TruthTrack – Fake News Prediction

A MINI PROJECT REPORT

Submitted by

RESHMA A (2116210701211)

RESHMA S (2116210701213)

in partial fulfillment for the award

of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING

COLLEGE ANNA UNIVERSITY,

CHENNAI

MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE,

CHENNAI

BONAFIDE CERTIFICATE

Certified that this Report titled "**TruthTrack – Fake News Prediction**" is the bonafide work of **"RESHMA A (2116210701211), RESHMA S (2116210701213)"** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Rakesh kumar. M

Assistant Professor,

Department of Computer Science and Engineering,

Rajalakshmi Engineering College,

Chennai – 602015

Submitted to Mini Project Viva-Voce Examination held on _____

Internal Examiner **External Examiner**

ABSTRACT

In urban environments, efficient management of street lighting systems is important for ensuring safety, reducing energy consumption, and optimizing maintenance efforts. Exploring IoT for smart streetlights, this project integrates automatic light control based on sunlight intensity and increases the intensity of the light when there is a motion in the street and also detects fault. It proposes an automated system designed to address the challenges of real-time street light fault detection, and efficient maintenance within urban areas. Leveraging Internet of Things (IoT) technology, the proposed system integrates smart sensors, wireless communication modules such as WIFI, and data analytics algorithms to monitor the operational status of street lights continuously. By detecting faults such as bulb failures, electrical issues, or physical damage in real-time, the system enables prompt responses to ensure uninterrupted illumination. Furthermore, the system incorporates LDR sensor for detecting the intensity of light and IR sensor to detect the motion in the , facilitating security maintenance interventions and optimizing resource allocation. Through centralized data processing and analysis, the system provides actionable insights to municipal authorities for proactive maintenance planning, thereby enhancing the overall reliability and sustainability of street lighting infrastructure in urban environments. The implementation of this automated system promises to significantly improve operational efficiency, reduce downtime, and enhance public safety in cities while contributing to energy conservation, manual operation and cost savings.

ACKNOWLEDGMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavour to put forth this report. Our sincere thanks to our Chairman **Mr. S.MEGANATHAN, B.E, F.I.E.**, our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN**, **Ph.D.,** for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time.

We express our sincere thanks to **Dr. P. KUMAR, Ph.D.**, Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide, **Dr. RAKESH KUMAR M, M.E., Ph.D.,** Assistant Professor Department of Computer Science and Engineering for his encouragement and guiding us throughout the project to build our project.

RESHMA A (210701211)

RESHMA S (210701213)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 GENERAL

Fake news prediction involves detecting false or misleading news using methods like logistic regression, machine learning, natural language processing, and social network analysis. Key indicators include linguistic anomalies, sensational content, and unusual sharing patterns. Datasets of labeled news articles and social media data aid in training models. The goal is to combat misinformation and protect the integrity of information ecosystems.

1.2 OBJECTIVE

The aim of the project is to develop a web-based system for predicting the authenticity of news articles, employing machine learning techniques. By implementing with a logistic regression model trained on a dataset of news articles, the system aims to accurately classify incoming articles as real or fake. The main objective is to provide users with a tool to verify the credibility of news sources and combat the spread of misinformation online. By accurately identifying and flagging fake news, the system seeks to restore trust in the authenticity of information circulated online, thereby mitigating the negative impact of misinformation on society.

- Enhancing Trust: To restore and enhance public trust in media and information sources by ensuring that the news consumed is accurate and reliable.
- Detection and indentification: To accurately detect and identify false or misleading information presented as news.
- Provide a logistic regression model to classify news articles as real or fake based on their content.
- Developed a web application using Flask to provide a user-friendly interface for submitting news articles.
- By reducing the prevalence of fake news and misinformation, the project aims to foster a more constructive and respectful online discourse, conducive to informed debate and democratic participation.

