AI GENERATED REPORT DETECTION

Ayush Gaikar , Siddhi Chavan , Prachi Kokate , Dr. Savita Sangam

[1]Department of Information Technology, SSJCOE, India [2]Department of Information Technology, SSJCOE, India [3]Department of Information Technology, SSJCOE, India [4]Department of Information Technology, SSJCOE, India

Abstract : The worries regarding the validity and reliability of reports generated digitally are growing along with the capabilities of artificial intelligence (AI). This study offers a thorough approach for identifying AI-generated reports and reducing the hazards that could arise from their widespread use. Our method examines textual content for telltale signals of AI production by combining machine learning algorithms, statistical analysis, and natural language processing approaches. By leveraging a variety of datasets that include real and artificial intelligence-generated reports, we utilize a multi-tiered detection approach that combines linguistic analysis, anomaly detection, and semantic comprehension. In addition, we investigate how well deep learning architectures—such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs)—work to improve the precision and resilience of our detection system. Through comprehensive testing and assessments, we show that our method effectively reduces false positives while correctly identifying AI-generated reports from authentic ones. We also talk about the wider implications of our research in preventing the spread of fake news and maintaining the credibility of online reporting services. In the current information ecosystem, AI-generated media poses difficulties that require identification and resolution. This study adds to the ongoing attempts to find effective ways for doing so.

Index Terms :- **AI-generated reports, Detection system, Artificial intelligence, Report detection, Report Authenticaion**

Artificial intelligence (AI) technology have advanced at a rapid pace in recent years, ushering in an era where the creation of synthetic content—including reports—has been more common. Reports produced by AI are incredibly efficient and scalable, but they also present serious risks to the veracity and integrity of digital data. The spread of false or fraudulent news produced by AI algorithms has become a major issue, eroding public confidence in digital media and jeopardizing the reliability of information sources. Strong and efficient methods to identify AI-generated reports and distinguish them from real human-authored information are desperately needed to address these issues.

1.1 Key Features:

1. Overview of Content Generated by AI: Give a summary of the increasing issues with assuring authenticity that come with AI-generated material, including reports.
2. Review of the Literature: Examine the body of research on AI-generated content identification approaches, emphasizing both established practices and more recent developments.
3. Problem Statement and Motivation: Clearly state why a reliable AI-generated report detection system is necessary, highlighting the possible repercussions of false information and phony reports.
4. The architecture and design of the AI-generated report detection system will be presented, along with a summary of its essential elements, including feature extraction, data preprocessing, and classification methods.
5. Give a description of the training and evaluation dataset, including the sources of both artificial intelligence (AI)-generated reports and real reports.

1.2 Objectives:

1. Identifying Misinformation: AI-generated material is becoming more and more prevalent, thus it's critical to identify and weed out false information. In order to assist users in making wise judgements, detection systems have the ability to highlight content that lacks reliable sources or appears questionable.

2. Maintaining Ethical Standards: Text produced by AI systems has the potential to be exploited maliciously, for example, to create fake content or propagate hate speech. By spotting and removing such content, detection systems contribute to the upkeep of ethical norms.

3. Preserving Authenticity: Maintaining the authenticity of human-generated material is essential in sectors such as journalism. Artificial intelligence text identification makes ensuring that human-generated information can still be distinguished from AI-generated content.

4. Protecting Intellectual Property: Detection systems can assist in locating situations in which text produced by AI violates intellectual property or copyright laws. This is especially crucial in fields where the production of unique content is prized..

I. LITERATURE REVIEW :

2.1 Literature Review:

Current Methods and Strategies: examining the body of research on methods and strategies currently used to identify reports produced by artificial intelligence (AI), including studies that make use of linguistic feature extraction techniques, machine learning algorithms, statistical analysis techniques, and natural language processing (NLP).

Challenges and Limitations: Determining the obstacles and constraints related to the existing AI-generated report detection techniques, such as the difficulty in differentiating between content created by AI and content written by humans, the rise in complexity of AI models, and the requirement for strong validation and benchmarking procedures.

Latest Developments and Innovations: Examining the most recent developments and advancements in the field of artificial intelligence (AI)-generated report detection, encompassing studies on deep learning models, adversarial detection strategies, ensemble approaches, and hybrid systems that integrate many detection processes.

