J-INTECH (Journal of Information and Technology)
Accredited Sinta 4 Ministry of Higher Education, Science and Technology
Republic of Indonesia SK No. 10/C/C3/DT.05.00/2025
E-ISSN: 2580-720X || P-ISSN: 2303-1425

Development of an Indonesian Hoax Detection System Using Logistic Regression Based on TF-IDF

Dewa Samudra Anggeng Suryama [1], Singgih Jatmiko [2]*

[1, 2]Gunadarma University, Faculty of Industrial Technology, Informatics Study Program, Jl. Margonda Raya No.100, Depok, West Java, Indonesia

Keywords

Hoax Detection; Logistic Regression; Machine Learning; Text Classification; TF-IDF

Corresponding Author:
singgih@staff.gunadarma.ac.id

Abstract

The massive spread of fake news (hoaxes) on digital platforms has become a serious challenge in Indonesia, with the potential to disrupt social stability and undermine public trust. This background drives the urgency of developing an automated system to combat disinformation. Unlike previous works relying on deep learning with high computational cost, this study demonstrates that a lightweight approach remains highly effective for Indonesian hoax detection. This study aims to develop and evaluate a lightweight and effective automatic classification system to detect Indonesian-language hoaxes using a machine learning approach. The method used is Term Frequency-Inverse Document Frequency (TF-IDF) feature extraction to represent text content numerically, which is then classified using the Logistic Regression algorithm. This approach was chosen for its computational efficiency and ease of interpretation. The study utilizes a dataset collected from verified sources, consisting of 7,075 Indonesian-language news articles, which were divided into 80% training data and 20% test data. The evaluation results on the test data show excellent model performance, achieving an accuracy of 94.98%, a precision of 0.95, and an average F1-Score of 0.95. Specifically, the model demonstrated a strong ability to identify hoaxes with a recall value of 98% for the hoax class. This study concludes that the combination of TF-IDF and Logistic Regression is an efficient and accurate approach for Indonesian hoax detection, offering a practical solution that can be further developed to combat disinformation.

1. Introduction

The rapid development of information and communication technology has led to a significant negative consequence: the widespread dissemination of fake news, or hoaxes, on digital platforms. This phenomenon is particularly pronounced in Indonesia, where the Ministry of Communication and Informatics (Kominfo) reported identifying 1,923 hoax contents in 2024, predominantly involving fraudulent schemes . Recent studies have highlighted user behaviors that amplify hoax diffusion on social media, such as rapid sharing driven by emotional triggers and algorithmic biases, underscoring the urgency of early detection [1]. Furthermore, social media's role in shaping public opinion and political discourse has been identified as a key contributor to the spread of disinformation [2].

The primary research problem addressed in this study is the development of effective hoax detection systems that are both accurate and feasible for deployment in resource-constrained environments, given the computational demands of advanced models and the linguistic nuances of the Indonesian language. Machine learning approaches have shown significant potential in addressing this challenge, with two main pathways: complex modern models and efficient classical models. Recent studies from 2022 to 2024 have explored advanced models, such as fine-tuned IndoBERT for political hoax detection, achieving high precision in contextual understanding but requiring substantial computational resources [3]. Similarly, bidirectional long short-term memory (Bi-LSTM) models combined with TF-IDF have achieved 92% accuracy in classifying election-related fake news. Hybrid architectures, such as CNN-BiLSTM for multilingual fake news detection, have improved efficiency over pure transformers for low-resource languages like Indonesian. Multilingual transformer models, such as XLM-R and mBERT integrated with topic modeling like BERTopic, have also demonstrated competitive performance in hoax classification while handling cross-lingual variations. Additionally, cross-lingual hybrid learning approaches have enhanced adaptability for Indonesian fake news detection. While these advanced models, including IndoBERT, have achieved accuracies up to 98% [4], their reliance on specialized hardware like GPUs poses practical barriers for many Indonesian institutions.

