

Intelligent News Veracity Classifier for COVID-19 Misinformation

Chetan. S ² Vinay patel.GL ¹

¹Assistant Professor, Department of MCA, BIET, Davanagere

² Student, 4th Semester MCA, Department of MCA, BIET, Davanagere

Abstract

The rapid spread of misinformation during the COVID-19 pandemic created serious challenges for public health communication, leading to confusion, fear, and mistrust among individuals worldwide. The widespread circulation of false and misleading news on social media and digital platforms highlighted the urgent need for automated and intelligent fact-checking systems. This study proposes an Intelligent News Veracity Classifier designed to detect and categorize COVID-19-related news articles as True, False, or Misleading using Machine Learning (ML) and Natural Language Processing (NLP) techniques. The system preprocesses textual data and applies TF-IDF vectorization to convert text into numerical features for model training. A Logistic Regression classifier is employed for prediction due to its efficiency and strong performance in text classification tasks. To address class imbalance and ensure fair model performance, the Synthetic Minority Oversampling Technique (SMOTE) is integrated into the training process. The solution is implemented as a web-based application with separate user and administrator modules. Users can register, log in, and verify news authenticity in real time, while administrators manage accounts and provide reliable health resources. By combining data-driven classification, secure user management, and trustworthy health information, the proposed system offers an intelligent and comprehensive approach to combating misinformation and promoting informed decision-making during global health emergencies.

Keywords: Machine Learning, Natural Language Processing (NLP), COVID-19 Misinformation, Fake News Detection, Logistic Regression, TF-IDF Vectorisation.

1. INTRODUCTION

The rapid advancement of digital technologies has significantly transformed the way information is created, distributed, and consumed across the globe. Social media platforms, online news portals, blogs, and messaging applications have become primary sources of information for millions of people. While this digital revolution has improved global connectivity and access to information, it has also given rise to a major challenge: the rapid spread of misinformation. During the COVID-19 pandemic, this issue became particularly severe, creating what was widely referred to as an “infodemic,” where an overwhelming amount of information — both accurate and false — circulated rapidly online.

A substantial volume of misleading and false news regarding COVID-19 infection rates, vaccines, treatments, preventive measures, and conspiracy theories spread widely across digital platforms. Such misinformation generated fear, confusion, and mistrust among the public, undermining efforts taken by health authorities and governments to manage the crisis effectively. Many individuals struggled to differentiate between authentic medical information and fabricated claims, leading to poor health decisions and reduced trust in scientific guidance. As people increasingly rely

on digital platforms for real-time updates, the ability to verify the authenticity of news content has become critically important.

Traditional manual fact-checking approaches are insufficient to handle the enormous volume of digital information produced daily. The speed at which misinformation spreads online far exceeds the capacity of human verification systems. Therefore, there is a pressing need for intelligent, automated systems capable of analyzing and validating online content efficiently and accurately. The Intelligent News Veracity Classifier for COVID-19 Misinformation is developed to address this challenge by leveraging Machine Learning (ML) and Natural Language Processing (NLP) techniques to automatically classify news articles based on their authenticity.

The proposed system is designed to detect and categorize COVID-19-related news articles into three categories: True, False, or Misleading. By applying advanced text-processing techniques, the system analyzes textual content to extract meaningful linguistic features. The textual data undergoes preprocessing steps such as lowercasing, punctuation removal, stop-word elimination, tokenization, and

normalization to ensure clean and structured input for model training. After preprocessing, the system transforms textual data into numerical representations using TF-IDF (Term Frequency–Inverse Document Frequency) vectorization. TF-IDF is a widely used statistical technique that assigns importance weights to words based on their frequency within a document relative to the entire dataset, enabling effective feature extraction from unstructured text.

The classification process is carried out using a Logistic Regression model. Logistic Regression is selected due to its strong performance in text classification problems, computational efficiency, and interpretability. It is particularly suitable for handling high-dimensional feature spaces generated through TF-IDF vectorization. The model is trained on labeled datasets to learn patterns that distinguish between authentic and false news articles. Once trained, the system can predict the authenticity of newly entered COVID-19-related news content in real time.

