

A Credibility Scoring Model for News Authenticity using SBERT/Logistic Regression

Dr. C. P. Divate¹, Mr. S. M. Patil², Shreyash Potdar³, Kedar Potdar⁴,
Tanmay Lad⁵, Nilesh Lokhande⁶, Raj Patil⁷

¹Dean, Dept of Computer Engineering, Shri Ambabai Talim Sanstha's Sanjay Bhokare Group of Institute
Miraj(poly), Maharashtra, India

²Lecturer, Dept of Computer Engineering, Shri Ambabai Talim Sanstha's Sanjay Bhokare Group of Institute
Miraj(poly), Maharashtra, India

^{3,4,5,6,7}Student Dept of Computer Engineering, Shri Ambabai Talim Sanstha's Sanjay Bhokare Group of Institute
Miraj(poly), Maharashtra, India

Abstract - The rapid spread of fake news on digital platforms has made it difficult to trust online information. Manual verification of news is slow and unreliable. This project presents **A Credibility Scoring Model for News Authenticity using SBERT and Logistic Regression** to automatically detect fake news. SBERT is used to capture the semantic meaning of news text, while Logistic Regression classifies the content as real or fake. The system also provides a credibility score showing prediction confidence. The model delivers accurate results and is integrated into an Android application for easy and real-time news verification.

Key Words: Fake News Detection, News Authenticity, Credibility Scoring, Sentence-BERT (SBERT), Logistic Regression, Natural Language Processing, Machine Learning, Text Classification

1. INTRODUCTION

Through websites, social media, and mobile applications. While this makes information easily accessible, it also increases the risk of fake and misleading news reaching a large audience. False information can create confusion, influence public opinion incorrectly, and sometimes lead to serious social and economic consequences. As a result, verifying the authenticity of news has become an important challenge.

Traditional methods of fact-checking rely heavily on human effort, which is slow and cannot handle the massive amount of online content generated every day. To overcome this problem, automated systems based on machine learning and natural language processing are being widely explored. These systems can analyze news content and determine its credibility more efficiently.

A Credibility Scoring Model for News Authenticity using SBERT and Logistic Regression, focuses on automatically identifying whether a news article is real or fake.

2. PROBLEM STATEMENT

The rapid expansion of digital news platforms and social media has made information easily accessible to users. However, this growth has also led to the widespread circulation of fake and misleading news. Such content can influence public opinion, create panic, and spread misinformation at a large scale. Manual verification of news authenticity is time-consuming and impractical due to the massive volume of online content.

Existing fake news detection systems often rely on keyword-based or shallow text analysis techniques. These approaches fail to understand the contextual meaning of news articles, resulting in inaccurate predictions. Additionally, many systems only classify news as real or fake without providing any confidence level, making it difficult for users to judge the reliability of the result. Therefore, there is a need for an intelligent, accurate, and user-friendly system that can evaluate news credibility effectively.

3. PROPOSED SOLUTION

To overcome the identified challenges, this project proposes a credibility scoring model for news authenticity using Sentence-BERT (SBERT) and Logistic Regression. SBERT is used to convert news text into meaningful semantic embeddings that capture contextual information. These embeddings are then classified using Logistic Regression to determine whether the news is real or fake.

In addition to binary classification, the system generates a credibility score based on prediction probability. This score helps users understand how confident the system is about the authenticity of the news. The trained model is integrated into an Android application, enabling real-time news verification through a simple and intuitive interface. The proposed solution is computationally efficient, scalable, and suitable for practical deployment, making it an effective tool for combating misinformation in digital media.

4. LITERATURE REVIEW

[1] **Shu et al. (2017)**: Studied the problem of fake news detection and highlighted how misinformation spreads rapidly through online platforms. Their work emphasized the need for automated systems to verify news credibility using machine learning techniques instead of manual fact-checking.

[2] **Castillo et al. (2011)**: Analyzed the credibility of online news content using traditional machine learning models such as Decision Trees and Support Vector Machines. Although these models showed reasonable performance, they relied mainly on surface-level features and lacked contextual understanding of text.

[3] **Devlin et al. (2018)**: Introduced BERT (Bidirectional Encoder Representations from Transformers), which significantly improved natural language understanding by capturing deep contextual relationships between words. This model demonstrated better performance compared to earlier word embedding techniques.

[4] **Reimers and Gurevych (2019)**: Proposed Sentence-BERT (SBERT), an extension of BERT designed to generate meaningful sentence embeddings efficiently. SBERT reduced computational complexity and proved effective for text classification and similarity tasks.

