International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05623
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Intelligent Detection and Feedback System for Plagiarism and AI – Generated Text

Dr. Atul Kumar Ramotra[1], Pakkurthi Ravi Kiran[2], Cheekati Shleshitha[3], Junuthula Shashidar[4]

[1]Associate Professor, Department of CSE (AI & ML), ACE Engineering College, Hyderabad, Telangana, India.
[2]Student of Department of CSE (AI & ML), ACE Engineering College, Hyderabad, Telangana, India.
[3]Student of Department of CSE (AI & ML), ACE Engineering College, Hyderabad, Telangana, India.
[4]Student of Department of CSE (AI & ML), ACE Engineering College, Hyderabad, Telangana, India.

Abstract

Over the past few years, exponential increases in digital content and development of generative Artificial Intelligence (AI) have introduced new issues in guaranteeing originality and authenticity of written materials. Conventional plagiarism scanners can identify only direct text copying or paraphrased equivalent but cannot identify AI-generated content, which typically imitates human writing. In an effort to bridge this gap, this project suggests a Plagiarism and AI-Generated Content Detection and Feedback System that will search textual data for both traditional plagiarism and artificially created results. The system combines Natural Language Processing (NLP) methods, semantic similarity indicators, and machine learning classifiers to separate human and artificially generated text. Besides detection, the system also gives users rich feedback, such as highlighting plagiarized text, possible AI-generated regions, and suggesting improvement in writing style and originality. Not just an academic integrity increase, as well as content genuineness, it is ethical writing practice as well because it leads users to originality. The solution can be used across all education systems, research fields, and industries based on content to provide stability and originality in the age of AI-driven content generation.

Keywords: Plagiarism Detection, AI Generated Content Detection, Natural Language Processing, Semantic Similarity, Text Authenticity, Feedback System.

1. Introduction

1.1Background and Motivation

With the advancement of Artificial Intelligence technologies, the creation of digital content has increased significantly across academic, professional, and social platforms. Generative AI tools such as language models can produce human-like text within seconds, which has transformed the way people write reports, essays, articles, and other forms of content. While these technologies provide many advantages in terms of efficiency and creativity, they also introduce serious challenges related to **content originality, plagiarism, and authenticity**.

Traditional plagiarism detection systems rely on keyword matching and document similarity techniques to identify copied or paraphrased text from existing sources. Although these systems are effective for detecting direct copying, they often fail to identify **AI-generated content**, which may appear original but is actually produced by machine learning models.

As AI writing tools become more sophisticated, it becomes increasingly difficult for educators, researchers, and organizations to distinguish between human-written and machine-generated text. This issue raises concerns about academic integrity, intellectual property rights, and the reliability of digital information.

To address these challenges, intelligent systems that combine machine learning and Natural Language Processing techniques are required. Such systems can analyse linguistic patterns, semantic similarity, and statistical features of text to determine whether content is original, plagiarized, or AI-generated.

This research proposes a **web-based Intelligent Detection and Feedback System for Plagiarism and AI-Generated Text**, which aims to improve the accuracy and transparency of text authenticity analysis.

1.2Introduction

Maintaining originality in written content has become increasingly difficult due to the rapid growth of AI- generated text technologies. Many existing plagiarism detection systems are unable to detect advanced forms of paraphrasing or automatically generated content. Additionally, most AI detection tools provide probability scores without explaining the reasoning behind the detection results.

The proposed system addresses these limitations by combining **plagiarism detection and AI-generated text identification into a single platform**. The system utilizes Natural Language Processing techniques for text preprocessing, machine learning algorithms for classification, and semantic similarity analysis to detect copied or modified content.

The web-based interface allows users to easily input text or upload documents and receive detailed analysis reports. The system highlights suspicious sections of text and provides feedback that helps users improve the originality and quality of their writing.

Important features of this system include:

- Detection of copied or paraphrased content using semantic similarity.
- Identification of AI-generated writing patterns using linguistic analysis.
- Interactive feedback to improve writing originality.
- A user-friendly web application for real-time text analysis.

By combining these technologies, the proposed system contributes to the development of more reliable tools for content authenticity verification.

