Minor Project Report on

Gen AI Based Agent for Telemedicine and Hospital

by

Reminkumar Dobariya &

Harsh Sangani

(22BCP155 & 22BCP141)

Under the Guidance of

Dr. Punit Gupta

Assistant Professor

Submitted to

Department of Computer Science and Engineering,

School of Technology,

Pandit Deendayal Energy University

November 2025

CERTIFICATE

This is to certify that the project report entitled **“Gen AI Based Agent for Telemedicine and Hospital”**, submitted by **Reminkumar Dobariya & Harsh Sangani**, has been conducted under the supervision of **Dr. Punit Gupta**, Assistant Professor, and is hereby approved for the partial fulfillment of the requirements for the award of the degree of Bachelor of Engineering in the Department of **Computer Science and Engineering** at Pandit Deendayal Energy University, Gandhinagar. This work is original and has not been submitted to any other institution for the award of any degree.

Sign:
Name of Guide: Dr. Punit Gupta
Designation: Assistant Professor
Department: Computer Science and Engineering
School of Technology
Pandit Deendayal Energy University

Sign:
Name of Guide:
Designation:
Department: Computer Science and Engineering
School of Technology
Pandit Deendayal Energy University

Acknowledgement

We would like to express our sincere gratitude to Dr. Punit Gupta, Assistant Professor, Department of Computer Science and Engineering, Pandit Deendayal Energy University, for his continuous guidance, encouragement, and insight throughout the duration of this project. His expertise in Artificial Intelligence and healthcare systems, as well as his timely feedback during weekly reviews, played a central role in shaping our ideas and refining our implementation. His support helped us understand both the technical and practical aspects of deploying AI in real-world healthcare environments.

We are grateful to our classmates and peers who contributed indirectly through discussions, suggestions, and collaborative learning. Their feedback helped us identify potential improvements in the system's design, user interface, and workflow. These interactions allowed us to evaluate our ideas from different perspectives and refine the solution accordingly.

We also acknowledge the contributions of the global open-source and research community. Several published studies, datasets, and technical resources played a crucial role in expanding our understanding of Generative AI, multimodal models, and telemedicine workflows. The insights derived from existing literature helped us build a strong theoretical foundation and implement our system in a technically sound manner.

Finally, we express our heartfelt gratitude to our families for their constant support, patience, and motivation throughout this multi-week development process. Their encouragement kept us focused and determined, even during the most challenging stages of integration, testing, and debugging. Without their support, completing this project to its present standard would not have been possible.

This project is the result of collective guidance, collaboration, and support from all the individuals and institutions mentioned above, and we are deeply thankful for their contributions.

Abstract

Telemedicine continues to expand as an essential component of modern healthcare, yet many platforms still depend heavily on manual workflows and limited automation. In this project, titled "**Gen AI Based Agent for Telemedicine and Hospital**" we developed a multimodal AI-driven telemedicine system that integrates conversational intelligence, dermatology image analysis, multilingual interactions, and audio capabilities (speech-to-text and text-to-speech). The system aims to enhance accessibility, streamline patient–doctor communication, and reduce clinician workload using responsibly designed AI components.

Our platform consists of a Patient Portal, enabling users to describe symptoms through text or voice, upload skin images, manage appointments, and receive AI-assisted guidance, and a Doctor Portal, where clinicians can view summarized conversations, examine dermatology insights, respond to cases, and issue prescriptions. The conversational assistant, powered by Gemini, handles multiple languages, maintains context, and provides safe, non-diagnostic responses. The dermatology model assists by identifying visual patterns in images and generating preliminary advisory insights.

The backend architecture uses Node.js and Express with PostgreSQL for structured data and MongoDB for chat logs. Real-time communication through Socket.IO ensures instant updates between patients and doctors. Testing across several weeks confirmed the system's stable performance in terms of response latency, image-processing time, API behavior, and chat reliability. The addition of multilingual and audio features further improved accessibility for patients with language or typing constraints.

Overall, this project demonstrates how Generative AI can meaningfully augment telemedicine when combined with thoughtful design, ethical safeguards, and user-centric workflows. The system offers a strong foundation for future expansion into advanced imaging, predictive analytics, and integration with hospital EHR systems.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1 Introduction

1.1 Background

Telemedicine has emerged as one of the most effective approaches to extending healthcare accessibility, particularly in rural and underserved regions. However, conventional telemedicine systems primarily depend on manual interactions and offer minimal intelligence. They often lack the ability to interpret patient symptoms, summarize medical conversations, or assist patients who struggle to describe their concerns effectively.

With advancements in Generative AI, especially Large Language Models (LLMs), there is an opportunity to transform telemedicine into a more interactive and supportive system. Generative AI can analyze patient queries, offer follow-up questions, generate summaries for doctors, and explain medical concepts in a simplified manner. The integration of computer vision enables preliminary image-based screening, particularly useful for dermatological concerns—one of the most common telemedicine consultation categories.