One of the significant challenges in misinformation detection is class imbalance within datasets. Real-world data often contain unequal distributions of true, false, and misleading articles. This imbalance can bias the machine learning model toward the majority class, reducing fairness and predictive reliability. To overcome this issue, the proposed system integrates SMOTE (Synthetic Minority Oversampling Technique). SMOTE generates synthetic samples of minority class instances to create a balanced dataset during the training phase. By addressing class imbalance, the model achieves improved accuracy, reduced bias, and enhanced generalization capability.

The system is implemented as a web-based application to ensure accessibility and practical usability. It consists of two main modules: an administrator module and a user module. The administrator is responsible for maintaining the integrity and reliability of the platform. Administrative functionalities include secure login, user management, updating verified hospital information, and maintaining reliable health recommendations. Through continuous monitoring and updating of health-related data, the administrator ensures that the system remains trustworthy and accurate.

The user module is designed with simplicity and usability in mind. Users can register and securely log

into the system to access its features. Once authenticated, users can enter any COVID-19-related news text into the system. The machine learning model analyzes the input and provides a real-time prediction indicating whether the news is True, False, or Misleading. In addition to verification services, users are provided with access to accredited hospital information and verified health advice. This integrated approach not only assists users in verifying news authenticity but also promotes public health awareness by directing individuals toward reliable resources.

By offering automated, data-driven fact verification, the Intelligent News Veracity Classifier represents a significant advancement in combating online misinformation. The system can potentially be integrated with social media monitoring tools, news aggregation systems, and public health communication networks to enable large-scale, real-time misinformation detection. Such integration would strengthen digital ecosystems by reducing the harmful effects of false information and supporting informed decision-making during public health emergencies.

The core objective of this project is to develop a reliable machine learning-based system capable of accurately identifying false or misleading COVID-19 news while promoting trustworthy health communication. The system aims to provide an easy-to-use platform for verifying news content, ensure balanced and unbiased predictions through advanced data-handling techniques, and empower administrators to manage health resources efficiently. Ultimately, the project contributes to the broader goal of using artificial intelligence responsibly to safeguard public understanding and mitigate the negative impacts of misinformation during global crises.

In conclusion, the COVID-19 pandemic highlighted the critical importance of accurate information dissemination in times of crisis. As digital platforms continue to serve as major information sources, intelligent automated systems become essential tools for preserving public trust and promoting informed behavior. By combining Natural Language Processing, Machine Learning, and web technologies, the Intelligent News Veracity Classifier provides a comprehensive and practical solution for detecting misinformation and strengthening reliable communication in the digital age.

II. RELATED WORK

01. A study by Hosam Alhakami et al. evaluated multiple machine learning and deep learning models for detecting COVID-19 fake news on social media. The researchers compared six traditional ML classifiers and two deep learning models using two publicly available datasets. Text representation techniques such as TF-IDF and GloVe embeddings were applied for feature extraction. The results showed that traditional machine learning models, particularly Logistic Regression and Naïve Bayes, outperformed deep learning models in terms of accuracy and false positive rate. The study concluded that classical ML approaches remain effective and computationally efficient for COVID-19 fake news detection.

02. A recent study by V. Rathinapriya and J. Kalaivani proposed an advanced deep learning-based fake news detection model for pandemic situations. Their approach integrates optimized feature selection using an Adaptive Artificial Rabbits Optimization (AARO) algorithm and introduces an Optimal Attention-based Multiscale DenseNet with LSTM (OAMDNet-LSTM) architecture. The model extracts textual features using GloVe, BERT, and TF-IDF, followed by weighted feature optimization and parameter tuning to enhance classification performance. Experimental results demonstrated improved accuracy and precision compared to conventional machine learning and optimization techniques, highlighting the effectiveness of deep neural architectures for pandemic-related misinformation detection.