---

2. Literature Review

2.1GPT-Sentinel: Distinguishing Human and ChatGPT Generated Content

This research focuses on developing a system capable of identifying whether a piece of text is written by a human or generated by an artificial intelligence model such as ChatGPT. The authors created a dataset called **OpenGPTText**, which contains both human-written texts and AI-generated versions of those texts. This dataset was used to train models to recognize subtle differences in writing patterns between human and AI-generated content.

The system analyzes various linguistic features such as sentence structure, vocabulary richness, and writing style. The study demonstrates that machine learning models can learn complex patterns in text and effectively differentiate between human-written and AI-generated content.

The results of the research showed that the detection models achieved accuracy levels above **97%**, proving the effectiveness of transformer-based deep learning models for AI-generated text detection.

2.1.1 Methodologies and Algorithms

The study uses advanced **Natural Language Processing techniques and transformer-based deep learning models**. Models such as **RoBERTa and T5** were trained on the OpenGPTText dataset. Feature extraction techniques were used to analyze lexical diversity, syntactic structures, and contextual relationships between words.

The training process involved optimization techniques such as **cross-entropy loss and gradient descent** to improve classification performance. Additionally, visualization methods such as **PCA and t-SNE** were used to analyze how the models separate AI-generated and human-written text.

2.2Smaller Language Models are Better Black-Box Machine Generated Text Detectors

This research investigates the ability of AI models to detect text generated by other AI models without having access to their internal architecture. The study explores whether smaller language models can effectively identify machine-generated text produced by large models such as GPT.

The researchers conducted experiments using datasets like **SQuAD and Writing Prompts** to test the detection performance. The results revealed that **smaller models such as OPT-125M performed better than larger models like OPT-6.7B** in identifying AI-generated text.

The research explains that smaller models are more sensitive to the patterns present in machinegenerated content and can therefore detect subtle differences between human-written and AIgenerated text.

2.2.1 Methodologies and Algorithms

The researchers used **transformer-based models such as GPT and OPT** to evaluate detection performance. Detection accuracy was measured using **Area Under Curve (AUC)** metrics.

The concept of **curvature analysis** was used to measure the sensitivity of models to machinegenerated text. Higher curvature values indicated that the model could better distinguish between human and AI-generated content.

2.3RAIDAR: Generative AI Detection via Rewriting

RAIDAR introduces a novel method for detecting AI-generated text using rewriting techniques. The central idea of this approach is that when an AI model is asked to rewrite text, it tends to modify **human-written text more significantly than AI-generated text**.

This difference occurs because AI-generated content is already optimized for machine-style writing. Therefore, when rewritten, it undergoes fewer changes compared to human-written text.

The researchers tested RAIDAR across multiple domains including **news articles, academic abstracts, student essays, programming code, and online reviews**. The results showed that RAIDAR significantly outperformed several existing AI detection methods.

2.3.1 Methodologies and Algorithms

The RAIDAR system works by measuring the **edit distance between original text and rewritten text**. The system calculates how many changes occur during rewriting.

If the rewritten text shows significant changes, it is likely to be human-written. If minimal changes occur, the text is likely AI-generated.

This method works effectively even for **black-box AI models such as GPT-3.5 and GPT-4** because it relies only on observable outputs.

2.4Assessing AI Detectors in Identifying AI-Generated Code

This research evaluates the performance of existing AI detection tools in identifying AI-generated programming code. The authors collected a dataset containing more than **5000 Python programming problems and solutions** from platforms such as Kaggle, LeetCode, and Quescol.

The researchers generated different versions of AI-written code by modifying prompts, removing comments, renaming variables, and adding test cases. These variations were used to test whether existing AI detectors could still identify machine-generated code.

The results showed that most existing AI detection tools perform poorly when analyzing programming code, achieving accuracy between **50% and 60%**.

2.4.1 Methodologies and Algorithms

The study evaluated several AI detection tools including:

- GPTZero
- DetectGPT
- GLTR
- Sapling AI Detector

The experiments involved generating code variations and analyzing how detection accuracy changed with different modifications.

2.5Textual Analysis and Detection of AI Generated Academic Texts

This research compares AI-generated academic abstracts produced by ChatGPT with human-written academic abstracts. The goal was to determine whether AI-generated text could be detected using machine learning techniques.

The study analyzed several linguistic features including **readability scores, stylometric features, sentiment analysis, semantic similarity, and vocabulary diversity**.