[5] **Hosmer et al. (2013)**: Discussed the effectiveness of Logistic Regression for binary classification problems, highlighting its simplicity, interpretability, and consistent performance. Due to these advantages, Logistic Regression remains widely used in real-world applications.

Research Objectives:

1. To study the impact of fake news on digital media platforms and understand the challenges in identifying news authenticity.
2. To design and develop a machine learning-based model for detecting fake and real news automatically.
3. To use Sentence-BERT (SBERT) for extracting meaningful semantic features from news text.
4. To implement Logistic Regression for efficient and accurate classification of news content.
5. To generate a credibility score that reflects the confidence level of the news authenticity prediction.
6. To evaluate the performance of the proposed model using standard accuracy and classification metrics.

5. METHODOLOGY

“A Credibility Scoring Model for News Authenticity using SBERT and Logistic Regression” is designed in six major modules. Each module focuses on a specific function, from data collection to final deployment, ensuring a structured and efficient approach to detecting fake news and generating credibility scores.

MODULE -1.Data Collection and Preprocessing:
The first module is critical as the quality of input data directly affects the performance of the entire system.

Core Functions:

- Collect a comprehensive dataset of real and fake news articles from multiple sources including news websites, and verified social media feeds.
- Ensure the dataset represents various categories and types of news, such as politics, sports, health, and technology, to make the model robust.

Key Functionalities:

- **Data Cleaning:** Removal of unwanted characters, symbols, HTML tags, and punctuation that may interfere with analysis.
- **Normalization:** Convert all text to lowercase to maintain uniformity across the dataset.
- **Noise Removal:** Remove stopwords, numbers, redundant spaces, and irrelevant words to retain only meaningful content.
- **Tokenization:** Break sentences into individual words or tokens to prepare for embedding generation.

Major Components:

- Raw data files (CSV, JSON, or TXT formats) containing labeled real and fake news articles.
- Preprocessing scripts developed in Python using libraries like NLTK, SpaCy, or Pandas.
- A clean, structured dataset ready for feature extraction, ensuring higher model accuracy.

MODULE -2.Feature Extraction using Sentence-BERT (SBERT):

This module converts the preprocessed text into numerical vectors that a machine learning model can interpret.

Core Functions:

- Transform textual news data into meaningful numerical representations that retain the semantic meaning.

- Capture contextual information from each sentence, allowing the model to understand nuances that indicate fake news.

Key Functionalities:

- **Sentence Embedding Generation:** SBERT converts each news article or sentence into a fixed-length dense vector.
- **Semantic Context Understanding:** Embeddings capture the relationship between words and the overall meaning of sentences.
- **Dimensionality Reduction:** SBERT embeddings provide compact and efficient representation of text while retaining important semantic information.
- **Similarity Measurement:** Enables comparison between news articles to detect patterns indicative of fake content.

Major Components:

- Pretrained SBERT model loaded through Python libraries like sentence-transformers.
- Embedding generation scripts for converting each article into a vector.
- Feature matrix storing embeddings for all news articles, used as input for the classifier.

MODULE -3. Model Training and Optimization using Logistic Regression:

This module focuses on training the machine learning classifier to distinguish real and fake news.

Core Functions:

- Learn patterns and relationships from the SBERT embeddings to classify news accurately.
- Optimize the classifier to maximize accuracy and reliability while keeping it interpretable

Key Functionalities:

- **Training the Model:** Logistic Regression is trained on the feature vectors derived from SBERT embeddings.
- **Hyperparameter Tuning:** Adjust parameters such as regularization strength and solver type to improve model performance.
- **Validation:** Test the model on unseen data to evaluate generalization and avoid overfitting.
- **Prediction Generation:** Provide binary outputs, labeling each news article as real or fake.

Major Components:

- Logistic Regression Classifier implemented using Python's scikit-learn library.
- Training and testing scripts for processing feature matrices and evaluating results.
- Saved trained model for deployment in real-time applications.

MODULE -4. Credibility Scoring and Prediction:

This module adds transparency to the system by providing a confidence measure along with classification results.

Core Functions:

- Generate a credibility score to indicate the likelihood that the news is real.
- Provide users with insights into the confidence of each prediction, enhancing trust in the system.

Key Functionalities:

- **Probability Calculation:** Use the output probability from Logistic Regression to derive the credibility score.
- **Interpretation of Scores:** Higher values indicate higher confidence in the authenticity of the news, whereas lower scores indicate potential misinformation.
- **Integration with Prediction:** Both classification and credibility scores are returned together to the user.
- **Threshold Management:** Set thresholds for real vs. fake classification based on probability scores to reduce misclassification.