03. A comprehensive systematic review conducted by A. R. Sanaullah, Anupam Das, Anik Das, Muhammad Ashad Kabir, and Kai Shu analyzed the applications of machine learning and deep learning techniques for detecting COVID-19 misinformation. The study reviewed 43 high-quality research articles selected through a PRISMA-based methodology from major databases such as Scopus and Web of Science. The authors examined the complete misinformation detection pipeline, including datasets, preprocessing techniques, feature extraction methods (e.g., TF-IDF, word embeddings), and classification algorithms. Furthermore, the review identified major challenges such as dataset imbalance, multilingual misinformation, and generalization issues, and highlighted future research directions for developing more robust and adaptive detection models. This survey provides a structured understanding of state-of-the-art approaches in COVID-19 misinformation detection.

04. A recent comprehensive review by Asma Ul Hussna, Md Golam Rabiul Alam, Risul Islam, Bader Fahad Alkhamees, Mohammad Mehedi Hassan, and Md Zia Uddin presents an in-depth analysis of COVID-19 misinformation detection on X (formerly Twitter). The study systematically reviewed 600 articles and closely analyzed 87 significant research works focusing on machine learning and deep learning approaches for fake news detection, as well as community-level analysis of misinformation spreaders. The authors observed that deep neural network-based models dominate recent research due to higher accuracy, although they often suffer from overfitting, vanishing gradient issues, and high computational cost. Furthermore, the paper highlights research gaps such as limited dataset sizes, insufficient community analysis, and the need for large-scale training datasets. This work provides valuable insights into both detection models and the behavioral dynamics of misinformation disseminators during the COVID-19 infodemic.

05. The study by Mohamad Nabeel and Christine Große proposes a convolutional neural network (CNN)-based framework for classifying COVID-19 disinformation on Twitter. Their work emphasizes integrating both creator/user-based features (such as account verification, URLs, and account age) and news content-based features (tweet text) to improve detection performance. Using a dataset of 12,469 labeled tweets, the CNN model achieved strong results, including 97.2% accuracy, 95.7% precision, and 99.8% recall, demonstrating the effectiveness of deep learning in identifying fake-labeled tweets. The authors also highlighted the importance of feature standardization and addressed challenges related to the interpretability of deep learning models. Their findings confirm that even relatively simple CNN architectures, when combined with well-engineered features, can achieve high classification performance in COVID-19 misinformation detection.

06. A recent study by Samira Malek, Christopher Griffin, Robert D Fraleigh, Robert Lennon, Vishal Monga, and Lijiang Shen proposed a comprehensive framework titled Misinformation Detection and Inoculation Process (MDIP) for combating health misinformation using large language models (LLMs). Published in *Journal of Medical Internet Research* (2026), the study integrates misinformation detection, topic modeling, thematic analysis, and refutation generation into a unified automated pipeline. The authors trained BERT, GPT-2, and T5 models for binary

misinformation classification, with BERT achieving 98% accuracy, and applied BERTopic for identifying misinformation themes. To enhance interpretability, the framework used prompt engineering to generate sentence-level topic descriptions and theme-based refutations, achieving high approval ratings from evaluators and 82% accuracy in theme detection. Unlike traditional approaches that focus only on detection, this study introduced an intervention-oriented system capable of producing inoculation-based counterarguments, thereby linking automated misinformation detection with actionable public health communication strategies.

07. A significant contribution to COVID-19 disinformation categorisation was made by Xingyi Song, Johann Petrak, Ye Jiang, Iknoor Singh, Diana Maynard, and Kalina Bontcheva through their proposed Classification-Aware Neural Topic Model (CANTM). Published in **PLoS ONE** (2021), the study introduced a novel framework that integrates BERT-based text representation with a Variational Autoencoder (VAE) to jointly perform classification and topic modeling. Unlike traditional topic models such as LDA, CANTM incorporates classification signals directly into the latent topic generation process, enabling robust performance even with limited annotated data. The authors also developed one of the largest manually annotated COVID-19 disinformation datasets derived from IFCN fact-checks and demonstrated that CANTM outperformed baseline models including BERT, SCHOLAR, NVDM, and LDA in both accuracy and topic coherence. Additionally, the model provided interpretable, class-associated topics that served as global explanations for classification decisions, facilitating large-scale monitoring and analysis of evolving misinformation trends.

