

NewsTrust: AI-Driven News Verification with Social Publishing Features

Akshita Sharma¹, Diksha Kumawat¹, Raj Mehta¹, Punit Kumar²

¹Students, Department of Artificial Intelligence, Poornima Institute of Engineering and Technology, Jaipur, India

²Assistant Professor, Department of Artificial Intelligence, Poornima Institute of Engineering and Technology, Jaipur, India

ABSTRACT

The proliferation of fake news poses a critical threat to democratic institutions, journalistic integrity, and public perception. This research presents NewsTrust, an innovative web-based platform designed to counter misinformation through automated news verification. Developed using the Django framework, the system leverages advanced Natural Language Processing (NLP) techniques in a two-tiered validation process. It combines real-time verification using live news APIs with semantic similarity analysis against a large-scale vector database of pre-embedded news articles. Based on computed similarity thresholds, the system autonomously approves, flags, or rejects user-submitted content. Additionally, NewsTrust facilitates user interaction, content creation, and engagement within a secure and trustworthy digital ecosystem.

Keywords: artificial intelligence, Machine Translation, Natural Language Processing, Semantic Similarity.

1. INTRODUCTION

1.1 Overview of Fake News Detection and the Role of AI.

The emergence of digital platforms has transformed the way news is consumed, while simultaneously facilitating the swift dissemination of false information, which poses a risk to public trust and the integrity of the media. Artificial intelligence presents a viable solution by employing natural language processing and similarity detection to assess the authenticity of news in real time. Although this enhances both accuracy and efficiency, it also brings forth apprehensions regarding excessive dependence on automated technologies. News Trust exemplifies the effective application of AI in the fight against misinformation.

1.2 Motivation and Objective

This research paper stems from the rising concern over misinformation today's digital world and the necessity for smart systems capable of verifying the authenticity of news before it circulates. As artificial intelligence technologies grow more powerful and widely available, there is a significant opportunity to utilize them in creating platforms that can automatically identify and manage false information.

1.3 Significance of the Study

This study emphasizes the need to strike a balance between technical training and linguistic proficiency in translation education. It also offers insights into developing a curriculum that aligns with industry expectations while enhancing students' technical abilities.

2. REQUIREMENTS

2.1 Integration of Critical Thinking Exercise

Create system elements that assess and judge the reliability of news articles by utilizing AI technologies. This plan involves setting up a two-step verification process: initially, employing a Live News API to find related news stories, and subsequently, using BERT embedding along with Vector DB similarity analysis. The classification system (approve, review, reject) highlights the capabilities and limitations of AI in identifying false information.

2.2 Focus on Contextual and Source Evaluation: Add training

Implement training sessions that focus on media literacy, evaluating sources, and analysing context. This will prepare students with the ability to spot deceptive narratives and supplement AI verification with human insight, particularly when cultural or contextual indicators are subtle or intricate.

2.3 Encouragement of Adaptive Learning through AI Feedback Loops:

Establish a responsive feedback system where the platform learns from past verifications and administrative choices.

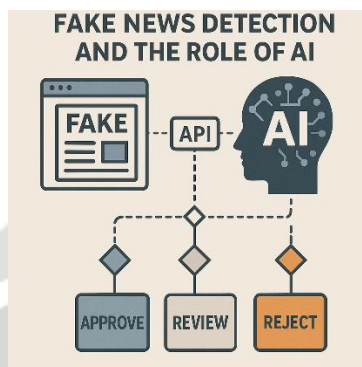


Fig.1 NewsTrust Platform Architecture

3. LITERATURE REVIEW

3.1 Existing Fake News Detection Systems.

Various tools have emerged to combat fake news, ranging from manual fact-checking platforms like Snopes and Politick to automated solutions. While machine learning has enhanced detection accuracy, many systems still face challenges in real-time application and large-scale misinformation management.

3.2 Limitations in Current Approaches

Current approaches frequently rely on antiquated training data and gaze algorithms, despite advancements in technology. Stated differently, it lacks the flexibility to adjust to changing linguistic trends and popular subjects. Additionally, a lot of systems are unable to pick up on minute variations in sound or intent. B. Differentiating deliberate disinformation from satire. Furthermore, faults may result from a heavy reliance on surface-level text elements.

3.3 Role of NLP and AI in News Verification

Semantic understanding of text has improved thanks to modern AI and natural language processing techniques like BERT, and GPT. These models enable systems to more accurately assess intent, context, and meaning. With the use of semantic similarity algorithms, these technologies allow real-time comparison with genuine sources in applications such as NewsTrust, leading to more dependable and scalable fake news identification.

4. METHODOLOGY

4.1 System Architecture. The platform is built using a modular architecture with three key layers:

Frontend: Developed using HTML, CSS, and JavaScript for a user-friendly interface that allows users to post and interact with news content.

Backend: Implemented with Django (Python-based web framework), responsible for handling user management, content moderation, and AI integration.

Database Layer: Uses SQLite/MySQL for storing user data, news articles, verification results, and system logs.

4.2 Fake News Detection Workflow

The process for detection and verification involves two main steps:

Comparison with Live News API: When a user submits news content, the system initially checks a News API (such as NewsData.io or News API) to find similar articles from reliable media sources. If credible articles are located, the news is categorized as "Verified."

AI-driven Semantic Similarity Assessment:

If no direct matches are identified, the news content is processed with BERT (Bidirectional Encoder Representations from Transformers) to grasp the underlying semantics. These embeddings are then assessed against a Vector Database (FAISS) that holds validated news vectors using cosine similarity. Depending on established similarity thresholds:

- High similarity: Approved
- Medium similarity: Sent to admin
- Low similarity: Rejected

This multi-layered system ensures scalability, real-time performance, and a strong balance between automation and manual oversight.

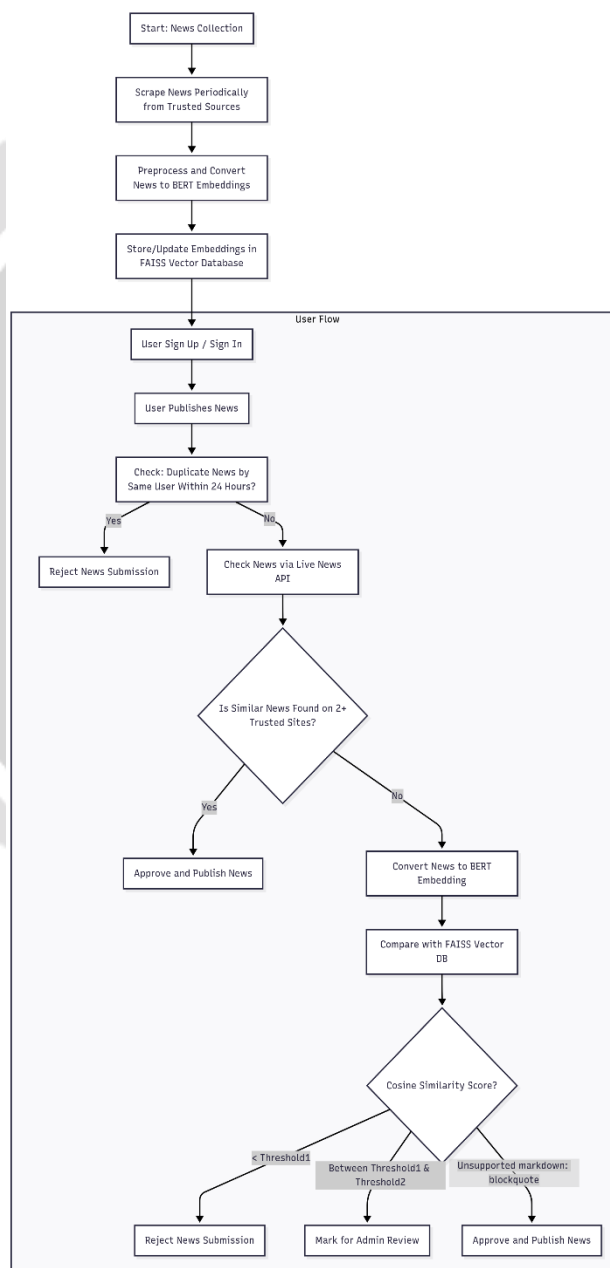


Fig. 2 Data Flow Diagram

In Fig. 2 the following Data Flow Diagram (DFD) illustrates the end-to-end flow of data in the NewsTrust verification process.

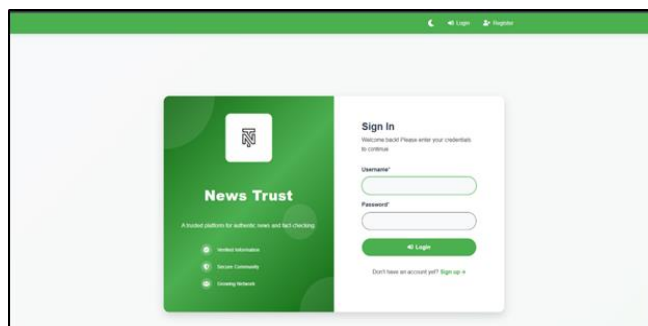


Fig 3. Home page of Fake News Detection

In above fig, web application for fake news detection where, the total backend is operated when the URL link is inserted or searched. The URL is generally copied from the browsers. The URL links must be related to news article.

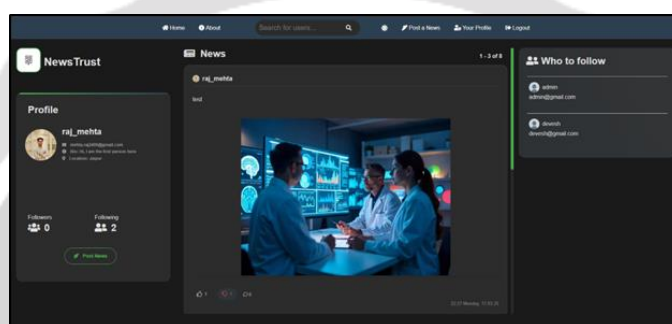


Fig 4. Profile Section of Fake News Detection

As in Fig 4, the Profile Section offers users a centralized view of their activity on the NewsTrust platform, including submitted articles, verification status, and engagement metrics. It also displays a credibility score, encouraging responsible sharing. This section enhances transparency and user accountability while ensuring easy access to personal content.

4.3 Application Logic and API Integration

The core logic of the platform resides in the Django backend and operates as follows:

User Submission Handling:

Users submit a news headline or a URL through the frontend interface. This triggers a backend API call that processes the content.

Live News API Comparison:

The system sends a GET request to third-party news APIs (e.g., NewsData.io, NewsAPI.org) with keywords extracted from the headline or URL content using NLP. If matching or highly similar articles are returned: The backend flags the news as Verified. Stores API metadata (title, source, date) in the database.

BERT Embedding and Vector Similarity:

If no credible news is returned by the API:

The article's text is passed through a pre-trained BERT model (Bert-base-uncased) to generate embedding.

Cosine similarity is calculated. If the similarity score exceeds a defined threshold (e.g., 0.75), it is marked likely True; otherwise, Likely Fake.

Result Update:

The final classification (Verified, Likely Fake, or Needs Review) is saved in the PostgreSQL database and reflected in the frontend in real time.

4.4 Database Design and Models

The database design of the NewsTrust platform follows a normalized relational structure, built using SQLite/MySQL and managed via Django’s Object-Relational Mapping (ORM). This robust schema is engineered to ensure data integrity, operational efficiency, and scalability, while seamlessly supporting the platform’s core functionalities — including user authentication, news submission, AI-driven verification workflows, content interaction, and system-level logging

Key considerations in the design include:

- **Normalization:**
The schema is carefully normalized to reduce data redundancy and optimize query performance. It allows the platform to handle high-volume data without compromising speed or consistency.

- **Foreign Key Relationships:**

Django models are interlinked through foreign keys, establishing strong referential integrity between entities such as users, news articles, verification outcomes, and user activity logs. This interconnection provides a structured and reliable data flow throughout the application.

- **Vector Storage for Semantic Search:**

A dedicated model is implemented to store high-dimensional BERT-based vector embeddings, generated during semantic similarity analysis. These embeddings are stored in an optimized format for efficient retrieval and comparison against a vector database, allowing precise AI-based fake news detection.

- **Timestamping and Auditing:**

All core models include auto-managed created at and updated at fields, enabling traceability and supporting features like real-time moderation, chronological sorting, and audit logging.

- **Modular and Scalable Design:**

The database structure is modular, offering flexibility to extend functionality in future releases — such as integrating a commenting system, reaction metrics (likes/dislikes), multi-tiered user roles, or multilingual content support.

Verification Logs and Credibility Tracking:

Additional tables track verification events and maintain **user credibility scores** based on their activity and sharing patterns. This contributes to the platform’s transparency and user accountability.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
News																	
1																	
2																	
3																	
4																	
5																	
6																	
7																	
8																	
9																	
10																	
11																	
12																	
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	
21																	
22																	
23																	
24																	
25																	
26																	
27																	
28																	
29																	
30																	

Fig 5. Database Storage-Datasets (Excel Sheet)

5. Conclusion

As social media continues to gain traction, an increasing number of individuals are turning to these platforms for news consumption rather than relying on traditional news outlets. However, the rise of social media has also facilitated the dissemination of misinformation, which adversely affects both individual users and society at large. This misinformation is often deliberately crafted to incite violence and anger within communities. The younger generation, in particular, is frequently subjected to significant mental health repercussions as a result.

In a time characterized by the swift spread of information, the challenge of differentiating credible news from misinformation has emerged as a pressing global issue. The NewsTrust platform offers a well-conceived and technologically advanced solution to this dilemma, merging contemporary web development techniques with AI-driven semantic analysis to promote responsible news consumption and distribution.

