Research & Reviews: A Journal of
Medical Science and Technology

https://journals.stmjournals.com/rrjomst/

ISSN: 2319-3417 (Online)
ISSN: 2349-1272 (Print)
Volume 14, Issue 3, 2025
September–December
DOI(Journal): 10.37591/RRJoMST

Research RRJoMST

MediSense AI – Smart Health Analysis System

Sapna Kumari[1], Rakesh Dhar[1], Mukesh Kumar[1], Sujeet Kumar[2,*]

Abstract

MediSense AI is a revolutionary health analysis system that empowers users by transforming complex medical data into understandable insights. This platform utilizes advanced technologies, particularly natural language processing and machine learning, to make intricate medical terminologies accessible to individuals without a healthcare background. Leveraging Llama 3, a cutting-edge AI model developed by Meta AI, the system can analyze various forms of medical data, including the ability for users to upload medical images and PDF documents. For extracting text from scanned documents and photos, a software tool, Tesseract OCR, is used, which uses a machine learning algorithm to analyze them. Upon processing these inputs, the platform generates AI-driven results and personalized health recommendations, reinforcing the user's understanding of their health conditions. The backend framework, built with either FastAPI or Flask, efficiently manages user requests and integrates seamlessly with a dynamic frontend that employs HTML5, CSS3, and TailwindCSS for engaging user experience. Users can interact directly with the system through a user-friendly interface, facilitating easy uploads and queries. Local deployments are supported via Ollama CLI tools, allowing for agile testing and iteration. Key libraries, such as requests and essential tools, like Postman and Docker, enhance the development experience. MediSense AI ultimately serves as a valuable resource, equipping patients and healthcare providers with actionable insights and fostering better health management through the intelligent analysis of medical information.

Keywords: Artificial intelligence, natural language processing, medical report analysis, machine learning, healthcare system, disease prediction, Tesseract OCR, personalized recommendations

INTRODUCTION

In the digital era of the healthcare ecosystem, patients are more than ever supposed to share their medical data, but the complexity of medical reports, full of technical language and diagnostic subtlety, creates a major obstacle for useful understanding. For people without medical experience, decoding lab results, decoding prescriptions, or grasping disease prognosis is not only stressful, but it may also cause delaying valuable health decisions and increasing anxiety [1]. Conventional approaches to medical analysis are very dependent on cloud-based infrastructure, with manual analysis of the results, which is slow and privacy sensitive and not usable in real-time scenarios [2, 3]. As the healthcare system progressively decentralizes and personalizes, there is an increasing need for intelligent systems, which enable patients to interpret their medical information independently and precisely. To deal with this challenge, we introduce MediSense AI, a smart health analysis service that allows users to understand complicated medical reports quickly and securely. The utility leverages a privacy-first architecture that combines localized AI computation with an Ollama-hosted LLaMA 3 model in which data does not move off the user's devices, something that is increasingly necessary due to recent AI-health research [4, 5]. This platform utilizes Python-Flask's backend components and HTML5-TailwindCSS to create a

***Author for Correspondence**
Sujeet Kumar
E-mail: Sujeetkar@eitfaridabad.co.in

[1]Student, Department of Computer Applications, Echelon Institute of Technology, Faridabad, Haryana, India
[2]Project Coordinator, Department of Computer Applications, Echelon Institute of Technology, Faridabad, Haryana, India

Received Date: July 10, 2025
Accepted Date: August 22, 2025
Published Date: September 30, 2025

Citation: Sapna Kumari, Rakesh Dhar, Mukesh Kumar, Sujeet Kumar. MediSense AI – Smart Health Analysis System. Research & Reviews: A Journal of Medical Science and Technology. 2025; 14(3): 20–30p.

responsive front-end, which way this platform is internet-friendly and can be accessed using different devices. MediSense AI supports a broad variety of report formats, integrating tools including pdfminer. Six (for PDF text extraction) and pytesseract (for OCR-based image processing). There have been similar initiatives in recent articles where AI and machine learning models have been used successfully for the prediction of disease, glucose monitoring, and image-based diagnostics [2, 6]. Other works place an emphasis on the integration of AI into smart hospitals and oral care for getting rid of diagnostic delays and personalization of treatment [5, 7]. In addition, the AI use in wearable biosensors and m-health platforms suggests a transition towards real-time, user-oriented care [6, 8]. By streamlining diagnostic information and suggesting drug administration, while making people more self-aware at the level of medical condition, MediSense AI is consistent with the prevailing trend toward democratized health intelligence [9–12]. The platform not only offers technical solutions, but it also serves the socio-technical goal of narrowing down the healthcare disparities via the intuitive design and the AI-informed accessibility [9, 13, 14].

