotional perspective, often found in fake news. Conversely, a lower pronoun count may indicate a more objective and factual reporting style, suggesting greater credibility.
- **Number of verbs-** A higher number of verbs can indicate action and engagement, often found in credible news. In contrast, an overabundance of verbs may suggest sensationalism or exaggeration, particularly in fake news.
- **Number of adjectives-** A high count of adjectives can imply subjective opinions or emotional language, which is common in sensationalized articles. Conversely, a lower count might indicate a more factual, straightforward reporting style, enhancing credibility.
- **Number of adverbs-** Excessive adverbs may signal attempts to add emphasis or emotion, often associated with less credible sources. Fewer adverbs can suggest a more objective tone, indicating potentially reliable news.
- **Number of stopwords-** Analyzing the ratio of stopwords (common words like "the," "is," "and") to meaningful words can provide insight into the text's clarity. A high ratio might suggest fluff or less substantive content, which could be a red flag for fake news.
- **Length of the title-** Short, catchy titles may indicate sensationalism, while longer, descriptive titles typically suggest more informative and credible content. Misleading or vague titles are often found in fake news.

- **Length of the text-** Short articles may lack depth and detail, often found in sensationalized fake news. Conversely, longer articles with detailed explanations and context are more likely to be credible, as they provide more information and support for claims.
- **Presence/Absence of Images-** Most reliable news sources tend to include images that support the content of the news or the story, especially for visually significant events and an absence of an image may suggest that these news sources are not doing their jobs well. In sharing fake news, most people tend to not add pictures as this can easily be detected through a reverse image search and the fake news can be detected. So most fake news sources do not tend to add images as adding an image serves as another layer of verification.
- **Number of Images-** It is possible that the news that contain more images are more reliable.
- **Sentiment of the news title and text-** Fake news articles often employ exaggerated language to grab readers attention and provoke strong reactions. By analyzing the sentiment of the news title and content, we can detect unusually positive or negative tones that deviate from neutral reporting.
- **Emotions of the news title and text-** Research shows that fake news often aims to elicit strong emotions such as fear, anger, or disgust to influence reader perceptions and encourage sharing. Also, fake news usually evoke more negative emotions than positive [5].
- **Emotion Intensity of the title and text-** Fake news writers usually employ intense emotions in their news to capture readers attention quickly, because high emotion intensity makes headlines and content more compelling. This makes readers to click, read, or share without critical evaluation. In addition, Giachanou et al. [2] found that incorporating emotion intensity features improved the performance of their credibility detection models, highlighting the effectiveness of emotion intensity in distinguishing fake content.
- **Flesch-Kincaid Readability Score of the Title and Text-** Horne et al. reported that fake news exhibits lower lexical diversity and complexity than real news [5]. On the other hand, Real news typically contains more complex sentences and vocabulary, reflecting professional journalism standards. This is therefore an important factor in differentiating fake news from real news.

4.1.2 Network Features.

- **Number of times a news article has been shared-** The number of times a news article is shared can help tell whether the news is fake or real. Fake news exploits sensationalism and polarizing topics to cause high spikes in shares. Real news tends to spread much more organically and therefore would not have astronomical shares or engagement. Hence analyzing the number of times a news article has been shared can help determine if it is a fake news or not.
- **Number of users that shared a news article-** The number of unique users who shared a news article can help determine whether a news item is fake or real. Fake news normally rely on orchestrated sharing from small number of users to be able to reach a wider audience quickly. Therefore, if a small number of users repeatedly share the same article, this could indicate that a fake news is being shared or there is a misinformation campaign. A high number of unique users sharing a story suggest that the new is legitimate and is gaining organic traction on social media.
- **Average number of followers for the users who shares a particular news-** The average number of followers for users who share a news item can help detect fake news. Fake news often spreads through accounts with lower follower counts as misinformation campaigns often use bots and fake accounts with low reach to boost the share numbers artificially. Real news, in contrast, is more likely to be shared by a mix of high and low follower accounts, especially from influencers or verified users who tend to endorse the credible content.
- **Average Node Similarity score among users who share the same news-** For each pair of users sharing the same news we would find the node similarity score, and then average it. Research has shown that Fake news often thrives in echo chambers (an environment or ecosystem in which participants encounter beliefs that amplify or reinforce their preexisting beliefs [15]). Users within these chambers will have high node similarity due to shared connections and interests.

5 EXPERIMENTS

5.1 Dataset Preparation

The data processing and cleaning phase involved several steps to ensure the data was suitable for analysis and feature extraction. Multiple libraries were utilized, including pandas, nltk, textstat, re, networkx, textblob, and ast. The initial step was to read all the JSON files into two CSV files. One contains the real news and the other contains the fake news. The CSV files,NewsContent.csv, News_User.csv, and User_user.csv contain the content of the news articles, the news user relationships, and the user id relationships, respectively.

