iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.56774

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

A Survey: Fake News Detection Mobile App Using Machine Learning

Snehal Thorave[1], Arunima Rathore[2], Pritamkumar Maji[3], Ashwin Kumar[4], Arya Magar[5]
[1]*Assistant Professor,* [2, 3, 4, 5]*Student, Department of Computer Engineering, Dhole Patil College of Engineering, Wagholi, Pune-412207, India*

Abstract: ***With the proliferation of digital information, the challenge of identifying and mitigating the spread of fake news has become increasingly crucial. This research paper provides a comprehensive review of existing technologies designed for the detection of fake news, with a particular focus on their adaptation to mobile applications. The study delves into various approaches, including natural language processing, machine learning algorithms, and deep learning models, highlighting their strengths and limitations. Additionally, it explores the integration of user-generated content analysis and social network features in enhancing the accuracy of fake news detection on mobile platforms. By synthesizing insights from recent advancements, this paper aims to inform the development of an effective mobile app for mitigating the impact of fake news in the digital landscape.***
Index Terms: fake news, news detection, machine learning, deep learning, Naïve Baye, Algorithms, Misinformation, Fact-checking

I. INTRODUCTION

In today's information age, where the boundaries between truth and falsehood are increasingly blurred, the proliferation of fake news and misinformation represents a formidable challenge to the integrity of our digital society. The rapid dissemination of false information through various online platforms has the potential to manipulate public opinion, sow discord, and erode trust in reliable sources of information. As a consequence, the urgent need for robust and accessible tools to combat the spread of fake news has never been more evident.

This research paper introduces effective solutions to tackle the scourge of fake news. Our focus centers on the development of a mobile application that harnesses the power of machine learning to detect and verify the authenticity of news articles, social media posts, and other online content. The envisioned application seeks to empower users with a practical and interactive tool to help them distinguish between credible news sources and deceptive or fabricated information.

The rapid proliferation of fake news is not a new concern, but the advent of digital technology and the ease of information dissemination have elevated this problem to unprecedented levels. Misinformation can spread like wildfire, often taking on a life of its own, influencing public opinion, and even impacting political processes. This phenomenon, coupled with the challenges of verifying the authenticity of content in real time, underscores the significance of our research project.

The proposed mobile application will leverage cutting-edge machine learning techniques, including natural language processing and advanced algorithms, to provide users with a mechanism for real-time assessment of the credibility of online content. It will offer features such as news article analysis, fact-checking, and source credibility evaluation, all accessible within the confines of a mobile device. This research paper will delve into a comprehensive exploration of the challenges posed by fake news, the existing state of the art in fake news detection, and the specific approach adopted in our research. We will detail the methodology used for data collection and machine learning model development.

II. PROBLEM STATEMENT

The project is concerned about distinguishing an answer that could be utilized to recognize and sift through locales containing counterfeit news for motivations behind assisting clients with trying not to be attracted by misleading content sources.

It is basic that such arrangements are recognized as they will end up being helpful to both pursuers and tech organizations engaged with the issue.

III. OBJECTIVES

1) Detection of fake news spread on any social media platforms.
2) Determining the credibility of the news the user has acquired.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

IV. LITERATURE REVIEW

The researchers in [1] "Detecting Fake News using Machine Learning: A Systematic Literature Review" stated that:

1) *Support Vector Machine:* This algorithm is mostly used for classification. This is a supervised machine learning algorithm that learns from the labelled data set.
2) *Naïve Bayes:* This algorithm is also used for classification tasks in machine learning & AI. This can be utilized to check whether the news is legitimate or counterfeit.
3) *Logistic Regression:* This classifier is employed when dealing with categorical predictions. For example, it has the capability to forecast or provide outcomes in terms of true or false.
4) *Random Forests:* In this classifier, there are different random forests that give a value and a value with more votes is the actual result of this classifier.
5) *Recurrent Neural Network:* This classifier proves beneficial in identifying fabricated news as well.
6) *Neural Network:* There are different algorithms of machine learning that are used to help in classification problems. One of these algorithms is the neural network.
7) *K-Nearest Neighbour:* It is a supervised machine learning algorithm employed to address classification issues. It retains information on every case, enabling the classification of new cases based on their similarity.
8) *Decision Tree:* This supervised algorithm of machine learning can help to detect the fake news. It dissects the dataset into various smaller subsets.