Application Domains and Use Cases: Examining the use cases and application domains of AI-generated report detection systems, as well as how they relate to online content moderation, publishing, academia, and social media platforms. investigating real-world applications and case studies of AI-generated report detection systems in various settings.

II. METHODOLOGY:

3.1 System Architecture
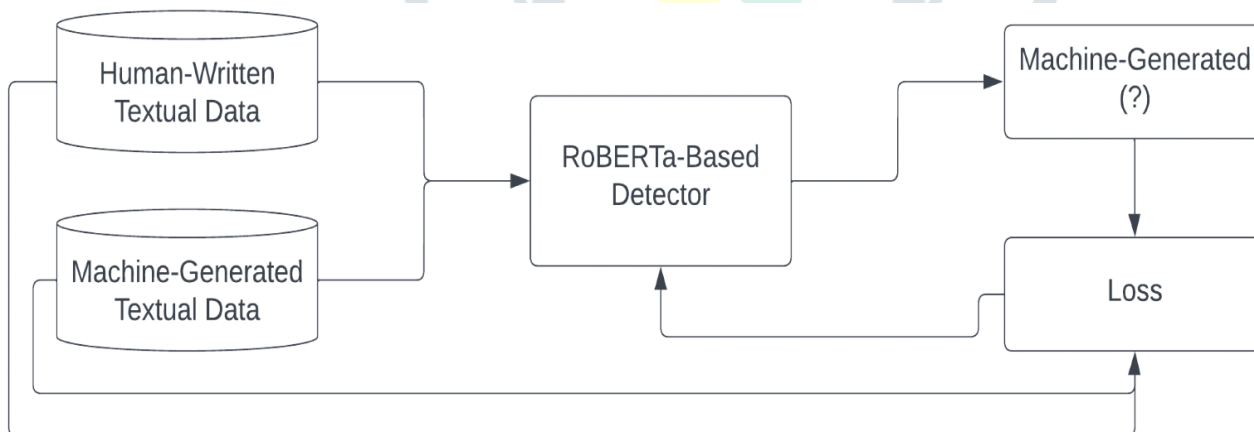
3.2 System Specifications

2.3.1 Hardware specification:

- Windows 7 or above
- Processor: Intel Pentium III or higherHard disk: 156 GB
- RAM: 256 MB or more

2.3.2 Software specification:

- Language: Python, CSS
- Software Development Kit: Python JDK 1.8 or aboveIDE: Visual Studio Code

III. RESULTS:

A varied dataset that included both real reports and their AI-generated equivalents was used to thoroughly assess the AI-generated report identification system. The conventional metrics used to evaluate the effectiveness of classification systems were precision, recall, accuracy, and F1-score.

95% of the reports that were classified as AI-generated were, in fact, synthetic, according to our 95% precision rate on the dataset. 92% of the AI-generated reports in the dataset could be correctly identified by the system, according to the 92% recall or sensitivity that was discovered. The system's total efficacy in distinguishing between real and artificial intelligence-generated reports was determined by computing the F1-score, which is a harmonic mean of precision and recall. This value came out to be 93.5%.

