# ENHANCING FAKE NEWS DETECTION THROUGH USER CREDIBILITY ANALYSIS

**Mr. AJEET KUMAR GUPTA[1]**
Assistant Professor,
Department of Computer Science & Engineering,
Darbhanga College of Engineering, Darbhanga, India.

**Mr. PRABHAT KUMAR[2]**
Assistant Professor,
Department of Electrical & Electronics Engineering,
Darbhanga College of Engineering, Darbhanga, India.

**Mr. SANJAY KUMAR[3]**
Assistant Professor,
Department of Electrical & Electronics Engineering,
Darbhanga College of Engineering, Darbhanga, India.

**ABSTRACT**
With the proliferation of information on the internet, the spread of fake news has become a significant societal concern. Traditional approaches to fake news detection primarily focus on analyzing the content of the news articles, neglecting the crucial role of user credibility in the dissemination of misinformation. This research proposes a novel approach to fake news detection by incorporating user credibility analysis as a key factor. The study leverages advanced natural language processing and machine learning techniques to assess the credibility of users sharing news content. Features such as the user's historical posting behavior, engagement patterns, and reliability of sources shared are considered. A comprehensive user credibility model is developed to evaluate the trustworthiness of information disseminated by individual users. To train and validate the proposed model, a diverse dataset comprising real and fake news articles, along with corresponding user information, is collected. The model is trained on a combination of content-based features and user-centric features, enabling it to discern patterns indicative of misinformation. The experimental results demonstrate that incorporating user credibility analysis significantly improves the accuracy and reliability of fake news detection compared to traditional content-based methods. By considering the credibility of users, the model can effectively identify and mitigate the impact of fake news propagated by users with a history of disseminating misinformation.

**Key Words: fake news detection, logistic regression, decision tree, random forest, support vector machine, user credibility**

## I. INTRODUCTION

The rapid evolution of digital communication platforms has transformed the way information is disseminated, creating both opportunities and challenges for society. One of the most pressing challenges is the proliferation of fake news, which can have profound consequences on public opinion, decision-making processes, and even democratic institutions. Conventional approaches to fake news detection often rely on content analysis, overlooking a crucial aspect of the dissemination process – the credibility of the users sharing the information. In the digital age, social media platforms and online forums serve as major conduits for the spread of news and information. Users play a pivotal role in shaping the online narrative by sharing, commenting, and endorsing content. However, not all users are equally reliable or trustworthy. Some individuals may have a history of sharing misinformation, while others consistently contribute to the dissemination of accurate and credible information. Recognizing the importance of user credibility in the fight against fake news, this research proposes a novel approach that integrates user-centric analysis into the detection process.

The key motivation for considering user credibility lies in the understanding that individuals with a history of spreading misinformation may have a higher likelihood of sharing fake news in the future. By assessing various aspects of a user's behavior, such as the consistency of their information sources, posting patterns, and engagement with reputable content, a model can develop a comprehensive model that goes beyond content analysis alone.

Traditional fake news detection methods often rely on linguistic and semantic analysis of the news articles themselves, assessing factors like language patterns, sentiment, and source reputation. While these approaches have proven valuable, they may not be sufficiently robust in an environment where misinformation is strategically designed to mimic genuine news. Integrating user credibility analysis adds an additional layer of sophistication to the detection process, considering the context in which information is shared and the reputation of the users involved.

This research builds upon existing literature in fake news detection and user behavior analysis, aiming to bridge the gap between content-based approaches and the inherent social dynamics of information dissemination. The proposed model seeks to enhance the overall accuracy and reliability of fake news detection, offering a more nuanced understanding of the complex interplay between content, users, and credibility in the digital information landscape.

In the subsequent sections, we will delve into the methodology employed for user credibility analysis, the dataset used for training and validation, and the results obtained through experiments comparing the proposed approach with traditional content-based methods. This research contributes to the ongoing discourse on combating fake news by emphasizing the importance of considering the credibility of users as an integral component of effective detection strategies.