Researchers in [2](Detecting Fake News using Machine Learning and Deep Learning Algorithms.) used the machine learning classifiers for detecting the fake news. Based on the researchers' experiments, it was found that "SVM and Naïve Bayes classifiers are the most effective in identifying fake news. These two outperform other classifiers in terms of the accuracy they deliver. A classifier with more accuracy is considered as a better classifier".

Researchers in[3] (Detecting Fake News with Machine Learning Method) concluded that "researchers have employed three machine learning techniques. These 3 techniques are: Naïve Bayes, Neural network and the SVM. The Naïve Bayes exhibited a precision rate of 96.08%. On the other hand, the other two methods that are neural network and SVM provided the accuracy of 90.90%."

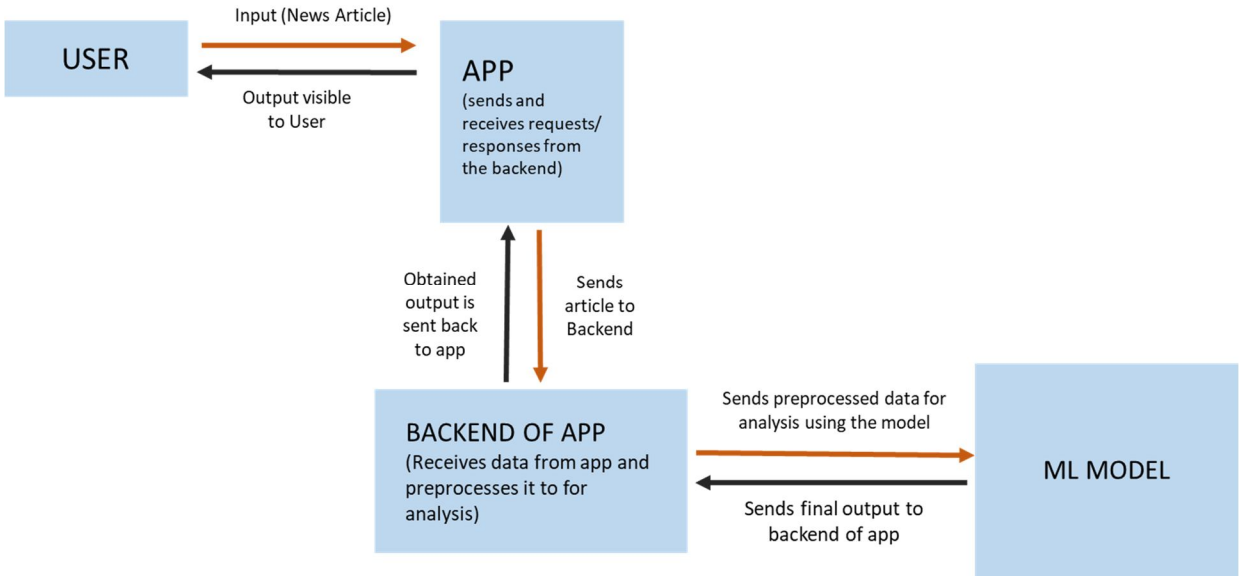
Researchers in [4](Fake News Detection Using Machine Learning Approaches) mention that "in the aforementioned research summary and system analysis, we concluded that most of the research papers used naïve bays algorithm, and the prediction precision was between 70-76%, they mostly use qualitative analysis depending on sentiment analysis, titles, word frequency repetition. In our strategy, we suggest incorporating an additional element into these methodologies—namely, POS textual analysis. This constitutes a quantitative approach involving the inclusion of numeric statistical values as features. We anticipate that augmenting these features and employing a random forest will contribute to further improvements in precision results. The features we propose to add in our dataset are total words (tokens), Total unique words (types), Type/Token Ratio (TTR), Number of sentences, Average sentence length (ASL), Number of characters, Average word length (AWL), nouns, prepositions, adjectives etc."