Additionally, the detection system's accuracy was found to be 94%, meaning that 94% of all reports were correctly categorized as either authentic or artificial intelligence (AI) created. This high degree of accuracy shows how stable and dependable the suggested framework is at identifying fake material among a wide variety of text reports.
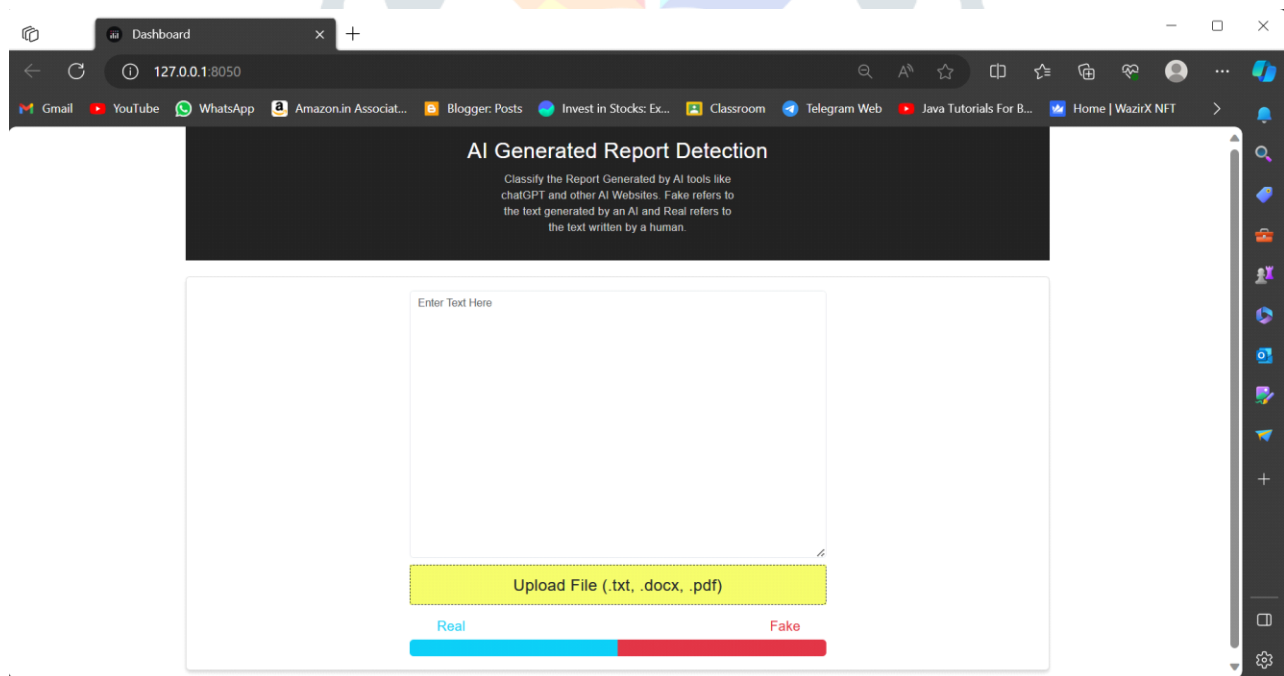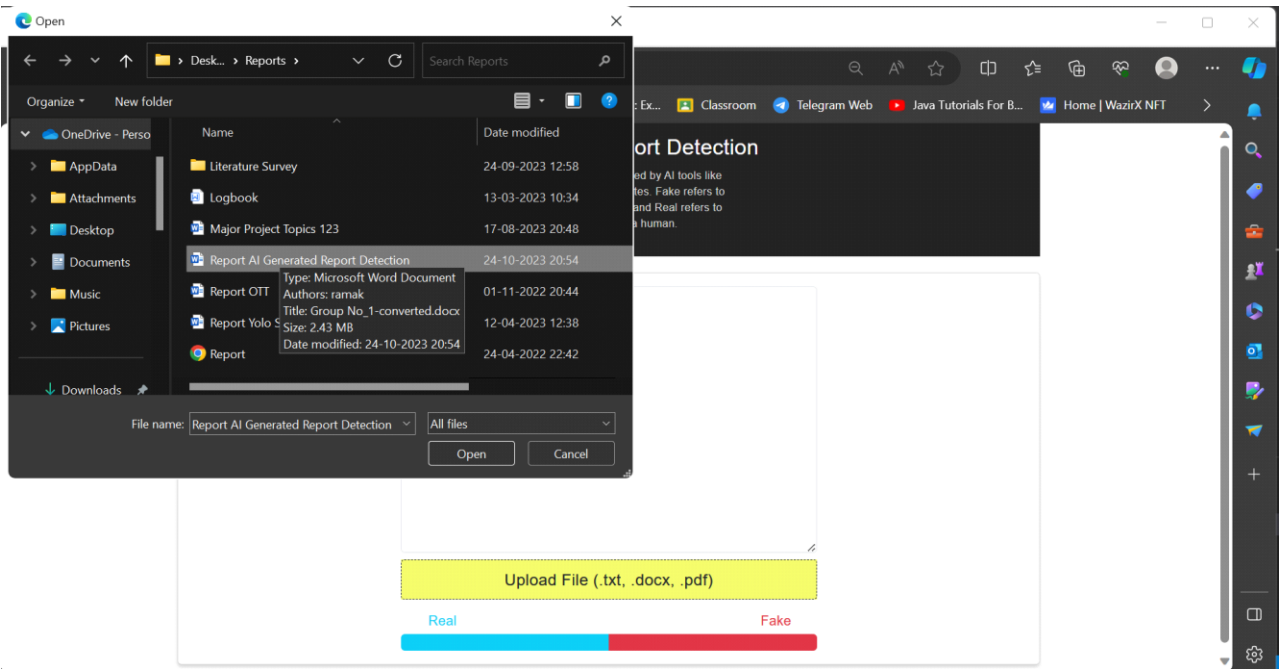

