International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 10 Issue 1, Jan-Feb 2026 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Newstruth: Distinguish Fake and Real News Using NLP

Dr. Srinivasa Reddy Kurukuntla[1], Adupala Sathwik Reddy[2], Bandla Bhavana[3], Chintha Saketh Pavan[4], Bantu Sairam[5], Chikati Kiran[6]

[1]Associate Professor, Department of ECE, Teegala Krishna Reddy Engineering College, Hyderabad, Telangana, India
[2,3,4,5,6]Student, Department of ECE, Teegala Krishna Reddy Engineering College, Hyderabad, Telangana, India

ABSTRACT

In the modern digital era, the rapid spread of misinformation and fake news across online platforms has become a major global concern. The proliferation of false information not only misleads the public but also influences opinions, behaviours, and even political or economic outcomes. To address this issue, Natural Language Processing (NLP) techniques are being increasingly utilized to automatically distinguish between real and fake news. This study focuses on developing an intelligent system that analyses linguistic patterns, semantic features, and contextual cues within news articles to identify their authenticity. Using machine learning and deep learning models such as Naïve Bayes, Support Vector Machines (SVM), and transformer-based architectures like BERT, the proposed system learns to classify news as real or fake based on textual content. The dataset undergoes preprocessing steps including tokenization, stop-word removal, stemming, and vectorization to ensure effective model training. The experimental results demonstrate that NLP-driven approaches significantly improve the accuracy of fake news detection, offering a robust and automated solution to combat misinformation. This research contributes to the development of reliable fact-verification tools that enhance digital media integrity and promote trustworthy information dissemination.

KEYWORDS: *Natural Language Processing, Naïve Bayes, Support Vector Machine, Random Forest.*

How to cite this paper: Dr. Srinivasa Reddy Kurukuntla | Adupala Sathwik Reddy | Bandla Bhavana | Chintha Saketh Pavan | Bantu Sairam | Chikati Kiran "Newstruth: Distinguish Fake and Real News Using NLP" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-10 | Issue-1, February 2026, pp.940-947, URL: www.ijtsrd.com/papers/ijtsrd100157.pdf

IJTSRD100157

Copyright © 2026 by author (s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0)

I. INTRODUCTION

In today's digital era, the rapid growth of online media has made information easily accessible to everyone. However, this convenience has also led to the widespread circulation of fake news - false or misleading information created to manipulate public opinion, spread misinformation, or gain attention. Distinguishing fake news from real news has therefore become a major challenge in maintaining the integrity of information shared on social platforms and news websites Natural Language Processing (NLP), a subfield of Artificial Intelligence (AI), provides effective tools and techniques to analyse and understand the textual content of news articles. By leveraging NLP, systems can automatically detect linguistic patterns, sentiment, writing styles, and contextual clues that differentiate fake news from real news. Machine learning algorithms combined with NLP features such as tokenization, TF-IDF, word embeddings, and sentiment analysis can help classify news articles accurately. The main goal of this paper is to build an automated model that can analyze textual data and predict whether a news article is genuine or deceptive. Such a system can play a crucial role in promoting trustworthy journalism, preventing misinformation, and ensuring that readers have access to reliable and fact-based content.

II. LITERATURE REVIEW

Wang (2017).[1] introduces LIAR, a large-scale benchmark dataset of 12.8K short political statements from PolitiFact, labelled across six fine-grained truthfulness categories and enriched with speaker metadata and evidence links, addressing the lack of realistic datasets for fake-news detection. The study shows that text-only linguistic features perform poorly due to lexical overlap and topic bias, and proposes a hybrid neural model that combines text with metadata such as speaker history and political affiliation, achieving better performance. The work highlights that effective fact-checking often requires external knowledge and evidence retrieval, establishing LIAR as a foundational benchmark that spurred research in metadata-aware, cross- domain, and evidence-based misinformation detection.

Shuet al. (2017)[2] present a comprehensive survey of fake news detection from a data-mining perspective, organizing the field into characterization, detection, and mitigation. The paper links psychological and social theories of misinformation spread to measurable social media signals and reviews content-based, context-based, and hybrid detection approaches, including features, models, and datasets. It highlights evaluation challenges such as dataset bias and poor cross-domain generalization, and outlines future directions including explainable AI, early detection, adversarial robustness, and scalable real-time systems.

