



Veritas Net: Hybrid Fake News Detection and Fact Verification System

Mr. S. Kumarakrishnan¹, M Fazil Ahamed², Reeman Infant D³, Karthikeyan C⁴, Sivanesan R⁵

¹Associate Professor, Dept. Of CSE, Sri Manakula Vinayagar Engineering College, Puducherry, India

^{2,3,4} UG Student, Dept. Of Cse, Sri Manakula Vinayagar Engineering College, Puducherry, India:

Email ID: skumarakrishnan@smvec.ac.in¹, fazilahamed2003@gmail.com², reeinf777@gmail.com³, ckarthikeyan024@gmail.com⁴, rsivanesan555@gmail.com⁵

Abstract

The rapid growth of digital media and social networking platforms has significantly accelerated the spread of fake news, leading to misinformation that affects public opinion, social harmony, and decision-making processes. Traditional fake news detection approaches primarily rely on binary classification and handcrafted features, which are often insufficient to handle semantic complexity, partial truths, and evolving misinformation patterns. To address these challenges, this project proposes an intelligent fake news detection platform based on VeritasNet, integrated with Automated Machine Learning (AutoML) and fact verification mechanisms. The proposed system employs advanced Natural Language Processing (NLP) techniques for semantic understanding of news content. Transformer-based deep learning models such as BERT are utilized to extract rich contextual representations from text, which are further processed by VeritasNet to perform deep pattern learning and uncertainty-aware classification. AutoML is incorporated to automatically select optimal model architectures and tune hyperparameters, ensuring robust performance while minimizing manual intervention. In addition to content-based analysis, the system includes a fact verification module that extracts key claims from news articles and validates them against trusted fact-checking sources and knowledge bases. To enhance transparency and user trust, Explainable Artificial Intelligence (XAI) techniques such as SHAP, LIME, and attention visualization are employed to interpret model predictions. Experimental results demonstrate that the proposed approach achieves improved accuracy, balanced precision–recall performance, and reliable confidence estimation compared to conventional models.

Keywords: Fake News Detection, VeritasNet, AutoML, Natural Language Processing, BERT, Fact Verification, Explainable AI

1. Introduction

The exponential growth of digital media platforms and social networking sites has transformed the way information is created, shared, and consumed. While this rapid dissemination of information offers numerous benefits, it has also led to the widespread circulation of fake news and misinformation. Fake news refers to deliberately fabricated or misleading information presented as legitimate news with the intent to deceive readers. The uncontrolled spread of such content can influence public opinion, disrupt social harmony, and impact critical domains such as politics, healthcare, and finance. Traditional approaches to fake news detection mainly rely on manual fact-checking or rule-based systems, which are time-consuming and incapable of handling the massive volume of online content generated daily.

Early machine learning techniques attempted to address this issue using handcrafted linguistic and statistical features; however, these methods often fail to capture deep semantic relationships and contextual nuances present in natural language. Moreover, most existing systems perform binary classification, labelling news as either real or fake, which does not adequately reflect partial truths or ambiguous claims commonly found in real-world news articles. Recent advancements in deep learning and Natural Language Processing (NLP), particularly transformer-based models such as BERT, have significantly improved text understanding by learning contextual representations. However, deep learning models often suffer from limited interpretability and require extensive manual tuning of architectures and



hyperparameters. To overcome these challenges, this project introduces VeritasNet, a truth-aware deep learning framework integrated with Automated Machine Learning (AutoML) for optimal model selection and performance enhancement.

2. Related works

2.1. An Enhanced Fake News Detection System with Fuzzy Deep Learning (Xu & Kechadi, 2024)

This paper introduces a system that integrates fuzzy logic with deep learning to handle the inherent uncertainty and ambiguity in fake news detection. It uses deep neural networks for classification while fuzzy logic manages uncertain features or decision thresholds. Drawbacks: While it addresses uncertainty, deep learning models often suffer from limited interpretability and require extensive manual tuning of architectures and hyperparameters[1].

2.2. A Comprehensive Survey on Fake News Detection: Methods and Challenges (Shah & Petal, 2025)

This survey presents a detailed overview of recent advancements in fake news detection, with particular emphasis on multimodal fusion techniques. Drawbacks: The focus is primarily on visual forgeries, which may not directly address text-based misinformation patterns that rely on semantic nuances rather than image manipulation[3].