1.2 Problem Definition

During the initial requirements elicitation, we identified several critical limitations in existing telemedicine systems:

- Lack of automated triage and symptom interpretation
- Poor integration of dermatology image analysis
- Absence of AI-generated summaries for doctors
- Manual overhead in patient-doctor communication
- Limited support for structured medical data representation

These limitations lead to delays, inconsistent communication, and increased workload for medical professionals.

1.3 Objectives

The primary objectives of this project were,

- Develop an intelligent conversational agent powered by the Gemini LLM to handle natural, health-related queries and facilitate patient interaction.
- Integrate a specialized skin-disease prediction model capable of automatically analyzing dermatological conditions from uploaded images.
- Create a seamless multimodal interface that combines text-based consultation with visual diagnostic capabilities.
- Reduce response time and enhance healthcare delivery efficiency through AI-assisted triage and the generation of preliminary patient insights.

1.4 Motivation

The motivation behind this project arises from the growing demand for intelligent, accessible, and efficient digital healthcare solutions. As Large Language Models like Gemini demonstrate exceptional ability in understanding and generating natural language, they offer a powerful foundation for enhancing patient interaction and addressing health-related queries. Similarly, advancements in medical imaging and AI-driven prediction models create opportunities for more accurate and timely assessment of dermatological conditions. By integrating conversational intelligence, automated skin-disease analysis, and a multimodal interface, this project aims to streamline early-stage triage, reduce response times, and support clinicians with meaningful, AI-generated insights that improve overall healthcare delivery.

Chapter 2 Literature Survey

2.1 Introduction

In order to understand how our proposed system fits into the current technological landscape, I reviewed existing research related to Generative AI, telemedicine platforms, large language models, and dermatology image analysis. The study helped me identify what today's systems can already do and where the major gaps still exist.

Generative AI has recently become a major part of modern healthcare research because of its ability to understand natural language and provide meaningful responses. Models like GPT and Gemini can read user queries, extract symptoms, summarize long interactions, and even explain medical concepts in simpler terms.

Many papers highlight that GenAI can reduce the workload on doctors by helping with repetitive tasks—such as initial patient interaction, summarizing their concerns, or giving general advice. However, these studies also warn that AI outputs must be safe, transparent, and monitored by human professionals. AI should support clinicians, not replace them.

2.2 Telemedicine Systems and Their Limitations

Telemedicine platforms have seen rapid adoption, offering features such as video consultations, appointment booking, and digital prescriptions. Despite these advantages, the literature identifies several limitations:

- Telemedicine heavily depends on patient articulation of symptoms, which can be unclear or incomplete.
- Doctors must manually review long chat histories, increasing consultation time.
- Existing platforms lack automated triage mechanisms or AI-driven symptom analysis.
- Dermatology consultations, which significantly rely on visual inspection, are difficult without automated image analysis.

These limitations highlight the need for integrating AI components—especially conversational agents and computer vision models—into telemedicine workflows to make remote consultations more efficient and informative.

2.3 Large Language Models for Medical Applications

Several LLM architectures were evaluated for their suitability in healthcare:

- GPT-4/5: Known for high-quality reasoning, contextual interpretation, and natural conversation.
- Gemini: Offers multimodal abilities and strong safety guardrails, making it suitable for medical conversations involving text and images.
- LLaMA models: Useful for fine-tuning and custom domain adaptation.
- ClinicalBERT/BioBERT: Effective for analyzing structured EHR data but limited in free-form patient interaction.

Based on this comparison and evaluations, Gemini was selected due to its balanced performance in conversational understanding, multimodal support, and stable prompt-following behaviour suitable for telemedicine contexts.

2.4 Dermatology Image Classification Techniques

Dermatology is one of the most common areas in telemedicine due to the visual nature of skin-related concerns. The literature shows that Convolutional Neural Networks (CNNs) remain the dominant approach for image-based classification of skin diseases. These models analyze texture, color distribution, lesion patterns, and irregularities to classify various conditions.

Studies emphasize that while CNN models are effective for early screening, their outputs must be treated as advisory and not diagnostic. This aligns with our project objective, where the skin disease model provides preliminary assessments and encourages patients to consult qualified medical professionals for confirmation.

2.5 Identified Research Gaps

Despite The survey highlights several gaps in current telemedicine and AI-assisted healthcare systems:

- Telemedicine platforms rarely integrate conversational AI with real-time workflows.

- Very few systems combine LLM-based conversation and computer-vision-based dermatology screening in a unified ecosystem.

- Existing solutions lack automated summaries to support clinicians during high patient volumes.

- There is minimal support for structured, AI-assisted triage analysis before the doctor’s involvement.

- Safety, auditability, and transparency in AI guidance are often insufficiently addressed.

These gaps illustrate a genuine need for a multimodal telemedicine system powered by Generative AI—a system capable of understanding patient queries, analyzing images, summarizing conversations, and supporting doctors throughout the consultation process.

Chapter 3 Methodology

3.1 Motivation behind the problem statement

The Access to timely and reliable healthcare remains a challenge for many people, especially when language barriers or accessibility issues prevent effective communication with healthcare providers. Traditional telemedicine systems often fall short in providing real-time, context-aware assistance for patient queries or early-stage dermatological assessments.