## II. OBJECTIVES

By emphasizing unique user traits, user credibility-based fake news detection seeks to improve the precision of detecting false information. This methodology develops algorithms that examine user behavior, engagement patterns, and historical reliability in an effort to reduce the spread of fake news on online platforms. The intention is to develop a personalized and context-aware fake news detection system, taking into account the fact that user credibility differs depending on the subject and domain. In the never-ending fight against fake news, the ultimate goal is to equip users, social media platforms, and content aggregators with tools that evaluate the reliability of information sources and the individual credibility of users.

## III. RELATED WORKS

Author [1] demonstrates through the analysis of publicly available false news data that news source information can serve as a powerful barometer of credibility. The results imply that the number of authors and an author's past history of being associated with fake news can be important factors in identifying fake news.

In a study by the author [2], 574 participants evaluated the credibility of the message, the site, and the sponsor for each of the four types of websites. The websites of news organizations received the highest credibility ratings, followed by those for e-commerce and special interests. The main causes of credibility evaluations were characteristics of the website, like design elements and in-depth material. There was a negative connection between self-reported and observed information verification behavior and a positive relationship between self-reported verification and internet/web experience. The results are applied to the theoretical development of web credibility perception.

The author [3] talks about automated techniques to evaluate the veracity of micro-blog entries about popular subjects. The techniques examine post text, external source citations, and user posting and re-tweeting behaviour. Using human assessments on a current sample of Twitter posts, the study assesses different approaches. Findings demonstrate quantifiable variations in the spread of messages, enabling automated categorization of messages as reliable or unreliable with accuracy and recall varying between 70% and 80%.

Author [4] evaluates the dissemination of propaganda, misinformation, and disinformation on online social networks using concepts from cognitive psychology. It finds metrics to deduce signs of dishonesty and gauges the spread of false information. Four primary questions must be addressed during the cognitive process: general acceptability, coherence, consistency, and believability. To help consumers make wise judgements, the authors suggest an algorithm to identify purposeful dissemination of erroneous information. The programme gauges news quality and credibility using collaborative filtering properties. Twitter was used for validation.

In the context of political fact-checking and fake news identification, author [5] presents an analytical analysis on news media jargon. We compare the language of real news with that of satire, hoaxes, and propaganda to uncover linguistic characteristics of untrustworthy material. Present a case study that utilizes PolitiFact.com 6-point factuality judgment scale to examine the viability of computerized political fact-checking. Experiments reveal that stylistic signals can aid in determining the veracity of writing, even though the subject of media fact-checking is still up to debate.

People's opinions about the reliability of information found on the internet in comparison to other media sources were examined in the study [6]. Newspaper material was deemed less reliable by 1,041 respondents, whereas information from the internet was regarded as credible as that from radio, television, and magazines. Media outlets differed in their level of credibility; news and entertainment were the most reliable. Web-based information was rarely confirmed by respondents, and the likelihood of verification depended on experience and the credibility of the material perceived. The paper addresses these findings in terms of theoretical understanding of cutting-edge communication technologies and investigates their social significance.

The impact of the digital transition on news markets and news quality is examined by author [7]. There is disagreement over what constitutes fake news when comparing various definitions of the term. According to survey data, conventional news publishers are highly trusted, while algorithm-driven news delivery methods like social media and aggregators are less trusted. Still, two thirds of customers use these platforms to get news. Reputable newspapers get sizable readerships, and the amount of genuine news that is consumed on these platforms outweighs the amount of bogus news. Platforms for distributing news powered by algorithms lower the barriers to entry and increase the market reach of both writers and readers. But they distinguish content creators from curators, frequently optimizing traffic and advertising income. This could result in failures of the news market by undermining the function of reliable editors as high-quality middlemen.

The comparative study [8] shows the efficiency of different machine learning classification for fake news detection. The study shows that Support Vector Machines (SVM) and Decision Trees perform better than other models in terms of accuracy and performance measures, and archive 97% of accuracy.