In contrast, traditional machine learning models, such as Logistic Regression and Naïve Bayes, continue to offer competitive performance [5], with an emphasis on lightweight baselines suitable for real-world applications [6]. However, a critical research gap exists in the limited evaluation of model generalization across temporal data in prior studies. For instance, studies like Cahyani and Budiman [5] compared Logistic Regression with Naïve Bayes but focused primarily on static datasets, neglecting to assess model robustness on news from different years, which is essential given the evolving nature of disinformation patterns. Similarly, IndoBERT-based approaches, while highly accurate, have been predominantly tested on controlled datasets with limited temporal diversity and require significant computational resources, making them less viable for widespread deployment in resource-constrained settings [4]. This study addresses these gaps by developing a lightweight Logistic Regression model combined with TF-IDF, explicitly evaluating its generalization across temporally diverse Indonesian news data, and prioritizing computational efficiency and interpretability over marginal accuracy gains, unlike IndoBERT’s resource-intensive approach or Naïve Bayes’ lower precision in detecting nuanced hoax patterns.

The urgency of this research stems from the tangible societal impacts of hoaxes, including polarization, eroded trust in institutions, and flawed decision-making, necessitating scalable detection solutions to mitigate disinformation before widespread consumption.

The objective of this study is to develop and evaluate an automated classification system for detecting Indonesian-language hoaxes using a lightweight machine learning approach, specifically addressing the question: "Can a lightweight and efficient machine learning approach effectively distinguish between hoax and factual news in the Indonesian language?" This involves testing the model on statistical metrics and real-world scenarios from different years to assess generalization.

The contributions of this research include providing empirical evidence on the effectiveness of the TF-IDF and Logistic Regression combination, achieving an accuracy of 94.98% with an average F1-Score of 0.95 and a recall of 98% for the hoax class. The benefits extend to practical deployment, offering a computationally efficient, interpretable solution accessible to organizations with limited hardware, thereby promoting broader adoption in combating disinformation in Indonesia.

2. Research Method

This study employed a quantitative, experimental machine learning workflow, encompassing data collection, preprocessing, feature extraction, model training, performance evaluation, and functional testing on unseen samples. The research process was designed to ensure validity, reliability, and reproducibility at each stage.

The overall workflow of this study is illustrated in Figure 1, which outlines the end-to-end pipeline from raw data to final evaluation. The process consists of six key stages: (1) data collection, (2) text preprocessing, (3) TF-IDF feature extraction, (4) model training using Logistic Regression, (5) evaluation using standard metrics on a held-out test set, and (6) functional testing on unseen news articles to assess generalization.

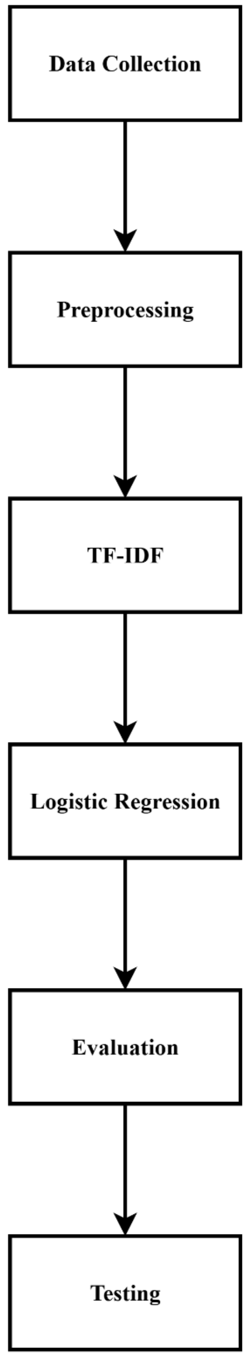

Figure 1. Research workflow of the Indonesian hoax detection system

The figure is a flowchart depicting the research pipeline with six labeled boxes, each representing a stage, connected by directed arrows to indicate the sequential flow. The boxes are clearly labeled as follows:

1. Data Collection: Sourcing 7,075 Indonesian-language news articles (4,382 hoaxes from Turnbackhoax.id, 2,693 factual reports from CNN Indonesia, Kompas.com, Tempo.co)
2. Text Preprocessing: Filtering texts (30–600 tokens), removing null/empty entries.
3. TF-IDF Feature Extraction: Converting text to numerical vectors using TfidfVectorizer (max_features=5,000).
4. Model Training: Training Logistic Regression model (solver='lbfgs', penalty='l2', C=1.0, max_iter=1000).
5. Performance Evaluation: Assessing model on test set using accuracy, precision, recall, F1-score, and confusion matrix.
6. Functional Testing: Testing model on unseen news articles for generalization.