Recent advancements in Large Language Models and AI-driven image analysis offer a unique opportunity to address these challenges. By incorporating multilingual support and audio-based interaction, patients can engage naturally with the system—whether through text or speech—in their preferred language. At the same time, automated skin-disease prediction can assist clinicians by streamlining preliminary assessments, reducing response times, and improving overall efficiency.

This project is motivated by the potential to create a more inclusive and responsive telemedicine platform that combines conversational intelligence, visual diagnosis, and multimodal interaction to enhance patient care and support healthcare professionals effectively.

3.2 System design and architecture

The project required combining multiple components—LLM-based conversation, dermatology image analysis, patient–doctor workflows, and real-time messaging—into one integrated platform. To make this possible, we followed a modular design approach where each major feature was developed as an independent unit but integrated through a common backend.

Our system architecture consists of three main layers:

1. **Frontend Layer:**
    - Built using Next.js
    - Includes separate portals for patients and doctors

- Supports chat interface, appointments, prescriptions, and reports

2. **Backend Layer:**
    - Implemented using Node.js and Express
    - Handles authentication, appointments, prescriptions, and notifications
    - Uses PostgreSQL for structured data
    - Uses MongoDB for chat logs and AI conversation history
3. **AI & Processing Layer:**
    - Gemini 2.5 Pro for conversational intelligence
    - CNN-based skin disease classifier for image analysis
    - Safety filters, prompt templates, and context management

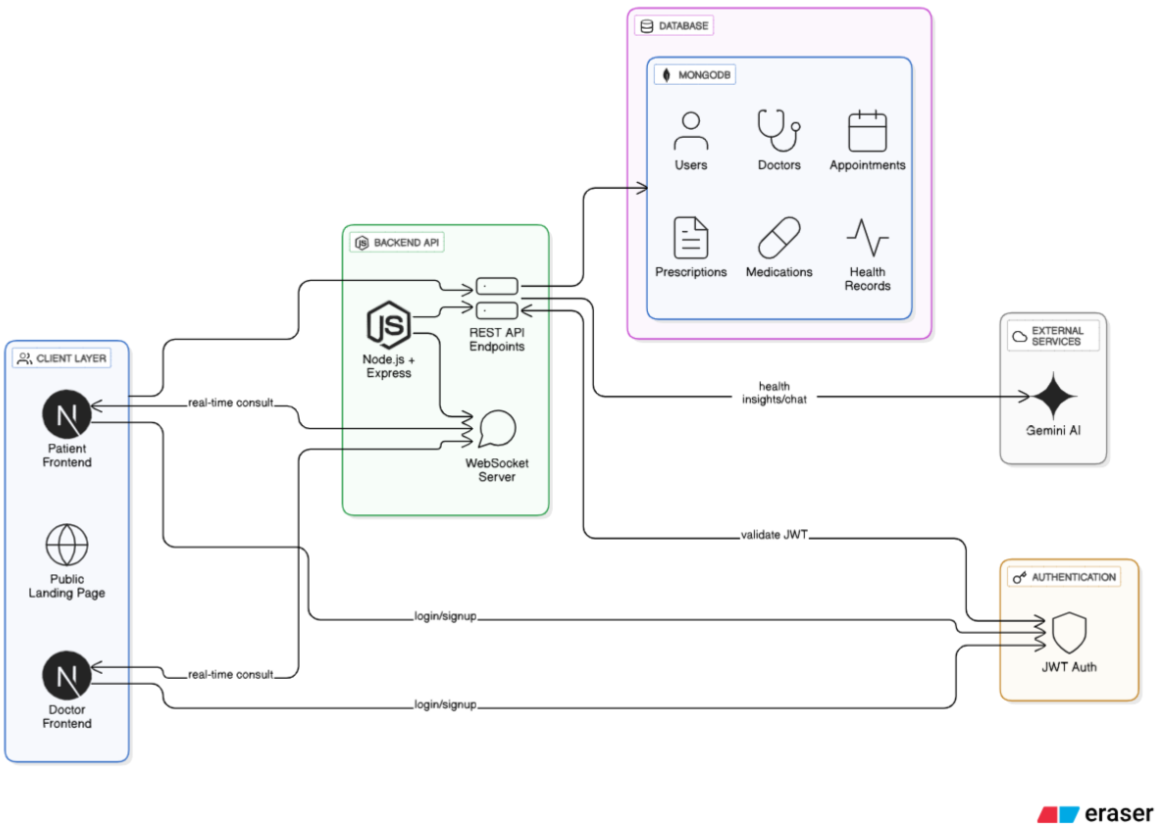

Figure 3.2 System Architecture

3.3 Workflow of the System

The workflow was designed to make the platform usable for both patients and doctors while ensuring AI remains an assistive component, not a diagnostic authority. The steps are as follows:

Figure 3.3 System Workflow

1. Patient logs in and opens the chat/consultation section.

2. Patient enters symptoms or uploads a skin image.

3. The LLM (Gemini) interprets the input, asks relevant follow-up questions, and provides a safe, context-aware response.

