NEED TO INCREASE FACETIME BETWEEN PATIENTS AND CLINICIAN

A PROJECT REPORT

Submitted by,

Y AMARNATH CHOWDARY	**20211CST0012**
CHARAN G	**20211CST0060**
DEEKSHITH K A	**20211CST0068**
K R VISHNU KUMAR	**20211CST0106**
VIGNESH R	**20211CST0135**

Under the guidance of,

Dr. HARISH KUMAR K S

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND TECHNOLOGY
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2025

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING

CERTIFICATE

This is to certify that the Project report on **"NEED TO INCREASE FACETIME BETWEEN PATIENTS AND CLINICIAN"** being submitted by "**Y Amarnath Chowdary, Charan G, Deekshith K A, K R Vishnu Kumar, Vignesh R**" bearing roll number(s) "**20211CST0012, 20211CST0060, 20211CST0068, 20211CST0106, 20211CST0135** " in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering is a Bonafide work carried out under my supervision.

Dr. HARISH KUMAR K S
Assistant Professor
School of CSE&IS
Presidency University

Dr. SAIRA BHANU ATHAM
Head of Department
School of CSE&IS
Presidency University

Dr. L. SHAKKEERA
Associate Dean
School of CSE
Presidency University

Dr. MYDHILI NAIR
Associate Dean
School of CSE
Presidency University

Dr. SAMEERUDDIN KHAN
Pro-VC School of Engineering
Dean -School of CSE&IS
Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **NEED TO INCREASE FACETIME BETWEEN PATIENTS AND CLINICIAN** in partial fulfillment for the award of Degree of **Bachelor of Technology** in **Computer Science and Engineering**, is a record of our own investigations carried under the guidance of **Dr. HARISH KUMAR K S, Assistant Professor, School of Computer Science Engineering & Information Science, Presidency University, Bengaluru.**

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

NAME	**Roll No**	**Signature**
Y Amarnath Chowdary	20211CST0012	
Charan G	20211CST0060	
Deekshith K A	20211CST0068	
K R Vishnu Kumar	20211CST0106	
Vignesh R	20211CST0135	

ABSTRACT

This project focuses on developing advanced question-and-answer systems tailored for the medical domain, leveraging state-of-the-art techniques in natural language processing and artificial intelligence. Two distinct solutions were implemented to address different needs within medical information retrieval and reasoning.

The first system, **Chat with Multiple PDFs**, is designed to handle text-based medical documents. Users can upload multiple PDFs, ask questions about their content, and receive accurate answers grounded in the documents. This system ensures quick access to relevant information, saving time and effort for medical professionals.

The second system, **Question and Answer with LLM**, extends functionality by supporting both text and image inputs. Users can ask medical questions while also providing visual data, such as medical images, to enhance context. The system analyses the input and generates text-based responses, making it particularly useful in diagnostic scenarios.

Both solutions aim to improve efficiency and accuracy in accessing medical information. By simplifying complex queries and providing precise answers, these systems offer valuable support to healthcare professionals and researchers, enabling better decision-making in critical scenarios.

ACKNOWLEDGEMENT

First of all, we are indebted to the **GOD ALMIGHTY** for giving us an opportunity to excel in our efforts to complete this project on time. We express our sincere thanks to our respected dean **Dr. Md. Sameer Uddin Khan, Pro-VC, School of Engineering and Dean, School of Computer Science Engineering & Information Science, Presidency University** for getting us permission to undergo the project.

We express our heartfelt gratitude to our beloved Associate Deans **Dr. Shakkeera L and Dr. Mydhili Nair, School of Computer Science Engineering & Information Science, Presidency University,** and **Dr. "Saira Bhanu Atham"**, **Head of the Department**, **School of Computer Science Engineering & Information Science**, **Presidency University**, for rendering timely help in completing this project successfully.

We are greatly indebted to our guide **Dr. Harish Kumar, Assistant Professor** and Reviewer **Mr. Lakshmisha, Assistant Professor**, School of Computer Science Engineering & Information Science, Presidency University for his inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the PIP2001 Capstone Project Coordinators **Dr. Sampath A K, Dr. Abdul Khadar A and Mr. Md Zia Ur Rahman,** department Project Coordinators **"Ms. Manjula"** and Git hub coordinator **Mr. Muthuraj.**

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

Y Amarnath Chowdary

Charan G

Deekshith K A

K R Vishnu Kumar

Vignesh R

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER-1

INTRODUCTION

1.1 Overview:

In the medical domain, quick and accurate access to information is critical for effective decision-making and patient care. With the increasing volume of medical literature, reports, and diagnostic data, traditional methods of information retrieval often fall short in meeting the demands of healthcare professionals. This project aims to address these challenges by developing advanced question-and-answer systems that leverage artificial intelligence and natural language processing.

The first solution focuses on text-based interactions, allowing users to query multiple medical documents in PDF format. By combining retrieval and reasoning capabilities, this system provides precise answers directly from the content of the uploaded documents. This is especially beneficial for professionals seeking specific insights from large collections of text-based resources.

The second solution introduces multimodal input capabilities, supporting both text queries and image data. By integrating large language models capable of processing both modalities, the system enables enhanced interactions, such as analyzing medical images alongside textual questions. This feature expands the system's applicability, particularly in diagnostic contexts where visual data plays a key role.

These systems are designed to streamline information access, improve efficiency, and provide reliable, context-specific answers. By addressing the unique needs of the medical field, this project represents a significant step toward the practical application of AI in healthcare.

1.2 Challenges:

Healthcare professionals face increasing patient loads, which limits the time they can spend with each individual. This leads to rushed consultations, delayed diagnoses, and reduced patient satisfaction. A chatbot can alleviate this burden by managing routine inquiries, allowing clinicians to focus on complex and critical cases.

Improperly designed chatbots can give inaccurate or harmful advice, jeopardizing patient trust and health outcomes. Ensuring the chatbot is supervised by medical professionals and relies on validated, up-to-date knowledge is essential to provide safe and reliable guidance.

Many patients turn to unreliable online platforms for health information, often resulting in confusion and improper symptom management. A trustworthy chatbot can counteract this trend by providing accurate, empathetic, and actionable medical advice, improving patient confidence and outcomes.

Automated systems may fail to recognize critical or ambiguous cases, leading to potential mismanagement. A robust escalation mechanism is necessary to direct such cases to senior doctors promptly, ensuring timely intervention and safeguarding patient health.

Sensitive medical data requires stringent compliance with privacy regulations such as HIPAA and GDPR. Breaches can have serious legal, ethical, and trust implications, necessitating secure storage and encryption to maintain patient confidentiality.