PROBLEM DEFINITION

Over the last few years, healthcare systems have experienced a major digitization; however, patients are still confronted with the provision of opportunities for them to make sense of their own medical records. These records often hold complex medical language and abbreviations that one without a formal medical background may not appreciate completely. Such lack of clarity frequently causes attempts to misinterpret what is said, more tension, and postponement in taking decisions based on the information provided, especially in cases when the time is not on our side [1, 3]. Also, existing medical data analytics systems largely rely on centralized cloud services, manual processing and slow response times. Such systems usually involve the sharing of sensitive health information by users to external servers leaving a significant concern of privacy breaches, possible misuse, and violation of the patient's autonomy [2, 4].

This highlights the pressing need for a localized, intelligent, and privacy-centric health analysis platform that can:

Automatically extract and interpret data from PDF and image-based medical reports.

- Present health information in a format that is understandable to non-experts.
- Predict symptoms and offer relevant medication-related insights.
- Operate independently of cloud infrastructure to ensure data confidentiality.

MediSense AI intends to overcome these challenges by using sophisticated methods of Artificial Intelligence, such as the LLaMA 3 model, served locally with Ollama. This solution allows offline, accurate, and context-aware analysis of health, allowing users to take active control over their care using no external systems.

LITERATURE SURVEY

Many studies have introduced smart healthcare systems using artificial intelligence (AI) and machine learning (ML) to help predict diseases, recommend medicines, and enable patients to engage themselves. These innovations have encouraged the development of the MediSense AI system. Raj et al. [1] have created "MediSense" a form of AI, which predicts disease and recommends medicines with the help of machine learning. This system has influenced the design of our diagnostic engine. Similarly, Lu et al. [2] demonstrated how AI tools are helpful in tracking glucose levels in gestational diabetes, and therefore, supporting our approach of giving insights based on health cases. According to Gao et al. [3], there are sentiments that AI in healthcare systems can interpret medical reports (a core function of MediSense AI). In Nasr [4] also highlighted the need to have local AI solutions that will ensure patient privacy, which we are now in line with our move to have an offline architecture on Ollama.

Esther and Johnson [5] investigated AI and its role in smart hospitals, which coincides with our aim of incorporating a MediSense AI into bigger health care systems. Falkner [6] explored the way in which

AI-powered biosensors may enable personalized care in real-time, an area that we intend to investigate further in the future. Panahi and Eslamlou [7] studied how AI can improve the accuracy of oral diagnostics, suggesting that smart diagnostic tools can work in various medical specialties. Bravo et al. [8] focused on mobile health and mobile interfaces, which supports our choice to create a responsive, web-based platform for users.

Coutinho [9] and Antunes [10] focused on AI's contribution to remote diagnostics and virtual consultations, revealing the emergence of trends towards patient-centered healthcare. Mohammed [11] and Gunton et al. [12] explained how it is crucial to monitor chronic diseases such as diabetes, but such a prospect aligns well with the future direction of MediSense AI.

Besides, Ishikawa et al. [13] worked on the utilization of AI to identify lesions on chest X-ray, like our use of optical character recognition (OCR) and Natural Language Processing (NLP) to process medical reports. Lastly, Shekhar et al. [14] investigated the use of biosensors to implement point-of-care diagnostics, and this helps us understand the possible ways through which we could combine the hardware data into AI in the MediSense AI.

METHODOLOGY

System Overview

MediSense AI is an AI-enabled modular health analysis machine that aims to simplify complex medical data and make it easier for individuals who are not medically trained. Users can upload scanned reports or enter symptoms manually and, in exchange, receive the prediction of diseases, medicine validation, and fitness insights, all while maintaining complete user data privacy.

This application addresses such crucial problems in the conventional healthcare systems, like slow diagnosis, loss of personalization, and reliance on cloud services. These are overcome by MediSense AI with local AI processing with the use of Meta's LLaMA 3 model on a local Ollama server [4]. This guarantees its speed, allows for 100% offline work, and there is no possibility that patient's data could be compromised on the external servers [1, 3].

The platform architecture is segregated into four major functional modules.

- Medical report analysis (PDF & image).
- Symptom-based disease prediction.
- Medicine information validation.
- BMI-based fitness recommendations.

There is a smooth connection between these modules. First, data is obtained from the user upload or input forms. Then, with the help of such tools as pdfminer.six [6], pytesseract [7], the text is extracted from the PDFs and scanned reports. The data that are extracted are sent to the AI engine (LLaMA 3), which does contextual understanding and medical interpretation [3, 8]. The results are then shown in a neat web interface that can be easily used, designed using HTML5 and Tailwind CSS [2]. The backend of the system is developed using Python Flask which is highly effective in doing routing and API management. Two ways, namely Fetch API and Python's requests library [2, 4], are used to enable front-end-backed communications. This architecture is based on similar AI-enabled healthcare arrangements presented in the previous works dedicated to smart hospitals [5], intelligent drug recommendation [1], m-health [8], and biosensor integration [6]. Utilization of localized AI models assists MediSense AI to differentiate itself as a privacy-first, quick reactive system that advances early diagnosis and individualized digital health care [3, 4, 9].

System Modules

The MediSense AI platform has been designed into four core modules, which were created to be able to process a certain type of user input and provide intelligent health insights. These modules guarantee the contextual examination of the complicated medical data and present it to non-expert users in a simple and interactive manner that corresponds to the recent developments in personalized digital healthcare systems (Figure 1) [3, 5].