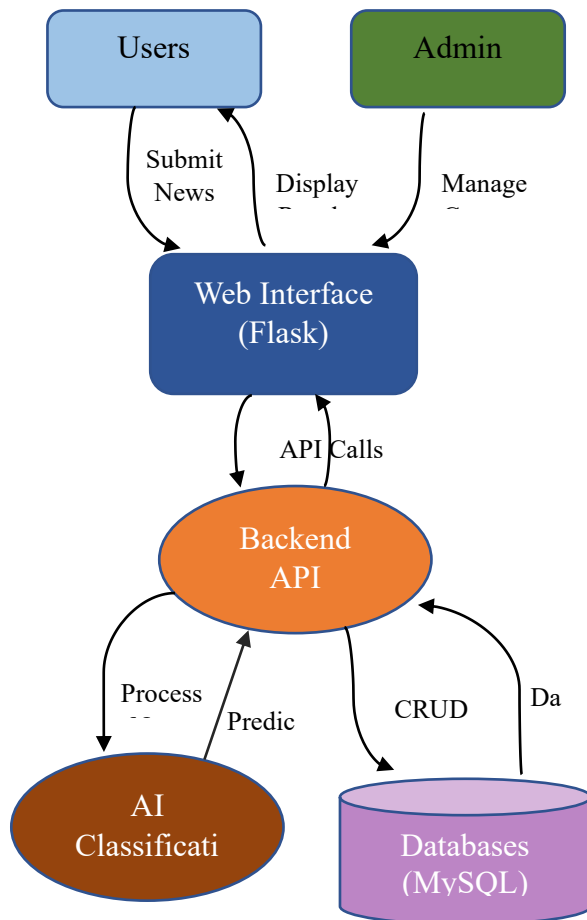
08. Saad et al. explore the mitigation of fake news propagation by leveraging the decentralized and transparent nature of blockchain technology. Their research emphasizes the use of smart contracts to validate the credibility of news sources and the integrity of the content shared across social media networks. By creating a verifiable record of information provenance, the proposed framework aims to deter the spread of misinformation and provide users with a mechanism to audit the history of news articles. This approach complements existing machine learning techniques by adding a layer of architectural trust that is resistant to centralized manipulation or unauthorized data tampering.

09. Juraev et al. demonstrate the effectiveness of combining Natural Language Processing (NLP) with

supervised Machine Learning (ML) classifiers to identify misinformation based on linguistic patterns. Their research highlights that preprocessing steps such as TF-IDF vectorization and lemmatization are critical for converting raw text into high-dimensional features suitable for classification. By evaluating models like Support Vector Machines (SVM) and Random Forest, the study achieves over 90% accuracy, particularly in politically sensitive and health-related categories. This textual-analysis-heavy approach provides a scalable method for detecting sensationalism and emotional manipulation, though the authors note that these models must be continuously updated to handle the evolving nature of digital content and regional language nuances.

10. Juraev et al. explore the synergy between Natural Language Processing (NLP) and supervised Machine Learning (ML) to combat the digital dissemination of fake news. Their methodology emphasizes the importance of rigorous text preprocessing—specifically tokenization, lemmatization, and TF-IDF vectorization—to extract high-dimensional linguistic features from unstructured news data. By comparing several classifiers, the study identifies Support Vector Machines (SVM) and Random Forest as superior models, achieving over 90% accuracy in detecting sensationalized content within political and health-related domains. While this research proves the scalability of content-based analysis, it also highlights the critical need for real-time model updates to address regional linguistic variations and the increasing sophistication of misinformation.