[5](A Comparative Study of Machine Learning and Deep Learning Techniques for Fake News Detection):

In this paper, an examination of fake news detection methods from existing literature is provided, "along with a comparison of seven traditional machine learning algorithms: LR, SVM, NB, DT, RF, XGB, and an ensemble that integrates all these algorithms, with two scenarios of word representation methods: statistical (sparse) word vector representation methods and context-free (dense) pre trained word representation models. Furthermore, this article conducts a comparison of eight sophisticated machine learning models, namely CNN, BiLSTM, BiGRU, CNN-BiLSTM, CNN-BiGRU, and various hybrid models, utilizing two types of text representation models—context-free and context-aware embedding models. Additionally, it assesses two advanced pretrained transformer-based models, namely BERTbase and RoBERTabase. They found that the ensemble of the classical ML methods with TF-IDF features outperformed the other methods on the LIAR dataset, including advanced ML models. Since the LIAR dataset comprises brief political statements, extracting valuable cues to distinguish fake news from genuine news poses a challenge. However, BERTbase alone was able to achieve an accuracy result that was on par with the best-performing models, demonstrating the power of such a model in capturing useful clues from a short piece of text. RoBERTabase performed the best on the PolitiFact dataset with an F1 score of 93.17. On the flip side, traditional machine learning approaches like SVM with TF-IDF features outperformed the deep learning models on the GossipCop dataset, surpassing even the official baselines. By depending solely on news text with minimal preprocessing, we outperformed the state-of-the-art results across the utilized datasets."

[6] A Comprehensive Review on Fake News Detection With Deep Learning
"This study has assessed methods for identifying fake news through the lens of natural language processing (NLP) and sophisticated deep learning strategies. The document introduced a classification system for fake news detection approaches, delving into various NLP methods and DL architectures while highlighting their strengths and weaknesses."

V. SYSTEM ARCHITECTURE



VI. PROPOSED OUTCOMES

The proposed outcomes of a fake news detection app using machine learning can be summarized as follows:

1) The primary goal should be accurate detection of fake news through advanced algorithms and machine learning techniques mainly Naïve Bayes in combination with advanced models like BERTbase and RoBERTabase.
2) The app should have a user-friendly interface, making it accessible to a wide range of users, regardless of their technological expertise. Clear indicators highlight potentially false or misleading content.
3) The app will provide users with the correct and credible source of information if available. Users can cross-verify the information from the original source.

VII. CONCLUSION

The literature review reveals a diverse range of machine-learning approaches employed in detecting fake news. Various classifiers, including Support Vector Machine (SVM), Naïve Bayes, Logistic Regression, Random Forests, Recurrent Neural Network, Neural Network, K-Nearest Neighbour, and Decision Tree, have been explored by different researchers. Notably, SVM and Naïve Bayes consistently emerge as effective classifiers, outperforming others in accuracy.

Researchers have tackled the challenges of fake news detection in different ways, considering factors such as social media posts, sentiment analysis, titles, word frequency repetition, and even adding features like part-of-speech (POS) textual analysis for a quantitative approach. The effectiveness of these methods is demonstrated through high accuracy rates, with Naïve Bayes achieving 96.08% accuracy in one study.

Comparisons between classical machine learning algorithms and advanced models, including deep learning techniques and transformer-based models like BERT and RoBERTa, highlight the nuanced nature of fake news detection. While ensemble methods with TF-IDF features perform well in certain datasets, advanced models like BERTbase and RoBERTabase demonstrate superior performance in specific contexts.