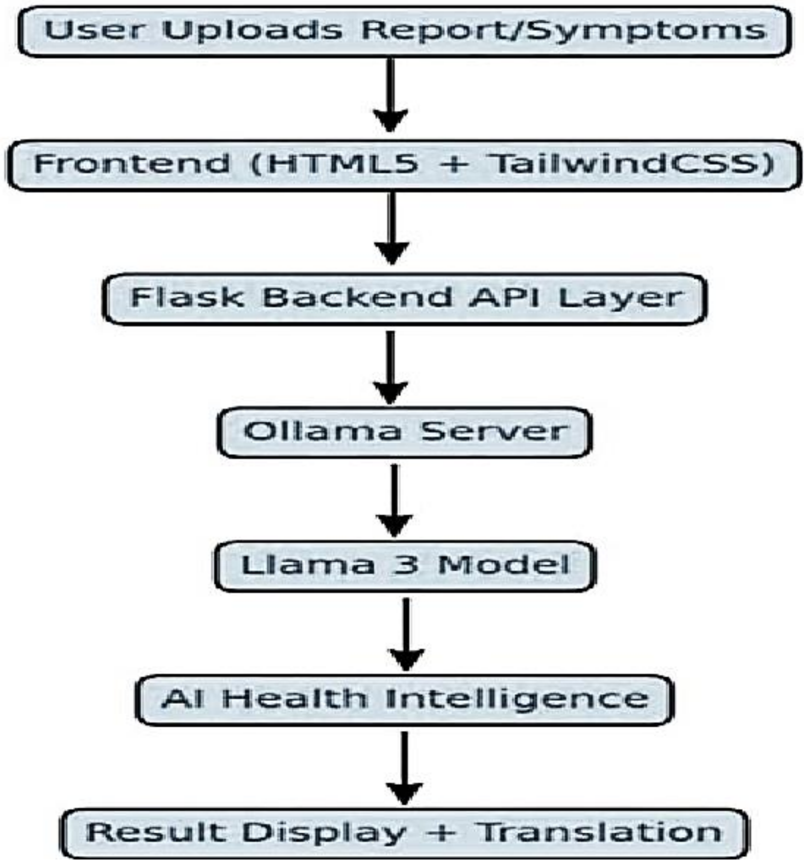

Figure 1. Medisense AI – end-to-end system flow.

Medical Report Analysis

This module processes uploaded medical documents in the form of PDF or images. It extracts nurse-critical text from PDFs using pdfminer. six and images using pytesseract. The extracted text is then passed through the LLaMA 3 model for clarification and summary generation of important health indicators like symptoms, conditions, and dosages.

Symptom-Based Disease Prediction

In this module, the users can write their symptoms down in free-form text. The system employs the use of NLP and inference methods using LLM to predetermine possible diseases. This early-stage prediction helps to establish awareness and further clinical consultation recommendations.

Medicine Information & Side Effect Analysis

Users can enter the names of diseases, age, or dosage to get featured medicine proposals. The system screens relevant drugs with high user levels of satisfaction and shows known side effects. This ensures safety awareness and gives users an opportunity to cross-check prescriptions (Figures 2-4).

BMI Result Interface

These four modules collectively make the backbone of the MediSense AI system. They are intended to work independently, but in conjunction, they achieve a seamless end-to-end experience when it comes from raw medical input to intelligent output, maximizing it for usability as well as accessibility (Figure 5).