2.3. Fake News Detection Using Deep Learning: A Systematic Literature Review (Sharma, Thakur, & Bhatt, 2021)

This systematic review surveys research from 2015–2020, focusing on the adaptation of CNN, RNN, LSTM, attention mechanisms, and Transformer models for misinformation detection. It organizes studies by input modality and data preprocessing strategies. Drawbacks: The review highlights persistent challenges such as the scarcity of high-quality annotated data, sample imbalance, vulnerability to adversarial manipulation, and the difficulty of interpreting large model predictions[4].

3. Proposed Solution

The proposed system presents an intelligent and trustworthy fake news detection platform that integrates VeritasNet, Automated Machine Learning (AutoML), fact verification, and Explainable

Artificial Intelligence (XAI) to overcome the limitations of existing approaches. The solution is designed to accurately classify news articles while handling semantic complexity, partial truths, and uncertainty in real-world information. The system begins with the collection of news content either through direct user input or by extracting text from a provided news URL. The collected data undergoes comprehensive preprocessing, including noise removal, tokenization, lemmatization, stop-word elimination, and sequence padding, to improve model efficiency and accuracy. Transformer-based NLP models such as BERT are then employed to generate contextual embeddings that capture semantic meaning and relationships within the news text. To enhance model performance and reduce manual intervention, an AutoML framework is incorporated. AutoML automatically evaluates multiple model architectures and optimizes hyperparameters such as learning rate, batch size, number of layers.

4. System Architecture and Design

The architecture of the proposed fake news detection system is designed as a modular and scalable framework that integrates deep learning, AutoML, fact verification, and Explainable AI (XAI). The system follows a layered architecture to ensure flexibility, maintainability, and efficient data flow between components[2].

4.1. Overall Architecture

The system architecture consists of multiple interconnected modules, starting from user input and ending with result visualization. Users can provide and optimizers. Based on evaluation metrics, the best-performing configuration is selected for deployment, ensuring robustness and adaptability to diverse datasets[5]. At the core of the system lies VeritasNet, a truth-aware deep learning architecture that performs deep pattern learning using LSTM-based layers and incorporates a reasoning mechanism to handle ambiguity and partial truth. Unlike traditional binary classifiers, VeritasNet categorizes news into multiple credibility levels such as Real, Mostly True, Half True, Mostly Fake, and Fake, providing a more realistic assessment of information reliability. Additionally, the system integrates a fact verification module that extracts key claims from news articles

and verifies them against trusted fact-checking sources and knowledge bases. To promote transparency and user trust, XAI techniques including SHAP, LIME, and attention visualization are applied to explain the model's predictions. The final output includes the predicted class, confidence score, and an interpretable explanation, making the system suitable for real-world misinformation detection and decision support[6].

4.2. Feature Extraction and Model Selection

The pre-processed text is passed to the Transformer-Based Feature Extraction Module, where models such as BERT generate contextual embeddings representing semantic relationships within the news content. These embeddings are then forwarded to the AutoML Engine, which evaluates multiple deep learning architectures and automatically tunes hyperparameters shown in Figure 1.1.

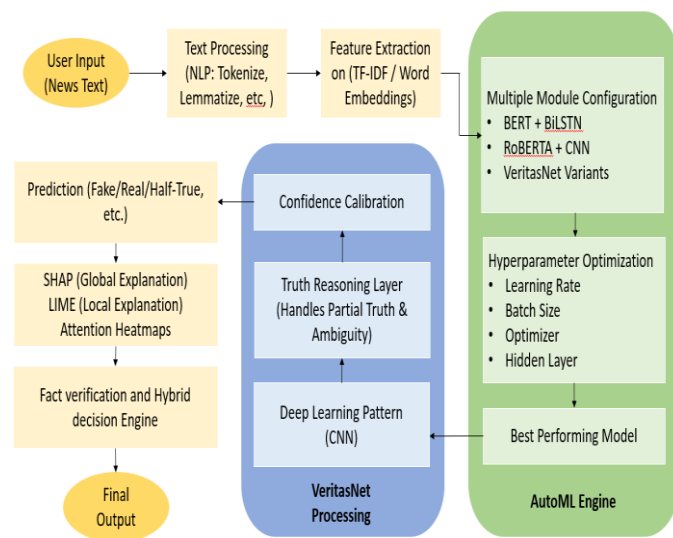


Figure 1.1 Architecture Diagram for Proposed Solution

5. Proposed Algorithm

Algorithm: VeritasNet with AutoML and Fact Verification.