4. If an image is uploaded, the CNN classifier processes it and returns a preliminary screening result.

5. All interactions are saved, and a summarized version is generated for doctors.

6. The doctor logs into their portal, reviews the AI summaries, dermatology outputs, and patient history.

3.4 Development Strategy

The Instead of attempting to build everything at once, the system was developed in clearly defined phases based on weekly progress reports. This incremental method allowed us to test each module before integrating it into the final platform.

Phase 1: Requirement & Architecture (Weeks 1–2)

- Understanding how telemedicine currently works
- Identifying problems like poor triage, long doctor review time
- Reviewing GenAI and dermatology literature
- Designing the complete architecture and dataflow

Phase 2: Prototype Development (Weeks 3–4)

- Initial UI for patient portal
- Database schema for users, appointments, and reports
- Basic backend API setup

Phase 3: Core Telemedicine Features (Weeks 6–7)

- JWT authentication, role-based access
- Doctor portal with appointments and profiles
- Real-time video call and messaging using Socket.IO
- Notification system

Phase 4: AI & Dermatology Integration (Weeks 7–9)

- Gemini-based conversation, prompt templates
- Summaries for doctor review
- Skin disease prediction model integration

Phase 5: Testing, Debugging & Refinement (Weeks 10–13)

- Edge-case handling (invalid input, missing fields)
- API tests using Postman
- Chat stability tests
- Presentation preparation and system polishing

3.5 Design Principles Followed

Throughout the project, we adhered to several key design principles to ensure a robust, user-friendly, and scalable telemedicine system:

- **Modularity:** Each component—authentication, chat, AI, and dermatology—was developed independently, allowing for easier maintenance and future upgrades.
- **Scalability:** The system was designed to accommodate future enhancements, such as multilingual support, additional medical models, or new features.
- **Safety:** AI outputs include clear disclaimers and are carefully designed to avoid providing formal medical diagnoses.
- **User-Centric Design:** Both patient and doctor portals were crafted for clarity, intuitive navigation, and a seamless user experience.
- **Real-Time Responsiveness:** WebSocket-based messaging ensures instant feedback, enabling smooth and interactive communication.

Chapter 4 Implementation

4.1 Patient Portal Implementation

The patient portal was one of the first components we developed, with a focus on simplicity, clarity, and ease of navigation to ensure that users of all backgrounds could interact with it without confusion. The portal enables patients to:

- Register and log in securely
- Describe their symptoms through the chat interface
- Upload images for skin disease analysis
- Book appointments
- Receive notifications and view previous prescriptions

The chat interface was designed to provide a natural, conversational experience. When a patient submits a query, it is processed by the Gemini LLM in the backend, which generates an appropriate response.

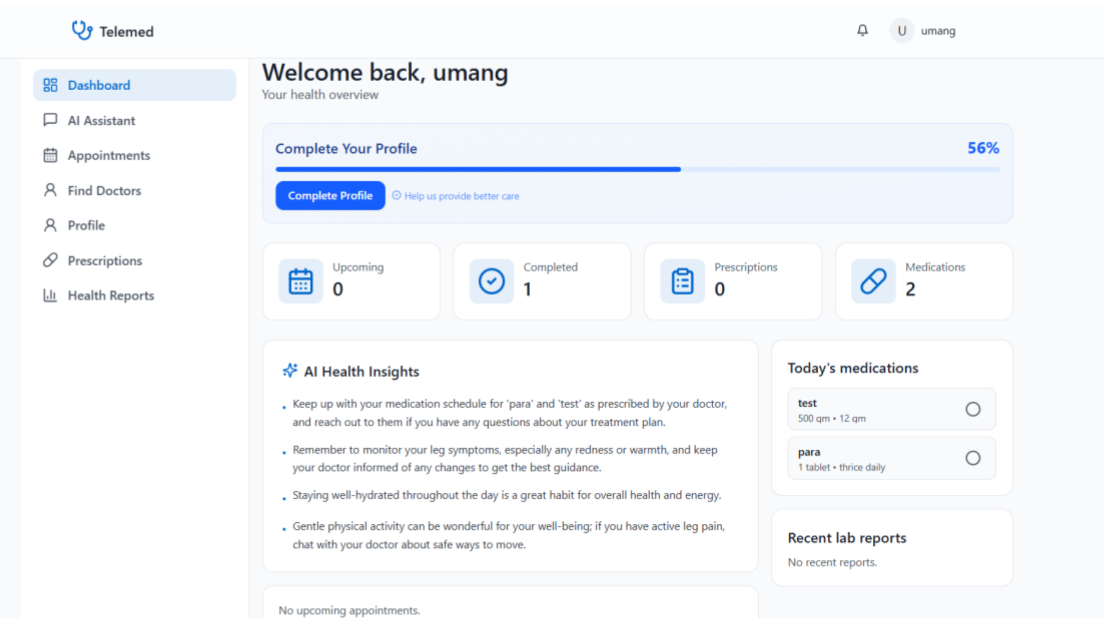

Figure 4.1 Patient Dashboard

4.2 Doctor Portal Implementation

The doctor portal was implemented after the patient portal and contains more structured information targeted toward clinicians. The goal was to give doctors a clear picture of ongoing consultations without overwhelming them.