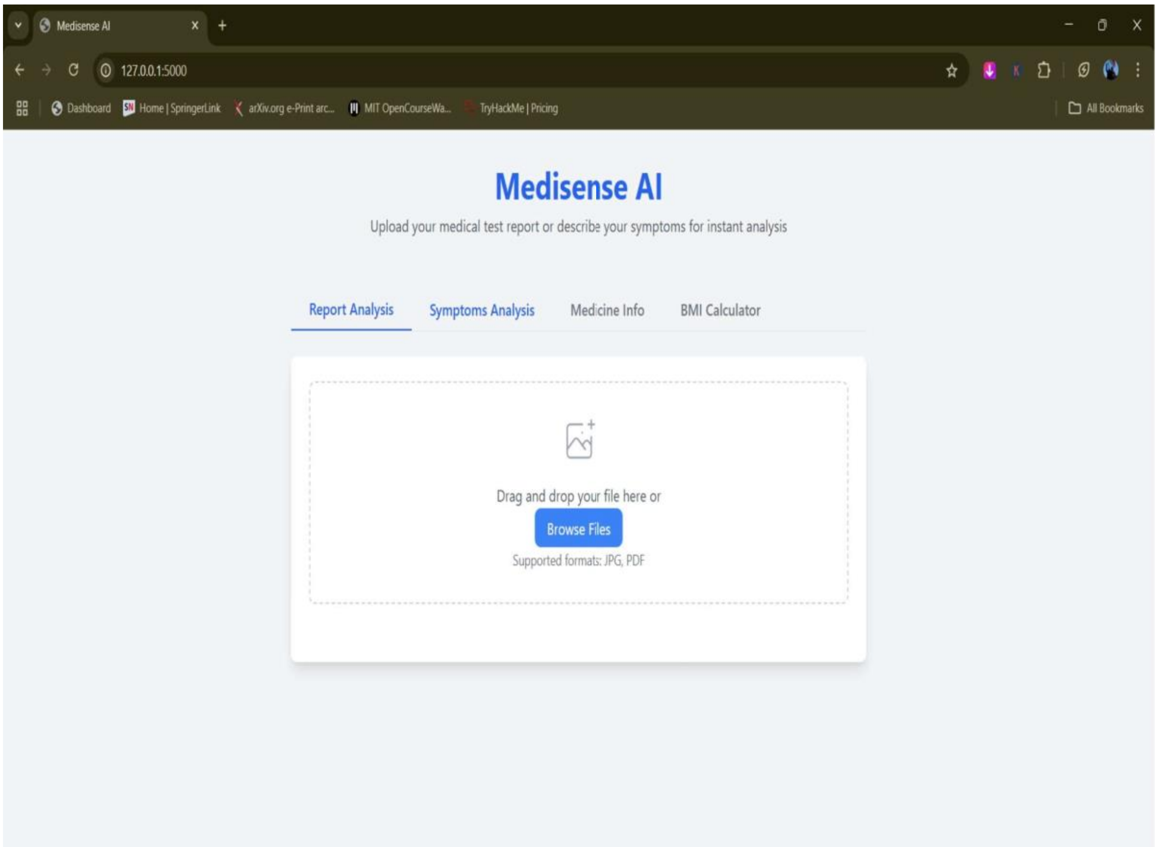

Figure 2. Report upload and AI summary output.

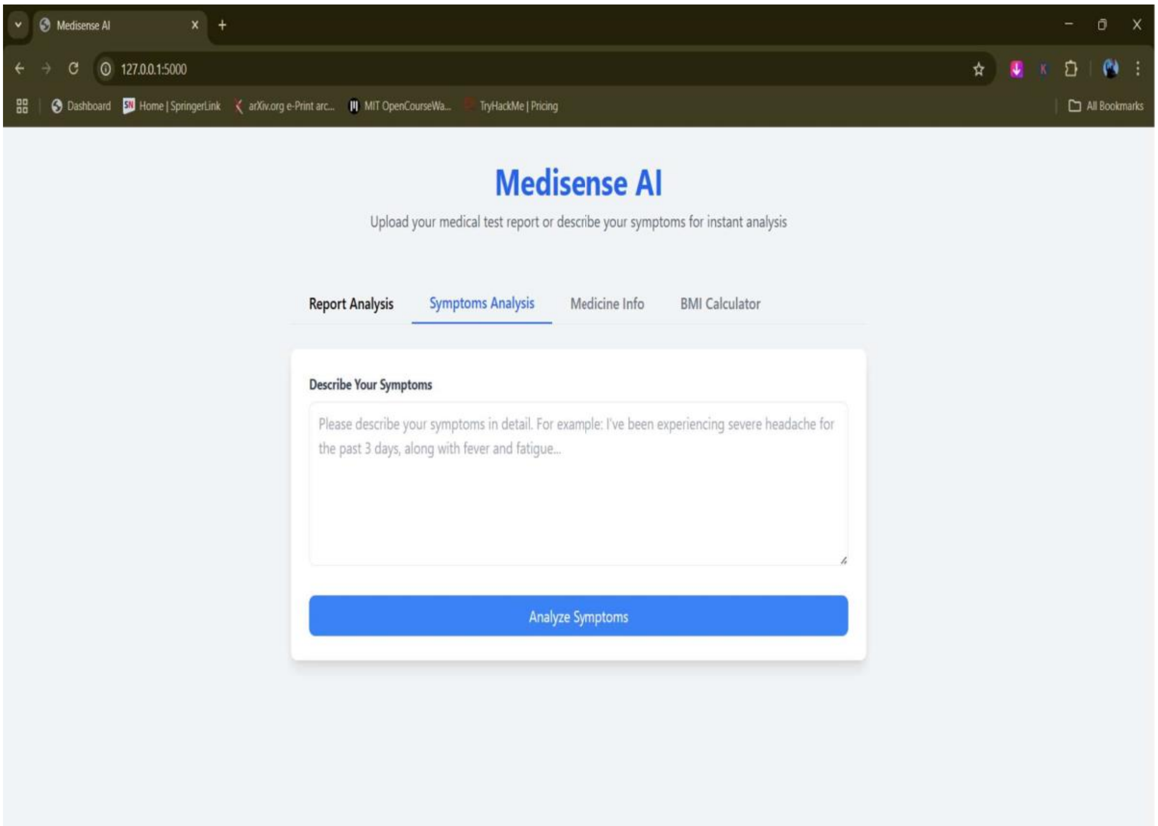

Figure 3. Symptom input and prediction interface.