Kaliyar et al. (2021).[3] propose FakeBERT, a hybrid fake-news detection model that combines BERT's contextual embeddings with parallel CNN blocks to capture both global context and local linguistic patterns. Evaluated on benchmark datasets such as LIAR and FakeNewsNet, FakeBERT consistently outperforms fine-tuned BERT and traditional LSTM/CNN baselines in accuracy and F1, especially for short and noisy text. The study shows that integrating transformer representations with task-specific inductive biases improves detection robustness, while also noting challenges like domain sensitivity and the need for explainable, evidence-aware models.

Qinet al. (2024).[4] address the challenge of poor cross- domain generalization in fine-tuned BERT models for fake news detection, showing that naive end-to-end fine-tuning often overfits to dataset-specific patterns. Through systematic evaluation of strategies such as gradual unfreezing, adapter-based parameter-efficient tuning, and contrastive objectives, the study demonstrates improved robustness and transfer performance across topics and time splits. The authors provide practical training guidelines and emphasize the need for generalization-focused benchmarks, while highlighting remaining risks of bias under domain shifts.

Nadeem et al. (2023).[5] introduce **SSM**, a multimodal fake news detection framework that combines stylometric features, semantic similarity to authoritative evidence, and image–text consistency to improve verification beyond surface-level text cues. By fusing these signals through attention-based classification, SSM outperforms text-only baselines on multimodal datasets, demonstrating the effectiveness of integrating writing style, evidence alignment, and visual context.

"Fake News Detection [6]: Comparative Evaluation of BERT-like Encoders and Autoregressive LLMs (encoder vs decoder evaluation)" that encoders are efficient and strong for supervised classification, while decoder LLMs excel in few-/zero-shot settings and generate richer explanations but risk hallucinations and prompt bias. The study demonstrates that calibration methods, prompt design, and evidence- aligned training improve reliability, and that a hybrid pipeline-encoder classifiers for prediction plus decoder LLMs for explanation and evidence checking-performs best.

Rucha sky et al. (2017).[7] propose **CSI (Capture–Score–Integrate)**, a hybrid deep model that jointly analyzes article content, user engagement behavior, and source characteristics to improve fake news detection. The Capture module models textual and temporal user-response patterns, the Score module learns user/source suspiciousness representations, and the Integrate module fuses these signals for final prediction, demonstrating that coordinated behavior patterns are key indicators of misinformation. Experiments show significant performance gains over text-only methods, and the work inspired later graph-based and behavior-aware detection approaches while highlighting practical issues such as evolving user behavior and privacy concerns.

Liuet al. (2018).[8] focus on **early fake news detection** by modeling propagation paths instead of relying on static content or final engagement statistics, arguing that early repost sequences and temporal patterns contain strong discriminative signals. Their approach encodes partial propagation trees using time-aware RNNs and path-level structural features, enabling accurate classification at early time horizons before news goes viral. Experiments show superior early-stage detection performance and robustness to adversarial spread patterns, while highlighting trade-offs between prediction speed and false alarm rates for real- world deployment.

Phanet al. (2023).[9] survey the growing use of **Graph Neural Networks (GNNs)** for fake news detection, highlighting how misinformation spreads over social networks where relational patterns like user interactions and community structure provide key signals. The review categorizes GNN-based approaches by graph construction strategies, propagation modeling, and learning objectives tailored to verification tasks. It shows that GNNs effectively capture user–content relationships and cross-post dynamics that flat feature models struggle to represent, making them powerful tools for misinformation analysis.

COVID-19 Misinformation Studies: Detection, Characterization, and Intervention [10] Misinformation brought together NLP, epidemiology, and public health to study how false health claims spread and how to counter them. These works introduced pandemic-specific datasets, analyzed topic shifts, bot/influencer roles, and built classifiers that leverage medical terminology and evidence from trusted health sources. Findings show misinformation strongly affected public behavior, and that domain knowledge plus interventions like authoritative corrections and platform labeling were key to effective mitigation.

Temporal and Longitudinal Modeling for Fake News Detection.[11] temporal and longitudinal modelling examine how fake news evolves over time, using signals like early burst patterns, corrections, and long-term author behaviour to improve detection. Approaches include time- series RNNs, temporal graph neural networks, and change- point detection to capture dynamic spread patterns. These time-aware methods enhance early detection and lower false alarms but require streaming systems and can introduce latency in real-time deployment.