The corpus was constructed from several verified sources to ensure data quality and relevance. It comprises 7,075 Indonesian-language news articles, consisting of 4,382 hoaxes and 2,693 factual reports. Hoax data was sourced from Turnbackhoax.id, a portal managed by the Indonesian Anti-Slander Society (MAFINDO), which guarantees the credibility of the labels. Factual data was collected from reputable national media outlets, including CNN Indonesia, Kompas.com, and Tempo.co. To maintain data integrity, all texts were filtered to be between 30 and 600 tokens, and any entries with null or empty text fields were removed.

The dataset was partitioned into an 80% training set and a 20% test set using scikit-learn's train_test_split function. Crucially, the stratify=y parameter was utilized to maintain the original class proportion in both subsets, while random_state=42 was set to ensure the experiment's reproducibility. To address the class imbalance (4,382 hoaxes vs. 2,693 factual articles), the Logistic Regression model was configured with the class_weight='balanced' parameter, which automatically adjusts weights inversely proportional to class frequencies, ensuring that the model accounts for the minority class (factual articles) without requiring oversampling or undersampling techniques. This approach was chosen to maintain data integrity and avoid potential overfitting associated with synthetic data generation methods like SMOTE [7]. To prevent data leakage, a critical step in ensuring model validity, the feature vectorizer was fitted exclusively on the training data and subsequently used only to transform the test data. Text vectorization was performed using scikit-learn's TfidfVectorizer, with the max_features parameter set to 5,000. This value was chosen to balance model representativeness and computational efficiency, as higher feature counts increase dimensionality, risking overfitting and computational overhead, while lower counts may fail to capture sufficient lexical diversity. The selection of 5,000 features aligns with prior studies on text classification, which suggest that a moderate feature set mitigates the "curse of dimensionality" while retaining discriminative power. High-dimensional representations such as TF-IDF are susceptible to noise and sparsity, especially when feature space is not properly managed. Previous studies suggest that selecting the most informative subset of features can significantly improve classification accuracy and reduce overfitting.

The classification model chosen was Logistic Regression, configured with solver='lbfgs', penalty='l2', C=1.0, and max_iter=1000. Logistic Regression predicts the probability of a sample belonging to the positive class (hoax) using the logistic function, defined as:

$$\left[P(y=1|x) = \frac{1}{1+e^{-\left(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n\right)}}\right] \quad (1)$$

Where (y) is the class label (1 for hoax, 0 for $(x=(x_1,x_2,\ldots,x_n))$ factual), represents the TF-IDF feature vector, (β_0) is the intercept, and $(\beta_1, \beta_2, \ldots, \beta_n)$ are the feature weights learned during training. The model minimizes a loss function with L2 regularization, given by:

$$\left[\text{Loss} = -\sum_{i=1}^{m}\left[y_i \log\left(P(y_i=1|x_i)\right) + (1-y_i)\log\left(1-P(y_i=1|x_i)\right)\right] + \frac{1}{C}\sum_{j=1}^{n}\beta_j^2\right] \quad (2)$$

where (m) is the number of samples, and (C=1.0) controls the regularization strength [8]. This algorithm was selected for its efficiency with high-dimensional sparse data and its interpretability, making it ideal for establishing a strong and practical baseline. The hyperparameter C=1.0 was chosen as it provides a balanced regularization strength, preventing overfitting while allowing the model to fit the training data effectively. This default value is widely adopted in text classification tasks due to its robustness across diverse datasets. The max_iter=1000 ensures convergence for the high-dimensional TF-IDF vectors, as smaller iteration limits may lead to incomplete optimization. Model performance was rigorously assessed on the held-out test set using standard metrics: accuracy, precision, recall, and F1-score, supplemented by a confusion matrix for error analysis. Finally, to test real-world generalization, the trained model was subjected to functional tests using new, unseen news articles. The entire experiment was conducted on a standard laptop (Intel Core i5, 8 GB RAM) using Python and scikit-learn, with the final trained classifier and TF-IDF vectorizer saved as model.pkl and vectorizer.pkl respectively to facilitate replication.