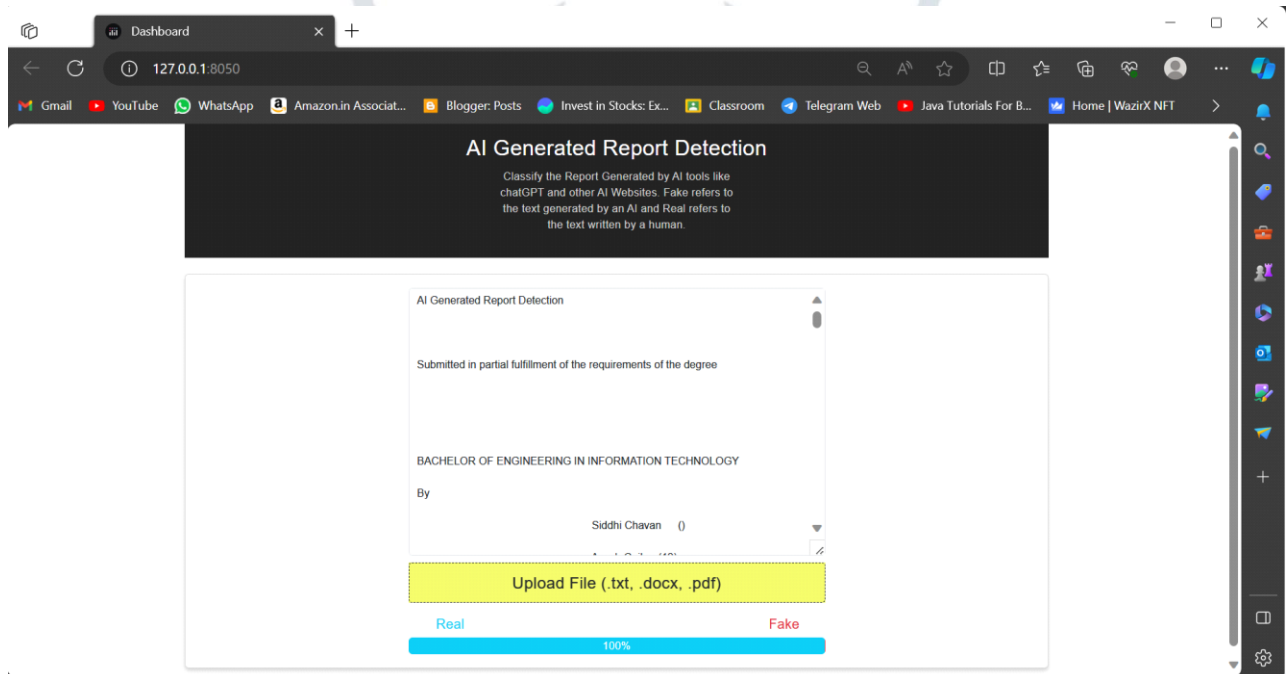Fig 4.1 Home Page

Fig 4.2 Select File
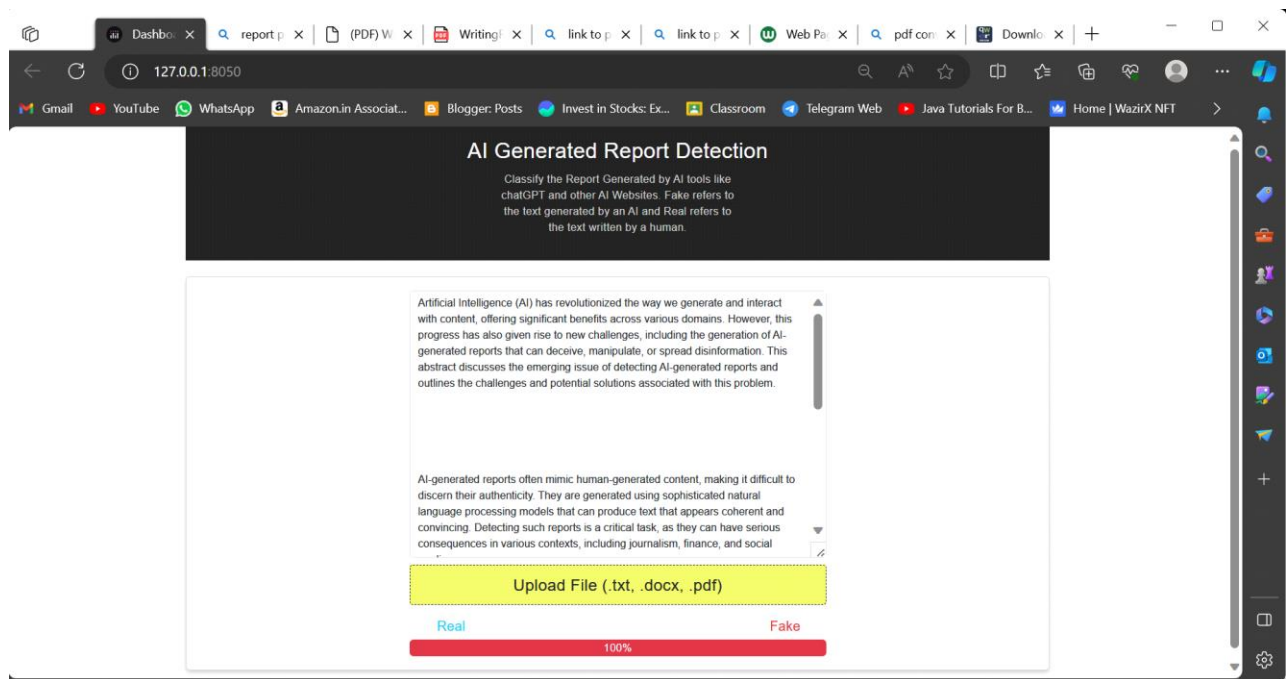

Fig 4.3 Generated by Human (Real)

Fig 4.4 Generated by AI (Fake)

CONCLUSION:

To meet the urgent requirement to guarantee the legitimacy and dependability of digital content, our research concludes with the presentation of an extensive framework for identifying reports generated by artificial intelligence. Our method shows promising accuracy in detecting fake reports with minimal false positives by utilizing a wide dataset and incorporating cutting-edge machine learning techniques. This framework has great potential to counter the spread of false information and protect the integrity of digital reporting platforms with further development and implementation. To improve the efficacy and uptake of AI-generated report detection systems going ahead and create a more open and reliable information ecosystem for all parties involved, researchers, industry players, and legislators must work together.

ACKNOWLEDGEMENTS :

REFERENCES:

1.Compilatio (2023). Comparison of the best AI detectors in 2023 (ChatGPT, YouChat...). https://www.compilatio.net/en/blog/best-ai-detectors. Accessed 12 April 2023