The main features include:

- a dashboard showing pending consultations,
- AI-generated summaries of patient interactions,
- access to dermatology predictions,
- prescription creation and management,
- real-time chat with patients,
- notifications for new messages and appointments.

Doctors can review the patient's conversation history and the AI-generated summary before responding. This directly reduces the time needed to sift through long chat threads.

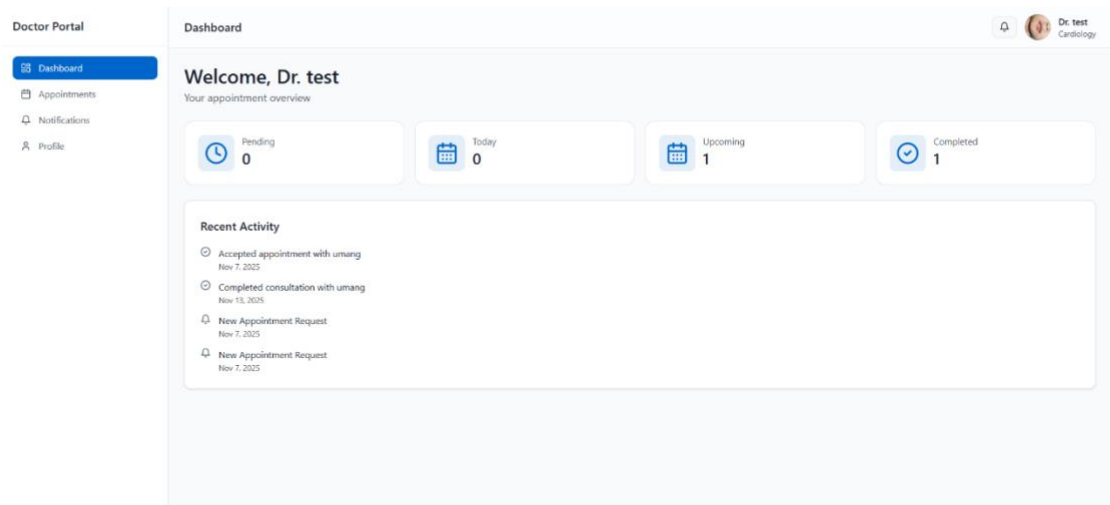

Figure 4.2 Doctor Dashboard

4.3 Backend API and Database Implementation

The backend forms the core of the system, managing all communication between the frontend, AI layer, and the database. It was built using Node.js and Express, with a well-structured folder layout for different modules.

Key backend components include:

- **Authentication Module:** Implemented in Week 6 (*weekly-report-6*), this module uses JWT with refresh tokens to ensure secure and persistent login sessions. Role-based access controls ensure that patients and doctors can only access features intended for them.

- **Appointment and Prescription Management:** CRUD APIs were developed to handle booking, updating, and listing appointments. Prescriptions can be generated and securely stored for future reference.

- **Chat and Conversation Logs:** All conversations between patients, doctors, and the AI are stored in MongoDB, preserving context for accurate responses, history tracking, and AI-generated summaries.

- **Notification Service:** Built during Weeks 7–9 (*weekly-report-7* to *weekly-report-9*), this service alerts users about unread messages, response updates, and appointment changes.

MongoDB serves as the sole database, storing both structured data (such as user profiles and appointments) and unstructured conversational content, providing flexibility, scalability, and efficient data retrieval.

4.4 AI Integration using Gemini

Integrating Gemini was one of the most crucial steps in the project. During Week 7 we created a dedicated AI service that manages:

- prompt templates for symptom checking,
- generating follow-up questions,
- summarizing full conversations for the doctor view,
- ensuring safe and non-diagnostic responses,
- context tracking across multiple messages.

We also added safety boundaries so the model never attempts to provide medical diagnoses. This ensures that the AI remains a supportive tool—not a replacement for a doctor.

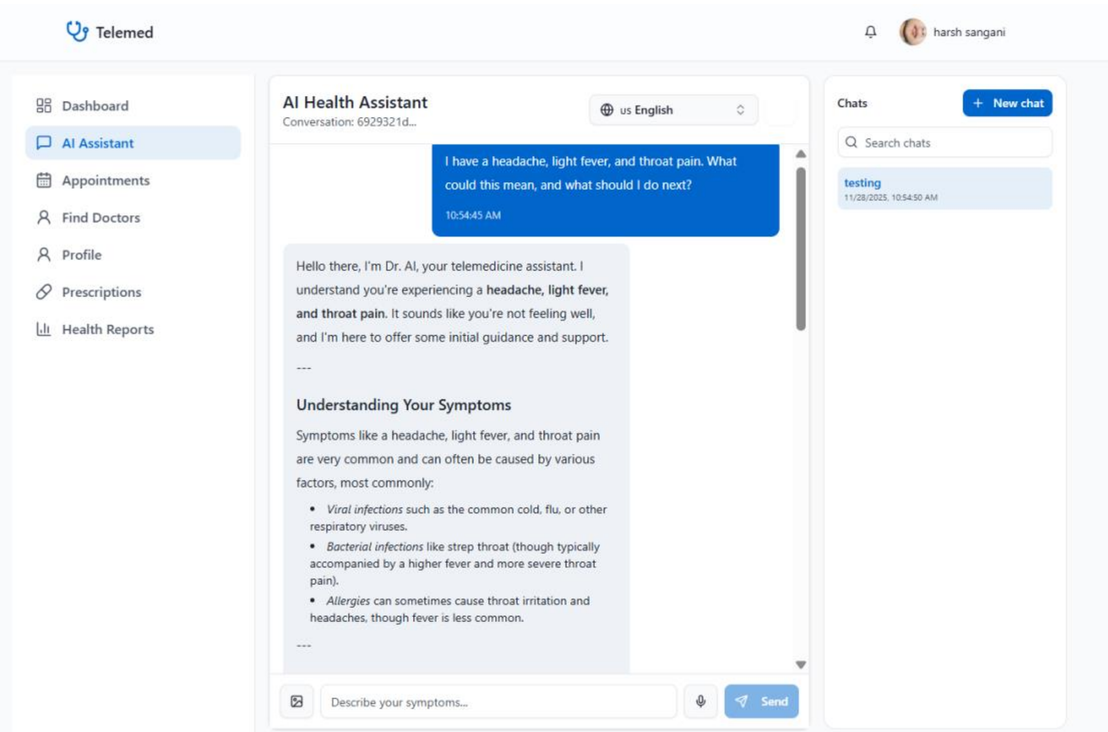

Figure 4.4 AI Interface

4.5 Skin Disease Prediction Module

To handle dermatology cases, we integrated a CNN-based skin disease classifier. Patients can upload images of affected skin areas, and the model provides a preliminary analysis. The output includes:

- the predicted skin condition,
- the confidence score,
- a patient-friendly explanation,
- whether doctor attention is recommended.

The classifier does not provide any final clinical diagnosis. Instead, it assists both the patient and the doctor by highlighting potentially concerning symptoms.