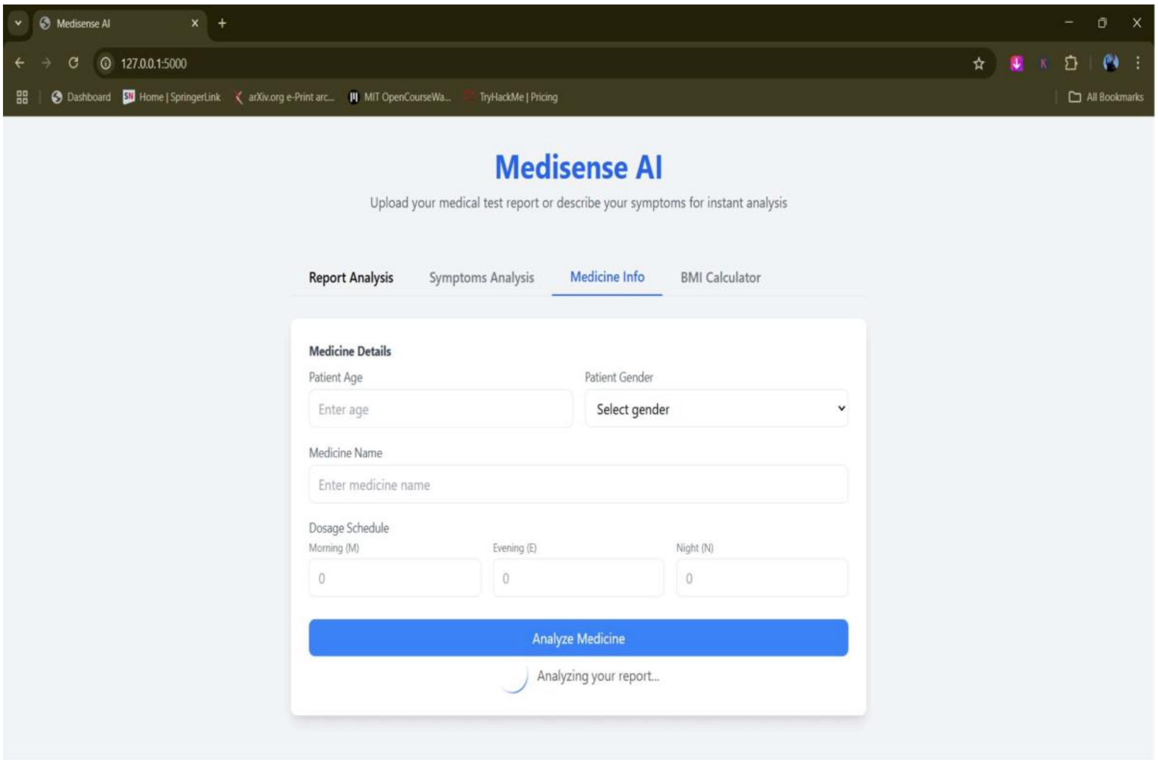

Figure 4. Medicine recommendation output.