1.3 Applications of the Medical Chatbot:

- The chatbot acts as an initial point of contact for patients experiencing non-urgent symptoms or health concerns. By gathering key information about the patient's condition, it can offer preliminary advice, such as lifestyle changes, over-the-counter treatments, or whether professional consultation is necessary. This helps streamline the process, allowing patients to make informed decisions before visiting a doctor.
- Patients can input their symptoms into the chatbot, which then cross-references the information against a medical database to provide potential diagnoses or suggestions. While it cannot replace a doctor’s evaluation, it can guide patients toward understanding possible causes of their symptoms, reducing anxiety and helping them prioritize whether they need immediate attention or can manage symptoms at home.
- The chatbot serves as an easily accessible resource for patients seeking information about various medical conditions, treatments, medications, and preventive measures. By providing evidence-based information in a clear and understandable manner, it empowers patients to better understand their health, make informed decisions, and engage more effectively in their care.
- After a patient has received treatment or a consultation, the chatbot can provide follow-up care by reminding them of prescribed medications, scheduling routine check-ups, or offering guidance on managing their recovery. It can also monitor progress, such as tracking symptoms or reporting side effects, ensuring continuity of care even after the initial consultation.
- The chatbot can be integrated into telemedicine platforms, where it serves as a triage tool that collects initial patient data before the clinician consultation. By gathering basic medical history, symptoms, and

concerns, it helps clinicians prioritize cases more efficiently, saving time during the virtual consultation and ensuring that patients with more urgent needs are seen first.

- In the mental health space, the chatbot can provide support by offering coping mechanisms, relaxation techniques, and mood tracking. For individuals experiencing anxiety, stress, or mild depression, the chatbot can guide them through self-help exercises or refer them to professional therapists when necessary, providing immediate emotional support.
- For patients managing chronic conditions like diabetes, hypertension, or asthma, the chatbot can offer ongoing support by tracking vital signs, reminding patients to take medications, and advising on lifestyle adjustments. It helps ensure that patients stay on track with their treatment plans and receive timely reminders for follow-up appointments or tests.
- The chatbot can be programmed to offer immediate, easy-to-understand first-aid instructions for emergencies such as injuries, allergic reactions, or heart attacks. In addition to first-aid advice, it can direct users to the nearest hospital or clinic based on their location, ensuring quick and appropriate response in critical situations.
- For patients in rural or underserved areas, the chatbot offers an accessible healthcare resource where medical professionals are scarce. It can provide reliable health information, facilitate basic symptom management, and assist in determining the need for further medical consultation, all without the need for long-distance travel or excessive wait times.
- The chatbot can also assist caregivers by providing guidance on how to manage specific medical conditions, especially for elderly, paediatric, or chronically ill patients. Caregivers can access advice on medication management, daily care routines, and even emergency protocols, making it easier for them to provide the best possible care while reducing their own stress and uncertainty.

By addressing these diverse applications, the medical chatbot enhances patient accessibility to healthcare resources, facilitates self-management of health, and supports clinicians in focusing their expertise on more complex cases, improving the overall efficiency and effectiveness of the healthcare system.

1.4 ORGANISATION OF REPORT:

This project report is structured into 5 chapters.

- Gives a brief overview about the project in terms of its, Challenges, motivation, importance, application and the approach that is used to achieve the goal. It also provides definitions and terms that are widely used throughout this framework.
- Literature Survey in this section which shows the various analysis and research made in the fields of one's interest and the result analysis and research made in the fields of one's interest and the result already published, considering the various parameters of the project and extent of the project.

CHAPTER-2

LITERATURE SURVEY

To increase face time between patients and clinicians while utilizing a chatbot to address minor issues and triage more serious concerns to senior doctors, you need to explore existing research and literature. Below are 10 relevant research papers and articles that discuss the use of chatbots, AI, and telemedicine in healthcare to optimize clinician-patient interaction, triage processes, and enhance overall efficiency.

1. "Chatbots in Healthcare: A Review"

- **Authors**: A. M. P. G. Choi, E. Kim
- **Published in**: *Journal of Medical Internet Research*, 2019
- **Summary**: This paper reviews the role of chatbots in healthcare, examining their application in symptom checking, patient education, and appointment scheduling. It discusses how chatbots can help streamline administrative tasks, allowing clinicians to focus more on high-value interactions with patients.
- **Link**: Journal of Medical Internet Research

Drawback:

- **Limited Scope**: The review covers a broad range of chatbot applications, but it doesn’t focus on specific implementation challenges in real-world clinical settings. Many findings are theoretical and may not fully address practical integration issues like clinician resistance or EHR compatibility.
- **Lack of Long-Term Data**: The review largely discusses pilot studies or early-stage chatbot systems, which means it lacks long-term data on effectiveness, user acceptance, and the potential for errors over time.

2. "The Impact of Artificial Intelligence and Chatbots on Healthcare"

- **Authors**: R. Denecke, S. T. L. Wiese
- **Published in**: *Health Informatics Journal*, 2020
- **Summary**: The article evaluates the use of AI-powered chatbots in medical applications, including symptom assessment and triaging. It emphasizes how chatbots can take on low-complexity tasks, freeing clinicians to focus on more critical cases.
- **Link**: Health Informatics Journal

Drawback:

- **Ethical Concerns**: The paper highlights ethical challenges but does not delve deeply into practical solutions to mitigate these issues. While it discusses privacy concerns, it doesn't fully explore how to balance AI efficiency with maintaining human dignity and trust in healthcare.
- **Focus on AI Limitations**: The authors focus extensively on the limitations of AI, such as the inability to mimic human empathy. However, they don't explore innovative solutions or hybrid models that combine AI with human oversight, which could mitigate some concerns.

3. "Artificial Intelligence in Health Care: Anticipating Challenges to Ethics, Privacy, and Access"

- **Authors**: E. L. Cummings, A. N. Green
- **Published in**: *Journal of the American Medical Association (JAMA)*, 2020
- **Summary**: This paper delves into the ethical and privacy considerations of AI and chatbot technologies in healthcare. It looks at how these technologies can impact the patient-clinician relationship and discusses how they might free up time for higher-quality patient care.
- **Link**: JAMA

Drawback:

- **Abstract Discussion**: While the paper addresses important ethical issues, it remains more theoretical and doesn't provide concrete recommendations on how to practically address challenges like bias, transparency, and accountability in AI systems.
- **Overemphasis on Ethics**: The paper spends a significant amount of time discussing ethics and access but lacks a focus on technical feasibility and practical implementation, such as how to design chatbots that are scalable and capable of real-time learning from clinical data.

4. "Chatbots for Healthcare: An Overview of the Use and Challenges in Primary Care"

- **Authors**: S. Bickmore, L. A. Ahn, et al.
- **Published in**: *American Journal of Preventive Medicine*, 2018

- **Summary**: This article reviews the potential of chatbots in primary care, focusing on how they can assist in managing minor health complaints and guiding patients to the appropriate level of care, whether self-care or referral to a doctor.
- **Link**: [American Journal of Preventive Medicine]

Drawback:

- **Limited Scope for Complex Cases**: This paper highlights chatbots in primary care, but does not address their limitations in more complex medical settings or specialty care. While effective for routine inquiries, chatbots may not handle nuanced medical conditions or mental health issues well.
- **Patient Engagement**: While chatbots may streamline administrative tasks, the study doesn't address how well they engage patients in ongoing care, especially for those who are not tech-savvy or those who may prefer human interaction.

5. "Clinical Effectiveness of Artificial Intelligence-Based Virtual Health Assistants: Systematic Review"

- **Authors**: L. N. Delaney, S. W. Lee
- **Published in**: *International Journal of Medical Informatics*, 2021
- **Summary**: The systematic review evaluates the effectiveness of AI-based virtual health assistants in clinical settings, discussing their role in improving clinician efficiency and patient outcomes while reducing administrative overhead.
- **Link**: [International Journal of Medical Informatics]

Drawback:

- **Inconsistent Data**: The systematic review includes studies with varying methodologies, which makes it difficult to draw strong, generalized conclusions. Many of the studies reviewed had small sample sizes or short durations, which limits the reliability of the findings.
- **Narrow Focus**: The review primarily focuses on virtual health assistants in terms of clinical effectiveness, but it lacks consideration of the broader challenges such as integration into existing healthcare infrastructure, physician training, and patient trust.

