

MediTrust: Intelligent AI Framework for Detecting Fraudulent Medical Fund Requests Using Document Understanding and Data Verification

Preethi.A¹

Rayapati Avinash², Kodidala Jashwanth², Kottapalli Naresh Kumar², Anantha Sai Kumar²

ABSTRACT

The rapid expansion of digital healthcare fundraising platforms has significantly improved access to financial assistance for patients facing critical medical conditions. Individuals can now share their medical needs online and receive support from a wide network of donors. However, this convenience has also introduced a serious challenge: the increasing prevalence of fraudulent medical fund requests.

In many instances, users upload manipulated or fabricated medical documents, including prescriptions, diagnostic reports, and hospital bills, to falsely claim financial assistance. Such activities not only exploit donor trust but also divert resources away from genuinely needy patients.

To address this issue, this paper proposes *MediTrust*, an intelligent AI-driven framework designed to detect fraudulent medical fund requests through advanced document understanding and data verification techniques. The system integrates Optical Character Recognition (OCR), Natural Language Processing

(NLP), and anomaly detection models to analyze and validate uploaded documents.

=The proposed approach enhances transparency and reliability in medical crowdfunding by automating the verification process and minimizing human error. Experimental observations indicate that MediTrust can effectively distinguish between genuine and suspicious cases, thereby improving trust among donors and ensuring fair distribution of funds.

KEY WORDS:

Medical Crowd funding, Fraud Detection, Document Understanding, AI Verification, CRAFT, Donut Model, Fuzzy Matching

I. INTRODUCTION

The evolution of digital platforms has transformed various aspects of healthcare, including the way individuals seek financial assistance for medical treatments. Medical crowdfunding platforms have emerged as a vital support system, enabling patients to raise funds quickly by reaching a large audience through the internet. This model has proven particularly beneficial in regions where healthcare costs are high and insurance coverage is limited.

Despite its advantages, the open nature of these platforms makes them vulnerable to misuse.

Fraudulent campaigns supported by forged or manipulated medical documents have become increasingly common. These fraudulent activities are often difficult to detect because they mimic genuine cases with convincing details.

Traditional verification methods rely heavily on manual review processes, where administrators examine submitted documents for authenticity. However, this approach is time-consuming and prone to human error, especially when dealing with large volumes of requests.

To overcome these limitations, there is a growing need for an intelligent system that can automate the verification process while maintaining high accuracy. MediTrust is developed with this objective, leveraging artificial intelligence to analyze medical documents in a structured and scalable manner.

II. RELATED WORK

Despite the increasing popularity of online medical fundraising platforms, there is no robust mechanism to verify the authenticity of uploaded medical documents in real time. Fraudsters often exploit this weakness by submitting manipulated prescriptions, altered billing details, or entirely fabricated reports.

Existing verification methods are either manual, time-consuming, or limited to basic checks that fail to capture deeper inconsistencies. This creates a significant risk for donors, who may unknowingly contribute to fraudulent campaigns.

Therefore, there is a critical need for an automated, scalable, and reliable solution that can accurately analyze medical documents and detect fraudulent patterns without compromising efficiency.

III. PROPOSED SYSTEM

The increasing reliance on online medical fundraising platforms has exposed a critical gap in document verification. Most platforms lack a robust and automated mechanism to validate the authenticity of uploaded records. Fraudsters

exploit this weakness by creating fake prescriptions, altering hospital bills, or reusing documents from unrelated cases. One of the major challenges lies in the complexity and variability of medical documents, as different hospitals follow different formats. Existing solutions are limited to basic checks such as document completeness or format validation, which are insufficient to detect deeper inconsistencies. Therefore, the need arises for an intelligent, automated, and scalable system capable of accurately classifying requests as genuine or fraudulent.