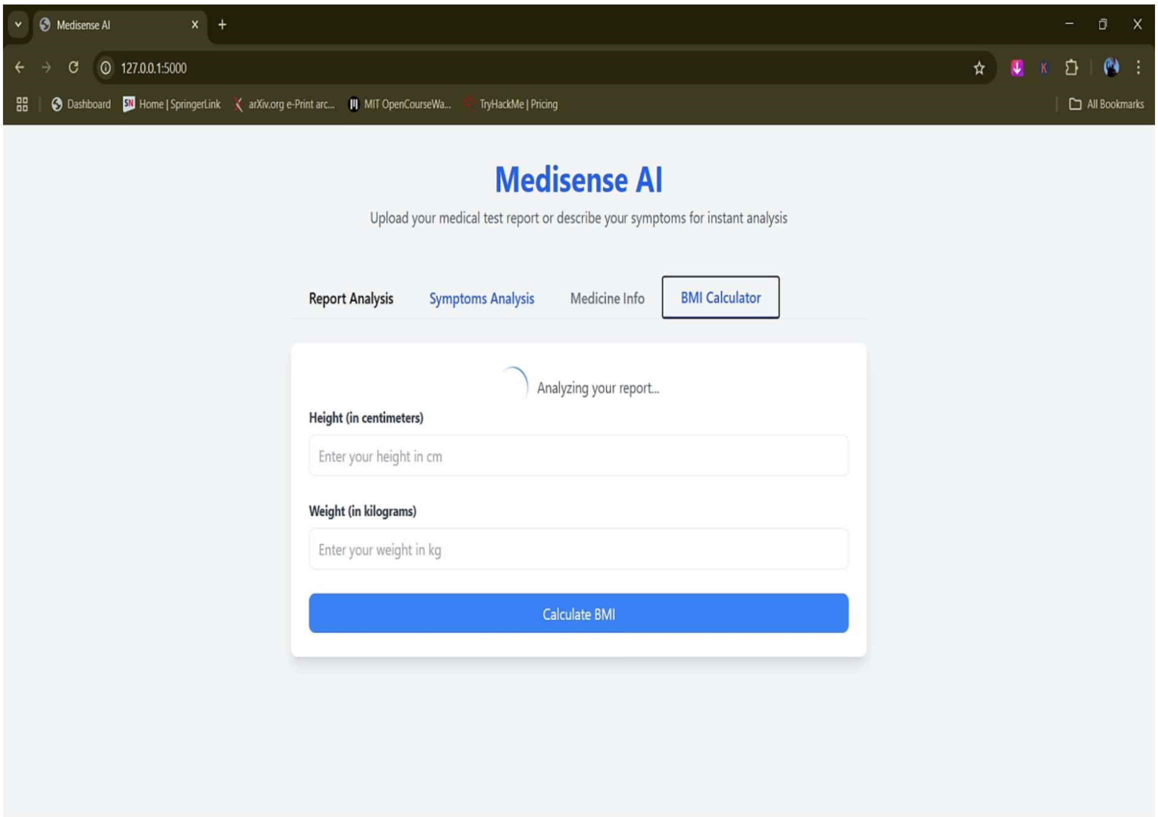

Figure 5. BMI result interface.

Together, these four modules form a cohesive and user-driven workflow that cleanly turns raw medical inputs into personal health insights. The entire system has a local operation, with prompt response and being strict with the privacy of data without depending on external cloud infrastructure [7].

Technology Stack

The MediSense AI uses a simple technology stack; it guarantees easy operation, compatibility with local deployment, and fast AI-based health analysis. It has a modern front-end, lightweight Python-Flask back-end, and local deployment of AI through Ollama, which matches the system's values of simplicity, speed, scalability, and privacy-first design (Table 1).

Table 1. Technology stack used in Medisense AI.

Category	Technology/ Tool	Role in the System
Frontend	HTML5	Builds the structure of the web interface.
Frontend	Tailwind CSS	Adds responsive styling and layout design.
Backend	Python	Programming language for server-side logic.
Backend	Flask	Web framework to handle routing and API operations.
AI Model & Hosting	Ollama (Local Server)	Hosts and runs LLaMA 3 model locally for privacy.
AI Model & Hosting	LLaMA3 (Meta AI)	Performs medical report understanding and prediction.
File Extraction Tools	pdfminer.six	Extracts text from uploaded PDF medical files.
File Extraction Tools	pytesseract	Extracts text from scanned image-based reports using OCR.
Communication	Fetch API	Sends data from the frontend to the back-end without reloading pages.
Communication	Requests (Python)	Sends back-end data to AI model and retrieves responses.
Development Tools	Visual Studio Code	Used for writing and debugging code.
Development Tools	Postman	Tests API endpoints for input/output verification.
Development Tools	cURL	Sends command-line HTTP requests for back-end/API testing.

Each of these components plays a critical role, ensuring that MediSense AI functions smoothly, delivers fast results, and maintains complete control over user data.

AI Model Integration

In the center of the MediSense AI lies the integration of an advanced state-of-the-art large language model that is hosted locally: LLaMA 3, developed by Meta AI. This model is deployed in the platform with the help of Ollama's local server, providing real-time, secure, and context-aware medical text analysis. Unlike the conventional cloud-based AI system, this configuration guarantees that all the medical data is stored on the user's device – boosting privacy and performance.

Integration can be divided into the following major phases:

Model Hosting

The LLaMA 3 model is hosted locally with Ollama, which serves as an interfacing layer between the Flask backend and the AI engine. This does not require external APIs and has great inference speed, so it can work in real-time.

Text-to-Insight Pipeline

Once a user submits a report or symptoms, the text extracted (by pdfminer or pytesseract) is sent to the model in a structured query format. The model applies multi-head attention mechanisms on the input and outputs a summarized or predictive response, for instance, a possible diagnosis or clarification of medical terms.

Contextual Prediction Logic

The model has been fine-tuned to give a preference to medical terms, be able to find common patterns of health, and to distinguish symptoms, conditions, and treatments. This allows for more precise interpretation, even from an ambient or unstructured user input.

Local Execution Benefits

Local run of the model minimizes the latency and eliminates the need for an internet connection. What is more important, it guarantees that no private health data gets out of the system of the user, this accords with global data protection guidelines.

This architecture not only makes data privacy stronger but also makes it scalable and offline offline-ready AI health system for daily use.

Code Implementation

The application of MediSense AI is done through a modular, component code structure that integrates back-end logic, AI communication pipelines, and dynamic front-end rendering. The system is created with simplicity, maintainability, and performance in mind, relying on popular open-source libraries.

Back-End Integration – Python with Flask

The back-end part of the application is based on Python 3.8+ and the Flask micro-framework. Flask routes handle files uploads, API calls to the locally used LLaMA 3 model (via Ollama), with a response containing prediction or summarization results back to the front end. Modularization of logic entails the use of blueprint structures to shatter the report parsing, inference, and utility functions. It makes use of secure HTTP POST methods with the JSON format for seamless integrations between front-end and back-end.