III. METHODOLOGY



The proposed Intelligent News Veracity Classifier for COVID-19 Misinformation follows a systematic machine learning pipeline for detecting and classifying COVID-19-related news into True, False, or Misleading categories. The methodology consists of dataset preparation, text preprocessing, feature extraction, data balancing, model training, evaluation, and deployment within a web-based system.

01. Dataset Collection

The dataset used for the Intelligent News Veracity Classifier consists of labeled COVID-19-related news articles stored in a structured CSV format. Each record contains two primary attributes: the textual news content and its corresponding authenticity label categorized as True, False, or Misleading. Initially, the dataset structure was validated to ensure the presence of required columns. An exploratory analysis was performed to examine class distribution and detect any imbalance among categories. This validation step ensures the integrity and suitability of data for further processing and model training.

02. Text Preprocessing

Text preprocessing was carried out to remove noise and improve the quality of input data. All text entries were converted into lowercase to maintain uniformity. URLs and hyperlinks were removed using regular expression techniques, followed by elimination of punctuation marks, special characters, and numerical values. Extra whitespace was normalized to maintain consistency. Duplicate records were removed to prevent bias, and entries containing missing values were discarded. These preprocessing steps enhance textual clarity and ensure clean, standardized input for feature extraction.

03. Feature Extraction using TF-IDF

After preprocessing, the cleaned text data was transformed into numerical format using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorization technique implemented through scikit-learn. TF-IDF measures the significance of words within a document relative to the entire dataset by assigning higher weights to informative words and reducing the influence of frequently occurring common terms. The vectorizer was configured with a maximum of 5000 features and English stop-word removal. This conversion produces structured numerical feature vectors suitable for supervised machine learning algorithms.

04. Handling Class Imbalance with SMOTE

Since real-world misinformation datasets typically suffer from unequal class representation, class imbalance was addressed using the Synthetic Minority Oversampling Technique (SMOTE). SMOTE was implemented using the imbalanced-learn library. This method generates synthetic samples for underrepresented classes by interpolating between existing minority instances. Applying SMOTE ensures balanced class distribution, improves prediction fairness, and reduces model bias toward majority categories. Applying SMOTE ensures balanced class distribution, improves prediction fairness, and reduces model bias toward majority categories.

05. Data Splitting Strategy

To evaluate the model effectively, the balanced dataset was divided into training and testing subsets using an 80:20 ratio. Stratified sampling was applied during splitting to maintain proportional class representation across both sets. The training dataset was used to build

the classification model, while the testing dataset was used to evaluate its performance on unseen data. This approach ensures reliable and unbiased performance assessment.

06. Model Selection and Training

The classification model selected for this system is Logistic Regression. Logistic Regression was chosen due to its computational efficiency, interpretability, and suitability for high-dimensional text classification problems. The model was configured with optimized parameters, including increased maximum iterations and balanced class weighting to improve convergence and fairness. The training process involved fitting the Logistic Regression model on the TF-IDF feature vectors derived from the training dataset.

07. Model Evaluation Metrics

The trained model was evaluated using multiple performance metrics, including accuracy, precision, recall, and macro F1-score. A confusion matrix was generated to analyze class-wise prediction performance and identify misclassification trends. The macro F1-score was particularly emphasized to ensure balanced evaluation across all three categories. These evaluation measures confirm the model's reliability and effectiveness in detecting misinformation.

08. Model Saving and Deployment

Upon successful training and evaluation, the trained model and TF-IDF vectorizer were saved using serialization techniques to enable reuse without retraining. These saved components were integrated into a Flask-based web application. When a user submits a news article, the system performs preprocessing, vectorization, and prediction in real time. The classification result is then displayed instantly through the web interface, enabling automated and accessible misinformation detection.

09. System Integration Architecture

The complete system follows a modular architecture consisting of a frontend user interface, a backend API developed using Python Flask, an AI classification engine, and a MySQL database for storing user and health-related information. The admin module manages user accounts, hospital details, and health tips, while the user module allows news verification and

access to trusted resources. This integrated framework ensures scalability, usability, security, and real-time performance.