## IV. METHODOLOGY

In the context of identifying fake news, it can typically be viewed as a two-category classification task. The objective is to categorize posts from social media platforms as either genuine or fake news. Specifically, if one takes a collection of multi-modal posts from these platforms, denoted as $P= P_1, P_2 ... P_n$, each post $P_i$. From individual publisher information the user credibility of that publisher $C_i$ has been calculated. Given the total number of posts as $N_p$, the goal is to derive a function f that maps to $P_i$ and $C_i$ an output set Y. This output set Y contains two values, 0 and 1. Here, 1 represents fake news, and 0 stands for real news. The overall flow of this study is shown in Fig. 1.

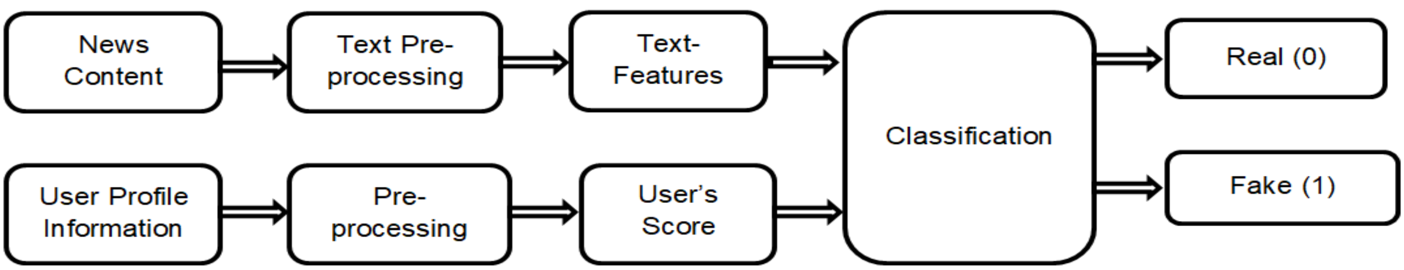


Fig. 1 overall flow of the study

In this paper two datasets have been taken, the first one is the user profile data created manually and second is post data taken as ISOT news dataset. User profile data contains the details of user profile and post contains the details of post.

**User profile dataset (Created manually)**

It was gathered from college students using Google Forms. Required fields included:

- User-id: typically used to uniquely identify users; short text
- User-name: display name of user
- Gender: (Female, male, other)
- Date of Birth (Month, day, year)
- Email address:

The optional fields were:

- Mobile no.
- Hometown
- Current residence
- Profile photo

The sample of the user profile is not being shown due to personal information.

**News dataset (ISOT news dataset)**

Dataset of news from ISOT [9-10] which include:

- Title
- Text
- subject
- Date published

| | title | text | subject | date | target |
|---|---|---|---|---|---|
| 10499 | 'Joe the Plumber' praises Trump, cites his 'be... | HOLLAND, Ohio (Reuters) - “Joe the Plumber,” t... | politicsNews | March 4, 2016 | 0 |
| 1076 | Factbox: Trump on Twitter (Oct 24) - Bob Corke... | The following statements were posted to the ve... | politicsNews | October 24, 2017 | 0 |
| 140 | Factbox: Trump on Twitter (December 14) - Stoc... | The following statements were posted to the ve... | politicsNews | December 15, 2017 | 0 |
| 32954 | HILLARY VOLUNTEER, Obama Supporter Mocks Ryan ... | Speaking of not having a soul..check out this ... | politics | Feb 28, 2017 | 1 |
| 28877 | Ben Carson Praises Time Trump Compared Him To... | The only thing sadder than losing a Republican... | News | March 16, 2016 | 1 |

Fig. 2 news dataset after merging and labeling

| | title | text | subject | date | target | UID |
|---|---|---|---|---|---|---|
| 10499 | 'Joe the Plumber' praises Trump, cites his 'be... | HOLLAND, Ohio (Reuters) - “Joe the Plumber,” t... | politicsNews | March 4, 2016 | 0 | 32 |
| 1076 | Factbox: Trump on Twitter (Oct 24) - Bob Corke... | The following statements were posted to the ve... | politicsNews | October 24, 2017 | 0 | 32 |
| 140 | Factbox: Trump on Twitter (December 14) - Stoc... | The following statements were posted to the ve... | politicsNews | December 15, 2017 | 0 | 65 |
| 32954 | HILLARY VOLUNTEER, Obama Supporter Mocks Ryan ... | Speaking of not having a soul..check out this ... | politics | Feb 28, 2017 | 1 | 36 |
| 28877 | Ben Carson Praises Time Trump Compared Him To... | The only thing sadder than losing a Republican... | News | March 16, 2016 | 1 | 12 |