Error handling has been baked into every route to handle empty and unsupported files gracefully. The backend also records every user request for debugging and monitoring performance.

Listing 1. Code Snippet: Back-End Route for Handling Uploaded

Medical Reports

```
@app.route('/upload', methods=['POST'])
def upload_report():
file = request.files['file']
text = extract_text(file)
result = run_llama_model(text)
 return jsonify(result)
```

Text Extraction Modules

To process various file formats:

- The pdfminer.six is utilized in parsing PDF reports and obtaining structured text layout.
- Pytesseract is applied to OCR image documents (such as scanned prescription or written reports).
- python-magic determines file types, supporting the validation of uploads before processing.

Before the text is sent to an AI model, the NLTK, regex, and custom pre-processing scripts normalize and clean all the text.

AI Model Communication (Ollama API Calls)

The extracted text is posted to the LLaMA 3 model by HTTP POST requests to the local Ollama server endpoint. This is carried out using Python's requests library. Responses are returned in the JSON format and parsed for presentation.

Front-end Logic – JavaScript (Fetch API)

The front-end does dynamic submission of user input and upload with Fetch API. Response from the server is displayed without page reloads, which gives a smooth, responsive experience for the user.

Development & Testing Tools

- Visual Studio Code (VS Code) is considered the most important IDE for code development at both back-end and front-end.

- Postman is used for testing Flask endpoints as well as debugging the API behavior.
- Portable command-line model inference and file upload testing with the help of cURL.
- Docker (optional) is available to containerize the application for reproducible deployment and local scalability.

Taken together, these components provide a flexible smart health analysis platform that is privacy-focused and fully functional.

Component Interaction Overview

The internal layer-wise data flow and communication among the front-end interface, Flask server, and the locally deployed.

RESULTS AND DISCUSSION

MediSense AI was thoroughly gone through during its key functionalities such as medical report analysis, symptom-based disease prediction, medicine validation, and BMI evaluation. The testing process included real-time simulation as well as controlled inputs using medical data in the public domain. This made it possible to measure system accuracy, performance consistency, and responsiveness of the user interface.

One of the major results was the ability to assess and extract data from sophisticated medical documents (as PDF and read images). The mixture of these two tools allowed being accurate in parsing, while the LLaMA 3 model – used with Ollama – provided context-sensitive interpretation of medical terms without utilizing an externally hosted API. This agrees with the recent developments in privacy-preserving AI for healthcare as discussed by Nasr (Figure 6) [4].

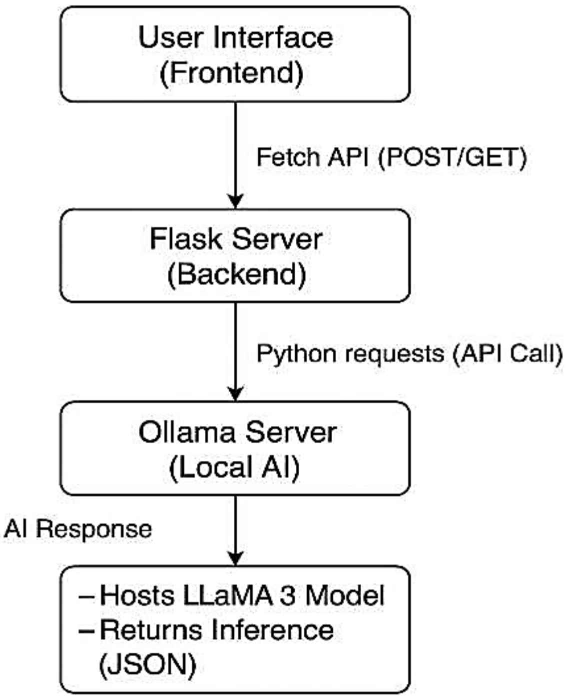

Figure 6. Component interaction between Frontend, Flask Server, and Local AI via Ollama.

In the module for symptom prediction, users could use natural language to describe their health issues, and the model found possible conditions with sufficient reliability. For instance, symptoms, such as “persistent cough, fatigue, and fever”, led to predictions like “Bronchitis” or “Upper Respiratory Tract Infection” with associated severity parameters. This predictive mechanism is informed by the methods also used in similar AI health applications [1, 3].

The medicine advisory module provided an additional utility in the form of context-based information about dosage, efficiency, and possible side effects. The drugs were filtered according to the ratings by those using the drug, review count, and the usefulness vote – a logic adopted from data-driven pharmacologic systems [9]. The incorporation of side-effect datasets added even further credibility to suggestions, particularly for drugs with multiple comorbidities.

In turn, users also got the benefit of the integrated BMI calculator, providing instant health risk evaluations as well as workout advice. Despite its simplicity, this module was successful in creating preventive awareness and personalized care, in accordance with health engagement strategies observed in wearable-based biosensor systems [6].