III. FLOWCHART

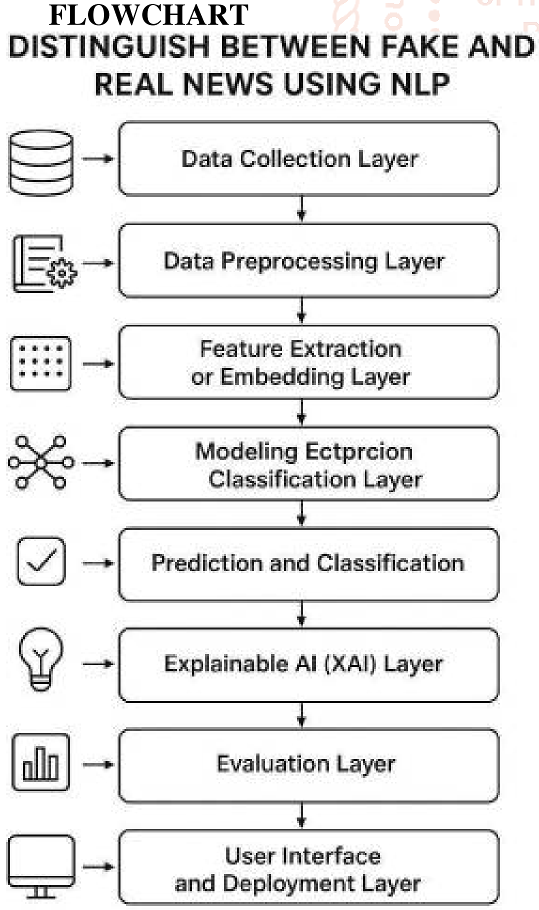

Figure 1: Flowchart of Fake / Real news Detection

1. **Data Collection Layer:** This is the first stage of the system where raw data is gathered from various sources such as online news portals, social media platforms, blogs, and fact- checking websites. The dataset includes both real and fake news articles to ensure balanced training. Each news item is labeled according to its authenticity. Purpose: • Collect diverse data across multiple topics (politics, sports, health, entertainment). •Build a rich dataset for model training and evaluation.
2. **Data Preprocessing Layer** Once data is collected, it is often messy, unstructured, and filled with unnecessary symbols or noise. The preprocessing layer cleans and prepares this text data for further analysis. Operations include:
   - Removing stop words, punctuation, numbers, URLs, and emojis.
   - Converting text to lowercase for uniformity.
   - Ensure that only meaningful textual information is passed to the model.
3. **Feature Extraction or Embedding Layer** Text data must be transformed into numerical vectors before being fed into machine learning models. This is achieved through feature extraction or embeddings. Techniques used: • TF- IDF, Word2Vec, Glove for traditional approaches. •BERT, Roberta, or DistilBERT for deep contextual embeddings. Purpose: • Convert words and sentences into dense vector representations. • Capture semantic meaning and context, enabling the model to understand relationships between words.
4. **Modelling or Classification Layer** This layer is the core engine of the system. It employs Transformer-based models (like BERT or Roberta) that can understand contextual dependencies and long-range relationships in text. Functions: • Fine-tune transformer models on labelled fake and real news datasets. • Learn linguistic cues like exaggeration, emotion, and bias that are common in fake news.
5. **Prediction and Classification** Layer In this stage, the trained model is used to classify unseen news data. When new input text is provided, it passes through preprocessing and embeddings, and the model predicts whether it is Fake or Real. Purpose: • Provide real-time classification with confidence scores. • Support automated misinformation detection and alerting.
6. **Explainable AI (XAI)** Layer To increase transparency and trust, this layer provides insights into why the model made a specific decision.

Using methods like attention visualization, SHAP, or LIME, the system highlights the keywords or sentences that influenced its classification. Example: If the sentence "Government giving ₹10 lakh to everyone for free" is detected as fake, the words "₹10 lakh" and "free" might be highlighted as suspicious indicators. Purpose: • Improve interpretability and accountability. • Help journalists or analysts verify model decisions.

7. **Evaluation Layer** After model development, this layer assesses system performance using quantitative metrics such as accuracy, precision, recall, and F1-score. Functions: • Validate model effectiveness on test datasets. • Detect bias or overfitting through confusion matrix analysis •Evaluate multilingual and domain-specific performance. Purpose: Ensure reliability, robustness, and fairness in predictions
8. **User Interface and Deployment Layer** Finally, the results are presented to end-users through a simple and interactive interface. This can be a web portal, dashboard, or mobile app where users can paste or upload news text for verification. Features: • Displays output as Real News, Fake News, or Needs Human Verification. • Shows confidence percentage and highlighted keywords (from XAI layer). • Provides APIs for integration into social media platforms or browser extensions. Purpose:

IV. PROPOSED SYSTEM:

The proposed system builds upon the limitations of existing models and introduces enhancements for better performance, scalability, and trust in multilingual fake news detection. It integrates hybrid transformer architectures, ensemble learning, explainable AI, and multimodal data handling.

1. Multilingual & Multimodal Data Collection

Fake news does not always appear in just plain text. It may include images, memes, and videos.

The proposed system collects:

- Textual data (articles, posts, blogs, tweets).
- Image-text combinations (memes, infographics).
- Multilingual news headlines and body content.

This multimodal approach improves robustness against misinformation campaigns that use mixed content.

2. Advanced Pre-processing Layer

Besides basic cleaning, tokenization, and normalization, the proposed system introduces:

- Language Identification (Lang-ID): Automatically detects the language of the text.
- Code-Mixing Handling: Deals with sentences mixing two or more languages (e.g., "Yeh news fake Hai bro, don't believe.").
- Image Pre-processing: For memes, embedded text is extracted using OCR (Optical Character Recognition).
- This ensures a wider range of inputs can be understood properly.

3. Embedding Layer (Hybrid Representations)

The system uses hybrid embeddings to capture meaning across multiple languages and content types.

Components include:

- mBERT/XLM-RoBERTa embeddings for textual data.
- FastText embeddings for low-resource languages.
- Image embeddings (from CNN-based models like ResNet for visual content).

4. Hybrid Transformer & Ensemble Layer

The heart of the proposed system.

Instead of relying on just one transformer, it combines multiple models:

- XLM-RoBERTa for deep multilingual understanding.
- DistilBERT for lightweight inference.
- ALBERT for memory-efficient training.

An ensemble strategy (majority voting or weighted averaging) combines predictions, boosting accuracy and robustness.

This ensures the system works well for both high-resource and low- resource languages.

5. Sentiment & Context Analysis Module

Fake news often manipulates emotions.

A sentiment analyser checks for unusually strong emotions (fear, anger, excitement) in the news.

Contextual verification ensures that sentences with factual consistency.

Example: A sentence like "Cure for COVID discovered in 2 days" gets flagged due to unrealistic context.

6. Knowledge Graph & Fact-Checking API Integration

- The system connects to knowledge bases (like Wiki data, DB pedia) and fact-checking APIs (like Google Fact Check Tools).
- This allows cross-verification of entities, events, and claims.
- Example: If a post says "Barack Obama is the current president of the USA," the knowledge graph can immediately flag it as false.

7. Evaluation Module
The proposed system evaluates models more thoroughly using:
- Cross-lingual accuracy (performance across different languages).
- Confusion matrix analysis (false positives vs false negatives).
- Real-time performance testing (to ensure quick detection on social media).

8. Explainability Layer
- Like the existing system, but enhanced.
- Uses attention visualization, SHAP (SHapley Additive Explanations), and LIME to explain predictions.
- Example: For a fake news headline, it shows why it was flagged (keywords, emotional tone, and fact-check mismatch).

9. Output Layer
Provides a final classification of the news as:
- Real News
- Fake News
- Possibly Misleading / Needs Human Verification

This layered output avoids overconfidence and supports fact-checkers.