Fig. 3 news dataset after allocating individual news with a user id (UID)

The sample of the news dataset can be seen in Fig.2. Since, there is no such dataset found in which news as well as user information are available, so the 21,417 true news and 23,481 fake news are distributed among the 85 users randomly as one-third users posting only real news, another third user posting a mixture of fake and real, and the remaining user posting only fake news. So, now after combining the user profile with the available news dataset have been prepared, shown in Fig. 3 and based on the user profile information user profile_score has been calculated for every user.

**User score calculation**

Complete any missing information, standardize user-id, user-name, and email addresses, and then use similarity metrics to determine how well they match to get a user profile score. Assess the degree of gender coherence among profile components, including the bio and profile photo, and allocate credibility ratings correspondingly. Determine the user's age by using their date of birth, take into account the area code on their cell phone, and compare it to their hometown. Evaluate the degree of resemblance between one's hometown and present town and allocate points correspondingly. As an extra measure of credibility, take into account the existence of a profile photo. Calculate an overall user profile score by adding the individual scores of each criterion and weighing them according to importance and relevance.

The post_score being calculated as following formula

$$Post\_score = e^{(real(i)\,/\,total\ real\ post)} - e^{(fake(i)\,/\,total\ fake\ news)}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (I)$$

Finally the final score of individual users calculated as

$$final\_score(i) = profile\_score(i) + c * post\_score(i) \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (II)$$

*Where c is the constant*

| | real | fake | score | post score | final score |
|---|---|---|---|---|---|
| 17 | 413 | 77 | 85.0 | 16.186249 | 33.389687 |
| 71 | 43 | 518 | 77.0 | -20.295751 | 4.028187 |
| 49 | 143 | 252 | 82.0 | -4.090597 | 17.432052 |
| 25 | 300 | 100 | 80.0 | 9.838286 | 27.378714 |
| 61 | 73 | 345 | 81.0 | -11.386876 | 11.709843 |

Fig. 4 sample of the user score, post score and final score

Fig. 4 shows the number of real and fake posts, user score, post score and final score of a user. From the figure it can be seen that if a user posts greater real news then the post score will be greater and also the final score of that user will be high. For example, user's 17 posts 413 real news and 77 fake news so the post score is only 16.18 and final score is 33.39 since the user has a good profile score of 85. In another example users 71 posts only 43 real news and 518 fake news so the post score is in negative as of -20.29 and final score is 4.03 since the user has a good profile score of 77.