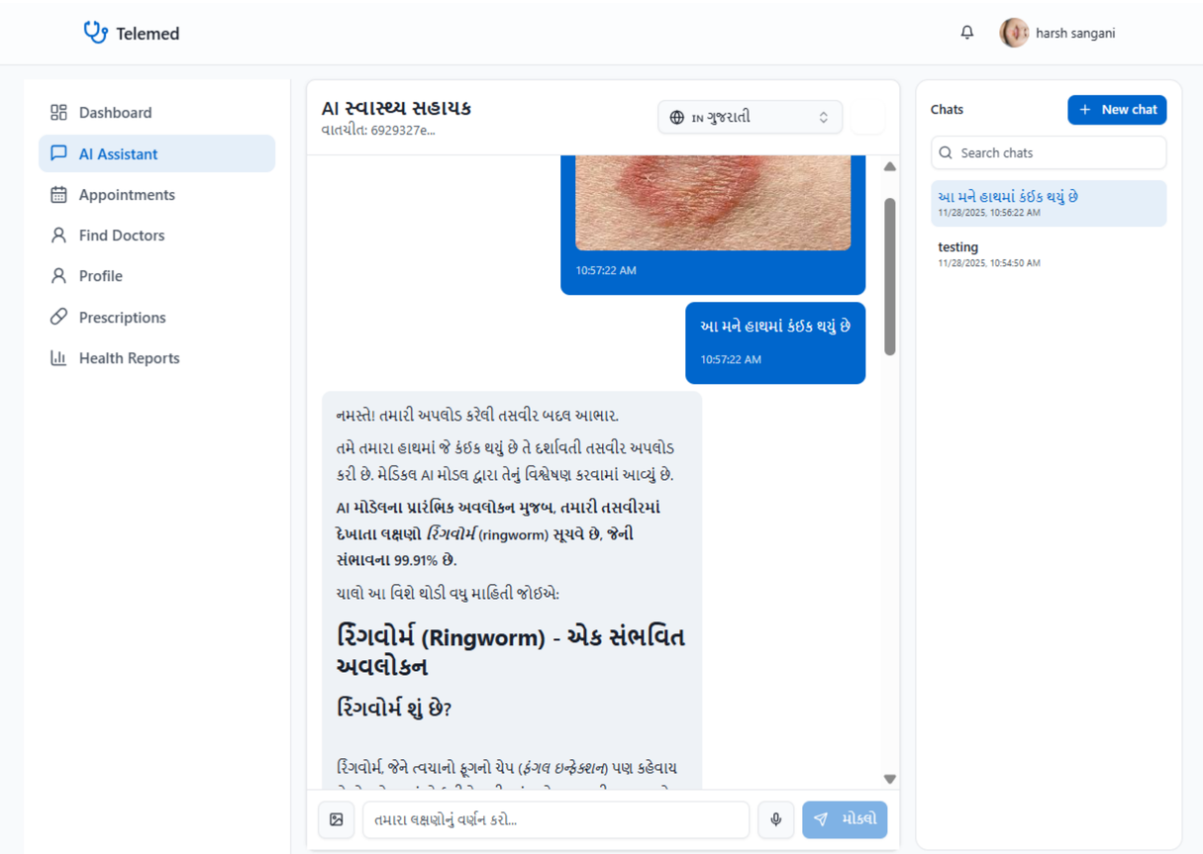

Figure 4.5 Photo Triage

4.6 Real-Time Communication and Video Calling

Real-time communication in our system is handled through Socket.IO, enabling instant messaging between the patient, doctor, and AI. Features such as typing indicators, read receipts, auto-scroll, and reconnection logic ensure a smooth and reliable chat experience. Doctors also receive alerts for urgent or abnormal patient messages.

In addition to messaging, we integrated secure video calling using WebRTC, allowing patients and doctors to conduct face-to-face consultations directly within the platform. The video module supports clear audio/video communication and essential in-call controls.

Both messaging and video calling functionalities were tested thoroughly to ensure stability across login/logout cycles, network changes, and prolonged usage.

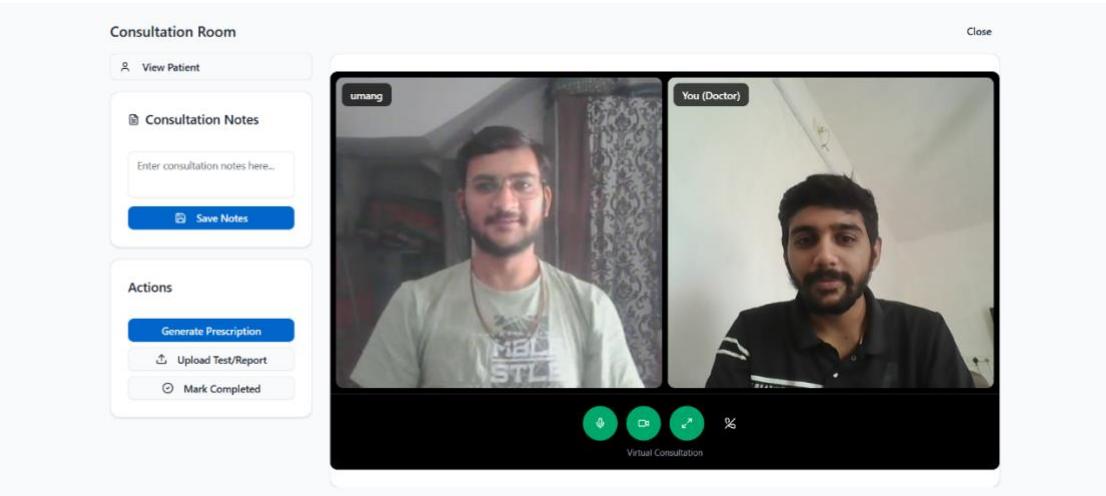

Figure 4.6 Real-time Interface

4.7 Final Integration and Testing

Once all modules were implemented, we integrated them into a single working platform. This phase included:

- checking API outputs and fixing mismatched fields (Week 10),
- cross-validating frontend forms with backend validation rules,

- ensuring summaries were accurately generated,
- verifying credential handling,
- performing edge-case and stress tests (Week 12),
- preparing final presentation and walkthrough (Week 13)

By the end of Week 13, the entire system was stable, fully functional, and ready for demonstration.

Chapter 5 Results & Discussion

5.1 Functional Results

The main objective of the project was to integrate Generative AI and dermatology analysis into a telemedicine workflow. By the time we completed the final integration, the platform successfully supported all major features we had planned.

Key functionalities achieved:

- **AI-driven conversation system:** The Gemini-based chatbot understood user symptoms, asked follow-up questions, and responded in a clear and safe manner.
- **Dermatology image analysis:** The CNN model analyzed uploaded skin images and provided preliminary insights with a confidence score.
- **Patient and doctor portals:** Both portals were fully functional with separate dashboards, consultations, notifications, and history views.
- **Appointment and prescription flow:** Patients could book appointments, and doctors could generate structured prescriptions.
- **Real-time messaging:** Communication between patient ↔ doctor ↔ AI worked smoothly with instant updates.

These results confirmed that the system operated as a complete end-to-end telemedicine solution rather than just a standalone AI tool.

5.2 Performance Evaluation

To evaluate system performance, we conducted focused tests on the AI responses, image processing latency, API behavior, and real-time communication.

Response Speed

- Gemini consistently produced replies in around **1.5–2.5 seconds**, making the chat experience smooth and interactive.
- The dermatology classifier took about **3–5 seconds** to analyze images, which is acceptable for preliminary screening.