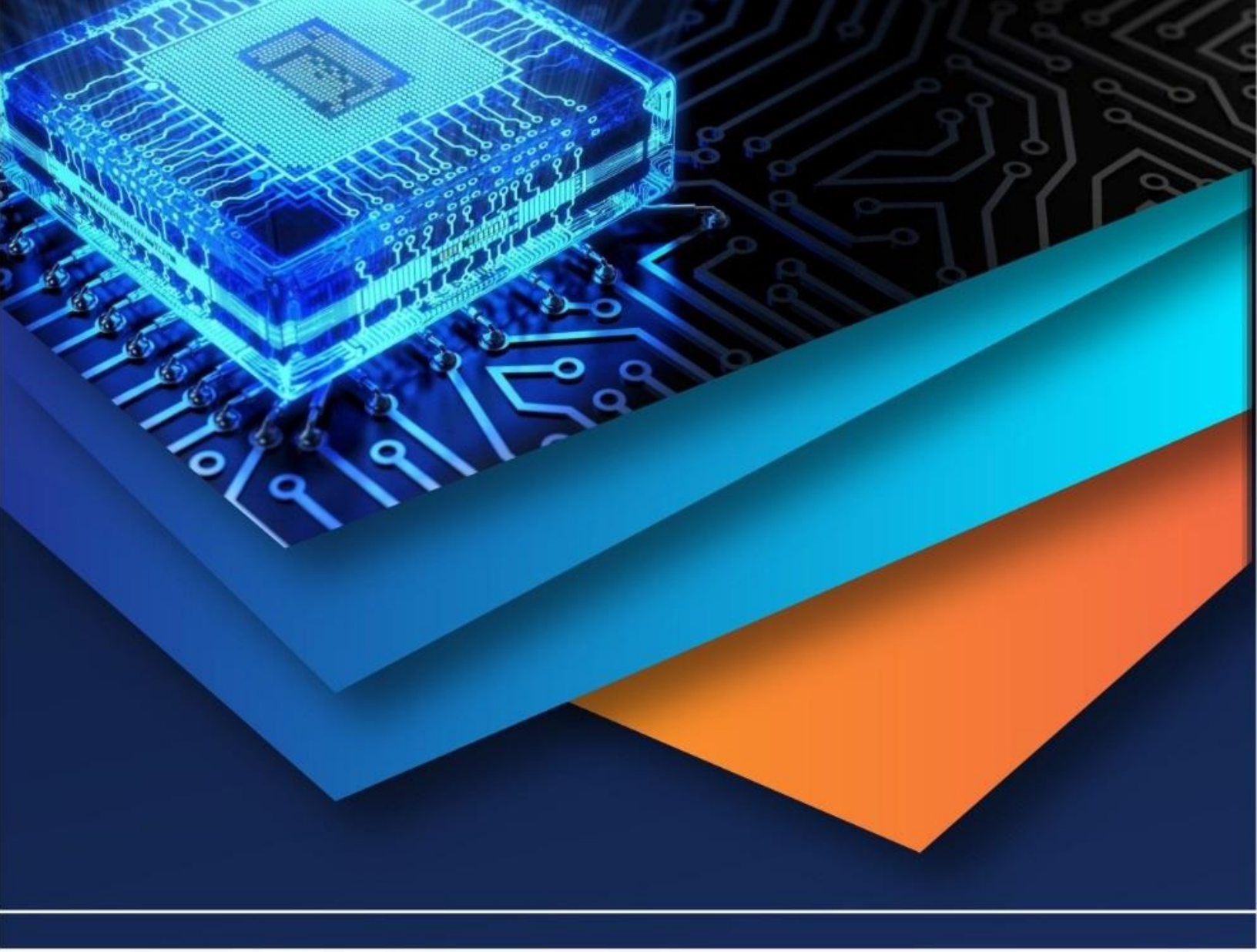
In essence, the literature review suggests that detecting fake news is a complex task, requiring a combination of approaches and consideration of contextual factors. The success of different models is often dataset-dependent, emphasizing the need for a nuanced and adaptable approach in the development of effective fake news detection systems.

VIII. FUTURE SCOPE

It might be possible to improve the performance in context-based information, in addition to news text, is considered and if other factors, such as style and sentiment, are included. Mostly, the experimental results indicate that no single technique can deliver the best performance scores across all used datasets.

REFERENCES

[1] Detecting Fake News using Machine Learning: A Systematic Literature Review Alim Al Ayub Ahmed (alim@jju.edu.cn) School of Accounting, Jiujiang University, Jiujiang, Jiangxi, CHINA.
[2] Detecting Fake News using Machine Learning and Deep Learning Algorithms. Abdullah-All-Tanvir, Mahir, E. M., Akhter S., & Huq, M. R. (2019).
[3] Detecting Fake News with Machine Learning Method. Aphiwongsophon, S., & Chongstitvatana, P. (2018) .
[4] Fake News Detection Using Machine Learning Approaches To cite this article: Z Khanam et al 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1099 012040
[5] A Comparative Study of Machine Learning and Deep Learning Techniques for Fake News Detection Jawaher Alghamdi, Yuqing Lin and Suhuai Luo
[6] A Comprehensive Review on Fake News Detection With Deep Learning M. F. MRIDHA, (Senior Member, IEEE), ASHFIA JANNAT KEYA, MD. ABDUL HAMID, MUHAMMAD MOSTAFA MONOWAR, AND MD. SAIFUR RAHMAN. Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka 1216, Bangladesh Department of Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia Corresponding author: M. F. Mridha (firoz@bubt.edu.bd)



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY