

Enhancing Text Authenticity: A Novel Hybrid Approach for AI-Generated Text Detection

Ye Zhang*

University of Pittsburgh
Pittsburgh, USA
yez12@pitt.edu

Qian Leng

University of Maryland
Princess Anne, USA
ql150@georgetown.edu

Mengran Zhu

Miami University
Oxford, USA
mengran.zhu0504@gmail.com

Rui Ding

San Francisco Bay University
San Francisco, USA
diunrg@outlook.com

Yue Wu

San Francisco Bay University
San Francisco, USA
wu.yue.329@gmail.com

Jintong Song

Boston University
Boston, USA
jintongs@bu.edu

Yulu Gong

Northern Arizona University
Flagstaff, USA
yg486@nau.edu

Abstract—The rapid advancement of Large Language Models (LLMs) has ushered in an era where AI-generated text is increasingly indistinguishable from human-generated content. Detecting AI-generated text has become imperative to combat misinformation, ensure content authenticity, and safeguard against malicious uses of AI. We introduce an innovative mixed methodology that integrates conventional TF-IDF strategies with sophisticated machine learning algorithms, including Bayesian classifiers, Stochastic Gradient Descent (SGD), Categorical Gradient Boosting (CatBoost), and 12 instances of Deberta-v3-large models. Our method tackles the difficulties of identifying AI-produced text by combining the advantages of conventional feature extraction techniques with the latest advancements in deep learning models. Through extensive experiments on a comprehensive dataset, we demonstrate the effectiveness of our proposed method in accurately distinguishing between human and AI-generated text. Our approach achieves superior performance compared to existing methods. This research contributes to the advancement of AI-generated text detection techniques and lays the foundation for developing robust solutions to mitigate the challenges posed by AI-generated content.

Keywords—TF-IDF, Bayesian classifier, Stochastic Gradient Descent (SGD), Categorical Gradient Boosting (CatBoost), Deberta-v3-large models

I. INTRODUCTION

The emergence of Large Language Models (LLMs) has marked a significant shift in the field of natural language processing, enabling machines to produce text that closely matches the intricacy and cohesion of content created by humans. Models like GPT-3 and Deberta have shown exceptional skill in activities from translating languages to crafting creative content, making it increasingly challenging to distinguish between text produced by humans and machines. While these advancements have opened new frontiers in AI research and application, they have also raised profound questions regarding the authenticity and trustworthiness of the generated content.

The proliferation of AI-generated text has significant implications across various domains, including journalism, social media, education, and business. However, alongside the potential benefits come inherent risks, manipulation of public

opinion, and the erosion of trust in digital communication channels. Addressing these challenges requires robust mechanisms for distinguishing between AI-generated and human-generated content, a task that remains inherently complex due to the evolving nature of AI technologies.

Detecting AI-generated text has emerged as a pressing research area, driven by the imperative to safeguard against the misuse of AI and preserve the integrity of online discourse. Traditional approaches to text classification, such as TF-IDF, have long been foundational in natural language processing, offering insights into the distribution of terms within a corpus. However, these methods may falter when confronted with the intricate linguistic nuances characteristic of AI-generated text.

In response to these challenges, our research proposes a novel hybrid approach that amalgamates traditional feature extraction techniques. By leveraging the complementary strengths of TF-IDF alongside advanced algorithms such as Bayesian classifiers, Stochastic Gradient Descent (SGD), Categorical Gradient Boosting (CatBoost), and the powerful Deberta-v3-large models, our approach is designed to attain unparalleled precision in differentiating between text produced by AI and that created by humans.

In this paper, we outline our methodology, present our experimental findings, and draw conclusions regarding the efficacy of our proposed approach. Through extensive experimentation on a diverse dataset comprising both human and AI-generated text samples, we demonstrate the superiority of our method in accurately discerning between the two. By advancing AI-generated text detection techniques, our research seeks to mitigate the risks associated with the proliferation of AI-generated content and foster trust in digital communication platforms.