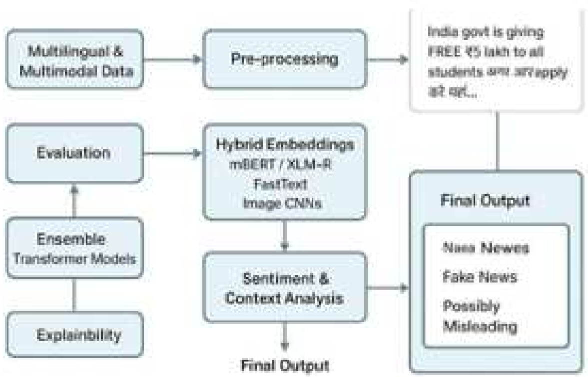

Figure 2: Proposed System Architecture

Blocks - Explanation

1. Multilingual & Multimodal Data (Input)
- Sources: news sites, blogs, Twitter/Facebook posts, WhatsApp forwards, memes (image+ text), videos.
- Languages: English, Hindi, Telugu, Spanish, etc. and code-mixed text (e.g., Hinglish).
- Data types: raw text, HTML, images (memes), screenshots, embedded links.

2. Pre-processing
- Language Identification (Lang-ID): detect language(s) per item.
- Cleaning: remove noisy tokens, URLs, duplicate posts, boilerplate.
- Code-mix handling: split/mask mixed language tokens. OCR for images: extract text from memes, screenshots.
- Normalization: lowercasing, diacritics handling, stemming/lemmatization where useful.

3. Hybrid Embeddings
- Textual: mBERT /XLM-RoBERTavectors for contextual multilingual embeddings.
- Low-resource fallback: FastText or language-specific embeddings for languages with few resources.
- Visual: CNN (ResNet/efficient models) embeddings for memes/images.
- Combined vector: concatenate or project text+image embeddings into unified representation.

4. Ensemble Transformer Models (Core Detection Engine)
- Multiple fine-tuned transformer variants run in parallel (e.g., XLM- R, DistilBERT, ALBERT).
- Ensemble method: majority voting, weighted averaging or stacking to reduce single-model bias.
- Benefits: higher robustness across languages, better generalization, lower susceptibility to adversarial phrasing.

5. Sentiment & Context Analysis
- Sentiment scoring to detect unusually extreme emotional content (fear, anger, urgency).
- Context checks for unrealistic claims (e.g., "cure in 2 days", "free₹10 lakh").
- Helps flag emotionally manipulative or contextually implausible items even if wording isn't obviously fake.

6. Knowledge Graph & Fact-Checking Integration
- Cross-verify named entities, dates, factual claims against trusted KBs (Wikidata, official APIs) and fact-check databases.
- If claim contradicts KB or matches debunked claims, raise confidence for "Fake."
- Useful for verifiable factual assertions (who, when, where, how much).

7. Explainability Module
- Attention visualization, LIME, or SHAP show which tokens/phrases/images influenced the decision.
- Provide human-readable reasons: highlighted keywords, supporting/contradicting facts from KB.

8. Evaluation & Monitoring
- Offline metrics: accuracy, precision, recall, F1 per language and overall.
- Confusion matrices to spot bias (e.g., over flagging a particular language).
- Live monitoring: inference latency, throughput, model drift detection, periodic re-training triggers.

9. Output Layer
- Final labels: Real News, Fake News, Possibly Misleading / Needs Human Verification.
- Metadata: confidence score, highlighted evidence, KB matches, timestamps, language tag.
- Delivery: dashboard, API, browser plugin, social-platform moderation queue.

10. Deployment & Scalability Considerations
- Use model distillation (Distil BERT) or quantization for low- latency inference.
- Horizontal scaling via microservices.
- Edge vs cloud trade-offs: sensitive data can be pre- filtered at edge; heavy models run in cloud.

11. Security & Privacy
- Encrypt stored embeddings and PII; follow data retention best practices.
- Anonymize user metadata where possible.
- Access control for fact-checker dashboards.

12. Feedback Loop & Human-in-the-Loop
- Flagged items sent to human verifiers; their labels used to retrain the model.

V. IMPLEMENTATION OF THE SYSTEM

The relevant implementation details are the following:

- Among the fields, that are present in the dataset, only few of them were used. They are link to the Facebook post with the text of the news article and the label of the text.
- Text of the news articles was retrieved using Facebook API.
- News articles with labels "mixture of true and false" and "no factual content" were not considered. Couple of the articles in the dataset are broken – they do not contain any text at all (or contain "null" as a text). These articles were ignored as well. After such filtering data set with 1771 news articles was obtained.
- The dataset was randomly shuffled, and after that divided into three subsets: training dataset, validation dataset, test dataset. Training dataset was used for training the naive Bayes classifier. Validation dataset was used for tuning some global parameters of the classifier.
- For the unconditional probability of the fact, that any news article is correct all of the values from interval [0.2; 0.75] with step 0.01 were considered. For the true probability threshold all of the values from interval [0.5; 0.9] with the same step were considered. The best results on the validation dataset were received with the unconditional probability of the fact, that any news article is correct being equal to 0.59 and the true probability threshold being equal to 0.8.