Input: News Text / News URL

Output: Credibility Label, Confidence Score, Explanation

BEGIN

- Read input news_text or news_url
- IF input is URL THEN

- Extract news_text from URL
- END IF
- Preprocess news_text
 - Remove special characters and punctuation
 - Convert text to lowercase
 - Remove stop-words
 - Perform tokenization
 - Apply lemmatization
 - Pad sequences
- Generate contextual embeddings
- embeddings ← BERT(news_text)
- Initialize AutoML engine
- AutoML process:
 - FOR each model configuration in search space DO
 - Train model using embeddings
 - Tune hyperparameters
 - Evaluate model using validation metrics
 - END FOR
 - Select best_model based on F1-score
 - Pass embeddings to VeritasNet
 - Perform deep pattern learning using LSTM
 - Apply truth reasoning layer
 - Compute confidence score
 - Extract key factual claims from news_text
 - Perform fact verification
 - FOR each extracted claim DO
 - Compare claim with trusted fact sources
 - Compute fact_match_score
 - END FOR
 - Generate final prediction
 - credibility_label ← VeritasNet output
 - confidence ← calibrated confidence score
 - Apply Explainable AI
 - Generate SHAP explanations
 - Generate LIME explanations
 - Produce attention heatmap
 - Display results
 - Credibility label
 - Confidence score
 - Fact verification result
 - Explanation highlights
 - END

6. Result

The proposed fake news detection system based on VeritasNet, AutoML, and fact verification was evaluated using a benchmark dataset to assess its effectiveness and reliability. Experimental results indicate that the system achieves high classification accuracy, demonstrating its ability to correctly distinguish between real and fake news articles. The precision values obtained are consistently high, showing that the model produces fewer false positives and avoids incorrectly labeling genuine news as fake. Recall results reveal that the system successfully identifies a large proportion of fake news instances, which is crucial for minimizing the spread of misinformation. The F1-score, which balances precision and recall, shows a noticeable improvement over traditional machine learning models and standalone deep learning approaches, particularly on imbalanced datasets. This confirms the robustness of the proposed approach. The integration of AutoML played a significant role in enhancing performance by automatically selecting the optimal model architecture and tuning hyperparameters. This resulted in stable and consistent performance across multiple experimental runs. Additionally, the inclusion of a fact verification module improved the reliability of predictions, especially for news articles containing partially true or ambiguous claims[8].

7. Evaluation Metrics

The performance of the proposed fake news detection system was evaluated using standard classification and reliability metrics. These metrics ensure a comprehensive, assessment of accuracy, robustness, and trustworthiness. Accuracy measures the overall correctness of the model by calculating the proportion of correctly classified instances among all predictions. Precision evaluates the reliability of fake news predictions by measuring the proportion of correctly identified fake news articles among all articles classified as fake. Recall (Sensitivity) measures the ability of the system to detect actual fake news articles, ensuring minimal misinformation escapes detection[7].

8. Existing Solution vs Proposed Solution

8.1. Existing System

Most existing fake news detection systems primarily rely on traditional machine learning or basic deep learning techniques. These systems typically perform binary classification, categorizing news articles only as real or fake. They depend heavily on handcrafted linguistic features such as word frequency, sentiment polarity, and n-grams, which are insufficient for capturing complex semantic relationships in natural language. However, they often act as black-box models, offering little to no explainability. Furthermore, most existing approaches do not incorporate fact verification, relying solely on textual patterns, which can lead to incorrect predictions for partially true or misleading content. The absence of confidence estimation and automated optimization further limits their reliability and scalability in real-world applications shown in Figure 1.2.