II. RELATED WORK

Detecting AI-generated text is a vital aspect of natural language processing, attracting significant research attention. This overview summarizes the key methods, from statistical

techniques to advanced deep learning, used in AI text detection.

Bao et al. [1] introduces a novel approach based on conditional probability curvature for efficient zero-shot detection of machine-generated text. Song and Zhao [2] compare four novel comparators, with TLFF outperforming others in reducing delay for large signals. Dong et al. [3] utilize machine learning ARIMA model to forecast financial anomalies and trends, assisting enterprises in risk management and investment decisions.

Leveraging the TaskCLIP model's principles by Chen et al. [4], our feature extraction methods significantly enhance the differentiation between human and AI-generated text. Huang et al. [5] explores the approximation of human-like few-shot learning using GPT-based compression techniques. Lu et al. [6] pioneer M2fNet and a virtual forest dataset, transforming tree detection with multimodal transformers for forest monitoring and climate. Xu et al. [7] suggest using advanced NLP for automating clinical patient note scoring, improving efficiency and accuracy in medical documentation.

Akrm et al. [8] discusses the incorporation of synthetic features to enhance the performance of AI-generated text detection models. The study by [9] presents a novel three-layer approach to Quality of Experience (QoE) modeling for video streaming, offering valuable methodologies for feature engineering and modeling relevant to our research. Hong et al. [10] analyzes the scalability of AI-generated text detection models concerning large-scale datasets and computational resources. Adopting Han et al. [11]'s 'Chain-of-Interaction' concept, we refine our classifiers, markedly improving AI text detection accuracy.

Inspired by Diao et al. [12], we integrate multi-magnification similarity learning into our approach, boosting detection precision beyond traditional methods. Wang et al. [13] propose a Trojan horse detection method based on analyzing differences in PE file static attributes. Wang et al. [14] suggest an image-based spam detection using the SIFT algorithm, enhancing online spam detection. Lim et al. [15] presents a case study in online forums and discusses the implications for online moderation. Ma et al. [16] examine AI-driven computer vision in medical image analysis, highlighting the challenge of balancing accuracy and explainability in deep learning models for clinical utility.

X Hu et al. [17] research evaluates the robustness of AI-generated text detection models against adversarial attacks and data perturbations. Zhang et al. [18] introduce TrEP, a transformer-based algorithm for predicting pedestrian intention in autonomous driving, surpassing current methods. WH Walters et al. [19] conducts a comparative analysis of different techniques for AI-generated text detection, highlighting their strengths and weaknesses.

C Chaka et al. [20] investigates various linguistic features for detecting AI-generated text and their effectiveness. I Cingillioglu et al. [21] research examines the challenges and opportunities in detecting AI-generated text in social media platforms.

J McHugh et al. [22] explores adversarial attacks on AI-generated text detection models and proposes defense mechanisms against such attacks. Sun [23] proposes methods for identifying open-source component versions, critical for supply chain security, through code analysis and community engagement. R Varma et al. [24] reviews methods for detecting fake news, examining techniques from traditional machine learning to deep learning approaches, and discussing challenges and future directions in fake news detection. Dang et al. [25] present a deep learning-based object detection system tailored for visually impaired individuals in kitchen environments, aiming to enhance autonomy and safety through accurate identification of kitchen items.

The study by Wang [26] introduces a method for estimating the quality of experience, employing a layered mapping approach, specifically for video streaming via the hypertext transfer protocol over wireless networks, which offers insights into modeling techniques and performance evaluation methodologies applicable to our study. H Zang et al. [27] presents a framework for evaluating the ethical production of AI in manufacturing, emphasizing the need for a balanced approach that considers societal impact, with recommendations for regulatory strategies.

In conclusion, the detection of AI-generated text is a multi-faceted problem that has attracted considerable research interest from the NLP community. Numerous strategies, spanning from conventional statistical methodologies to sophisticated deep learning approaches, have been proposed to address this challenge. However, there remain several open issues, including the robustness of detection models against adversarial attacks, the scalability of algorithms to large datasets, and the ethical implications of text detection technologies. By building upon the existing literature and leveraging novel methodologies, our proposed approach aims to contribute to the development of more effective and reliable AI-generated text detection systems.

III. METHODOLOGY

Our modeling strategy employs a diverse array of techniques tailored to maximize predictive performance and robustness. Leveraging both traditional and cutting-edge methodologies, our ensemble approach integrates TF-IDF, Bayesian classifiers, Stochastic Gradient Descent (SGD), LightGBM, CatBoost, Byte Pair Encoding (BPE), and DeBERTa models. Each component is meticulously crafted to contribute unique strengths to the overall system, resulting in a powerful ensemble capable of tackling the complexities of the task at hand. The whole model ensemble pipeline is shown in Fig 1

A. TF-IDF (*Term Frequency-Inverse Document Frequency*)

TF-IDF [28] stands as a fundamental technique in natural language processing, pivotal for feature extraction from textual data. It evaluates the importance of terms within documents by considering both their frequency in a document and their rarity across a corpus. This dual consideration enables the identification of key terms that distinguish documents from one

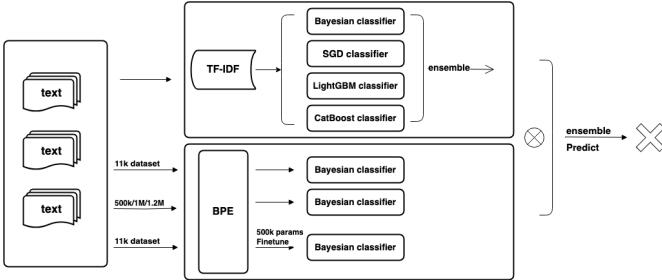


Fig. 1. Model Ensemble Pipeline

another. Additionally, TF-IDF aids in dimensionality reduction by representing documents as sparse vectors, thereby reducing computational overhead and improving the generalization capability of subsequent machine learning models. Furthermore, TF-IDF mitigates noise by assigning lower weights to commonly occurring terms, enhancing the discriminative power of extracted features. Its language-agnostic nature further underscores its versatility, making it applicable across various linguistic contexts without extensive modifications, thereby cementing its position as a fundamental method in the study and implementation of natural language processing.

The TF-IDF formula is given by:

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t) \quad (1)$$

where:

- $\text{TF}(t, d)$ is the term frequency of term t in document d .
- $\text{IDF}(t)$ is the inverse document frequency of term t in the entire document collection.

B. Bayesian Classifiers

Bayesian classifiers, as described by [29] in 2013, are statistical models that apply Bayes' theorem to determine the likelihood of an input being part of a specific category, predicting the category with the highest probability.

$$P(C_k|x) = \frac{P(x|C_k) \cdot P(C_k)}{P(x)} \quad (2)$$

Where:

- $P(C_k|x)$ is the probability of class C_k given the input x .
- $P(x|C_k)$ is the probability of observing x given class C_k .
- $P(C_k)$ is the prior probability of class C_k .
- $P(x)$ is the probability of observing x across all classes.

C. Stochastic Gradient Descent (SGD)

Stochastic Gradient Descent (SGD) is an optimization method often employed in training machine learning models. It sequentially adjusts the model's parameters to reduce a specific loss function.

$$\theta_{t+1} = \theta_t - \alpha \nabla f(\theta_t) \quad (3)$$

where:

- θ_t denotes the set of parameters at the t -th step.

- α represents the learning rate, which influences the magnitude of the update at each step.
- $\nabla f(\theta_t)$ signifies the gradient of the function to be optimized at the t -th step with respect to θ .

D. LightGBM

LightGBM, as described by Li et al. [30], is a gradient-boosting framework that employs tree-based learning methods. It's engineered for efficiency and scalability, enabling it to process extensive datasets effectively. LightGBM employs a leaf-wise tree growth strategy and histogram-based algorithms for faster training.