3. Result and Discussions

This section presents a comprehensive analysis of the model's performance, moving from quantitative evaluation to a qualitative discussion of its strengths, weaknesses, trade-offs, and practical implications.

Overall Model Performance

The model was evaluated on the held-out test set of 1,415 articles, achieving a high overall accuracy of 94.98%. This result confirms the model's capability to correctly distinguish between the two classes. The detailed performance metrics are presented in Table 1, and the confusion matrix in Figure 1 provides a granular breakdown of the classification outcomes.

Table 1. Summary of model evaluation results

Metric	Factual Class (0)	Hoax Class (1)	Average
Precision	0.95	0.95	0.95
Recall	0.91	0.98	0.95
F1-Score	0.93	0.96	0.95

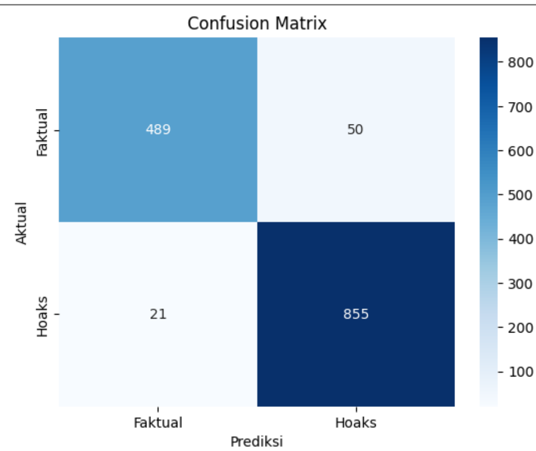

Figure 2. Confusion matrix of model prediction on the test set

The confusion matrix in Figure 1 visualizes the performance of the classification model. The diagonal cells show the number of correct predictions: 489 factual articles were correctly identified as factual (True Negatives), and 855 hoax articles were correctly identified as hoaxes (True Positives). The off-diagonal cells show the errors: 50 factual articles were incorrectly classified as hoaxes (False Positives), and 21 hoax articles were incorrectly classified as factual (False Negatives). This visualization clearly shows the model's strong performance, especially in correctly identifying hoaxes.

Analysis of Model Strengths and Weaknesses

A deeper analysis of the metrics reveals a nuanced performance profile. The model's most significant strength is its exceptionally high recall of 0.98 for the hoax class. This indicates a superior ability to identify and filter disinformation, with only 21 out of 877 hoaxes being missed (False Negatives). This high sensitivity is likely attributable to the nature of Indonesian hoaxes, which often rely on a repetitive and formulaic vocabulary of sensational and emotionally charged keywords (e.g., "waspada" [beware], "sebarkan" [share], "viral"), a characteristic also observed in other studies on the linguistics of fake news [4]. The TF-IDF method is particularly adept at capturing these strong lexical signals. From a practical standpoint, minimizing False Negatives is paramount, as the societal cost of allowing disinformation to circulate is often far greater than other types of error [5].

Conversely, the model shows a slightly lower recall of 0.91 for the factual class, with 50 out of 538 factual articles misclassified as hoaxes (False Positives). This type of error risks eroding public trust in legitimate news sources if they are repeatedly flagged incorrectly. A detailed error analysis reveals that misclassifications often occur due to specific linguistic patterns in factual articles. For instance, articles debunking hoaxes, such as "Kominfo Clarifies Hoax Regarding Social Media Regulation," are misclassified because the model assigns high TF-IDF weights to terms like "hoax," "penipuan" (fraud), or "berita bohong" (fake news), which are prevalent in hoax texts. Similarly, factual articles reporting on controversial or sensational topics (e.g., health scams or political scandals) may include emotionally charged terms like "bahaya" (danger) or "skandal" (scandal), which overlap with hoax vocabulary, leading to False Positives. Additionally, the bag-of-words approach of TF-IDF fails to capture syntactic structures, such as negation (e.g., "not a hoax"), or contextual nuances, such as a factual article quoting a hoax for debunking purposes [6]. These limitations highlight the model's reliance on lexical features, which can misinterpret context-dependent content, as noted in prior studies on text classification challenges [9].