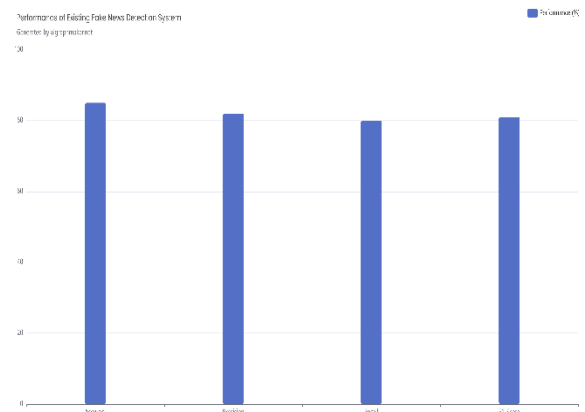


Figure 1.2 Accuracy Percentage of Existing Solution

8.2. Proposed System

The proposed system introduces VeritasNet, an advanced truth-aware fake news detection framework integrated with Automated Machine Learning (AutoML), fact verification, and Explainable AI (XAI). Unlike existing binary classifiers, the proposed solution performs multi-class credibility classification shown in Table 1.1, handling partial truths and uncertainty more effectively. Transformer-based models such as BERT are used for deep semantic feature extraction, followed by reasoning mechanism to analyze contextual patterns and truthfulness. AutoML automatically selects optimal model architectures and tunes hyperparameters, eliminating manual effort and improving

performance consistency shown in Figure 1.3.

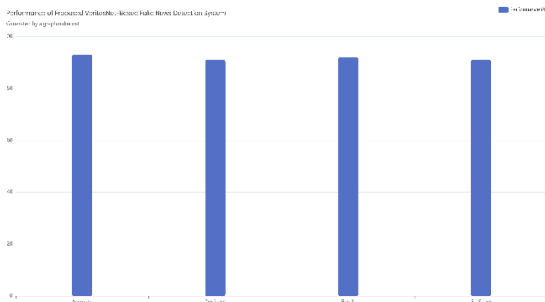


Figure 1.3 Accuracy Percentage of Proposed Solution

Table 1 Difference on Existing and Proposed Solution

Aspect	Existing System	Proposed System
Technology Used	Manual or outdated technology	Modern automated technology
Processing Speed	Slow response time	High-speed processing
Accuracy	High chance of human errors	Very accurate with minimal errors
Scalability	Difficult to expand	Easy to scale and upgrade
User Experience	User-unfriendly interface	User-friendly interface
Reliability	Less reliable	Highly reliable
Decision Making	Manual decision making	Data-driven decision making

Conclusion

This project presented an intelligent and trustworthy fake news detection system based on VeritasNet, integrated with Automated Machine Learning (AutoML), fact verification, and Explainable Artificial Intelligence (XAI). The proposed approach effectively addresses the limitations of traditional fake news detection methods by combining deep semantic understanding with automated optimization and factual validation. By leveraging transformer-based NLP models such as BERT, the system successfully captures contextual and semantic relationships within news content. The integration of AutoML enables automatic selection of optimal

model architectures and hyperparameter tuning, resulting in consistent and improved performance across different datasets. VeritasNet's truth-aware reasoning mechanism allows the system to handle uncertainty and partial truths, providing multi-class credibility classification rather.

Reference

- [1]. Xu, C., & Kechadi, M. (2024). An Enhanced Fake News Detection System With Fuzzy Deep Learning. <https://doi.org/10.1109/ACCESS.2024.3418340>
- [2]. Sharma, N., Thkur, R., & Bhatt, H. (2021). Fake News Detection Using Deep Learning: A Systematic Literature Review. <https://doi.org/10.1109/ICCMC51019.2021.9418349>
- [3]. Zheng, Y., Qian, Y., Huang, L., & Tian, Q. (2022). Exposing Deepfake Face Forgeries with Guided Residuals. <https://doi.org/10.1109/TMM.2022.3164685>
- [4]. Shah, S., & Petal, S. (2025). A Comprehensive Survey on Fake News Detection: Methods and Challenges. <https://doi.org/10.1109/ACCESS.2025.3372469>
- [5]. Rai, N., Kumar, D., Kaushik, N., Raj, C., & Ali, A. (2025). Fake News Detection Using Transformer-Based Deep Learning Models. <https://doi.org/10.1109/ICCCS.2025.11457723>
- [6]. Vardanyan, E. (2025). Detecting Fake News in Financial Domain Using Hybrid Neural Networks. <https://doi.org/10.1007/s00500-024-08967-4>
- [7]. Giachanou, A., & Zhang, G. (2024). A Survey on Multimodal Fake News Detection Using Visual and Textual Features. <https://doi.org/10.1109/ACCESS.2024.3368891>
- [8]. Xu, X., Li, X., Wang, T., & Jiang, Y. (2024). Emotion-Aware Fake News Detection Using LSTM. <https://doi.org/10.1109/ACCESS.2024.3372254>