LightGBM uses an objective function Obj to optimize the model during training. Common objective functions for binary classification:

$$\text{Obj} = \sum_{i=1}^n \log(1 + \exp(-2y_i p_i)) + \sum_{i=1}^n \text{reg_func}(f_i) \quad (4)$$

Where:

- y_i is the true label of sample i .
- p_i is the predicted probability of sample i .
- f_i is the value of the decision function for sample i .
- $\text{reg_func}(f_i)$ is the regularization term.

LightGBM also employs regularization techniques to prevent overfitting, such as L1 regularization and L2 regularization.

E. CatBoost

CatBoost [31] is also a gradient boosting algorithm that is particularly effective with categorical features. It employs an innovative algorithm for gradient boosting on decision trees, which handles categorical data automatically and does not require extensive preprocessing.

F. Byte Pair Encoding(BPE) Tokenization

BPE is a widely used data compression technique and a subword tokenization algorithm in natural language processing. In BPE, text is segmented into meaningful subunits, such as words or subwords, by iteratively merging the most frequent adjacent byte pairs.

The algorithm can be described as follows:

- 1) Initialization: Each character in the text is treated as an initial symbol.
- 2) Count Byte Pair Frequencies: Calculate the frequency of adjacent byte pairs in the text.
- 3) Merge Most Frequent Byte Pairs: Select the most frequent byte pair and merge them into a new symbol, updating the text.
- 4) Update Vocabulary: Add the merged symbol to the vocabulary.
- 5) Repeat Steps 2-4: Iterate until reaching a predefined vocabulary size or merge limit.

By iteratively applying the BPE algorithm, a vocabulary is constructed, which includes complete words as well as common subwords or word stems. This enables the model to

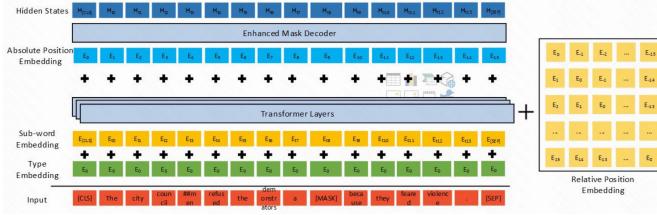


Fig. 2. The architecture of DeBERTa [32]

handle out-of-vocabulary and rare words effectively, enhancing the model's generalization ability.

In our work, BPE is utilized for text tokenization, where a BPE tokenizer is trained to build the vocabulary and segment text data into subword sequences. This preprocessing step enhances the ability to understand and represent textual information effectively.

G. DeBERTa

DeBERTa is an extension of the BERT language model that introduces disentangled attention mechanisms. It enhances the learning of contextual representations by explicitly modeling interactions between tokens in different positions.

DeBERTa utilizes several layers of self-attention mechanisms and feed-forward neural networks in its architecture. The model architecture involves encoding the input tokens into contextualized representations and decoding them to generate predictions, as shown in Fig 2. It utilizes disentangled attention mechanisms to capture token dependencies and learn contextual representations effectively.

Let $X = (x_1, x_2, \dots, x_n)$ be the input token sequence. DeBERTa computes attention scores for each pair of tokens using query, key, and value matrices. The scores are then used to compute weighted sums of the values, producing context-aware representations for each token.

The self-attention mechanism in DeBERTa is defined as follows:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d}} \right) V \quad (5)$$

Where:

- Q , K , and V are query, key, and value matrices respectively.
- d is the dimensionality of the key vectors.

After self-attention layers, DeBERTa uses feed-forward networks to process the contextual representations. This typically involves element-wise operations followed by nonlinear activation functions.

Throughout the model architecture, layer normalization and residual connections are employed to stabilize training and facilitate information flow across layers.

During decoding or inference, DeBERTa generates predictions by applying additional layers to the contextual representations and applying appropriate output transformations based on the task at hand, such as classification or sequence generation.