1.3 EXISTING SYSTEM

Through the inserted data, the algorithm will first learn to distinguish between bogus and authentic news. After understanding the distinction, the system will learn to make judgments based on the data presented. Fake news tracker programmes monitor the collection, analysis, and visualisation of fake news. The bogus database displays no news channel names, but the genuine dataset displays individual headquarters for each station. Manipulating the concept of dataset fraudulent channels are exploiting an unregistered news portal. As a result, using the original dataset, one may compare and explicitly identify them.The data analysis also involves a number of dangers. The proper usage of data assessment in relation to references must be considered. During data analysis, there are some assessment elements that Python does not recognise, which creates the data clarity difficulty.

1.4 PROPOSED SYSTEM

The approach to developing a fake news prediction model involves several key steps. First, we perform text transformation using either a Count Vectorizer, which converts a collection of text documents into a matrix of token counts, or a TF-IDF Vectorizer, which converts the text into a matrix of TF-IDF features that reflect the importance of words in the document relative to the entire corpus. We then decide whether to use headlines or the full text of news articles for the analysis, as each option may yield different insights and effectiveness in classification. Next, we extract optimal features by selecting the most frequently occurring words and/or phrases, converting all text to lowercase to maintain uniformity, and removing common stop words such as "the," "when," and "there" to focus on more meaningful words. We also ensure that only words appearing at least a certain number of times in the dataset are included, to filter out less informative words. Finally, we implement a Linear Regression model to classify news as fake or real, where the model predicts a continuous score that, when compared against a threshold, determines the classification.

ADVANTAGES

- Information was very clear and understandable.
- It gives accurate predictions which is very clear to the user.
- User friendly and faster time compatibility.

CHAPTER 2
LITERATURE SURVEY

[1] This paper explores the development of a fake news detection system using sentiment analysis for the Indonesian language. It builds on previous research that used the VADER method for English and achieved 78% accuracy with Random Forest. The study extracts extreme sentiment scores from news headlines and employs Random Forest, Support Vector Machine, and Naive Bayes for classification, with Random Forest achieving the best results at an average accuracy of 62% and a peak of 66%. The findings indicate progress in Indonesian fake news detection through sentiment analysis, with future improvements suggested to enhance system performance.

[2] This paper aims to develop an automated, reliable, and effective system for detecting fake news to reduce misinformation and promote accurate information. It compares the effectiveness of Random Forest (RF) and K Nearest Neighbor (KNN) algorithms using a dataset of 2000 news articles labeled as fake or real. The results show that KNN outperforms RF, achieving 81.20% accuracy. The study concludes that KNN is a more effective algorithm for fake news detection, with statistically significant differences between the two methods ($p=0.001$). This research provides valuable insights into the effectiveness of machine learning algorithms in identifying fake news.

[3] This paper tackles the pervasive issue of fake news, recognizing its potential to cause significant harm in society. Introducing a machine learning-based system, the research aims to combat the dissemination of false information. As fake news detection gains prominence, businesses are increasingly investing in this sector, either for internal use or as a service offering. Leveraging machine learning and deep learning methodologies, the study explores various approaches to discerning the authenticity of news articles. Through an extensive review, the paper underscores the prevalence of Support Vector Machine (SVM) and Long Short-Term Memory (LSTM) as favored algorithms in this domain.

[4] In today's digital age, the internet has become a primary source of news for many, leading to the rapid dissemination of fake news. With an unprecedented volume of data being generated, estimated at 120 zettabytes per second , various technologies are evolving to manage this influx. As people increasingly rely on online platforms like Instagram, Facebook, and Wikipedia for news, misinformation can spread swiftly across the globe. Recognizing this challenge, our paper employs machine learning (ML) techniques to detect fake news.

[5] In the realm of Indian politics, the proliferation of false news poses a significant threat to information integrity. This paper introduces a machine learning-driven approach aimed at discerning fraudulent news by analyzing textual content. Leveraging a pre-defined dataset, our model effectively distinguishes between authentic and fraudulent news items, surpassing existing detection algorithms. Tested on benchmark datasets, our methodology showcases superior performance, offering a robust solution to combat false news in Indian politics and beyond. With its versatility across topics and languages

[6] The rise of fake news on social media poses a significant threat, prompting the need for effective countermeasures. This study utilizes machine learning algorithms to identify and control the spread of misinformation. Employing techniques like Passive-Aggressive Classifier, Random Forest, Gradient Boosting, and Decision Tree, the research achieves notable success, with an accuracy of 88.66% using the Passive Aggressive Classifier.