The results showed that machine learning models could successfully identify AI-generated academic content with high accuracy.

2.5.1 Methodologies and Algorithms

The study used machine learning models including:

- Support Vector Machine (SVM)
- Logistic Regression
- Random Forest

The models were trained using **DistilBERT embeddings**, which convert text into numerical vectors representing semantic meaning.

The best-performing models achieved an **F1 score of up to 99%**, outperforming traditional AI detection tools.

2.6Differentiating Chat Generative Pretrained Transformer from Humans: Detecting ChatGPT Generated Text and Human Text Using Machine Learning

This study proposes a hybrid machine learning framework called **TSA-LSTM-RNN** to classify humanwritten and AI-generated text. The system combines feature extraction, deep learning models, and optimization algorithms to improve detection accuracy.

The research focuses on identifying patterns in word usage, sentence structure, and contextual relationships between words.

The results demonstrated that the proposed model achieved higher accuracy than traditional machine learning classifiers.

2.6.1 Methodologies and Algorithms

The framework uses multiple feature extraction techniques including:

- TF-IDF
- Word2Vec embeddings
- Count Vectorization

A **Long Short-Term Memory (LSTM) Recurrent Neural Network** was used to analyze sequential patterns in text. The **Tunicate Swarm Algorithm** was used for hyperparameter optimization.

2.7A New Online Plagiarism Detection System based on Deep Learning

This research proposes a plagiarism detection system using deep learning techniques to identify copied or paraphrased content more accurately than traditional systems.

The system is capable of detecting multiple types of plagiarism including:

- Direct copying
- Paraphrased plagiarism
- Translation plagiarism
- Idea plagiarism

The results showed that the proposed system achieved **98.33% accuracy**, outperforming traditional plagiarism detection tools such as Turnitin and iThenticate.

2.7.1 Methodologies and Algorithms

The system uses deep learning models such as:

- Doc2Vec
- Convolutional Neural Networks (CNN)
- SLSTM models

The framework consists of three main layers:

1. Preprocessing Layer
2. Learning Layer
3. Detection Layer

ISSN: 2583-6129
DOI:10.55041/ISJEM05623

2.8curacy of Existing Algorithms and Models

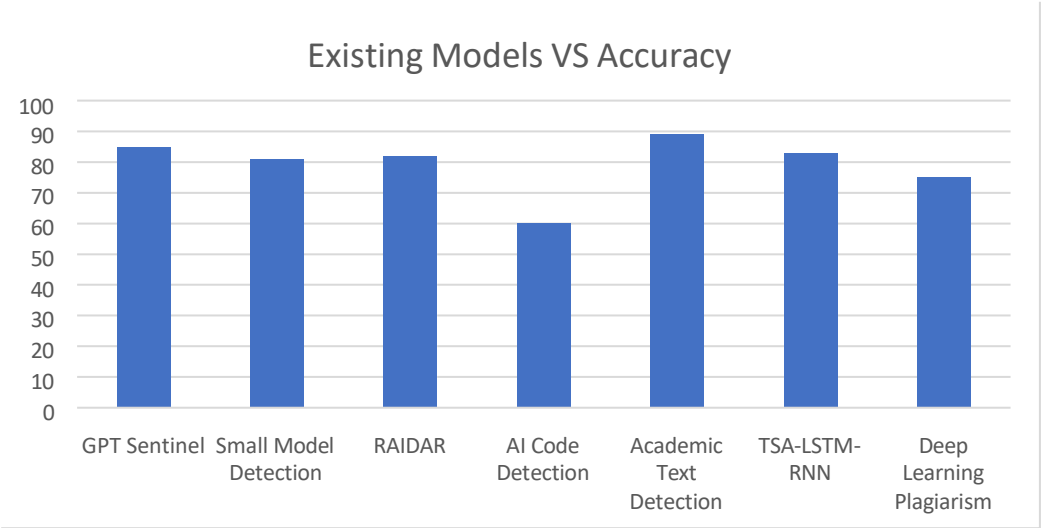

Figure: Accuracy comparison of existing AI detection and plagiarism detection models