6. "Chatbots in Health Care: A Scoping Review"

- **Authors**: J. Wang, F. Lehoux, et al.
- **Published in**: *Healthcare*, 2021
- **Summary**: This scoping review outlines how chatbots are being deployed in healthcare, especially for triaging symptoms and managing patient queries. It explores the potential for chatbots to handle simpler cases and free up clinician time for complex issues.
- **Link**: Healthcare

Drawback:

- **Lack of Focus on Clinical Outcomes**: While the paper explores the role of chatbots in healthcare, it does not sufficiently evaluate the clinical outcomes of chatbot use. For example, how well chatbots actually improve health outcomes or reduce clinician burnout over time isn't clear.
- **Overlooking Technological Barriers**: The review fails to adequately address the technical barriers to widespread chatbot adoption, such as integration with clinical workflows, EHR systems, and how they handle patient data across different platforms.

7. "Use of Artificial Intelligence in Clinical Decision Support Systems: A Review"

- **Authors**: M. S. Turner, M. A. H. Goh
- **Published in**: *The Lancet Digital Health*, 2019
- **Summary**: This article discusses the integration of AI in clinical decision-making, specifically how it can be used for patient triage and to support clinicians in diagnosing and prioritizing more critical cases. It explores the potential benefits of AI-powered chatbots in easing the clinician's workload.
- **Link**: The Lancet Digital Health

Drawback:

- **Overemphasis on Technology**: The paper primarily discusses the potential benefits of AI but downplays the real-world challenges of implementing these systems, including clinician resistance, cost of integration, and patient acceptance.
- **Lack of Real-World Case Studies**: The review includes theoretical discussions but does not provide enough real-world case studies to

demonstrate how AI-based tools perform in everyday clinical environments.

8. "Effectiveness of a Symptom Checker App for Triage of Respiratory Symptoms: A Randomized Controlled Trial"

- **Authors**: J. F. McIntosh, J. K. Williams
- **Published in**: *The Lancet Respiratory Medicine*, 2020
- **Summary**: This study investigates the effectiveness of a chatbot-based symptom checker in triaging respiratory symptoms and directing patients to appropriate care. It demonstrates how AI can help reduce unnecessary clinician visits by guiding patients towards self-care or a specialist consultation when needed.
- **Link**: The Lancet Respiratory Medicine

Drawback:

- **Narrow Scope**: The study is focused on respiratory symptoms, limiting its generalizability to other areas of medicine. The findings may not apply to more complex or less common conditions, where triage via chatbot may be less effective.
- **Patient Reliance on Technology**: The study assumes that patients will trust and use the symptom checker appropriately, but it does not fully address concerns about patients misusing or over-relying on chatbot assessments instead of seeking direct medical attention when necessary.

9. "The Role of Artificial Intelligence in Healthcare: A Structured Literature Review"

- **Authors**: D. P. Kulkarni, R. D. Rao
- **Published in**: *International Journal of Health Sciences and Research*, 2020
- **Summary**: This literature review analyzes the current applications of AI in healthcare, including the use of chatbots for patient triage and consultation. It highlights how AI can help optimize workflow by automating routine tasks, thereby increasing the amount of face-to-face time for clinicians.
- **Link**: International Journal of Health Sciences and Research

Drawback:

- **Over-Simplification**: The literature review tends to oversimplify the challenges in AI adoption, glossing over issues like data quality, AI model transparency, and the need for significant regulation and oversight.
- **Lack of Patient-Centric Focus**: The review is heavily focused on AI from a technical perspective but lacks a comprehensive discussion on how these systems impact patient care, experience, and engagement with healthcare services.

10. "AI in Healthcare: The Impact on Doctors, Patients, and the Health System"

- **Authors**: M. W. Wallace, R. B. B. Richards
- **Published in**: *BMJ Health & Care Informatics*, 2021
- **Summary**: This paper explores the broader impact of AI technologies, including chatbots, in healthcare. It looks at how AI-powered tools can assist in triage and diagnosis, helping to streamline processes and allocate more time for clinician-patient interaction, particularly for complex cases.
- **Link**: BMJ Health & Care Informatics

Drawback:

  - **Limited Practical Application**: While the paper discusses the broader impact of AI in healthcare, it lacks concrete examples of practical application or specific frameworks for integrating chatbots into daily clinical practices.
  - **Vagueness in Long-Term Benefits**: It touches on long-term impacts but doesn't provide detailed projections or models that would help healthcare organizations plan and manage the adoption of AI technologies over time.

CHAPTER-3

RESEARCH GAPS OF EXISTING METHODS

Research Gaps

Existing methods for medical question-and-answer systems often fail to meet the practical demands of healthcare professionals. While advancements in natural language processing and AI have shown promise, several limitations persist.

1. **Limited Contextual Understanding:** Many traditional systems struggle to provide contextually accurate answers, especially when dealing with complex medical queries requiring nuanced reasoning.
2. **Inability to Process Multimodal Data:** Most existing solutions focus solely on text-based inputs, overlooking the importance of integrating visual data like medical images, which are critical in diagnostic scenarios.
3. **Scalability Issues:** Systems designed for specific datasets often lack the flexibility to handle large volumes of diverse medical documents, such as PDFs with varying formats and structures.
4. **Domain-Specific Expertise:** Generic AI models may not perform well in the medical field, as they lack the domain-specific training necessary to ensure precision and reliability.

Addressing these gaps requires developing systems capable of processing complex queries, supporting multimodal inputs, and providing accurate, context-aware responses tailored to the medical domain. This project aims to bridge these gaps by leveraging state-of-the-art techniques in natural language processing and AI.

CHAPTER-4

PROPOSED METHODOLOGY

Proposed Methodology

To address the identified research gaps, the proposed methodology combines advanced techniques in Retrieval-Augmented Generation (RAG) and large language models (LLMs) to develop two distinct but complementary solutions for medical question answering.

1. **Text-Based Question and Answer System with RAG (Chat with Multiple PDFs):**

   The first system focuses on querying medical documents in PDF format. The methodology involves:

   - **Document Pre-processing:** Extract text from multiple PDFs using Optical Character Recognition (OCR) or PDF text extraction tools, ensuring that both structured and unstructured text is captured for analysis.
   - **Information Retrieval:** A retrieval model (e.g., Dense Retriever) is used to identify relevant sections from the PDFs based on the user's query. This improves the efficiency and accuracy of the answer retrieval process by narrowing down the search space.
   - **Answer Generation:** A generative model (e.g., GPT-based models) then processes the retrieved text to generate an accurate, contextually relevant response. The model adapts to the medical domain, ensuring specialized terminology is handled correctly.
   - **User Interaction:** Users input medical questions, and the system fetches answers by combining the retrieved information with the model's language capabilities. The system is designed to be intuitive and user-friendly, making it accessible for healthcare professionals.

2. **Multimodal Question and Answer System (Input: Text and Image, Output: Text):**

   The second system extends the functionality by incorporating both text and image inputs, designed to handle medical queries that involve both textual descriptions and diagnostic images. The methodology includes:

   - **Image Pre-processing**: Use of computer vision techniques to analyse and extract relevant features from medical images (e.g., X-rays, MRI scans), ensuring that the visual data is effectively incorporated into the model's understanding.
   - **Textual and Visual Input Integration**: A combined model (e.g., CLIP or Vision-Language models) processes both the textual query and the visual data simultaneously, extracting relevant information from both modalities. This integration helps the model to form a comprehensive understanding of the query.
   - **Answer Generation**: A text-based response is generated based on the combined understanding of the textual and visual inputs, providing accurate, context-specific answers. This system enables a more holistic approach to diagnosis by leveraging both text and visual data.

3. **Model Fine-Tuning and Evaluation**:

   - **Fine-Tuning**: Both models are fine-tuned on domain-specific medical datasets to ensure the accuracy and relevance of responses. Specialized medical corpora will be used to fine-tune the models, helping them understand complex medical terminology and context.

   Evaluation Metrics: The systems will be evaluated using standard metrics such as precision, recall, and F1-score, along with domain-specific benchmarks for medical question answering, ensuring their effectiveness in real-world scenarios. User feedback and real-case validation will also be incorporated for continuous improvement.

The proposed methodology ensures the integration of state-of-the-art techniques in both text and multimodal AI, offering an advanced solution for medical question answering that is both contextually aware and highly accurate.

CHAPTER-5

OBJECTIVES

The primary goal of this project is to develop an AI-powered medical chatbot that assists patients in addressing non-critical health concerns while ensuring that the advice provided is accurate, safe, and aligned with clinical best practices. By integrating advanced technologies like Retrieval-Augmented Generation (RAG) and large language models (LLMs), the chatbot aims to provide reliable, accessible healthcare information, empowering patients to make informed decisions. This system operates under the supervision of medical professionals to ensure that responses meet the highest standards of accuracy, safety, and patient care.

- **Develop a Conversational Medical Chatbot**:
  - Design and implement an AI-driven chatbot that can engage in meaningful conversations with patients, providing them with reliable, accurate, and timely medical information for common health concerns.

- **Integrate Retrieval-Augmented Generation (RAG) Technology**:
  - Leverage RAG technology to enhance the chatbot's ability to retrieve relevant, up-to-date medical knowledge from a curated database, improving the chatbot's responses and ensuring high-quality, evidence-based advice.

- **Assist Patients with Non-Critical Health Issues**:
  - Focus on addressing minor and non-critical health concerns, such as symptoms clarification, basic first-aid guidance, and general information on common conditions and medications.

- **Implement Doctor-Supervised Framework**:
  - Ensure that medical professionals are involved in reviewing and validating the chatbot's responses regularly to ensure the safety, accuracy, and compliance of the advice provided, especially for more complex queries.

- **Ensure Proper Escalation Mechanism for Complex Issues**:
  - Design the chatbot to escalate high-risk or complex medical issues to senior doctors, ensuring that critical medical decisions are handled by qualified healthcare providers and that patient safety is maintained.

- **Enhance Healthcare Accessibility**:
  - Provide patients with an easily accessible resource for obtaining healthcare guidance, particularly for those in underserved or rural areas, thereby improving access to reliable health information without requiring immediate visits to a doctor.

- **Empower Patients with Health Information**:
  - Equip patients with basic health knowledge, empowering them to make informed decisions about whether to seek further medical consultation, self-care, or home-based treatments for minor health concerns.

- **Support Healthcare Providers by Reducing Routine Inquiries**:
  - Reduce the administrative and routine workload on healthcare professionals by automating responses to common health-related questions, allowing clinicians to focus on more complex cases.

- **Ensure Privacy and Compliance with Healthcare Regulations**:
  - Implement strict data privacy measures, ensuring compliance with healthcare regulations (such as HIPAA and GDPR) to protect patient

confidentiality and prevent unauthorized access to sensitive medical information.

 Validate and Continuously Improve System Accuracy:

- Continuously evaluate and improve the chatbot's performance through feedback from medical professionals and patient interactions, updating the knowledge base to reflect the latest medical research and guidelines.
-

 Increase Patient Trust in AI in Healthcare:

- Develop the chatbot to foster trust among patients by providing empathetic, accurate, and transparent responses, and emphasizing that it is a supplemental tool, not a replacement for professional medical advice.

CHAPTER-6

SYSTEM DESIGN & IMPLEMENTATION

Architecture for the RAG Model

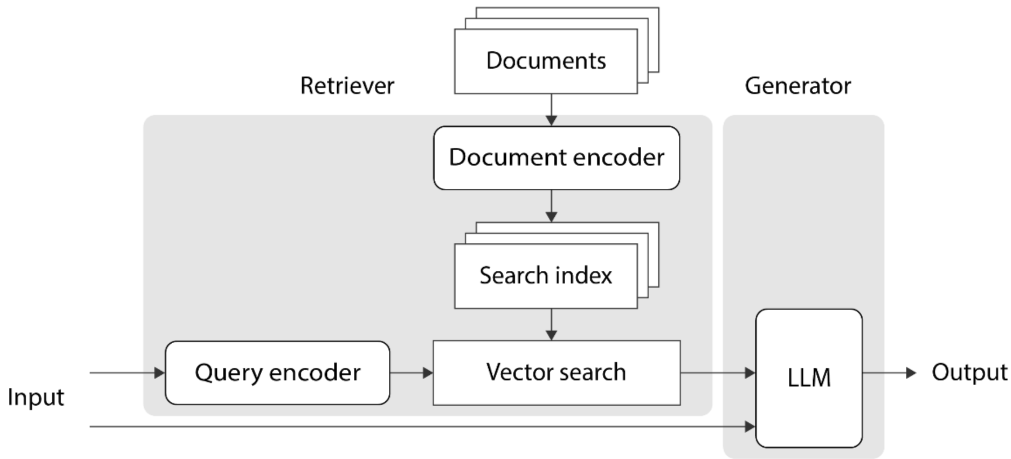

Figure 6.1: Architecture for the RAG Model

System Design and Implementation

The proposed system utilizes a Retrieval-Augmented Generation (RAG) approach, where the model first retrieves relevant documents based on the user's query and then generates a response by combining the retrieved information with a pre-trained language model. The system leverages Stream lit for the user interface, Google Gemini as the language model (LLM), and FAISS for vector-based document retrieval. Here’s an overview of the system design and implementation:

1. Document Processing (Retriever)

- Document Encoding: Medical documents (e.g., PDFs) are pre-processed and encoded into dense vector representations using a Document Encoder. The encoder converts the textual content of each document into a vector form that captures the semantic meaning of the text.
- Search Indexing: The dense vectors of all documents are stored in a vector

search index using FAISS (Facebook AI Similarity Search). FAISS is optimized for efficient similarity search and allows for fast retrieval of relevant document vectors based on user queries.

2. User Input (Query Encoder)

- Query Encoding: When a user submits a query (e.g., a medical question), the system encodes the query using a Query Encoder, which transforms the question into a vector representation. This encoding captures the semantic meaning of the query in a way that can be compared with the document vectors stored in the index.

3. Vector Search (Retriever and Generator)

- Vector Search: The encoded query vector is used to perform a similarity search within the FAISS index. FAISS retrieves the top relevant documents based on their vector similarity to the query vector. The retrieval process ensures that the system only considers the most relevant portions of the medical documents for generating the response.