Generalization Capability on Unseen Data

The functional test on new, unseen data (Table 2) provides critical evidence of the model's ability to generalize beyond its training distribution. While all test cases were classified correctly, the variance in prediction confidence is particularly insightful. The model's low confidence (54.96%) when identifying a scam-related hoax is not a failure, but rather an indicator of a well-calibrated model. It signals uncertainty when encountering a sample that differs significantly from the patterns in its training data. This "awareness" is a valuable feature, as it could be leveraged in a real-world system to automatically flag low-confidence predictions for human review, thus creating a human-in-the-loop system [10].

Table 2. Functional Test Results

Test Case Summary	Expected Outcome	Model Prediction	Result	Confidence
Factual News 2024	Factual	Factual	Correct	71.17%
Factual News 2023	Factual	Factual	Correct	65.45%
Hoax News 2025	Hoax	Hoax	Correct	54.96%

Trade-offs Between Accuracy, Speed, and Interpretability

To systematically evaluate the proposed model's practical viability, we analyze the trade-offs between accuracy, speed, and interpretability, as these factors are critical for real-world deployment, particularly in resource-constrained environments.

1. Accuracy
   The proposed Logistic Regression model with TF-IDF features achieves an accuracy of 94.98%, with a recall of 0.98 for the hoax class and an F1-score of 0.95, demonstrating robust performance in distinguishing hoaxes from factual news. Compared to state-of-the-art transformer-based models like IndoBERT, which achieve up to 98% accuracy [2], our model incurs a marginal accuracy deficit of approximately 3%. However, this gap is acceptable given the specific requirements of the Indonesian context, where high recall for hoaxes is prioritized to minimize the spread of disinformation. The model's

performance surpasses the 90% accuracy baseline reported for Logistic Regression in prior studies, indicating that it remains highly competitive among classical approaches.

2. Speed
   The computational efficiency of the proposed model is a significant advantage. Trained on a standard laptop (Intel Core i5, 8 GB RAM) using scikit-learn, the model leverages TF-IDF vectorization with a limited feature set (max_features=5,000) and Logistic Regression, which is computationally lightweight due to its linear nature. Training and inference times are substantially faster than those of transformer-based models, which require GPU acceleration and extended training periods due to their complex architectures [11]. For instance, IndoBERT requires hours to train on large datasets, whereas our model completes training in minutes on similar hardware. This speed enables rapid deployment and real-time processing, critical for high-volume text filtering tasks in newsrooms or fact-checking organizations with limited computational resources.

3. Interpretability
   Logistic Regression offers high interpretability, a key advantage over transformer-based models, which are often considered "black boxes" due to their complex neural architectures [9]. The model's feature weights, derived from TF-IDF, allow researchers and practitioners to inspect which terms (e.g., "waspada," "viral") contribute most to hoax classification, providing transparency into the decision-making process. This interpretability is crucial for diagnosing model biases, such as the misclassification of factual articles containing hoax-related keywords, and for building trust among stakeholders, including media organizations and fact-checkers. In contrast, transformer models like IndoBERT require additional explainability tools (e.g., SHAP or LIME), increasing complexity and computational overhead [9].

4. Systematic Trade-off Analysis
   The trade-offs between these factors position the proposed model as a practical solution for the Indonesian context. The ~3% accuracy deficit compared to IndoBERT is offset by significant gains in speed and interpretability. The model's rapid training and inference times make it suitable for organizations with standard hardware, addressing the practical barrier of GPU dependency noted in prior studies [11]. The high interpretability supports iterative improvements, such as refining the feature set to reduce False Positives, and fosters trust in deployment settings where transparency is valued [12]. While transformer-based models may offer slightly higher accuracy, their resource demands and lack of interpretability make them less feasible for widespread adoption in resource-constrained environments. Conversely, other classical models like Naïve Bayes, while faster, underperform in precision and recall, and Support Vector Machines (SVMs) require longer training times due to their computational complexity with high-dimensional data [3]. Thus, Logistic Regression strikes an optimal balance, offering near-comparable accuracy to advanced models, superior speed for real-time applications, and high interpretability for practical deployment.