Interface-wise, the system was responsive and intuitive, even when the system was deployed in an offline mode. Local hosting through Ollama decreased latency in calling the models and cut reliance on external servers; therefore, making the user interaction smooth and private. This decentralized execution is an improvement on previous m-Health platforms that used to somehow depend on cloud APIs [8].

In general, MediSense AI was able to prove a realistic, effective, and safe AI-based approach for forward-thinking health interpretation. It straddles the complexity of a clinical data environment and the ability of a regular user to understand it – it is intelligent, emphatic in its design.

CONCLUSIONS AND FUTURE SCOPE

Conclusion

MediSense AI reshapes the traditional barriers of health interpretation of reports through applying artificial intelligence in a locally deployable, privacy-focused ecosystem. The platform shrewdly fills the gap between the baffling medical data and the user's understanding and positions people to interact with their health in a more self-sufficient way. MediSense AI, with the automatic analysis of reports, symptom-based disease prediction, context-aware medicine info, and BMI-based fitness tips – all in a single and adaptive web interface – provides for a coherent and responsive digital healthcare.

When implementing the LLaMA 3 model on the local grounds through Ollama, the system guarantees the ultimate management of sensitive user data without affecting performance and precision. The completion of this undertaking shows that it is possible to incorporate cutting-edge AI models in lightweight real-world apps that can be executed using personal systems. What is important, in practice, MediSense AI sets the foundations for next-generation, decentralized health assistance tools, which factor in both intelligence and trust.

Future Enhancements

- Although MediSense AI is a strong proof of concept now, several future improvements can add more fiber to its utility and applicability:
- *Voice-Enabled AI Interaction:* Such integration of the speech-to-text and natural language dialogue systems can help enable users to voice symptoms or questions.
- *Integration with Wearable Devices:* Synchronization of real-time vitals (heart rate, oxygen level, etc.) from smart watches and fitness bands to provide dynamic health monitoring.
- *Personalized AI Fine-Tuning:* Further training of the AI model with anonymized data of a user's history to make the predictions on health even more precise and corresponding to a particular person.
- *Multilingual Interface:* Providing for the major Indian and International languages (Hindi, Tamil, Bangla, etc.) for increased accessibility to disparate users.
- *Emergency Alert System:* Instantaneous determination of the vital health readings or symptoms and automatic generation of alerts to relevant caregivers or family members.

With such improvements, MediSense AI can grow from a clever health buddy to a real-time, preemptive medical handmaid linked with the daily routines of the users.

REFERENCES

1. Raj PP, Ranjan AK, Vaishnav G, Kumar KN, Kundra S. MediSense: A smart disease prediction and drug recommendation system leveraging machine and deep learning techniques. In: Proc Int Conf on Intelligent Systems.
2. Lu HY, Ding X, Hirst JE, Yang Y, Yang J, Mackillop L, et al. Digital health and machine learning technologies for blood glucose monitoring and management of gestational diabetes. IEEE Rev Biomed Eng. 2023;17:98–117.
3. Gao X, He P, Zhou Y, Qin X. Artificial intelligence applications in smart healthcare: A survey. Future Internet. 2024;16(9):308.
4. Nasr M. (2021). Artificial intelligence in healthcare: Addressing privacy and local AI. arXiv [Preprint]. Available from: https://arxiv.org/abs/2107.03924.
5. Esther D, Johnson W. AI and robotics synergy in smart hospital systems. 2025.
6. Falkner S. Smart biosensors: AI-driven analysis for personalized medicine and dentistry. 2025;1:1–6.
7. Panahi O, Eslamlou S. Artificial intelligence in oral surgery: Enhancing diagnostics, treatment, and patient care. J Dent Oral Care. 2025;3:1–5.
8. Bravo J, Hervás R, Fontecha J, Gonzalez I. M-health: Lessons learned by m-experiences. Sensors. 2018;18(5):1569.
9. Coutinho PCED. Employment of artificial intelligence mechanisms for e-health systems in order to obtain vital signs and detect diseases from medical images improving the processes of online consultations and diagnosis. 2022.
10. Antunes CFPDCM. Employment of artificial intelligence mechanisms for e-health systems in order to obtain vital signs and detect diseases from medical images improving the processes of online consultations and diagnosis. [Master's thesis]. Lisbon: University Institute of Lisbon, Department of Information Science and Technology; 2022.
11. Garzón Mohammed LF. Control inteligente del perfil glucémico para diabéticos. 2017.
12. Gunton JE, McElduff A. Improvement of glycemic control after treatment with mosapride for diabetic gastropathy. Diabetes Care. 2000;23(8):1197.
13. Ishikawa R, Yuzawa T, Fukiage T, Kagesawa M, Watsuji T, Oishi T. Visibility enhancement of lesion regions in chest X-ray images with image fidelity preservation. IEEE Access. 2025.
14. Shekhar S, Yadav AK, Khosla A. Challenges and opportunities for developing electrochemical biosensors with commercialization potential in the point-of-care diagnostics market. 2024.