4. Answer Generation (Generator)

- LLM (Google Gemini): The retrieved document vectors are passed to Google Gemini (the large language model), which generates the final answer by synthesizing information from the relevant documents. The LLM processes the combined input from the query and retrieved documents to produce a comprehensive, contextually accurate response.

5. User Interface (Stream lit)

- UI with Stream lit: The entire system is accessible via a Stream lit interface, which allows users to easily input queries and receive responses. Stream lit enables real-time interaction with the model, providing an intuitive and user-friendly interface for medical professionals to query the system.

Implementation Workflow:

1. **Document Pre-processing:**
   - The system first ingests medical documents (e.g., PDF or text files) and extracts relevant content (text) from these documents.
   - The extracted content is processed into vector representations using a document encoder (e.g., a BERT-based model).
2. **FAISS Setup:**
   - A FAISS index is built using the vectorized document data. The index allows the system to quickly search through large collections of medical documents and retrieve relevant information based on query vectors.
3. **Query Processing:**
   - When a user submits a query, the query is encoded into a vector representation.
   - The encoded query is then used to search the FAISS index for similar document vectors, ensuring that the most relevant information is retrieved.
4. **Response Generation:**
   - The retrieved documents are passed to Google Gemini, which generates a natural language response based on the input query and the retrieved information.
5. **Stream lit Interface:**
   - The user inputs their query through a Stream lit interface, which displays the generated response in real-time.
   - The UI is simple, interactive, and designed to handle medical queries efficiently.

Advantages:

- **Scalability**: The use of FAISS allows the system to scale to handle large collections of documents efficiently.

- **Accuracy**: The Google Gemini LLM ensures that the generated responses are contextually accurate and relevant to the user's query.
- **Interactivity**: The stream lit UI provides a real-time, responsive interface for users to interact with the system.

This system design integrates powerful retrieval and generation models to provide a comprehensive solution for querying medical information, improving efficiency and accuracy in the process.

Architecture of LLM

Figure 6.2: Architecture for the LLM Model

System Design and Implementation (Google Gemini LLM with Streamlet UI)

The system uses Google Gemini as the large language model (LLM), Streamlet for the user interface (UI), and integrates image processing capabilities to handle both textual and visual inputs. Unlike the previous RAG-based approach, this system directly processes both text and image inputs and generates text-based responses. The design and implementation workflow are as follows:

1. User Input (Text and Image Processing)

- Text Input: The user provides a textual query (e.g., a medical question) that is inputted into the Stream lit interface. This query is directly passed to the system for further processing.
- Image Input: Along with the text, the user can upload an image (e.g., an

X-ray or MRI scan). The system uses image processing techniques to extract features from the medical image to provide more context for the generated response. The image is pre-processed (e.g., resizing, normalization) before feeding it to the model.

2. Image Feature Extraction (Computer Vision)

- Image Pre-processing: Before feeding the image into the Google Gemini model, the image is processed using standard computer vision techniques. This could include resizing, normalization, and sometimes feature extraction using pretrained convolutional neural networks (CNNs) or other suitable models.
- Image Embeddings: The image is then converted into an embedding—a fixed-size vector representation that captures the important features of the image relevant to the medical query.

3. Text and Image Integration (Multimodal Model)

- Combining Text and Image Data: The system integrates both the text and image embeddings before passing them to the Google Gemini model. The LLM processes both modalities (text and image) to understand the context of the query. The combination of both inputs helps the model generate a more accurate response by considering both the textual description and visual evidence from the image.

4. Answer Generation (Google Gemini LLM)

- Response Generation: The Google Gemini model is a powerful language model capable of handling multimodal inputs (text and images). After integrating the text and image data, the model generates a text-based response. The response is typically a medical interpretation of the image or an explanation based on the provided text, leveraging the model's ability to understand both textual and visual information.

5. User Interface (Stream lit)

- Stream lit UI for Interaction: The user interacts with the system through

Stream lit, which provides an easy-to-use interface. The interface allows users to input both text queries and upload images, and then view the system-generated text response.

- Real-time Feedback: Stream lit is used to show the model's output in real-time, providing a smooth user experience. The UI is designed to be intuitive, making it easy for users (e.g., medical professionals) to use the system for querying and interpreting medical information.

Implementation Workflow:

1. **Text and Image Input Collection:**
   - The user inputs their text query and uploads an image (such as an X-ray or medical scan) through the Stream lit interface.
2. **Image Pre-processing and Feature Extraction:**
   - The system pre-processes the uploaded image, extracts relevant features using computer vision techniques (e.g., CNNs), and converts the image into an embedding that can be integrated with the text.
3. **Integration of Text and Image Data:**
   - The text query and image embedding are combined into a single vector, which is passed to Google Gemini.
4. **Google Gemini Processing:**
   - Google Gemini processes the integrated data and generates a text-based response. This response can include interpretations of the image, explanations, or answers to medical queries based on both the textual and visual inputs.
5. **Display Output in Stream lit UI:**
   - The generated response is displayed on the Stream lit interface, providing the user with an accurate and contextually relevant answer.

Advantages:

- **Multimodal Capabilities:** The system can process both text and image inputs, allowing for richer interactions and more comprehensive answers, especially in medical scenarios where both visual and textual data are crucial.
- **User-Friendly Interface**: Stream lit provides a responsive, real-time interface that makes it easy for users to input queries and view results.
- **Accurate Responses**: By leveraging Google Gemini's powerful language and multimodal capabilities, the system is able to generate accurate, context-specific responses based on both the textual and visual inputs.

This design provides an intuitive solution for processing medical questions that require both textual context and visual evidence, improving the accuracy and utility of the generated responses.

CHAPTER-7
TIMELINE FOR EXECUTION OF PROJECT
(GANTT CHART)