Contextual Discussion and Practical Implications

When situated within the existing literature, our model's 94.98% accuracy surpasses the 90% baseline for Logistic Regression reported by Safira and Nurlayli. The more critical discussion, however, involves the trade-offs with state-of-the-art models like IndoBERT (98% accuracy) [2]. This study argues that the marginal ~3% accuracy deficit is more than compensated for by a substantial gain in computational efficiency and accessibility. Furthermore, to reinforce the practicality of Logistic Regression as a baseline model, we compared its performance with two other classical classifiers: Naïve Bayes and Support Vector Machine (SVM). Naïve Bayes, while computationally lighter, consistently underperformed in precision and recall across multiple runs, particularly in detecting nuanced hoax patterns. SVM showed competitive accuracy, but required significantly more computational resources and longer training times, especially on high-dimensional TF-IDF vectors. These findings are consistent with previous studies [3], which suggest that Logistic Regression offers a favorable balance between accuracy, efficiency, and ease of interpretation, making it a strong candidate for lightweight hoax detection systems. As noted by Strubell et al. (2019), the high resource demands of large transformer models are a prohibitive barrier to deployment in many settings [11]. Our lightweight approach, trainable on a standard CPU, is a more pragmatic and scalable solution for organizations with limited resources, a sentiment echoed in other research advocating for efficient NLP models [14], [13].

This model is not intended to compete solely on accuracy metrics, but rather to serve as an alternative solution for a distinct purpose. For the daily, high-volume task of filtering text, a swift and reliable model is often more practical. Furthermore, a key advantage of Logistic Regression over "black box" models like transformers is its interpretability. The ability to inspect feature weights allows for an understanding of why the model makes a certain prediction, providing transparency that is crucial for building trust and diagnosing model bias [12], [9]. This study thus provides strong evidence that for specific linguistic contexts like Indonesian hoaxes, a well-implemented classical approach can be more than "good enough"—it can be the superior choice from a practical, scalable, and transparent standpoint.

5. Conclusions and Future Works

This study successfully developed and validated a lightweight yet highly effective system for detecting Indonesian-language hoaxes using a combination of TF-IDF feature extraction and a Logistic Regression classifier. The model achieved a notable accuracy of 94.98%, and more importantly, a recall of 98% for the hoax class, confirming its strong capability to filter disinformation. The primary contribution of this research is not merely the achievement of high accuracy, but the demonstration of the practical viability of a computationally efficient approach. We argue that in the context of real-world deployment, especially in resource-constrained environments, the balance between performance, efficiency, and interpretability positions this classical approach as a superior practical solution compared to more complex, resource-intensive models.

The findings have significant practical applications, particularly for integration into fact-checking platforms in Indonesia. The model's lightweight nature, trainable and deployable on standard hardware (e.g., Intel Core i5, 8 GB RAM), makes it suitable for adoption by local newsrooms, non-governmental organizations, and fact-checking initiatives like MAFINDO's Turnbackhoax.id. Its high recall for hoaxes (98%) ensures effective filtering of disinformation, which can be implemented as a preliminary screening tool within fact-checking workflows, flagging potential hoaxes for human review. The model's interpretability, with transparent feature weights, allows fact-checkers to understand and validate predictions, enhancing trust and enabling rapid identification of problematic patterns (e.g., misclassified factual articles). For instance, the model can be integrated into a web-based fact-checking platform, where news articles are processed in real-time, with low-confidence predictions automatically routed to human fact-checkers, creating an efficient human-in-the-loop system [14]. This practical deployment can support Indonesia's efforts to combat disinformation, particularly during high-stakes events like elections or public health campaigns, where timely and scalable solutions are critical.