IV. TECHNOLOGIES USED

The proposed COVID-19 News Veracity Classification System integrates multiple modern technologies and frameworks to enable automated misinformation detection, real-time news verification, and secure web-based deployment. The following tools and technologies were used in the development and implementation of the system:

1. Python Programming Language

Python serves as the primary programming language for implementing the entire system, including data preprocessing, machine learning model development, backend integration, and deployment. Python was selected due to its readability, extensive ecosystem of libraries, and strong support for machine learning and Natural Language Processing (NLP) applications.

2. Flask Web Framework

The web application interface was developed using Flask, a lightweight and flexible Python web framework. Flask is responsible for handling HTTP requests, routing, user authentication, API interactions, and integration of the trained machine learning model into the web environment. It enables real-time prediction of COVID-19 news authenticity through an interactive web interface.

3. Natural Language Processing (NLP) Techniques

Natural Language Processing techniques are employed to analyze and process textual news content before classification. Text preprocessing operations such as tokenization, stop-word removal, lowercase conversion, and noise removal are implemented using built-in Python libraries and regular expressions. These steps ensure that raw textual data is transformed into clean

and structured input suitable for machine learning models.

4. Data Science and Machine Learning Libraries

Several data science and machine learning libraries are used to implement the classification pipeline:

- pandas: Utilized for dataset loading, cleaning, transformation, and manipulation of tabular data.
- NumPy: Used for numerical computations and array-based operations during preprocessing.
- scikit-learn: Provides tools for TF-IDF vectorization, dataset splitting, model training, performance evaluation, and classification report generation.
- imbalanced-learn: Used to implement SMOTE (Synthetic Minority Oversampling Technique) for handling class imbalance in the dataset.
- Joblib: Used for saving and loading the trained machine learning model and TF-IDF vectorizer for deployment.

5. Machine Learning Algorithm

The system employs Logistic Regression as the primary supervised learning algorithm for news classification. Logistic Regression is well-suited for high-dimensional textual data and provides interpretable classification results. It ensures efficient training and reliable performance for multi-class text classification problems.

6. Database Management System

MySQL is used as the relational database management system for storing user credentials, hospital information, health tips, and system data. The database enables secure storage, efficient retrieval, and structured management of application data. It supports CRUD (Create, Read, Update, Delete) operations required for both admin and user modules.

7. Frontend Technologies

The user interface of the system is developed using HTML, CSS, and basic JavaScript. These technologies provide a responsive and user-friendly experience for both administrators and end users. The frontend enables user registration, login, news submission, and viewing of hospital and health-related information.

8. Development Environment

Visual Studio Code is used as the Integrated Development Environment (IDE) for coding, debugging, and managing project files. It offers extensions, syntax highlighting, and version control integration that enhance development productivity.

9. Model Development and Testing Environment

During the development phase, Jupyter Notebook was used for exploratory data analysis, preprocessing experiments, feature engineering, model testing, and performance visualization. It allowed interactive experimentation and efficient comparison of model results before deployment.

10. System Deployment

The application can be deployed on a web server or cloud hosting platform to ensure accessibility and scalability. Deployment enables real-time interaction between users and the trained machine learning model, allowing instant verification of COVID-19 news content.

V Result



News Veracity Checker Page

VI. CONCLUSION

The Intelligent News Veracity Classifier for COVID-19 Misinformation presents an effective and practical solution to address the rapid spread of false and misleading information during global health crises. By leveraging Natural Language Processing techniques and supervised machine learning approaches, the system successfully classifies COVID-19-related news articles into True, False, or Misleading categories with reliable performance. The integration of text

preprocessing, TF-IDF feature extraction, SMOTE-based class balancing, and Logistic Regression ensures accurate and unbiased prediction results.

The developed system not only focuses on misinformation detection but also provides additional value through its web-based architecture. The platform enables real-time news verification, secure user authentication, and administrative management of hospital information and health tips. This dual functionality enhances both technological effectiveness and practical usability. By incorporating tools such as scikit-learn for model development and a Flask-based deployment framework, the system demonstrates scalability, efficiency, and ease of integration into real-world environments.