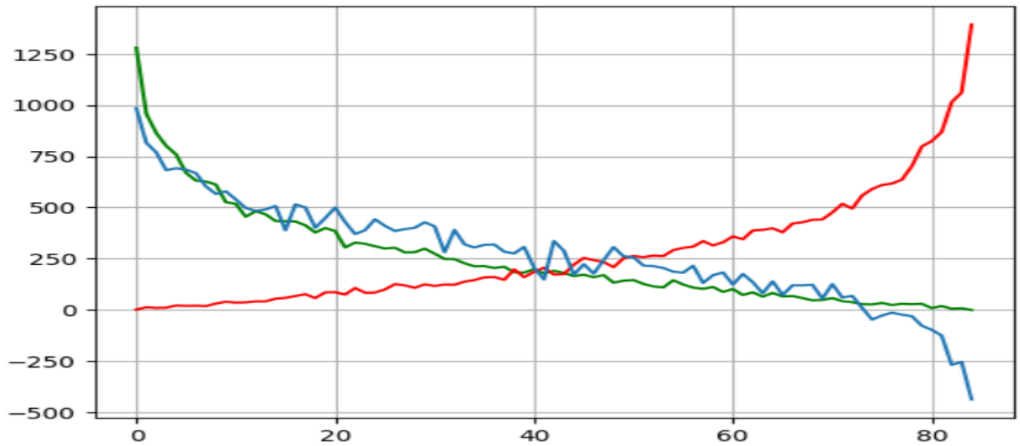


Fig. 5 graphical representation of user real and fake post and final score

Fig. 5 illustrates the nature of the number of real and fake posts and their final score of the users. In this graph x-axis shows the user id and y-axis shows the values. Red color is for the number of fake news and the green color shows the number of real news. The blue colour shows the final score of the users. As it can be seen, if a user has fewer fake news stories, the final score will be high, and vice versa.

From the news dataset combine all the column title, text and subject together and apply different NLP text preprocessing techniques. Preparing and cleaning unprocessed text data for analysis is known as text pre-processing. Typical tasks include removing stop words, tokenizing (dividing text into words or phrases), converting text to lowercase, removing special characters, and stemming or lemmatization (reducing words to their root form). Effective text pre-processing also involves managing problems like missing data, encoding, and eliminating unnecessary information. The objective is to produce a clean, standardized text corpus that will enable precise and insightful analysis.

A technique called Term Frequency-Inverse Document Frequency (TF-IDF) converts unprocessed textual data into numerical vectors in order to extract features from the text has been used for text content of the news. Every term is given a score according to how frequently it appears in a document compared to how frequently it appears in all documents. This produces a feature matrix that illustrates the importance of terms in particular documents. High TF-IDF scores are useful for text-based machine learning models because they highlight terms that are significant within a document but less prevalent overall in the dataset.

The extracted features and the final score of the user posting that news are considered as independent variable and the label of the news as dependent variable. This study uses different machine learning supervised classification techniques for detection if news is real or fake. The comparative results of all the techniques with user credibility analysis and without considering user credibility are shown in table-1.

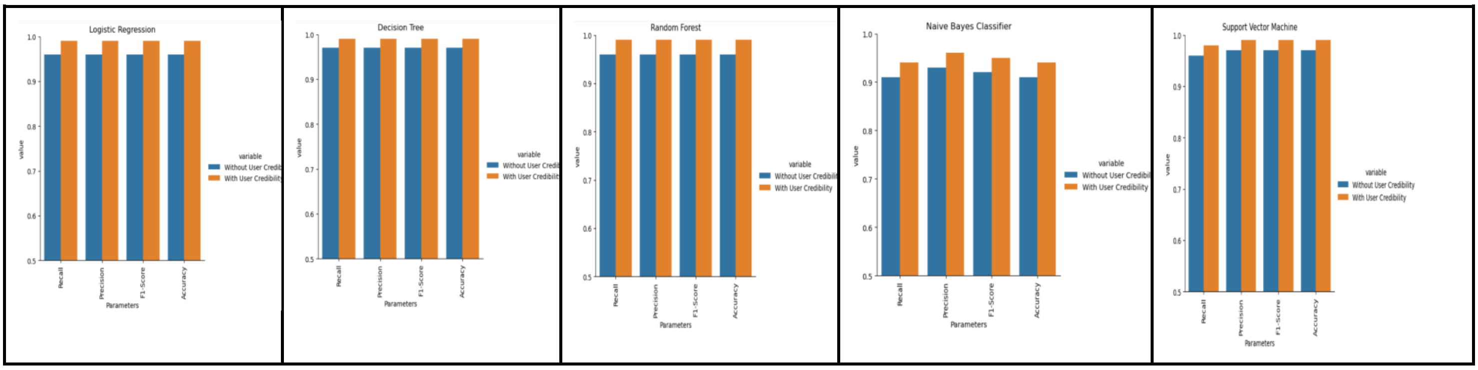


Fig. 6 graphical representation of user real and fake post and final score

Table 1. Comparative study of different techniques with and without user credibility analysis