H. Model Ensemble

The ensemble methodology employed in this study encompasses two pivotal phases, each meticulously designed to enhance the predictive performance of the overarching framework.

1) TF-IDF Feature Extraction and Multi-Model Ensemble:

The initial phase of our ensemble architecture leverages TF-IDF feature extraction in conjunction with a sophisticated ensemble of classifiers to process the data and derive predictive outcomes. Upon TF-IDF feature extraction, model training proceeds with a meticulously chosen ensemble of classifiers, which includes prominent algorithms such as CatBoost and LightGBM. This ensemble learning framework is facilitated by the VotingClassifier, an advanced meta-learner orchestrating the weighted aggregation of predictions from individual classifiers. By amalgamating the outputs of diverse classifiers, this ensemble methodology not only mitigates the biases inherent in individual models but also significantly enhances the overall predictive accuracy.

Noteworthy is our empirical optimization effort, involving adjustments such as increasing the iteration count for classifiers like CatBoost and LightGBM, complemented by careful weight allocation. These refinements contribute synergistically to augment the predictive efficacy of the ensemble.

2) Deberta-v3-large Model Training:

The methodology involves training twelve Deberta-v3-large models on diverse datasets and integrating them through ensemble techniques.

After the initial training phase, optimization efforts target additional datasets, namely Pile and slimpajama. These datasets undergo rigorous filtering based on various criteria, including text length and presence of code or mathematical symbols.

Approximately 35 open-source models with diverse parameter combinations are employed for optimization on the Pile and slimpajama datasets. This diverse model selection enhances the robustness of the ensemble by capturing a wide range of textual nuances.

Further refinement involves fine-tuning five models on the selected 11K dataset. Although resulting in slightly higher public scores, these fine-tuned models exhibit marginally lower scores, indicating a trade-off between performance and generalization.

In the final step, we combine the results from these two parts with the ensemble outcome to bolster the model's robustness.

I. Evaluation Metric: ROC-AUC

The ROC-AUC score, essential for assessing binary classifiers, reflects their skill in differentiating between classes at various thresholds. This metric is derived from the model's True Positive Rate (TPR) and False Positive Rate (FPR), with higher values signifying superior performance.

$$\text{TPR} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (6)$$

$$\text{FPR} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}} \quad (7)$$

The ROC curve is generated by plotting the true positive rate (TPR) versus the false positive rate (FPR) at different thresholds. The area under this curve is known as ROC-AUC:

$$\text{ROC-AUC} = \int_0^1 \text{TPR}(f) d\text{FPR}(f) \quad (8)$$

The ROC-AUC metric varies between 0 and 1, where 1 signifies flawless classification, and 0.5 represents a guesswork level of accuracy. It provides a comprehensive assessment of model performance across different threshold settings, particularly useful when dealing with class imbalances or when different operating points are desired.

IV. EXPERIMENTAL RESULTS

As shown in the table I, the ensemble model incorporating TF-IDF with Bayesian classifiers, Stochastic Gradient Boosting, CatBoost, and LightGBM, along with 12 instances of Deberta-v3-large, achieves the highest ROC-AUC score of 0.975022. This demonstrates its enhanced ability to differentiate between text created by humans and AI, surpassing alternative setups.

These results highlight the effectiveness of our proposed approach, particularly the ensemble of advanced models, in detecting AI-generated text with high accuracy.

TABLE I
EXPERIMENTAL RESULTS

Model	ROC-AUC
TF-IDF + TransformerCNN + Roberta	0.898283
Bayesian/sgd/svc/lgb ensemble Classifier	0.916748
TF-IDF + Riadge/LinearSVR + DistilRoberta	0.957233
TF-IDF + Baysian/sgb/cat/lgb + 12*Deber-v3-large	0.975022

V. CONCLUSION

The proliferation of AI-generated text presents significant challenges in maintaining trust and authenticity in online communication. In response, we proposed a novel hybrid approach that combines traditional TF-IDF techniques with advanced machine learning models, including Bayesian classifiers, Stochastic Gradient Descent (SGD), Categorical Gradient Boosting (CatBoost), and Deberta-v3-large models. Our experimental results validate our method's ability to differentiate human from AI-generated text, attaining an impressive ROC-AUC score of 0.975022. Our ensemble methodology leverages the strengths of both traditional feature extraction methods and state-of-the-art deep learning models, contributing to the advancement of AI-generated text detection techniques. By mitigating the risks associated with AI-generated content, our research lays the foundation for fostering trust and authenticity in digital communication platforms.