API Stability

During Weeks 12 and 13,

we tested multiple endpoints using Postman:

- /auth/login handled both valid and invalid credentials reliably.
- /appointments properly rejected duplicate or incomplete entries.
- /notifications updated correctly during login/logout cycles.

These tests helped us identify and fix issues such as missing validation, inconsistent status codes, and field mismatches.

Real-Time Communication

Socket.IO performance remained stable:

- Messages were delivered instantly.
- Typing indicators and read-status updates worked as expected.
- The system handled repeated login/logout events without breaking the connection.

5.3 Discussion

From the results, it became clear that combining LLM-based conversation with telemedicine workflows can significantly improve both patient and doctor experience. Patients benefited from:

- clearer explanations of their symptoms,
- immediate AI assistance even before the doctor review,
- faster triage,
- and guidance on whether an issue required medical attention.

Doctors, on the other hand, could:

- rely on AI-generated summaries to quickly understand patient concerns,
- view dermatology predictions without manually interpreting images,
- manage prescriptions and consultations more efficiently.

The system does not replace medical professionals but rather complements them by automating repetitive tasks and structuring patient communication. This makes telemedicine more scalable and improves overall accessibility.

The combination of text-based conversation and image-based screening also adds a multimodal dimension to telemedicine that most current platforms lack.

Conclusion

The goal of this project was to design and develop a Gen AI–based multimodal telemedicine system that could support both patients and doctors through intelligent conversation, preliminary dermatology analysis, and structured healthcare workflows. Over the course of thirteen weeks, we successfully implemented a platform that integrates a conversational LLM, a CNN-based skin disease prediction model, real-time messaging, and user-specific dashboards for patients and doctors.

Through multiple rounds of testing and refinement, the system demonstrated reliable performance in understanding user queries, generating safe and context-aware responses, analyzing skin images, and managing appointments and prescriptions. The AI-generated summaries also helped reduce the manual effort required by doctors during consultations.

The project highlighted the practical value of combining Generative AI with telemedicine. While the system is not intended to make medical diagnoses, it effectively supports triage, improves accessibility for patients, and provides clinicians with structured and meaningful insights. Overall, the project met its objectives and showcased how AI can enhance digital healthcare services when used responsibly and collaboratively.

Future Scope

The While the system is functional and meets the primary goals of this minor project, there are several areas where it can be expanded and improved:

1. **Integration with Hospital EHR/EMR Systems:** Connecting the platform to real hospital databases would allow automatic retrieval of patient records, lab results, and historical data.

2. **Advanced Medical Imaging Models:** Beyond dermatology, the platform can integrate models for chest X-rays, retinal imaging, or dental screening.

3. **Role-Based Advanced Analytics:** Doctors could benefit from analytics dashboards displaying patient trends, frequently reported symptoms, and treatment outcomes.

4. **HIPAA/GDPR-Level Data Security:** Additional encryption, audit trails, and compliance mechanisms can be introduced to meet global healthcare security standards.

Each of these improvements can help transform the system into a fully deployable telemedicine solution suitable for hospitals and clinics.

Appendices

The appendices contain supporting material, intermediate outputs, and additional documentation generated throughout the project. These items provide evidence of the development and testing processes.

Appendix A – Weekly Progress Reports

Includes Week 1 to Week 13 reports submitted during the project timeline.

Appendix B – System Architecture Diagrams

All architecture, workflow, and AI integration diagrams referenced in Chapters 3 and 4.

Appendix C – API Testing Logs

Postman test results for authentication, appointments, chat, and notification routes.

Appendix D – Screenshots of Interfaces

Snapshots of:

- Patient dashboard
- Doctor dashboard
- Real-time chat
- Dermatology analysis module
- Appointment and prescription views

Appendix E – Sample Prescriptions and Summaries

AI-generated summaries, doctor notes, and sample prescriptions stored during testing.

References

[1] S. A. Holmes, T. Samuel, and S. A. Rabbani, "Generative AI in Healthcare: Opportunities and Ethical Considerations," *Frontiers in Digital Health*, Vol. 4, 2025. Available: https://www.frontiersin.org/journals/digital-health

[2] T. Samuel and T. Soetan, "AI-Powered Virtual Assistants in Telemedicine: A Review of Applications and Challenges," *Journal of Medical Informatics*, Vol. 12, No. 1, 2025. Available: https://jmi.org/

[3] S. A. Rabbani, M. A. Qamar, and H. Khan, "Generative Artificial Intelligence in Healthcare: Transforming Patient Interaction and Clinical Decision Support," *BioMedInformatics*, Vol. 3, No. 2, 2025.
Available: https://www.mdpi.com/journal/biomedinformatics

[4] A. Mehta and R. Verma, "Generative Artificial Intelligence in Healthcare: Applications, Implementation Challenges, and Future Directions," *Journal of Medical Systems*, Vol. 49, No. 2, 2024.
Available: https://www.mdpi.com/2673-7426/5/3/37

[5] S. A. Holmes, V. Faria, and E. A. Moulton, "Generative AI in healthcare: challenges to patient agency and ethical implications," *BioMedInformatics*, Vol. 3, No. 2, 2025. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12213482/