[7] In response to the rampant spread of misinformation online, this paper advocates for the use of AI and ML techniques to automate news story classification. By examining textual properties, various machine learning algorithms are trained and evaluated for their effectiveness in distinguishing fake news from real. The study presents multiple detection methods adaptable to diverse contexts, aiming to enhance accuracy in news classification tasks.

CHAPTER 3

SYSTEM DESIGN

3.1 DEVELOPMENT ENVIRONMENT

3.1.1 HARDWARE SPECIFICATIONS

This project uses minimal hardware but in order to run the project efficiently without any lack of user experience, the following specifications are recommended

PROCESSOR	Intel Core i5
RAM	4GB or above (DDR4 RAM)
GPU	Intel Integrated Graphics
HARD DISK	6GB
PROCESSOR FREQUENCY	1.5 GHz or above

Table 3.1.1 Hardware Specifications

3.1.2 SOFTWARE SPECIFICATIONS

The software specifications in order to execute the project has been listed down in the below table. The requirements in terms of the software that needs to be pre- installed and the languages needed to develop the project has been listed out below.

FRONT END	HTML, CSS, JavaScript
BACK END	Python, Flask (a micro web framework for Python)
Web Server	Flask (for running the Python web application)
SOFTWARES USED	Visual Studio, Jupyter Notebook

Table 3.1.2 Software Specifications

3.2 SYSTEM DESIGN

3.2.1 ARCHITECTURE DIAGRAM

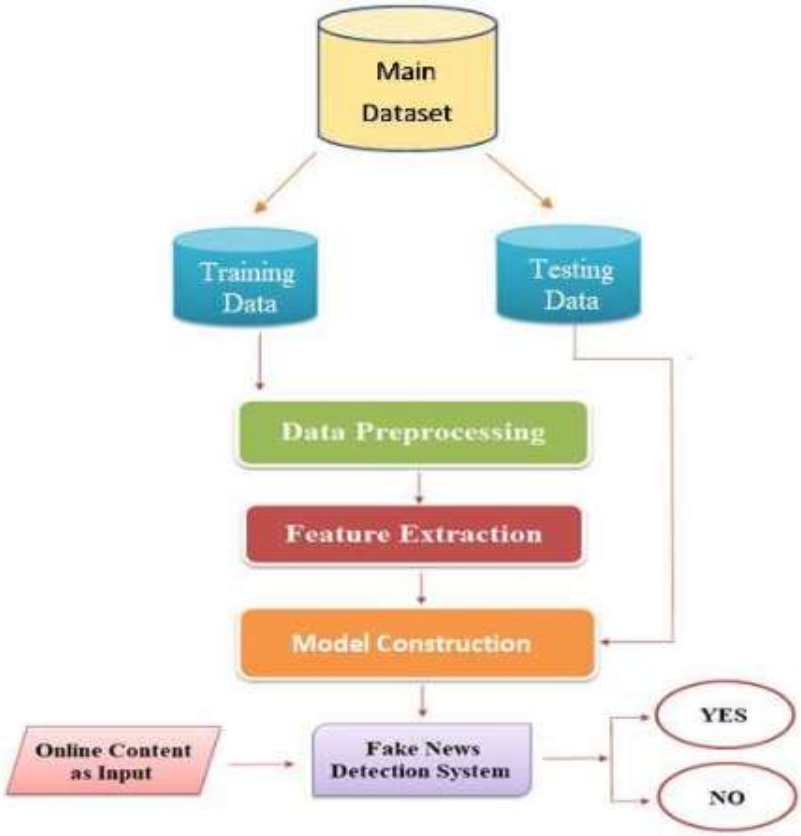

Fig 3.2.1 Architecture Diagram

This paper explains about the development of the system . The first part is static which works on machine learning classifier. We trained the model with 4 different classifiers and chose the best classifier for final execution. The second part is dynamic which takes the keyword/text from user and searches online for the truth probability of the news. In this paper, we have used Python and its Sci-kit libraries . Python has a huge set of libraries and extensions, which can be easily used in Machine Learning. Sci-Kit Learn library is the best source for machine learning algorithms where nearly all types of machine learning algorithms are readily available for Python, thus easy and quick evaluation of ML algorithms is possible. We have used flask and machine learning for the web based deployment of the model, provides client side implementation using HTML, CSS and Javascript.