| Algorithm | User Credibility Used | Precision | Recall | F1 Score | Accuracy |
|---|---|---|---|---|---|
| **Logistic Regression** | No | 0.96 | 0.96 | 0.96 | 0.96 |
| | **Yes** | **0.98** | **0.98** | **0.98** | **0.98** |
| **Decision Tree** | No | 0.97 | 0.97 | 0.97 | 0.97 |
| | **Yes** | **0.99** | **0.99** | **0.99** | **0.99** |
| **Random Forest** | No | 0.96 | 0.96 | 0.96 | 0.96 |
| | **Yes** | **0.98** | **0.98** | **0.98** | **0.98** |
| **Naive Bayes** | No | 0.92 | 0.90 | 0.91 | 0.91 |
| | **Yes** | **0.94** | **0.93** | **0.92** | **0.93** |
| **Support Vector Machine** | No | 0.97 | 0.96 | 0.96 | 0.97 |
| | **Yes** | **0.99** | **0.99** | **0.99** | **0.99** |

The performance of different machine learning algorithms in terms of precision, recall, F1-score, and accuracy is shown in Fig. 6. The performance of different approaches without and with considering user credibility is shown in blue and orange, respectively. It can be clearly seen that the performance of different algorithms has been enhanced after considering user credibility.

## VI. CONCLUSION

Accurate identification of fake news has been greatly increased by including user credibility indicators into machine learning classification models. Models provide a more comprehensive picture of the validity of user-generated material by taking into account variables including user-id consistency, email address validation, gender matching, age-related believability, and geographic information alignment. The model's capacity to discern between reliable and unreliable sources of information is improved by this all-encompassing strategy, which combines user profile data with conventional textual and contextual features. Combining user credibility indicators with cutting-edge machine learning algorithms improves the accuracy of detecting fake news while also strengthening decision-making processes for thwarting disinformation and maintaining the integrity of information distribution platforms.

## REFERENCE

[1] Sitaula, N., Mohan, C. K., Grygiel, J., Zhou, X., & Zafarani, R. (2020). Credibility-Based Fake News Detection. Lecture Notes in Social Networks, 163–182. https://doi.org/10.1007/978-3-030-42699-6_9

[2] Flanagin, A. J., & Metzger, M. J. (2007, April). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, *9*(2), 319–342. https://doi.org/10.1177/1461444807075015

[3] Castillo, C., Mendoza, M., & Poblete, B. (2011, March 28). *Information credibility on twitter*. https://doi.org/10.1145/1963405.1963500

[4] Kumar, K. P. K., & Geethakumari, G. (2014, September 24). *Detecting misinformation in online social networks using cognitive psychology*. Human-centric Computing and Information Sciences. https://doi.org/10.1186/s13673-014-0014-x

[5] Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017, January 1). *Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking*. https://doi.org/10.18653/v1/d17-1317

[6] Flanagin, A. J., & Metzger, M. J. (2000, September 1). *Perceptions of Internet Information Credibility*. Journalism & Mass Communication Quarterly. https://doi.org/10.1177/107769900007700304

[7] Martens, B., Aguiar, L., Gómez-Herrera, E., & Müller-Langer, F. (2018, January 1). *The Digital Transformation of News Media and the Rise of Disinformation and Fake News*. Social Science Research Network. https://doi.org/10.2139/ssrn.3164170

[8] Gupta A K, Kumar P (2018), "A Comparative Analysis Of Machine Learning Techniques For Fake News Detection", IJRAR - International Journal of Research and Analytical Reviews (IJRAR), E-ISSN 2348-1269, P- ISSN 2349-5138, Volume.5, Issue 4, Page No pp.345-348

[9] Ahmed H, Traore I, Saad S. “Detecting opinion spams and fake news using text classification”, Journal of Security and Privacy, Volume 1, Issue 1, Wiley, January/February 2018.

[10] Ahmed H, Traore I, Saad S. (2017) “Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. In: Traore I., Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. ISDDC 2017. Lecture Notes in Computer Science, vol 10618. Springer, Cham (pp. 127- 138)