5.1.1 Libraries Used for Data Cleaning and Feature Extraction: The analysis used various libraries that facilitated both data cleaning and feature extraction:

- **Pandas**: Essential for loading and manipulating the data
- **Numpy**: Utilized for numerical operations and managing arrays of data.
- **NLTK (Natural Language Toolkit)**: This library was used for tokenization, part-of-speech tagging, and sentiment analysis. The Vader sentiment analyzer from NLTK was used to extract sentiment scores for text and titles.
- **TextBlob**: Another NLP library used to extract named entities and perform basic sentiment analysis as a backup to NLTK.
- **Textstat**: Used for calculating readability scores such as the Flesch-Kincaid grade level.
- **NetworkX**: For analyzing social networks between the users and computing the average node similarity.
- **Re (Regular Expressions)**: used for text cleaning and pattern matching.

- **TQDM**: Added to display progress bars for loops to monitor the efficiency of operations.
- **WordNetLemmatizer (from NLTK)**: Used for lemmatizing words, reducing them to their base or root form, which is needed for analysis.
- **ast**: This library was used for safely evaluating strings.

We use various machine learning models to evaluate performance in classifying the news as either "Real" or "Fake." The detailed setup is described below:

- The data for the training of the model was loaded from `ExtractedFeatures.csv` using `pandas`.
- The labels were converted to binary. "Real" was assigned 0 and "Fake" was assigned 1.

Five different machine learning models were used for classification:

- **Logistic Regression**: with L2 regularization and class weighting to balance the dataset.
- **Random Forest**: A robust ensemble method with class weighting to handle class imbalances.
- **Decision Tree**: A simple and interpretable model, also configured with class weighting.
- **Gradient Boosting**: An ensemble method is known for reducing bias and variance.
- **XGBoost**: A highly efficient implementation of gradient boosting, optimized for performance.
- The dataset was divided into five folds using `StratifiedKFold` cross-validation to maintain class distribution across all folds.
- Each model was trained and evaluated five times, once for each fold, to ensure reliable performance metrics.

6 RESULTS

The following metrics were used to evaluate the performance of these models

- **Precision**: Measures the accuracy of positive predictions of fake news.
- **Recall**: Measures the ability of the model to accurately remember the positive outcomes or fake news.
- **F1 Score**: The harmonic mean of precision and recall.
- **Accuracy**: The overall correctness of predictions.
- **AUROC (Area Under the Receiver Operating Characteristic Curve)**: Indicates the model's ability to distinguish between classes.
- **Average Precision**: Evaluates the precision-recall trade-off.

The performance metrics of the classifiers evaluated in this experiment are shown in Table 1. The classifiers were evaluated on precision, recall, F1 score, accuracy, AUROC, and average precision.

Logistic Regression achieved a precision of 0.7387, a recall of 0.7083, and an F1 score of 0.7180, resulting in an overall accuracy of 0.7250 and with AUROC scores of 0.7809. Random Forest exhibits a notable improvement over Logistic Regression, achieving a precision of 0.8305, recall of 0.8167, and an F1 score of 0.8230, with an overall accuracy of 0.8250. Its AUROC of 0.8851 and average precision of 0.8786 shows that it performs better than logistic regression and is a more robust model. The Decision Tree classifier achieved a precision of 0.7882, recall of 0.7417, and an F1 score of

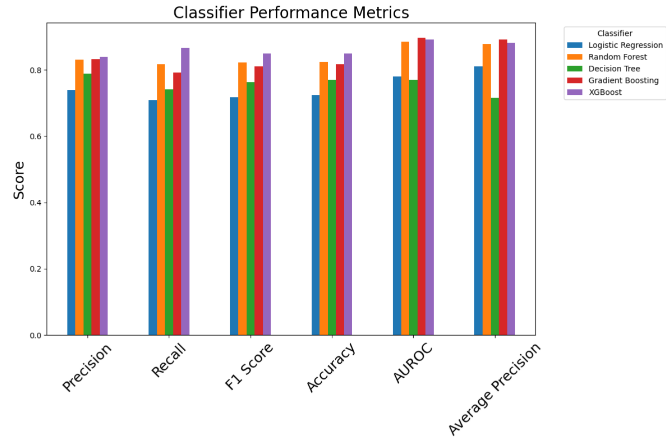

Figure 1: Confusion Matrix - Random Forest

0.7634, with an overall accuracy of 0.7708. Its AUROC of 0.7708 and lower average precision of 0.7154 indicate comparatively weaker performance than Random Forest. Gradient Boosting and XGBoost both delivered high performance, with Gradient Boosting achieving a precision of 0.8332, recall of 0.7917, and an F1 score of 0.8106, and XGBoost attaining a precision of 0.8391, recall of 0.8667, and an F1 score of 0.8504. Both models exhibited strong accuracy, at 0.8167 and 0.8500, respectively. The AUROC values of 0.8972 for Gradient Boosting and 0.8913 for XGBoost. Overall, Random Forest and the boosting techniques (Gradient Boosting and XGBoost) emerged as the top-performing models, Logistic Regression demonstrated solid but comparatively less optimal performance, a typical outcome for a simpler linear model.

7 CODE AVAILABILITY

The code for reproducibility of this work can be found here: link to GitHub repository.