3.2.2 DATA FLOW DIAGRAM

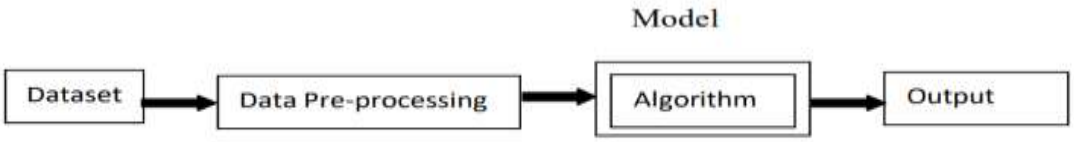

Figure 4.1 : Data Flow Diagram

The DFD takes an input-process-output view of a system i.e., data objects flow into the software, are transformed by processing elements, and resultant data objects flow out of the software. The dataset contains real and fake news information. Then the information is fed to algorithm .Thus news is analyzed as fake or real

PRE-PROCESSING:

The first step collecting data from a CSV file containing labeled news articles. The dataset was sampled to reduce size and filled missing values. Each article's author and title were combined into a single 'content' field. We applied text preprocessing techniques including lowercasing, removing non-alphabet characters, and stemming words. Stopwords were removed to retain meaningful words. The processed text was vectorized using TF-IDF to convert it into numerical format suitable for machine learning.

TRAINING SET:

The training set for this project was carefully curated by splitting the preprocessed dataset, ensuring 80% of the data was allocated for training. This set included a balanced mix of real and fake news articles to provide comprehensive learning examples. Text preprocessing involved lowercasing, removing non-alphabetic characters, and applying stemming and stopword removal to ensure that only meaningful words were retained. The preprocessed text was then transformed into numerical features using the TF-IDF vectorizer. This vectorized training data was fed into a logistic regression model, allowing it to learn distinguishing features and patterns between real and fake news. The model was iteratively trained and tuned for optimal performance, and cross-validation was employed to verify its generalization capability. The trained model was subsequently validated on the remaining 20% of the data to assess its accuracy and effectiveness in real-world scenarios.

CHAPTER 4

PROJECT DESCRIPTION

4.1 MODULE DESCRIPTION

1. DATA PRE-PROCESSING:

This module mainly consists of the preliminary step of collecting data from a CSV file containing labeled news articles. The dataset was sampled to reduce size and filled missing values. Each article's author and title were combined into a single 'content' field. We applied text preprocessing techniques including lowercasing, removing non-alphabet characters, and stemming words. Stopwords were removed to retain meaningful words. The processed text was vectorized using TF-IDF to convert it into numerical format suitable for machine learning

2. TRAINING SET:

The training set for this project was created by splitting the preprocessed data into training and testing subsets. The training set comprises 80% of the data, ensuring a balanced representation of both real and fake news. The content was vectorized using TF-IDF to convert the textual information into numerical features. This vectorized data was then used to train a logistic regression model, allowing it to learn patterns and relationships that distinguish fake news from real news. The model's performance was evaluated on the testing set to ensure accuracy and reliability before deployment.

3. TRAINING MODEL:

For the training phase of the model, we utilized logistic regression, a robust and efficient algorithm for binary classification tasks like distinguishing between fake and real news. The process began by transforming the preprocessed news articles into numerical features using the TF-IDF vectorizer, which captures the importance of words in the context of each document.

4. TESTING MODEL:

During the testing phase, the trained logistic regression model was evaluated to assess its performance and reliability in detecting fake news. This evaluation was conducted using a separate portion (20%) of the preprocessed dataset, designated as the testing set.The testing set consisted of news articles that the model had not been exposed to during training. Each article in the testing set was vectorized using the same TF-IDF vectorizer used during training, ensuring consistency in

feature representation.The model's predictions were compared against the ground truth labels in the testing set to determine its accuracy and effectiveness.