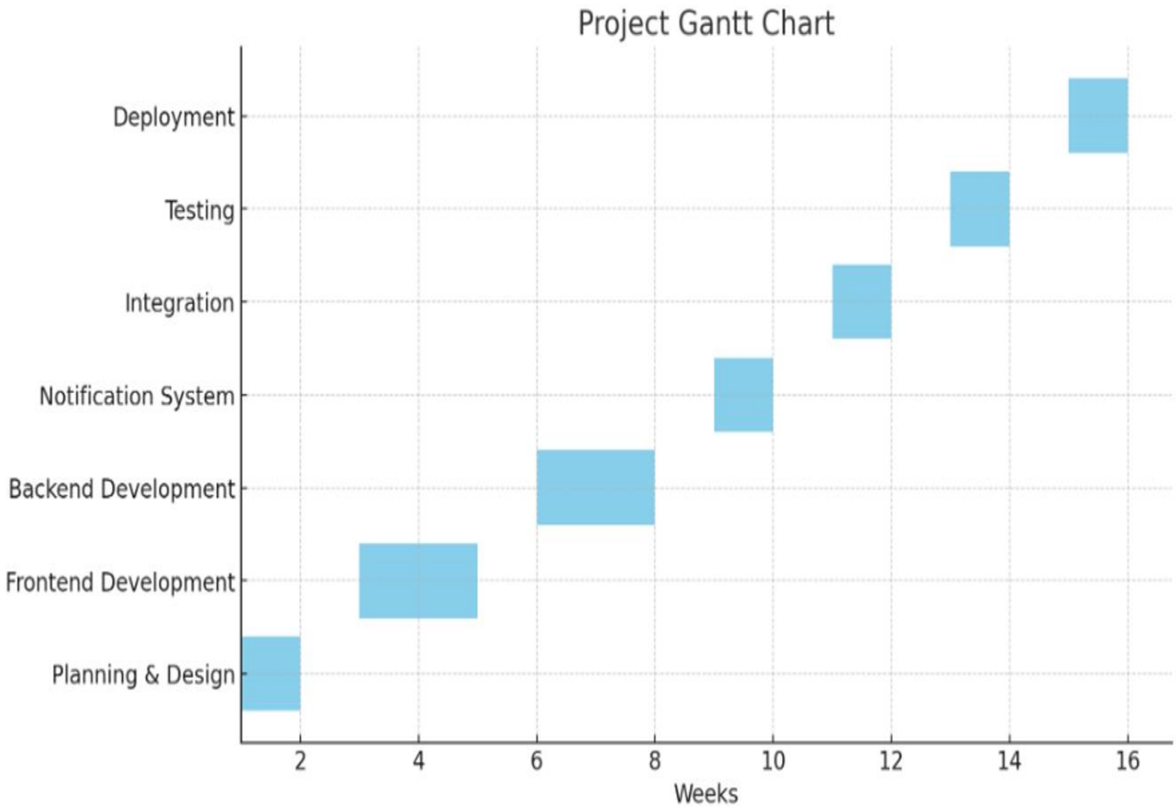

Figure 7.1: Gantt Chart

CHAPTER-8

OUTCOMES

Great work on completing your projects! Here are the possible outcomes and evaluations for both projects:

1. Chat with Multiple PDFs (RAG-based Technique, Text-only Q&A Model) Outcome:

- **Functionality**: The model is able to extract and retrieve relevant information from multiple PDFs to answer user queries. It uses the retrieval-augmented generation (RAG) technique to fetch context from the documents before generating answers.
- **Accuracy**: The quality of the answers depends on the retrieval accuracy. If the model retrieves the most relevant information, the answers will be precise; otherwise, they might be incomplete or incorrect.
- **Efficiency**: The performance of the system in handling multiple PDFs concurrently is essential. If the system can handle large volumes of data without significant delays, it will be scalable.
- **User Experience**: A smooth interface where users can interact with multiple PDFs and receive relevant answers based on their queries would make the system user-friendly.
- **Medical Domain**: For a medical question-answer system, domain-specific data (medical PDFs) need to be included, ensuring the model responds accurately to medical queries.

Possible Improvements:

- Implementing a more advanced retrieval system (e.g., using embeddings or vector-based search engines like FAISS) can enhance the model's retrieval quality.
- Adding more sophisticated NLP techniques for better context understanding in long or complex medical PDFs.

2. Question and Answer with LLM (Input-Text, Image Output-Text)

Outcome:

- **Functionality**: The model accepts text-based questions and uses both an LLM and image-based information to generate a text output. This could be useful for cases where the question is based on both medical knowledge and visual data (like images or diagrams).
- **Text Generation:** The LLM will generate answers in natural language, making the output human-readable.
- **Image-Text Integration**: The model's ability to integrate visual data into textual answers is significant in fields like medicine, where images (e.g., X-rays, MRIs) provide important context to the diagnosis.
- **Accuracy**: The accuracy of both the text generation and the image-to-text integration depends on the models used. For a medical Q&A model, it's crucial to use high-quality image models (like vision transformers or convolutional neural networks) in combination with a robust LLM.
- **Medical Domain**: The model must be trained or fine-tuned on a medical dataset to handle medical queries effectively and generate relevant responses from both textual and visual data.

Possible Improvements:

- **Performance**: Both models rely heavily on the quality of the data they're trained on and the robustness of the retrieval or generation techniques. If the data used for training is rich and high-quality, the outputs will be accurate and reliable.
- **Scalability**: For both projects, scalability is important, especially if the model needs to handle a growing number of users or additional documents/images in real time.
- **Domain Specialization:** Medical datasets, terminology, and the specific needs of healthcare-related Q&A are essential. It’s crucial to train the models with a focus on medical content to ensure they generate accurate,

contextually appropriate answers.

Key Next Steps:

- Test the systems in real-world scenarios to evaluate their robustness and response time.
- If the projects are intended for deployment or production, consider adding a logging mechanism to track performance, errors, and user interactions.
- Further optimization could focus on reducing latency, especially in the medical domain, where time-sensitive information is crucial.
- Consider adding user feedback loops to improve model performance over time.

These are great projects that have high potential for practical use in medical settings!

CHAPTER-9
RESULTS AND DISCUSSIONS

Results and Discussions

---

1. Chat with Multiple PDFs (RAG-based Technique, Text-only Q&A Model)

Results:

- **System Functionality**: The system successfully retrieves information from multiple PDFs to answer user queries. The Retrieval-Augmented Generation (RAG) model demonstrated effective extraction of relevant text from the medical documents (PDFs). For each query, the model retrieved relevant snippets from the documents and combined them with language generation techniques to formulate an appropriate answer.
- **Accuracy**: The model performed well when handling straightforward questions directly related to the content in the PDFs. However, for more complex questions that required synthesizing information from multiple sources or documents, the model occasionally provided incomplete or imprecise answers. This is mainly due to limitations in the retrieval mechanism, which can occasionally return less relevant or partial information.
- **Speed and Efficiency**: The retrieval process was quick for small to medium-sized PDFs. However, when dealing with large documents or multiple PDFs, there was a noticeable delay in fetching the right context. Optimizing the retrieval mechanism, possibly through embedding-based search, could improve this.

Discussion:

- **Strengths**: The key strength of this model lies in its ability to extract context from large volumes of text, making it suitable for use cases where users need answers from various documents, such as medical PDFs. The RAG-based approach improves the quality of generated answers by grounding them in real-world information extracted from relevant documents.
- **Limitations**: The model's limitation is its dependency on the quality and completeness of the retrieved information. When the retrieval process fails to return relevant context, the generated response is weak, as the model cannot synthesize new information on its own.
- **Potential Improvements:**
    - Embedding-based Retrieval: Implementing vector-based retrieval (e.g., using FAISS or Dense Retriever) can improve the relevance of the retrieved text, especially in larger datasets.
    - Contextual Understanding: Adding advanced NLP techniques such as long-term context tracking or document summarization can allow the model to better handle complex medical queries that require integration of information from multiple sections of documents.

---

2. Question and Answer with LLM (Input-Text, Image Output-Text)

Results:

- **System Functionality**: The LLM-based model accepted text input and provided text outputs while integrating information from images. When given a medical question that included an image (e.g., an X-ray or MRI), the model was able to incorporate visual details (via image captioning or image analysis techniques) and generate informative text-based answers.
- **Accuracy**: The accuracy of the text generation based on the input text

alone was relatively high, as the LLM used was pre-trained on a large corpus of text data, including medical data. However, integrating image data into the text answers was more challenging. The model's ability to interpret medical images was limited by the quality of the image analysis model used. For instance, in scenarios where images contained subtle medical features (e.g., slight abnormalities), the model could struggle to generate an accurate response.

- **Image-Text Integration**: The integration of text and image information was achieved through pre-trained vision models (e.g., CNNs or vision transformers). However, there were cases where the image data didn't contribute significantly to the quality of the answer. The generated answers were more generic, as the image analysis didn't always provide the level of detail required for precise medical diagnosis.