8 CONCLUSION

In conclusion, this study demonstrates the potential of combining content-based, user-based, and network-based features to improve the accuracy of fake news detection models. By utilizing data from FakeNewsNet, we were able to capture not only the textual characteristics of news articles but also the network structure among users who share these articles. Our findings suggest that integrating diverse feature sets can enhance the performance of classification models, helping to distinguish fake news from credible information more effectively.

This study also contributes valuable insights to the fight against misinformation by emphasizing the importance of features beyond text alone. Future work can explore the use of transformer architectures, which have shown promising results in natural language processing, as well as the integration of image or video features for news articles with attached media to enhance detection accuracy.

Classifier	Precision	Recall	F1 Score	Accuracy	AUROC	Average Precision
Logistic Regression	0.7387	0.7083	0.7180	0.7250	0.7809	0.8103
Random Forest	0.8305	0.8167	0.8230	0.8250	0.8851	0.8786
Decision Tree	0.7882	0.7417	0.7634	0.7708	0.7708	0.7154
Gradient Boosting	0.8332	0.7917	0.8106	0.8167	0.8972	0.8925
XGBoost	0.8391	0.8667	0.8504	0.8500	0.8913	0.8822

Table 1: Performance metrics for different classifiers.

REFERENCES

[1] Alessandro Bessi and Emilio Ferrara. 2016. Social bots distort the 2016 us presidential election online discussion. *delete*. https://api.semanticscholar.org/CorpusID:20990413.

[2] Anastasia Giachanou, Paolo Rosso, and Fabio Crestani. 2019. Leveraging emotional signals for credibility detection. In (July 2019), 877–880. ISBN: 978-1-4503-6172-9. DOI: 10.1145/3331184.3331285.

[3] Andrew Guess, Jonathan Nagler, and Joshua Tucker. 2019. Less than you think: prevalence and predictors of fake news dissemination on facebook. *Science Advances*, 5, 1, eaau4586. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aau4586. DOI: 10.1126/sciadv.aau4586.

[4] Christopher G. Harris. 2012. Detecting deceptive opinion spam using human computation. In *HCOMP@AAAI*. https://api.semanticscholar.org/CorpusID:3652977.

[5] Benjamin D. Horne and Sibel Adali. 2017. This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. (2017). https://arxiv.org/abs/1703.09398 arXiv: 1703.09398 [cs.SI].

[6] Srijan Kumar and Neil Shah. 2018. False information on web and social media: a survey. (2018). https://arxiv.org/abs/1804.08559 arXiv: 1804.08559 [cs.SI].

[7] Srijan Kumar, Robert West, and Jure Leskovec. 2016. Disinformation on the web: impact, characteristics, and detection of wikipedia hoaxes. In (Apr. 2016). ISBN: 9781450341431. DOI: 10.1145/2872427.2883085.

[8] Myle Ott, Yejin Choi, Claire Cardie, and Jeffrey T. Hancock. 2011. Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Dekang Lin, Yuji Matsumoto, and Rada Mihalcea, (Eds.) Association for Computational Linguistics, Portland, Oregon, USA, (June 2011), 309–319. https://aclanthology.org/P11-1032.

[9] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2017. Automatic detection of fake news. (2017). https://arxiv.org/abs/1708.07104 arXiv: 1708.07104 [cs.CL].

[10] Vahed Qazvinian, Emily Rosengren, Dragomir R. Radev, and Qiaozhu Mei. 2011. Rumor has it: identifying misinformation in microblogs. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. Regina Barzilay and Mark Johnson, (Eds.) Association for Computational Linguistics, Edinburgh, Scotland, UK., (July 2011), 1589–1599. https://aclanthology.org/D11-1147.

[11] Anu Shrestha and Francesca Spezzano. 2022. Characterizing and predicting fake news spreaders in social networks. *International Journal of Data Science and Analytics*, 13, (May 2022). DOI: 10.1007/s41060-021-00291-z.

[12] Craig Silverman. 2015. Lies, Damn Lies, and Viral Content: How News Websites Spread (and Debunk) Online Rumors, Unverified Claims, and Misinformation. Funded by The Tow Foundation and the John S. and James L. Knight Foundation. Tow Center for Digital Journalism. https://www.cjr.org/tow_center_reports/lies_damn_lies_and_viral_content.php.

[13] Kate Starbird. 2017. Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, 11, 1, (May 2017), 230–239. DOI: 10.1609/icwsm.v11i1.14878.

[14] V.S. Subrahmanian et al. 2016. The darpa twitter bot challenge. *Computer*, 49, 6, (June 2016), 38–46. DOI: 10.1109/mc.2016.183.

[15] Wikipedia contributors. 2024. Echo chamber (media) — Wikipedia, the free encyclopedia. [Online; accessed 28-October-2024]. (2024). https://en.wikipedia.org/w/index.php?title=Echo_chamber_(media)&oldid=1244003932.

[16] Xinyi Zhou and Reza Zafarani. 2019. Network-based fake news detection: a pattern-driven approach. *ACM SIGKDD explorations newsletter*, 21, 2, 48–60.