5. USER INTERFACE MODULE:

Fig 4.1.1 FRONT END OF FAKE NEWS FOR PROVIDING INPUT

The User Interface Module for our project entails designing a streamlined input form where users can input the author's name and the title of the news article they wish to verify. Upon clicking the "Predict" button, the system processes the input and displays the prediction result in a dedicated section.

Fig 4.1.2 FRONT END OF FAKE NEWS FOR PREDICTING

This section clearly indicates whether the news is classified as fake or real, accompanied by a corresponding confidence score or percentage that quantifies the likelihood of the prediction.

CHAPTER 5

IMPLEMENTATION AND RESULTS

5.1 IMPLEMENTATION

1) Pre-processing: The text goes through preprocessing procedures like tokenization, stopword removal, and sentence splitting after data collection. The text data is ready for additional analysis and summary after these procedures.

2). Cleaning: Cleaning up the text data is necessary to highlight attributes that we're going to want our machine learning system to pick up on. Cleaning (or pre-processing) the data typically consists of a number of steps.

a) Remove punctuation: Punctuation can provide grammatical context to a sentence which supports our understanding. But for our vectorizer which counts the number of words and not the context, it does not add value, so we remove all special character.

b) Remove stopwords: Stopwords are common words that will likely appear in any text. They don't tell us much about our data so we remove them

c) Stemming: Stemming helps reduce a word to its stem form. It often makes sense to treat related words in the same way. It removes suffices, like "ing", "ly", "s", etc. by a simple rule-based approach. It reduces the corpus of words but often the actual words get neglected. eg: Entitling, Entitled -> Entitle

3) Feature Generation: We can use text data to generate a number of features like word count, frequency of large words, frequency of unique words, n-grams etc. By creating a representation of words that capture their meanings, semantic relationships, and numerous types of context they are used in, we can enable computer to understand text and perform Clustering, Classification etc

4) Vectorizing Data: Vectorizing is the process of encoding text as integers i.e. numeric form to create feature vectors so that machine learning algorithms can understand our data

5) Confusion Matrix: A confusion matrix is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known. It allows the visualization of the performance of an algorithm. A confusion matrix is a summary of prediction results on a

classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix. The confusion matrix shows the ways in which your classification model is confused when it makes predictions. It gives us insight not only into the errors being made by a classifier but more importantly the types of errors that are being made .

Total	Class 1 (Predicted)	Class 2 (Predicted)
Class 1 (Actual)	TP	FN
Class 2 (Actual)	FP	TN

6) Logistic regression used for Classification: Logistic regression can be used for classification by setting a threshold on the predicted output to distinguish between classes. For fake news prediction, the model learns to assign a continuous score to each news article. By selecting an appropriate threshold, scores above it can be classified as real news, while those below are classified as fake news.

7) Integration with Backend: Developed a backend application using Python and a web framework like Flask to handle user requests and serve predictions. Load the trained model into the backend application and implement prediction logic to classify news articles as real or fake.

8) User Interface Development: Designed a user-friendly web interface where users can input news articles for classification. Implemented a frontend components using HTML, CSS, and JavaScript to create input forms and display prediction results.

9) Testing and Validation: Conduct thorough testing of the entire system, including backend functionality, frontend user interface, and end-to-end prediction accuracy**.**

5.2 RESULTS

The results are presented to the user in a clear and informative manner. The prediction outcome, whether the news article is classified as real or fake, is prominently displayed on the user interface. Additionally, the system provides a confidence score or probability, indicating the level of certainty associated with the prediction. To further elucidate the model's performance, the system presents a confusion matrix, showcasing the true positives, true negatives, false positives, and false negatives. This matrix offers insights into the model's accuracy, precision enabling users to assess its effectiveness in distinguishing between real and fake news. The user interface is designed to be intuitive and interactive, allowing users to easily interpret the results and make informed decisions based on the prediction outcome. Visual feedback elements such as color-coded messages or icons enhance the user experience, making the results easily understandable at a glance. Overall, the result presentation in the fake news detection system aims to empower users with valuable insights and facilitate their critical evaluation of news articles for authenticity.