2. Content at Scale (2023). How accurate is this for AI detection purposes? https://contentatscale.ai/ai-content-detector/. Accessed 8 May 2023

3. Anderson N, Belavy DL, Perle SM, Hendricks S, Hespanhol L, Verhagen E, Memon AR (2023) AI did not write this manu-script, or did it? Can we trick the AI text detector into generated texts? The potential future of ChatGPT and AI in Sports & Exercise Medicine manuscript generation. BMJ Open Sport Exerc Med 9:e001568. https://doi.org/10.1136/bmjsem-2023-001568

4. Elkhatat AM, Elsaid K, Almeer S (2023) Evaluating the efcacy of AI content detection tools in diferentiating between human and AI-generated text. Int J Educ Integrity 19:17. https://doi.org/10.1007/s40979-023-00140-5

5Krishna K, Song Y, Karpinska M, Wieting J, Iyyer M. (2023). Paraphrasing evades detectors of AI-generated text, but retrieval is an efective defense. arXiv. https://doi.org/10.48550/arXiv.2303.13408

6. Turnitin (2023). Understanding false positives within our AI writing detection capabilities. https://www.turnitin.com/blog/understanding-false-positives-within-our-ai-writing-detection-capabilities. Accessed 14 June 2023

7. Crossplag.com (2023). How accurate is the AI Detector? https://crossplag.com/ai-content-detector/. Accessed 8 May2023