2.9Comparative Analysis

Name of Paper	Year	Algorithms Used	Key Findings	Limitations
GPT-Sentinel: Distinguishing Human and ChatGPT Generated Content.	2023	RoBERTa, T5	Detects AI text with high accuracy	Requires large dataset
Smaller Language Models are Better Black-box Machine-Generated Text Detectors	2022	OPT Models	Smaller models perform better	Limited generalization
RAIDAR: GENERATIVE AI DETECTION VIA REWRIT ING.	2023	Edit Distance Analysis	Works for black-box models	Requires rewriting step
Assessing AI Detectors in Identifying AI-Generated Code: Implications for Education	2024	GPTZero, DetectGPT	AI code detection evaluated	Low accuracy
Textual Analysis and Detection of AI Generated Academic Texts	2023	SVM, Random Forest	High detection accuracy	Dataset limitations
Differentiating Chat Generative Pretrained Transformer from Humans: Detecting ChatGPT-	2023	LSTM + Optimization	Better pattern recognition	Computational cost

Generated Text and Human Text Using Machine Learning				
A New Online Plagiarism Detection System based on Deep Learning	2021	CNN, Doc2Vec	Detects multiple plagiarism types	Requires training data

2.10 Research Gaps

Despite significant advancements in plagiarism detection and AI-generated text analysis, several research gaps still remain.

Most existing systems focus only on either plagiarism detection or AI-generated text detection rather than combining both functionalities. Additionally, many detection systems lack transparency and fail to provide meaningful feedback to users about why a piece of text was flagged.

Another challenge is the limited availability of diverse datasets for training AI detection models. Many systems are trained using specific datasets, which reduces their ability to generalize across different writing styles and domains.

The proposed system attempts to address these research gaps by integrating plagiarism detection and AI- generated text identification into a single platform while also providing user-friendly feedback and explainable results.

Abbreviations

NLP – Natural Language Processing ML – Machine Learning
AI – Artificial Intelligence
TF-IDF – Term Frequency Inverse Document Frequency CNN – Convolutional Neural Network
LSTM – Long Short-Term Memory

3. Conclusion

The rapid development of AI writing technologies has made it increasingly challenging to maintain authenticity and originality in digital content. Traditional plagiarism detection systems alone are not sufficient to address these challenges because they cannot reliably identify AI-generated text.

This research proposed an **Intelligent Detection and Feedback System for Plagiarism and AIGenerated Text**, which combines Natural Language Processing, semantic similarity analysis, and machine learning techniques to detect both copied and AI-generated content.

The system provides a user-friendly web interface where users can upload text and receive detailed originality reports with highlighted suspicious sections and feedback. The results show that integrating plagiarism detection with AI-generated text analysis can significantly improve content authenticity verification.

Future improvements may include incorporating deep learning models, supporting multilingual text detection, and deploying the system on cloud platforms for large-scale analysis.

References

[1] Y. Chen, H. Kang, V. Zhai, L. Li, R. Singh, and B. Raj, "GPT-Sentinel: Distinguishing Human and ChatGPT Generated Content," Carnegie Mellon University, 2023.

[2] F. Mireshghallah, J. Mattern, S. Gao, R. Shokri, and T. Berg-Kirkpatrick, "Smaller Language Models are Better Black-Box Machine-Generated Text Detectors," 2023.

[3] C. Mao, C. Vondrick, H. Wang, and J. Yang, "RAIDAR: Generative AI Detection via Rewriting," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2024.

[4] W. H. Pan, M. J. Chok, J. L. S. Wong, Y. X. Shin, Y. S. Poon, Z. Yang, D. Lo, and M. K. Lim, "Assessing AI Detectors in Identifying AI-Generated Code: Implications for Education," in *Proceedings of the 46th International Conference on Software Engineering (ICSE-SEET)*, 2024.

[5] A. Al Medawer, "Textual Analysis and Detection of AI-Generated Academic Texts," Bachelor's Thesis, Mid Sweden University, 2023.

[6] I. Katib, F. Y. Assiri, H. A. Abdushkour, D. Hamed, and M. Ragab, "Differentiating Chat Generative Pretrained Transformer from Humans: Detecting ChatGPT-Generated Text and Human Text Using Machine Learning," *Mathematics*, MDPI, 2023.

[7] E. M. Hambi and F. Benabbou, "A New Online Plagiarism Detection System Based on Deep Learning," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 11, no. 9, 2020.