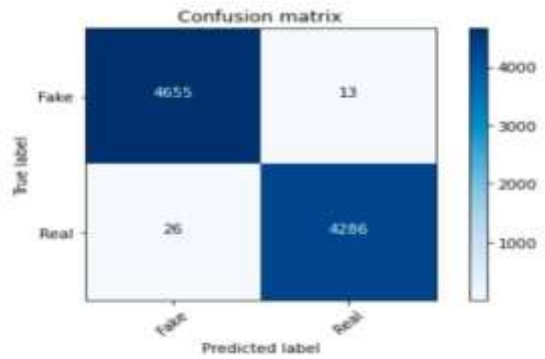

Fig 5.5.1 Confusion matrix

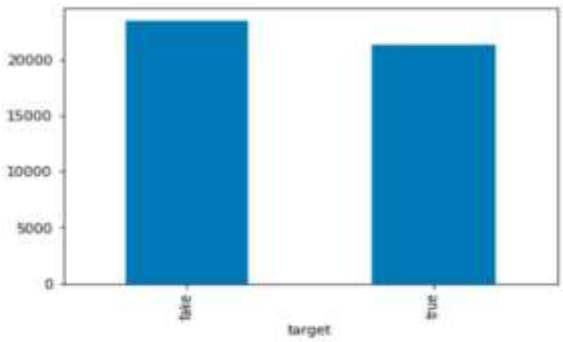

Fig 5.5.2 Analysing fake and real news in a bar graph

Fig 4.1.5.2 FRONT END OF FAKE NEWS FOR PREDICTING

This section clearly indicates whether the news is classified as fake or real, accompanied by a corresponding confidence score or percentage that quantifies the likelihood of the prediction

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENTS

6.1 CONCLUSION

Due to increasing use of internet, it is now easy to spread fake news. A huge number of persons are regularly connected with internet and social media platforms. There is no any restriction while posting any news on these platforms. So some of the people takes the advantage of these platforms and start spreading fake news against the individuals or organizations. This can destroy the repute of an individual or can affect a business. Through fake news, the opinions of the people can also be changed for a political party. There is a need for a way to detect these fake news. Machine learning classifiers are using for different purposes and these can also be used for detecting the fake news. The classifiers are first trained with a data set called training data set. After that, these classifiers can automatically detect fake news. The data we used in our work is collected from the Kaggle.com and contains news articles from various domains to cover most of the news rather than specifically classifying political news. The learning models were trained and parameter-tuned to obtain optimal accuracy. Fake news detection has many open issues that require attention of researchers.

6.2 FUTURE ENHANCEMENTS

Fake news is categorized as any kind of cooked-up story with an intention to deceive or to mislead. In this paper we are trying to present the solution for fake news detection task by using Machine Learning techniques. Many events have resulted to a rise in the prominence and spread of phony news. The widespread impacts of the massive onset of fake news can be seen, humans are conflicting if not outright poor detectors of fake news. With this, endeavours are being made to automate the task of fake news detection. The most mainstream of such actions include blacklisting of sources and authors that are unreliable. Even though these tools are useful, but in order to produce a progressive complete end to end solution, we are required to represent for tougher cases where reliable sources and authors are responsible for releasing fake news. Here, the purpose of this project was to build a model that help us to recognize the language patterns that can be used to classify fake and real news with the help of ML (machine learning) techniques. The outcomes of this project shows the capability of ML to be fruitful in this task. We have tried to build a model that helps in catching many intuitive indications of real and fake news as well as in the visualization of the classification decision. Now-a-days fake news is such a big problem that it is affecting our society as well as our facts and opinions.