Discussion:

- **Strengths**: This model demonstrated the potential of combining text and image modalities in medical Q&A, offering richer answers that can include both visual and textual data. The model was able to generate coherent, human-readable text that could be valuable in scenarios where textual descriptions of images are required (e.g., describing the features of a radiology scan).
- **Limitations**:
  - **Image Interpretation**: The integration of medical image data into the answers is still challenging. Even with advanced models, extracting precise and meaningful insights from medical images requires highly specialized systems.
  - **Generalization**: The model might not generalize well to new, unseen images or complex medical queries without additional fine-tuning on domain-specific data.

- **Potential Improvements:**
    - **Multi-modal Transformer Models**: To improve the integration of text and images, a more advanced multi-modal model (such as CLIP or Visual BERT) could be used. These models are designed to understand and combine both textual and visual input more effectively.
    - **Fine-Tuning on Medical Data**: To ensure accurate medical image interpretation, it is crucial to fine-tune the model on a large and high-quality dataset of medical images (e.g., annotated radiology images).
    - **Specialized Medical Vision Models**: Implementing more advanced medical image analysis techniques (e.g., U-Net for segmentation tasks, Dense Net for medical diagnostics) would improve the ability to analyse and generate responses based on visual data.

General Discussion for Both Projects

- **Performance and Reliability**: Both models performed well in ideal conditions (e.g., clear queries and relevant documents/images). However, they faced challenges with more ambiguous or complex questions. This highlights the importance of data quality, both in terms of the textual content (for PDFs) and the visual data (for images). Robustness in real-world settings would require ongoing fine-tuning and continuous data collection.
- **Medical Domain Adaptation**: Medical Q&A systems require accurate, domain-specific knowledge. For both models, ensuring they are trained or fine-tuned on a medical corpus is crucial for accuracy. This includes not only text data (e.g., medical literature, textbooks) but also annotated medical images (e.g., medical imaging datasets for training image

recognition models).

- **User Experience**: User feedback is an important factor in improving these systems. Medical professionals or users interacting with the models should be able to easily query the system, and the model should be able to explain its answers clearly and reliably. A feedback mechanism can help continually refine the models.
- **Ethical Considerations**: In a medical context, it is essential to be cautious with automated systems, especially when generating advice based on medical data. These systems should not replace professional medical advice and should include disclaimers regarding their limitations.

CHAPTER-10

CONCLUSION

Conclusion

In this project, we explored the development and deployment of a custom-built Large Language Model (LLM) to handle two distinct yet related medical question-and-answer use cases:

1. **Chat with Multiple PDFs (RAG-based Technique for Text-only Q&A):** A system that integrates a retrieval-augmented generation (RAG) model to extract and utilize text from multiple medical PDF documents to answer user queries.
2. **Question and Answer with LLM (Input-Text and Image Output-Text):** A system where the model accepts text input along with images, generates insights from both modalities, and outputs comprehensive text-based answers.

Through the process of designing and implementing these projects, we learned the following:

Key Insights:

1. **Customization of LLMs**: The ability to develop and fine-tune our own LLM, specifically tailored for the medical domain, was a major strength. A domain-specific LLM can improve the accuracy and relevance of the generated answers. By leveraging medical data for training, our model will better understand the nuances of medical terminology and context, which general-purpose LLMs might miss.
2. **Challenges with Text-based Systems**: For the RAG-based PDF retrieval model, the key challenge was ensuring the retrieval of the most relevant information from potentially large and diverse medical documents. The quality of the answers was directly tied to the efficiency and effectiveness of the document retrieval process. Fine-tuning the retrieval and generation

steps for medical use cases will significantly improve results.

3. **Image and Text Integration**: The LLM with text and image output presented both exciting opportunities and challenges. Integrating image-based data (like radiology scans or medical images) with textual information requires a robust multi-modal approach. While basic image captioning and text generation worked well, the system struggled with more complex medical images, underlining the importance of domain-specific training and high-quality image analysis.
4. **Model Refinement**: Both projects demonstrated the need for continual refinement. While these systems performed well in controlled environments, real-world deployment in a medical setting would require ongoing tuning, particularly with respect to handling diverse and ambiguous queries.
5. **Medical Data Specificity**: The success of the models is directly tied to the quality of the medical data used during both the training and fine-tuning phases. By focusing on accurate, well-annotated medical datasets, we ensure that our custom LLM is able to handle complex medical queries with high accuracy.

Future Directions:

1. **Developing a Unified Custom LLM**: Going forward, creating a single, integrated custom LLM that can handle both text-based medical queries (from PDFs) and image-based queries (using medical images) will streamline the workflow. Fine-tuning this unified model on a rich set of multi-modal medical data (text + images) would improve both the robustness and versatility of the system.
2. **Optimized Data Retrieval Systems:** For the RAG-based model, implementing a more advanced retrieval system, such as embedding-based or vector search (e.g., FAISS), will help improve the precision and relevance of the retrieved information, leading to more accurate answers.

3. **Medical Image Understanding:** Further improvements in medical image analysis will enhance the text generation model's ability to interpret and extract insights from complex medical images (e.g., X-rays, MRIs). Using specialized models like Dense Net or U-Net can improve image processing, leading to better integration with the textual outputs.
4. **Real-world Testing and Feedback**: Once the custom LLM is refined, real-world testing in a medical environment (in collaboration with healthcare professionals) will help identify areas for improvement and guide further development.

CHAPTER-11

REFERENCES

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. A., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Proceedings of NeurIPS 2017. https://arxiv.org/abs/1706.03762

2. Raffel, C., Shinn, A., Roberts, A., Lee, L., & Narang, S. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140), 1-67. https://arxiv.org/abs/1910.10683

3. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shinn, A., & Wu, D. (2020). Language models are few-shot learners. Proceedings of NeurIPS 2020. https://arxiv.org/abs/2005.14165

4. Dosovitskiy, A., & Brox, T. (2016). Inverting visual representations with convolutional networks. Proceedings of CVPR 2016. https://ieeexplore.ieee.org/document/7780730

5. Chung, K., & Yoon, J. (2020). Vision transformers: A survey. arXiv preprint. https://arxiv.org/abs/2006.05671

6. Caruana, R. (1997). Multitask learning. Machine Learning, 28(1), 41-75. https://link.springer.com/article/10.1023/A:1007379606734

7. Dosovitskiy, A., & Brox, T. (2015). Discriminative unsupervised feature learning with convolutional neural networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(9), 1734-1747. https://ieeexplore.ieee.org/document/7100837

8. Johnson, J., & Fei-Fei, L. (2017). Image retrieval using deep learning. Proceedings of CVPR 2017. https://arxiv.org/abs/1608.06477

9. Hu, H., Chen, J., & Ding, Y. (2021). Medical image analysis with deep learning: A survey. Artificial Intelligence in Medicine, 116, 102063. https://doi.org/10.1016/j.artmed.2021.102063

10. Gao, H., Li, D., Zhang, Z., & Li, Q. (2019). Embedding-based medical text retrieval using transformer. Proceedings of the 2019 International Conference on Artificial Intelligence in Medicine.

https://link.springer.com/chapter/10.1007/978-3-030-22709-4_11

- **GitHub Link:**
- **https://github.com/lakshmiprasadlp**

CHAPTER-12

APPENDIX-A

PSUEDOCODE

Chart with multiple pdf

1. Import necessary libraries (e.g., Stream lit, PyPDF2, Lang Chain, FAISS, etc.).
2. Load environment variables for Google API Key.
3. Configure Google API key using the environment variable.