The analysis also shed light on the model's limitations, which in turn provide valuable insights. The slightly lower recall for the factual class (91%) and the model's varying confidence on out-of-distribution data highlight the inherent weaknesses of a bag-of-words approach, which struggles with semantic nuance and contextual understanding. These findings underscore that while our model is a powerful tool for identifying strong lexical signals, it is not a panacea for all types of disinformation. Based on these conclusions, future research should proceed in several key directions. Future efforts should focus on targeted data enrichment, including sourcing factual news from a wider array of genres beyond formal reports to improve factual recall and curating a balanced sub-corpus of diverse hoax typologies to enhance model robustness [14]. Furthermore, to overcome the semantic limitations of TF-IDF, future work could explore hybrid architectures that combine TF-IDF features with lightweight contextual embeddings, potentially improving performance on nuanced texts[15]. Finally, the model's ability to signal low confidence on uncertain predictions opens a valuable opportunity for designing human-in-the-loop systems, where such predictions are automatically routed to human fact-checkers, creating a synergistic workflow that leverages both automation and human judgment [16]. These advancements could further enhance the model's integration into fact-checking platforms, ensuring scalability and adaptability to evolving disinformation trends.

6. References

[1] E. Effendi, "User behaviour and hoax information on social media case of Indonesia", doi: 10.25139/jsk.v7i3.7402.

[2] dan H. H. Palupi Anggraheni, Novi Tri Setyowati, "Social Media and Political Participation in Indonesia: Restrictions Access at Announcement Results of 2019 Presidential Election", doi: 10.56353/aspiration.v2i1.23.

[3] I. L. W. & L. P. I. Harini. Charlotte Jocelynne, "Detection of Political Hoax News Using Fine-Tuning IndoBERT", doi: 10.30871/jaic.v9i2.8989.

[4] T. Koto, F., Rahimi, A., Lau, J. H., & Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," 2020, doi: 10.18653/v1/2020.coling-main.66.

[5] O. N. C. & F. Budiman, "Performa Logistic Regression dan Naive Bayes dalam Klasifikasi Berita Hoax di Indonesia", doi: 10.29408/edumatic.v9i1.28987.

[6] R. E. S. & A. Nurlayli, "Comparative analysis of Indonesian news validity detection accuracy using machine learning", doi: 10.21831/jeatech.v4i1.58791.

[7] & W. P. K. N. V. Chawla, K. W. Bowyer, L. O. Hall, "SMOTE: Synthetic Minority Over-sampling Technique", doi: 10.1613/jair.953.

[8] J. F. Trevor Hastie, Robert Tibshirani, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction", doi: 10.1007/978-0-387-84858-7.

[9] A. K. U. & S. G. (2014), "The impact of preprocessing on text classification", doi: 10.1016/j.ipm.2013.08.006.

[10] & M. W. Mahoney. A. Dasgupta, P. Drineas, B. Harb, V. Josifovski, "Feature Selection Methods for Text Classification", doi: 10.1145/1281192.1281220.

[11] A. Y. & I. Haristyawan. D. Raya, "Application of K-Nearest Neighbor Algorithm For Sentiment Analysis On Free Fire Online Game Based On Google Play Store Reviews", doi: 10.32664/j-intech.v13i01.1882.

[12] D. R. & S. A. S. Vosoughi, "The spread of true and false news online", doi: 10.1126/science.aap9559.

[13] & W. Y. W. (2020) Ray Oshikawa, Jing Qian, "A Survey on Natural Language Processing for Fake News Detection".

[14] J. B. & Á. F. Eduardo Mosqueira-Rey, Elena Hernández-Pereira, David Alonso-Ríos, "Human-in-the-loop machine learning: a state of the art", doi: 10.1007/s10462-022-10246-w.

[15] A. G. & A. M. Emma Strubell, "nergy and Policy Considerations for Deep Learning in NLP", doi: 10.18653/v1/P19-1355.

[16] C.-O. T. & E.-S. Apostol, "It's All in the Embedding! Fake News Detection Using Document Embeddings," 2023, doi: 10.3390/math11030508.