REFERENCES

[1]I. F. Rozi, R. Arianto and H. H. Mahdyan, "Fake News Detection Using Sentiment Analysis Approach in Indonesian Language," 2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA), Surabaya, Indonesia, 2023, pp. 206-211, doi: 10.1109/ICAMIMIA60881.2023.10427891. keywords: {Support vector machines;Sentiment analysis;Machine learning algorithms;Tokenization;Fake news;Random forests;Testing;Classification;Fake News Detection;Sentiment Analysis;Text Mining},

[2]K. S. Sri Saranya and A. H. Juliet, "Comparison of Random Forest with K-Nearest Neighbors to Detect Fake News with Improved Accuracy," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 2543-2547, doi: 10.1109/IC3I59117.2023.10397893. keywords: {Radio frequency;Measurement;Machine learning algorithms;Social networking (online);Fake news;Random forests;Software development management;Fake news detection;KNN;Random Forest;Dataset;Result;Accuracy;Innovative news},

[3]Ashish, Sonia, M. Arora, Hemraj, A. Rana and G. Gupta, "An Analysis and Identification of Fake News using Machine Learning Techniques," 2024 11th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2024, pp. 634-638, doi: 10.23919/INDIACom61295.2024.10498879. keywords: {Deep learning;Support vector machines;Machine learning algorithms;Reviews;Fake news;Long short term memory;Business;News Data Analysis;Machine Learning;Deep Learning;Long Short Term Memory;Artificial Intelligence},

[4]R. Vunnava, L. Bodla, M. K. Dehury and B. K. Mohanta, "Performance Analysis of ML Techniques in Identification of Fake News," 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2022, pp. 276-281, doi: 10.1109/ICSCDS53736.2022.9760905. keywords: {Social networking (online);Multimedia Web sites;Machine learning;Encyclopedias;Computer architecture;Internet;Performance analysis;Machine Learning;Classification;Fake News;Internet;Social Media},

[5]T. Babu, R. R. Nair, A. Challa, R. Srikanth, S. S. Aravindan and S. S, "Fake News Detection using Machine Learning Algorithms," 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), Bangalore, India, 2023, pp. 1-7, doi: 10.1109/ICCAMS60113.2023.10525971. keywords: {Machine learning algorithms;Limiting;Detectors;Forestry;Graph theory;Security;Labeling;Fake News;Self-Learning;Response Generation;Natural Language Processing;Context Free Grammar;Stochastic Gradient Decent},

[6]L. B. Rananavare and S. Chitnis, "Fake News Identification for Web Scrapped Data using Passive Aggressive Classifier," 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 1-6, doi: 10.1109/NMITCON58196.2023.10275818. keywords: {Machine learning algorithms;Social networking (online);Weapons;Forestry;Inspection;Classification algorithms;Fake news;Decision Tree;Fake News Detection;Gradient Boosting;Linear

Regression;Passive Aggressive Classifier;Random Forest;TF-IDF;Web Scrapping},

[7]S. A. A, P. R, P. J. S. Kailash and D. Ramesh, "Analysis of Online Fake News Using Machine Learning Techniques," 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), Vellore, India, 2024, pp. 1-5, doi:

[8]10.1109/ic-ETITE58242.2024.10493794. keywords: {Logistic regression;Machine learning algorithms;Streaming media;Market research;Real-time systems;Fake news;Task analysis;Internet;AI;ML;Random Forest;Bagging;Business;Online news;Regression;Ensemble models},

[9]D. Patel, S. Bansal, J. Arora, R. Rani, A. Dev and A. Sharma, "Analysis and Visualization of Fake News Detection Dataset through Machine Learning Models," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10112024. keywords: {Training;Social networking (online);Computational modeling;Voting;Data visualization;Organizations;Predictive models;Fake News Prediction;K Means Clustering;Logistic Regression;Random Forest Classification},

[10] V. U S, "Detecting the Presence of Fake News Content on Social Networking Platforms using ML & NLP," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 1520-1525, doi: 10.1109/ICAC3N56670.2022.10074210. keywords: {Deep learning;Social networking (online);Training data;Prototypes;Process control;Market research;Digital communication;Fake news detection;social networking platforms;machine learning;DNN;deep learning;natural language processing;KNN;PA},

[11] V. Gupta, R. S. Mathur, T. Bansal and A. Goyal, "Fake News Detection using Machine Learning," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 84-89, doi: 10.1109/COM-IT-CON54601.2022.9850560. keywords: {Measurement;Support vector machines;Machine learning algorithms;Social networking (online);Natural language processing;Web sites;Fake news;Natural Language Processing;Random Forest;Fake News Detection;Logistic Regression;k-Nearest Neighbor;Machine Learning;SVC},