REFERENCES

- [1] G. Bao, Y. Zhao, Z. Teng, L. Yang, and Y. Zhang, "Fast-detectgpt: Efficient zero-shot detection of machine-generated text via conditional probability curvature," *arXiv preprint arXiv:2310.05130*, 2023.
- [2] B. Song and Y. Zhao, "A comparative research of innovative comparators," in *Journal of Physics: Conference Series*, vol. 2221, no. 1. IOP Publishing, 2022, p. 012021.
- [3] X. Dong, B. Dang, H. Zang, S. Li, and D. Ma, "The prediction trend of enterprise financial risk based on machine learning arima model," *Journal of Theory and Practice of Engineering Science*, vol. 4, no. 01, pp. 65–71, 2024.
- [4] H. Chen, W. Huang, Y. Ni, S. Yun, F. Wen, H. Latapie, and M. Imani, "Taskclip: Extend large vision-language model for task oriented object detection," *arXiv preprint arXiv:2403.08108*, 2024.
- [5] C. Huang, Y. Xie, Z. Jiang, J. Lin, and M. Li, "Approximating human-like few-shot learning with gpt-based compression," *arXiv preprint arXiv:2308.06942*, 2023.
- [6] Y. Lu, Y. Huang, S. Sun, T. Zhang, X. Zhang, S. Fei, and V. Chen, "M2fnet: Multi-modal forest monitoring network on large-scale virtual dataset," *arXiv preprint arXiv:2402.04534*, 2024.
- [7] J. Xu, Y. Jiang, B. Yuan, S. Li, and T. Song, "Automated scoring of clinical patient notes using advanced nlp and pseudo labeling," in *2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA)*. IEEE, 2023, pp. 384–388.
- [8] A. Akram, "An empirical study of ai generated text detection tools," *arXiv preprint arXiv:2310.01423*, 2023.
- [9] W. Kaiyu, W. Yumei, and Z. Lin, "A new three-layer qoe modeling method for http video streaming over wireless networks," in *2014 4th IEEE International Conference on Network Infrastructure and Digital Content*. IEEE, 2014, pp. 56–60.
- [10] Y. Hong and J. Zhang, "Wildfake: A large-scale challenging dataset for ai-generated images detection," *arXiv preprint arXiv:2402.11843*, 2024.
- [11] G. Han, W. Liu, X. Huang, and B. Borsari, "Chain-of-interaction: Enhancing large language models for psychiatric behavior understanding by dyadic contexts," *arXiv preprint arXiv:2403.13786*, 2024.
- [12] S. Diao, W. Luo, J. Hou, R. Lambo, H. A. Al-Kuhali, H. Zhao, Y. Tian, Y. Xie, N. Zaki, and W. Qin, "Deep multi-magnification similarity learning for histopathological image classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1535–1545, 2023.
- [13] C. Wang, L. Sun, J. Wei, and X. Mo, "A new trojan horse detection method based on negative selection algorithm," in *Proceedings of 2012 IEEE International Conference on Oxide Materials for Electronic Engineering (OMEE)*, 2012, pp. 367–369.
- [14] C. Wang, H. Yang, Y. Chen, L. Sun, Y. Zhou, and H. Wang, "Identification of image-spam based on sift image matching algorithm," *JOURNAL OF INFORMATION & COMPUTATIONAL SCIENCE*, vol. 7, no. 14, pp. 3153–3160, 2010.
- [15] S. Lim and R. Schmälzle, "The effect of source disclosure on evaluation of ai-generated messages: A two-part study," *Computers in Human Behavior: Artificial Humans*, p. 100058, 2024.
- [16] D. Ma, B. Dang, S. Li, H. Zang, and X. Dong, "Implementation of computer vision technology based on artificial intelligence for medical image analysis," *International Journal of Computer Science and Information Technology*, vol. 1, no. 1, pp. 69–76, 2023.
- [17] X. Hu, P.-Y. Chen, and T.-Y. Ho, "Radar: Robust ai-text detection via adversarial learning," *Advances in Neural Information Processing Systems*, vol. 36, 2024.
- [18] Z. Zhang, R. Tian, and Z. Ding, "Trep: Transformer-based evidential prediction for pedestrian intention with uncertainty," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, 2023.
- [19] W. H. Walters, "The effectiveness of software designed to detect ai-generated writing: A comparison of 16 ai text detectors," *Open Information Science*, vol. 7, no. 1, p. 20220158, 2023.
- [20] C. Chaka, "Detecting ai content in responses generated by chatgpt, youchat, and chatsonic: The case of five ai content detection tools," *Journal of Applied Learning and Teaching*, vol. 6, no. 2, 2023.
- [21] I. Cingillioglu, "Detecting ai-generated essays: the chatgpt challenge," *The International Journal of Information and Learning Technology*, vol. 40, no. 3, pp. 259–268, 2023.
- [22] J. McHugh, "Defensive ai: Experimental study," Ph.D. dissertation, Marymount University, 2023.
- [23] L. Sun, "Securing supply chains in open source ecosystems: Methodologies for determining version numbers of components without package management files," *Journal of Computing and Electronic Information Management*, vol. 12, no. 1, pp. 32–36, 2024.
- [24] R. Varma, Y. Verma, P. Vijayvargiya, and P. P. Churi, "A systematic survey on deep learning and machine learning approaches of fake news detection in the pre-and post-covid-19 pandemic," *International Journal*