Major Components:

- Prediction module that combines Logistic Regression output with score calculation logic.
- Scripts for mapping probability outputs to a human-readable credibility score (e.g., 0-100%).
- Data structures to store predicted labels and associated scores.

MODULE -5. System Deployment and Android Application Integration:

This module ensures that the trained model is accessible to users in real-time through a practical interface.

Core Functions:

- Deploy the trained model using a REST API, allowing external applications to access prediction and credibility scoring.
- Integrate the model into an Android application for easy user interaction.

Key Functionalities:

- API Development: Develop endpoints that accept news text as input and return prediction results and credibility scores.
- Android App Interface: Provide a simple interface for users to enter news content and view results.
- Real-Time Prediction: Ensure low-latency responses so users get immediate feedback.
- Error Handling: Manage invalid inputs and system errors gracefully.

Major Components:

- REST API built using Python frameworks like FastAPI or Flask.
- Android application coded in Java/Kotlin with frontend interface for user input.
- Integration scripts to connect the app with backend API endpoints.

MODULE -6. Testing, Evaluation, and Maintenance:

The final module ensures that the system performs accurately, reliably, and consistently over time.

Core Functions:

- Test the system on unseen data to validate performance.
- Monitor accuracy and reliability over time and adapt the system for new news trends.

Key Functionalities:

- Accuracy Evaluation: Measure performance using metrics like accuracy, precision, recall, and F1-score.
- Usability Testing: Ensure the Android app is intuitive and provides clear results.
- Maintenance: Regularly update the model and dataset to handle evolving fake news patterns.
- Error Analysis: Identify cases where the model misclassifies and refine preprocessing, embeddings, or classifier parameters.

Major Components:

- Evaluation scripts for performance metrics.
- User feedback loop integrated into the app for reporting errors.
- Maintenance plan for updating datasets and retraining the model periodically.

6. RESULTS AND DISCUSSION

The proposed credibility scoring model was evaluated using a labeled dataset of real and fake news articles. The dataset

was divided into training and testing sets to measure the model's performance on unseen data. The results demonstrate that the integration of Sentence-BERT (SBERT) with Logistic Regression provides reliable and meaningful news authenticity classification.

Key Results:

- The model achieved good classification accuracy, showing its ability to effectively distinguish between real and fake news articles.
- Precision and recall values indicate that the system correctly identifies fake news while minimizing false predictions.
- The use of SBERT significantly improved contextual understanding compared to traditional keyword-based approaches.
- Logistic Regression ensured fast and stable predictions, making the system suitable for real-time usage.

Credibility Score Analysis:

- Along with classification, the system generates a credibility score based on prediction probability.
- News articles classified as real generally produced higher credibility scores, indicating strong confidence.
- Fake news articles resulted in lower credibility scores, helping users identify unreliable content.
- This scoring mechanism improves transparency and enhances user trust in the system.

Discussion:

- SBERT enables deep semantic understanding of news content, allowing the model to detect subtle differences in meaning.
- Logistic Regression provides an interpretable and efficient classification framework.
- The combination of these techniques offers a balanced solution between performance and computational efficiency.
- Integration with an Android application demonstrates the practical applicability of the system in real-world scenarios.

7. CONCLUSION

This project presented a credibility scoring model for news authenticity using Sentence-BERT (SBERT) and Logistic Regression. The proposed system effectively addresses the challenge of fake news detection by understanding the semantic meaning of news content rather than relying only on keyword-based analysis. SBERT enables accurate feature extraction by capturing contextual information, while

Logistic Regression provides efficient and interpretable classification results.

The system not only classifies news as real or fake but also generates a credibility score that indicates the confidence level of the prediction. This improves transparency and helps users make informed decisions while consuming online news. The integration of the trained model with an Android application makes the solution practical and accessible for real-time use.

8. REFERENCES

[1] S. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," *IEEE Intelligent Systems*, vol. 32, no. 1, pp. 22–36, 2017.

[2] C. Castillo, M. Mendoza, and B. Poblete, "Information Credibility on Twitter," in *Proceedings of the 20th International World Wide Web Conference (WWW)*, 2011, pp. 675–684.

[3] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of NAACL-HLT*, 2019, pp. 4171–4186.

[4] N. Reimers and I. Gurevich, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2019.

[5] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed., Wiley, New York, 2013.

[6] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2013.

[7] A. Thakur, "Fake News Detection using Machine Learning," *International Journal of Engineering Research & Technology (IJERT)*, vol. 9, no. 6, pp. 45–49, 2020.