[12] D. K. Sharma and S. Garg, "A Framework for Automatic Fake Content Identification," 2021 International Conference on Simulation, Automation & Smart Manufacturing (SASM), Mathura, India, 2021, pp. 1-5, doi: 10.1109/SASM51857.2021.9841096. keywords: {Training;Deep learning;Costs;Pandemics;Transformers;Reliability;Fake news;Fake news;feature selection Online misinformation;Text classification},

[13] D. Rohera et al., "A Taxonomy of Fake News Classification Techniques: Survey and Implementation Aspects," in IEEE Access, vol. 10, pp. 30367-30394, 2022, doi: 10.1109/ACCESS.2022.3159651.
keywords: {Fake news;Social networking (online);Classification algorithms;Biological system modeling;Task analysis;Support vector machines;Analytical models;Social media;fake news classification;machine learning;LSTM;NB},

[14] A. Altheneyan and A. Alhadlaq, "Big Data ML-Based Fake News Detection Using Distributed Learning," in IEEE Access, vol. 11, pp. 29447-29463, 2023, doi: 10.1109/ACCESS.2023.3260763.
keywords: {Social networking (online);Fake news;Blogs;Task analysis;Voting;Big Data;Behavioral sciences;Machine learning;Ensemble learning;Big data;machine learning;fake news;ensemble learning;social media},

[15] J. Wilson, S. E. Hosseini and S. Pervez, "Identification of Fake News in Social Media Using Sentimental Analysis," 2023 IEEE Industrial Electronics and Applications Conference (IEACon), Penang, Malaysia, 2023, pp. 220-224, doi: 10.1109/IEACon57683.2023.10370300. keywords: {Support vector machines;Text mining;Sentiment analysis;Logistic regression;Machine learning algorithms;Social networking (online);Supervised learning;social media;fake news;sentiment analysis;machine learning},

[16] A. Kanavos, I. Karamitsos, A. Mohasseb and V. C. Gerogiannis, "Comparative Study of Machine Learning Algorithms and Text Vectorization Methods for Fake News Detection," 2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA), Volos, Greece, 2023, pp. 1-8, doi: 10.1109/IISA59645.2023.10345953. keywords: {Support vector machines;Machine learning algorithms;Social networking (online);Classification algorithms;Naive Bayes methods;Fake news;Task analysis;Machine Learning;Text Mining;Information Retrieval;Fake News Detection},

[17] A. Khalil, M. Jarrah, M. Aldwairi and Y. Jararweh, "Detecting Arabic Fake News Using Machine Learning," 2021 Second International Conference on Intelligent Data Science Technologies and Applications (IDSTA), Tartu, Estonia, 2021, pp. 171-177, doi: 10.1109/IDSTA53674.2021.9660811. keywords: {Training;Deep learning;Machine learning algorithms;Data science;Internet;Noise measurement;Fake news;Arabic news corpus;fake news detection;machine learning;underfitting;overfitting},

[18] M. Aljabri, D. M. Alomari and M. Aboulnour, "Fake News Detection Using Machine Learning Models," 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), Al-Khobar, Saudi Arabia, 2022, pp. 473-477, doi: 10.1109/CICN56167.2022.10008340. keywords: {Support vector machines;Radio frequency;Machine learning algorithms;Phishing;Text categorization;Feature extraction;Natural language processing;Fake news;Features extraction;Machine Learning;NLP},

[19] S. Nagesh, A. N. Sampath, P. Guruprasad, S. S. S and J. G, "A Machine Learning Approach for Detecting Fake News in Hindi," 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Gautam Buddha Nagar, India, 2023, pp. 69-74, doi: 10.1109/UPCON59197.2023.10434587. keywords: {Support vector machines;Logistic regression;Computational modeling;Machine learning;Libraries;Fake news;Graphical user interfaces;Hindi Fake News;Machine Learning;Natural Language Processing;IndicBERT;MuRIL;StanfordNLP},