By utilizing BERT-based semantic similarity algorithms alongside real-time data from Live News APIs, the system adopts a hybrid methodology that harmonizes computational intelligence with human insight. When a user submits a news article, the platform first seeks to identify verified sources through reliable APIs. If a direct match is not found, the system applies deep learning methods to assess the semantic meaning of the submitted content against a curated vector database of verified news articles. This multi-layered verification approach enhances detection accuracy and adapts to the changing patterns of misinformation that often escape keyword-based or superficial analysis. From a technical perspective, NewsTrust is constructed with a modular and scalable architecture that guarantees maintainability, responsiveness, and extensibility. The frontend is crafted using HTML, CSS, and JavaScript to create an intuitive user interface, while the backend, powered by Django, manages user accounts, news submissions, content moderation, and the integration of AI logic. The platform employs PostgreSQL and Django ORM for its database, ensuring a structured, normalized, and scalable approach to data storage. Additionally, it features an intelligent profile section that fosters user accountability by showcasing engagement statistics, verification status, and a credibility score.

The platform's design incorporates features such as timestamping, vector storage for semantic embeddings, and foreign key-linked models, ensuring that it is not only efficient but also prepared for audits and future developments. Users can interact with the platform across various devices.

References

- [1] Jiawei Zhang, Bowen Dong, Philip S. Yu, "FAKE DETECTOR: Effective Fake News Detection with Deep Diffusive Neural Network" published in August 2019
- [2] Steni Mol T S and Shreeja P Sin, "Fake News Detection on Social Media-A Review" published in April 2020.
- [3] Monther Aldwairi and Ali Alwahedin, "Detecting Fake News in Social Media Networks" published in 2018.
- [4] Natali Ruchansky, Sungyong Seo and Yan Liu, "CSI: A Hybrid Deep Model for Fake News Detection".
- [5] Z. Nie, "Study on the application of training modes to English teachers in universities in the Internet+ era," *Boletin Tecnico/Technical Bulletin*, Vol. 55, no. 11, pp. 504-511, 2017.
- [6] K. Liu, "An Empirical Study on English O2O Education System in Higher Education Based on MOOC Training," *Revista de la Facultad de Ingenieria*, vol. 32, no. 14, pp. 784789, 2017.
- [7] Y. Hao and J. Su, "A Study on the Flipped Classroom Teaching Model of University English Considering Educational Information," *Revista de la Facultad de Ingenieria*, Vol.
- [8] Y. Zhang, "Improvement and optimization of functional modules of university English distance education based on intercultural theory," *Revista de la Facultad de Ingenieria*, vol. 32, No. 15, pp. 174-177, 2017. 32, no. 14, pp. 41-46, 2017.
- [9] C.J. Lin, "An Online Peer Assessment Approach to Support Mind Mapping Flip Learning Activities in University English Writing Courses," *Journal of Computers in Education*, Vol. 6, No. 3, pp. 385-415, 2019.
- [10] E. Pakoci, B. Popovic, and D.J. Pekar, "Improving Serbian Speech Recognition Using Sequence-Trained Deep Neural Networks," *SPIRAS Proceedings*, Vol. 3, No. 58, p. 53, 2018.
- [11] W. Bo, K. Li, F. Ge, et al., "An end-to-end deep learning approach to simultaneous phonetic cloning and acoustic modeling for robust speech recognition," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 11, No. 8, 2016. S. 1289-1300, 2017.
- [12] P. Glauner, J. A. Meira, and, P. Valtchev, R. State, and F.

- Bettinger, "Challenges of Non-Technical Loss Detection Using Artificial Intelligence: A Survey," *International Journal of Computational Intelligence Systems*, Vol. 10, No. 1, pp. 135-142. 760, 2017.
- [13] P. Havea, N. Meratnia, and M. Bahrepour, "AI-based event detection in wireless sensor networks," *University of Twente*, vol. 85, no. 6, S. 1553–1562, 2017.
- [14] L. Caviglione, M. Gaggero, J. F. Lalande, W. Mazurczyk, and M. Urbanski, "Uncovering hidden,communications of mobile malware via energy consumption and AI,"*IEEE Transactions on Information Forensics and Security*,, vol. 11, Nr. 4, S. 799–810, 2017.
- [15] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, "Neuroscience-inspired artificial intelligence," *Neuron*, Vol. 95, No. 2, p. 245–258, 2017. Kai Shu, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective".
- [16] Junaed Yonus Khan, Md. Tawkat Islam Khondaker, Anindya Iqbal, Sadia Afroz, "A Benchmark Study on Machine Learning Methods for Fake News Detection" published in May, 2019.
- [17]. Xinyi Zhou, Reza Zafarani, Kai Shu, Huan Liu, "Fake News: Fundamental Theories, Detection Strategies and Challenges" published in January 2019.
- [18]. Nichlar Olivirea, Pedro Pisa, Martin Lopez, Diogo Mattos, "Identifying Fake News on Social Networks Based on Natural Language Processing: Trends and Challenges" published in December 2020.
- [19]. Srishti Agrawal, Vaishali Arora, Ruchika Arora, Pronika Chawla, Madhumita Kathuria. "FAKE NEWS DETECTION USING ML".