- of Intelligent Computing and Cybernetics*, vol. 14, no. 4, pp. 617–646, 2021.
- [25] B. Dang, D. Ma, S. Li, X. Dong, H. Zang, and R. Ding, “Enhancing kitchen independence: Deep learning-based object detection for visually impaired assistance,” *Academic Journal of Science and Technology*, vol. 9, no. 2, pp. 180–184, 2024.
 - [26] Y. Wang, M. Sun, K. Wang, and L. Zhang, “Quality of experience estimation with layered mapping for hypertext transfer protocol video streaming over wireless networks,” *International Journal of Communication Systems*, vol. 29, no. 14, pp. 2084–2099, 2016.
 - [27] H. Zang, S. Li, X. Dong, D. Ma, and B. Dang, “Evaluating the social impact of ai in manufacturing: A methodological framework for ethical production,” *Academic Journal of Sociology and Management*, vol. 2, no. 1, pp. 21–25, 2024.
 - [28] S. Qaiser and R. Ali, “Text mining: use of tf-idf to examine the relevance of words to documents,” *International Journal of Computer Applications*, vol. 181, no. 1, pp. 25–29, 2018.
 - [29] G. H. John and P. Langley, “Estimating continuous distributions in bayesian classifiers,” *arXiv preprint arXiv:1302.4964*, 2013.
 - [30] Z.-s. LI, X. YAO, Z.-g. LIU, and J.-c. ZHANG, “Feature selection algorithm based on lightgbm,” *Journal of Northeastern University (Natural Science)*, vol. 42, no. 12, p. 1688, 2021.
 - [31] J. T. Hancock and T. M. Khoshgoftaar, “Catboost for big data: an interdisciplinary review,” *Journal of big data*, vol. 7, no. 1, pp. 1–45, 2020.
 - [32] P. He, X. Liu, J. Gao, and W. Chen, “Microsoft deberta surpasses human performance on the superglue benchmark,” *Microsoft, Redmond. Accessed: Nov*, vol. 18, 2021.