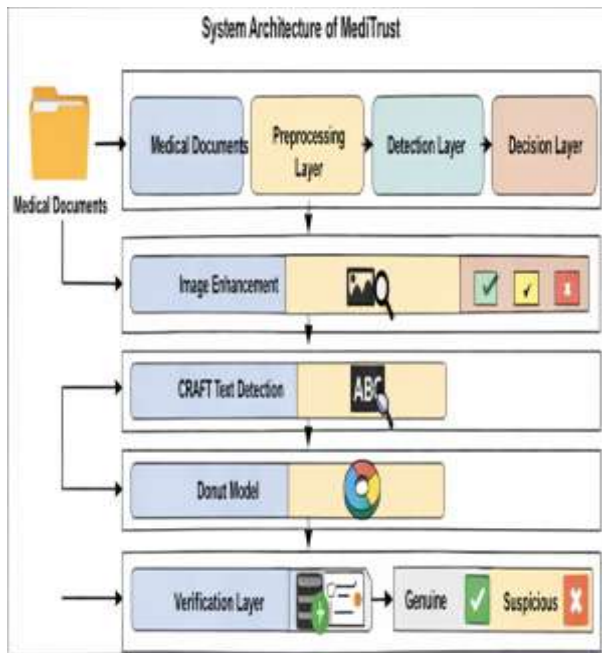
IV. SYSTEM ARCHITECTURE

The architecture of MediTrust is structured into multiple interconnected layers, each responsible for a specific function within the verification pipeline. This modular design ensures flexibility, scalability, and ease of integration with existing platforms.

The first layer, known as the input layer, handles the ingestion of user-submitted documents. It supports multiple file formats and ensures that all inputs are preprocessed before analysis. Preprocessing steps include noise reduction, image enhancement, and format normalization to improve the accuracy of text extraction.

The second layer, the processing layer, forms the core of the system. It includes OCR modules for text extraction and NLP components for entity recognition and semantic analysis. This layer transforms raw data into meaningful information that can be used for verification.

The third layer, the intelligence layer, incorporates machine learning models and validation logic. It performs both rule-based checks and anomaly detection to evaluate the authenticity of the data. This layer is critical in identifying complex fraud patterns that may not be evident through simple analysis.



The final layer, the decision layer, consolidates the results from previous stages and generates the final classification output. It also provides explanations for auditing and improving system transparency.

This layered architecture not only improves processing efficiency but also allows the system to scale as the number of users increases, making it suitable for real-world deployment

V. METHODOLOGY

The MediTrust framework is designed as a multi-stage pipeline that systematically processes and verifies medical documents. Each stage plays a critical role in ensuring the accuracy and reliability of the final decision.

The process begins with the document acquisition stage, where users upload various medical records such as prescriptions, diagnostic reports, and billing statements. These documents may be in different formats, including images or scanned PDFs. To handle this diversity, the system employs optical character recognition techniques to extract textual information and convert it into a structured format suitable for analysis.

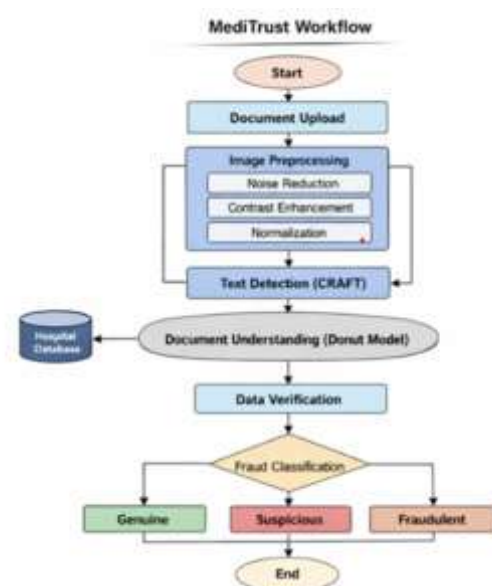
Once the text is extracted, the system applies natural language processing methods to identify and categorize key entities. These include patient

details, hospital names, doctor credentials, diagnosis descriptions, and financial information. Entity recognition helps in organizing the data and preparing it for further validation.

The next stage involves data verification, where the extracted information is evaluated using a combination of rule-based and data-driven approaches. Rule-based validation checks for logical consistency, such as matching patient names across documents, verifying date sequences, and ensuring that cost estimates align with standard medical practices.

In parallel, anomaly detection models are used to identify patterns that deviate from normal behavior. These models are trained on datasets containing both genuine and fraudulent cases, enabling them to detect subtle irregularities that may not be captured by predefined rules. For example, unusually high billing amounts or inconsistent diagnosis models treatment relationships flagged as suspicious.

Finally, the system aggregates the results from all validation stages and assigns a credibility score to each request. Based on this score, the request is classified into categories such as genuine, suspicious, or fraudulent. This classification assists platform administrators in making informed decisions quickly and efficiently.



VI. MODEL IMPLEMENTATION

The implementation of MediTrust relies on a combination of advanced models, each contributing to a specific aspect of the system. The CRAFT model is used for text detection, where it identifies individual character regions and links them to form meaningful text segments. This approach allows the system to handle complex layouts commonly found in medical documents.

The Donut transformer model is responsible for document understanding and information extraction. Unlike traditional OCR-based methods, Donut directly processes document images and generates structured outputs, capturing both the content and context of the document. This capability significantly enhances the system's ability to interpret complex data.

In addition, fuzzy matching algorithms are used to validate the extracted information. These algorithms measure the similarity between strings, allowing the system to detect inconsistencies even when there are minor variations or typographical errors. Together, these models form a robust implementation that enables accurate and efficient fraud detection.

VII. RESULTS AND DISCUSSION

The evaluation of MediTrust demonstrates its effectiveness in identifying fraudulent medical fund requests with a high degree of accuracy. By automating the verification process, the system significantly reduces the time required to analyze each request compared to manual methods.

One of the key findings is that the integration of rule-based validation with machine learning techniques provides a balanced approach to fraud detection. Rule-based methods ensure that obvious inconsistencies are quickly identified, while anomaly detection models capture more complex patterns that may otherwise go unnoticed.

The system also improves operational efficiency by filtering out clearly fraudulent cases, allowing human reviewers to focus on borderline or ambiguous scenarios. This not only reduces workload but also enhances the overall quality of decision-making.

However, certain challenges remain, such as handling highly sophisticated fraud attempts and adapting to new document formats. Continuous training of the models and integration of updated medical datasets can further improve system performance over time.

VIII. ADVANTAGES

MediTrust offers several advantages that make it a valuable solution for fraud detection in medical crowdfunding platforms. The system is fully automated, reducing the need for manual intervention and minimizing human error. Its high accuracy ensures reliable detection of fraudulent requests, while its scalability allows it to handle large volumes of data efficiently. Additionally, the system provides faster processing times, enabling timely verification of fund requests and improving overall platform efficiency.

IX. APPLICATIONS

The MediTrust framework can be applied across various domains within the healthcare ecosystem. It is particularly useful for medical crowdfunding platforms, where it can help verify the authenticity of fund requests. The system can also be integrated into insurance claim verification processes to detect fraudulent claims. In hospital billing systems, MediTrust can assist in auditing and ensuring the accuracy of financial records. Furthermore, it can be used in digital healthcare systems to enhance transparency and accountability.

X. CONCLUSION

MediTrust offers a comprehensive and intelligent solution to the growing problem of fraud in medical crowdfunding platforms. By combining

advanced AI techniques with structured document verification processes, the system ensures that financial assistance is directed toward genuine patients.

The proposed framework enhances transparency, builds donor trust, and reduces the risk of misuse. Its scalable architecture makes it suitable for large-scale deployment, while its modular design allows for future enhancements.

In the future, the system can be extended by integrating real-time hospital databases, blockchain-based document verification, and more advanced deep learning models. These improvements can further strengthen the reliability and accuracy of the framework, making MediTrust a robust solution for secure and trustworthy medical fundraising.

XI. FUTURE WORK

Future enhancements to the MediTrust system may include the integration of blockchain technology to ensure secure and tamper-proof verification processes. Real-time connectivity with hospital databases through APIs can further improve the accuracy of data validation. Additionally, incorporating advanced anomaly detection techniques using deep learning could enhance the system's ability to identify complex fraud patterns. The development of mobile-based verification solutions may also expand the accessibility and usability of the system in real-world scenarios.

XII. REFERENCES

1. S. Hussain et al., "Blockchain-Based Trust Management in Crowdfunding," IEEE Access, 2026.
2. S. Lee et al., "Fraud Detection Using Feature Selection," IEEE Access, 2025.
3. O. K. Alia et al., "IHSAN Crowdfunding Platform," IEEE Access, 2024.
4. K. O. Asamoah et al., "Blockchain-Based Loan Platform," IEEE Access, 2023.
5. I. Matloob et al., "Sequence Mining for Fraud Detection," IEEE Access, 2022.
6. IEEE, "Standards for Artificial Intelligence and Data Verification Systems," IEEE Standards Association, 2022.
7. R. Smith, "An Overview of the Tesseract OCR Engine," *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, 2007.
8. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *NAACL-HLT*, 2019.
9. M. Schuster and K. Nakajima, "Japanese and Korean Voice Search," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012.
10. Z. Akata et al., "CRAFT: Character Region Awareness for Text Detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.