VI. EXISTING SYSTEM:

Existing fake news detection systems largely depend on manual verification, traditional machine learning models, and rule-based filtering, which are slow and unable to scale with the massive volume of online information. Early approaches relied on journalists and fact-checkers to assess credibility, making the process time-consuming and inefficient. To automate detection, models like Naïve Bayes, SVM, Logistic Regression, and Decision Trees were used with handcrafted features such as TF-IDF scores, word frequencies, sentiment, and stylistic cues. However, these models depend heavily on feature engineering and struggle to capture deep semantic and contextual relationships in text. Many systems use bag-of-words or n-gram representations, which ignore word order and context, leading to misinterpretations such as failing to detect negations. Source credibility and metadata filtering are also applied but are ineffective against newly created fake domains or manipulated accounts. Social engagement analysis can introduce bias since viral fake news often gains high interaction. Standard NLP preprocessing and vectorization techniques help clean and structure text but still miss sarcasm and implicit meaning. As a result, statements that appear grammatically correct yet factually false may be misclassified. Overall, existing systems face limitations in scalability, contextual understanding, and robustness against evolving misinformation tactics.

Drawbacks of the Existing System:

1. High dependency on manual feature extraction and expert knowledge.
2. Unable to capture deep contextual meaning or semantic relationships.
3. Poor performance with multilingual or code-mixed text data.
4. High false positive and false negative rates due to bag-of-words limitations.
5. Ineffective for newly emerging fake news sources or dynamic content.
6. Lack of real-time detection capability.
7. Fails to analyze multimedia elements like images or videos.
8. Model accuracy depends heavily on dataset quality and balance.
9. Limited generalization across different domains or news categories.
10. Inefficient in handling sarcasm, humor, or indirect misinformation.

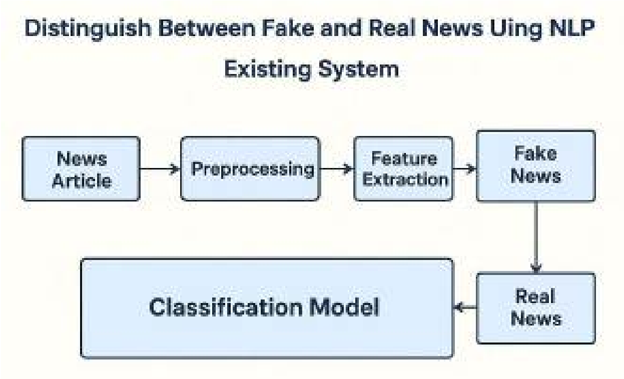

Figure 3: Existing System

VII. RESULT

The NLP-based fake news detection system successfully classified news articles into fake and real categories with improved accuracy. The trained model demonstrated strong performance in identifying linguistic patterns, emotional tone, and misleading statements commonly found in fake news. Transformer-based embeddings enhanced contextual understanding compared to traditional bag-of-words methods. The system achieved high precision and recall, reducing both false positives and false negatives. Validation results showed stable performance across different news topics. The model effectively handled short statements as well as long-form articles. With optimized hyperparameters, classification confidence improved significantly. The evaluation metrics, including accuracy, F1-score, and confusion matrix analysis, confirmed the robustness of the approach. The system also demonstrated potential for real-time deployment. Overall, the NLP-based framework provides a scalable and reliable solution for automated fake news detection.