Experimental results indicate satisfactory accuracy and macro F1-score performance, confirming the reliability of the proposed model across multiple classification categories. The use of SMOTE to address class imbalance further improves fairness and reduces bias toward majority classes. Comprehensive testing procedures validate that the system operates securely, consistently, and efficiently under different conditions.

Overall, this project contributes to combating digital misinformation by providing an automated, intelligent, and accessible news verification platform. It supports public health awareness, informed decision-making, and trust in credible information sources. The system serves as a foundational step toward developing more advanced AI-driven fact-checking solutions capable of handling emerging misinformation challenges in future global crises.

REFERENCES

- [1] H. Alhakami, W. Alhakami, A. Baz, M. Faizan, M. W. Khan, and A. Agrawal, "Evaluating intelligent methods for detecting COVID-19 fake news on social media platforms," *Electronics*, vol. 11, no. 15, p. 2417, Aug. 2022, doi: 10.3390/electronics11152417.
- [2] V. Rathinapriya and J. Kalaivani, "An intelligent feature selection-based fake news detection model for pandemic situation with optimal attention based multiscale densenet with long short-term memory layer," *Applied Soft Computing*, vol. 176, p. 113158, May 2025, doi: 10.1016/j.asoc.2025.113158.
- [3] A. R. Sanaullah, A. Das, A. Das, M. A. Kabir, and K. Shu, "Applications of machine learning for COVID-19 misinformation: A systematic review," *Cognitive Computation*, vol. 14, no. 6, pp. 1–22, 2022, doi: 10.1007/s12559-022-10032-0.
- [4] A. Ul Hussna, M. G. R. Alam, R. Islam, B. F. Alkhamees, M. M. Hassan, and M. Z. Uddin, "Dissecting the infodemic: An in-depth analysis of COVID-19 misinformation detection on X (formerly Twitter) utilizing machine learning and deep learning techniques," *Heliyon*, vol. 10, no. 18, p. e37760, Sep. 2024, doi: 10.1016/j.heliyon.2024.e37760.
- [5] M. Nabeel and C. Große, "Classifying COVID-19 disinformation on Twitter using a convolutional neural network," in *Proceedings of the 8th International Conference on Information Systems Security and Privacy (ICISSP 2022)*, SCITEPRESS – Science and Technology Publications, 2022, pp. 264–272, doi: 10.5220/0010774800003120.
- [6] S. Malek, C. Griffin, R. D. Fraleigh, R. Lennon, V. Monga, and L. Shen, "Intervention in health misinformation using large language models for automated detection, thematic analysis, and inoculation: Case study on COVID-19," *Journal of Medical Internet Research*, vol. 28, p. e75500, Jan. 2026, doi: 10.2196/75500.
- [7] X. Song, J. Petrak, Y. Jiang, I. Singh, D. Maynard, and K. Bontcheva, "Classification aware neural topic model for COVID-19 disinformation categorisation," *PLoS ONE*, vol. 16, no. 2, p. e0247086, Feb. 2021, doi: 10.1371/journal.pone.0247086.
- [8] A. Peda Gopi, G. Sai Avanthi, M. Himani, M. Leela Swarna, and B. Achitha, "Fake News Detection on Social Media Using Block Chain," *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 11, no. 10, pp. 142-152, Oct. 2024. ISSN: 2349-5162.
- [9] D. A. Juraev, "Fake News Detection Using Natural Language Processing and Machine Learning," *The Chitransh Academic & Research Journal*, vol. 1, no. 4, pp. 275-299, Sept. 2025. ISSN: 3049-303X. DOI: 10.5281/zenodo.17239827.
- [10] D. A. Juraev, "Fake News Detection Using Natural Language Processing and Machine Learning," *The Chitransh Academic & Research Journal*, vol. 1, no. 4, pp. 275-299, Sept. 2025. ISSN: 3049-303X. DOI: 10.5281/zenodo.17239827.