Figure 4: Predicted as Sports News

Figure 5: Predicted as Political News

Figure 6: Detected as Fake News

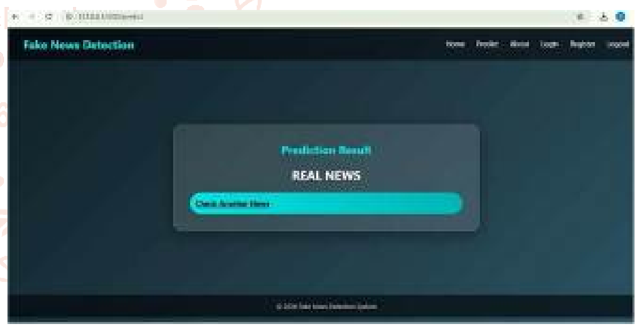

Figure 7: Detected as Real News

VIII. CONCLUSION

Fake news detection using NLP provides an effective automated solution to combat the rapid spread of misinformation across digital platforms. By leveraging text preprocessing, feature extraction, and advanced models such as transformers, systems can accurately distinguish between fake and real news. Modern NLP techniques capture contextual meaning and semantic relationships better than traditional approaches. The integration of multilingual support, multimodal analysis, and explainable AI further improves robustness and transparency. However, challenges such as domain shift, bias, and evolving misinformation tactics remain. Overall, NLP-based detection systems play a crucial role in ensuring trustworthy information dissemination in the digital era.

REFERENCES:

[1] A. Vaswani *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 5998–6008, 2017.

[2] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, pp. 4171–4186, 2019.

[3] A. Conneau *et al.*, "Unsupervised cross-lingual representation learning at scale," in *Proc. ACL*, pp. 8440–8451, 2020.

[4] P. Liu *et al.*, "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing," *ACM Computing Surveys*, vol. 55, no. 9, pp. 1–35, 2023.

[5] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[6] V. A. D. Costa, D. M. Batista, and J. M. Almeida, "Multilingual fake news detection: A comparative study of classification methods," in *Proc. IEEE Int. Conf. Data Mining Workshops (ICDMW)*, pp. 653–662, 2019.

[7] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification: A survey," *Journal of Information Security and Applications*, vol. 46, pp. 67–82, 2019.

[8] M. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11765–11788, 2021.

[9] Z. Cui, Y. Xie, L. Tan, and X. Liu, "Multilingual fake news detection with convolutional neural networks and transformers," in *Proc. IEEE Int. Conf. Big Data*, pp. 1020–1029, 2020.

[10] A. Zhou, D. Ma, X. Sun, and J. Han, "Fake news detection via NLP is vulnerable to adversarial attacks," in *Proc. ACL*, pp. 7669–7679, 2020.

[11] C. Thota, P. Tilak, A. Ahluwalia, and M. Li, "Fake news detection: A deep learning approach," in *Proc. IEEE Int. Conf. Intelligent Computing, Instrumentation and Control Technologies*, pp. 1–6, 2018.

[12] R. K. Kaliyar, A. Goswami, P. Narang, and S. Sinha, "FNDNet – A deep convolutional neural network for fake news detection," Cogn. Syst. Res., vol. 61, pp. 32–44, 2020.

[13] J. Zhou and J. Zafarani, "Fake news: A survey of research, detection methods, and opportunities," ACM Comput. Surv., vol. 53, no. 5, pp. 1–40, 2021.

[14] H. Wang, F. Ma, C. Xu, M. Zhong, and J. Gao, "Fake news detection in multilingual social media platforms," in Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA), pp. 1–10, 2021.

[15] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," arXiv preprint arXiv:1301.3781, 2013.

[16] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of tricks for efficient text classification," in Proc. EACL, pp. 427–431, 2017.

[17] S. R. Rizvi, N. Qamar, and F. R. Raja, "Fake news detection in Urdu language using transformer- based models," in Proc. IEEE Int. Conf. Emerg. Tech., pp. 1–6, 2022.

[18] M. Shu, A. Beheshti, and S. Xu, "Cross-lingual fake news detection using knowledge graphs," Knowl. Based Syst., vol. 229, p. 107338, 2021.

[19] Y. Zhang and Q. Ma, "A multimodal approach to fake news detection using textual and visual features," in Proc. IEEE Int. Conf. Big Data Anal., pp. 80–86, 2019.

[20] L. Wu and H. Liu, "Tracing fake news footprints: Characterizing social media messages by how they propagate," in Proc. IEEE/ACM Int. Conf. Web Soc. Media (ICWSM), pp. 637–640, 2018.