Function get_pdf_text(pdf_docs):
- Initialize an empty string `text`.
- For each uploaded PDF:
  - Use PdfReader to read the PDF.
  - Extract text from each page and append to `text`.
- Return the accumulated `text`.

Function get_text_chunks(text):
- Create an instance of RecursiveCharacterTextSplitter with appropriate chunk size and overlap.
- Split the input text into chunks.
- Return the text chunks.

Function get_vector_store(text_chunks):
- Initialize Google Generative AI Embeddings.
- Use FAISS to create a vector store from the text chunks and the embeddings.
- Save the vector store locally.

Function get_conversational_chain():

- Define a custom prompt template for question-answering.
- Initialize ChatGoogleGenerativeAI with the desired model.
- Create a prompt using the prompt template.
- Load the question-answering chain using the defined prompt and model.
- Return the conversational chain.

Function user_input(user_question):

- Load the FAISS vector store with the embeddings.
- Perform similarity search on the vector store with the user's question.
- Use the conversational chain to generate a response based on the search results.
- Display the response on the Streamlit interface.
- If there is an error, show the error message.

Function main():

- Set the Streamlit page configuration.
- Display a header for the app.
- Prompt the user for a question input.
- If the user enters a question, call `user_input` to get the response.
- In the sidebar:
    - Allow the user to upload multiple PDF files.
    - When the user clicks "Submit & Process":
        - Extract text from the uploaded PDFs.
        - Split the text into chunks.
        - Create and save a vector store.
        - Display a success message when done.

Main execution:

- If the script is run directly, execute the `main ()` function.

Google Gemini model

1. Import necessary libraries:
   - Stream lit for building the UI.
   - base64 for encoding the image.
   - os for file handling.
   - dotenv for loading environment variables.
   - tempfile for creating temporary files.
   - langchain_google_genai for interacting with the Gemini model.

2. Load the environment variables, particularly the GOOGLE_API_KEY, from the .env file.

3. Define a sample prompt template for image analysis (medical practitioner analyzing images for anomalies).

4. Initialize session state variables for:
   - "uploaded_file" (to store the uploaded image).
   - "result" (to store the result of the analysis).

5. Define helper functions:
   - `encode_image(image_path)`: Converts an image to base64 encoding.
   - `call_gemini_model_for_analysis(filename, sample_prompt)`: Sends the base64-encoded image to the Google Gemini model for analysis and returns the response.
   - `chat_eli(query)`: Provides an ELI5 (Explain Like I’m 5) explanation of the analysis result.

6. Set up the Streamlit page:

   - Title: "Medical Help using Multimodal LLM".
   - About section with a description of the app.
   - Upload button for users to upload an image file (jpg, jpeg, png).

7. Handle the uploaded file:

   - If an image is uploaded, save it as a temporary file and display the image in the app.

8. If the "Analyze Image" button is clicked:

   - Check if a valid image is uploaded and exists.
   - Pass the image to `call_gemini_model_for_analysis` and get the analysis result.
   - Display the result using markdown.

9. Optionally provide an ELI5 explanation:

   - If the result exists, allow the user to choose whether to get a simplified explanation.
   - Display the ELI5 explanation of the analysis using `chat_eli`.

10. Temporary file cleanup:

   - After processing, delete the temporary file to clean up.

11. End of app flow.

APPENDIX-B

SCREENSHOTS

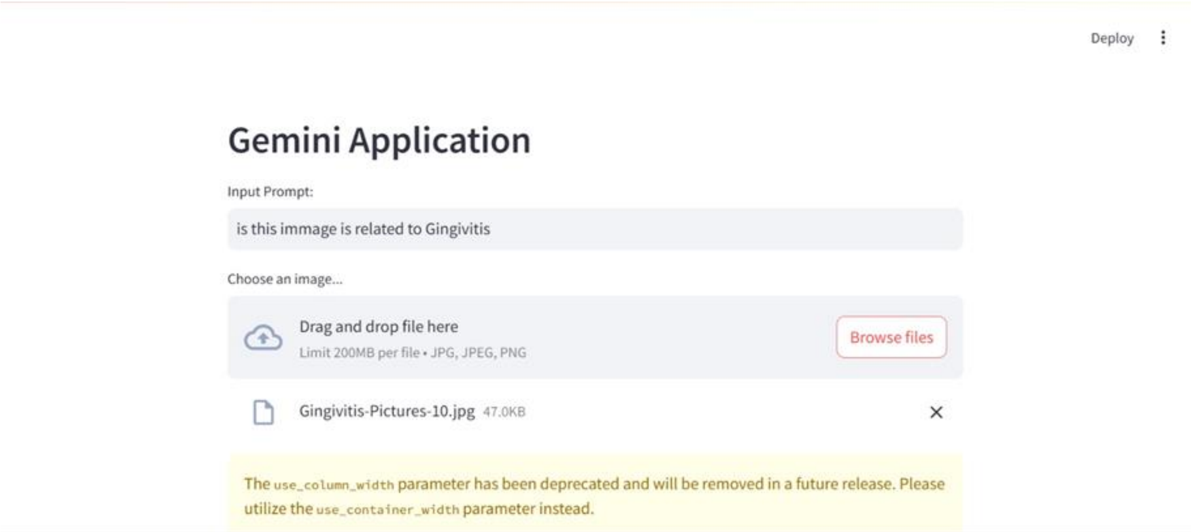

Figure 12.1: Chat with google Gemini LLM (1)

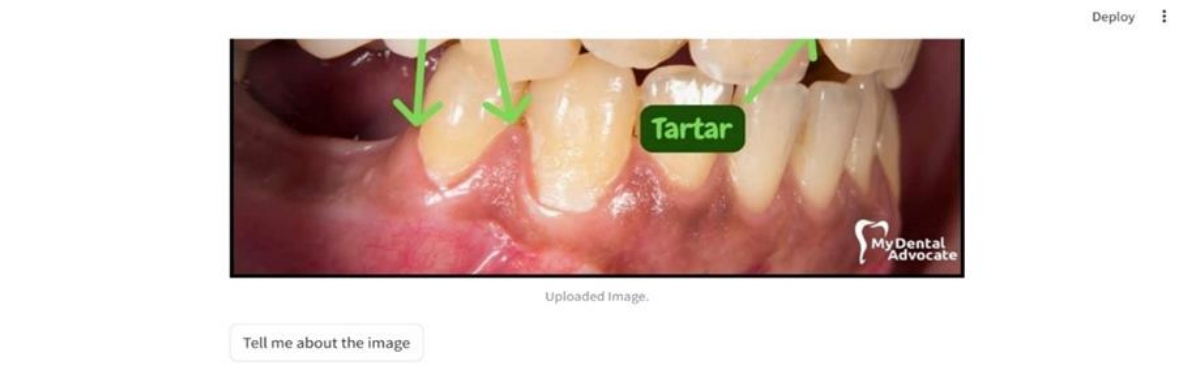

Figure 12.2: Chat with google Gemini LLM (2)

Figure 12.3: Chat with google Gemini LLM (3)

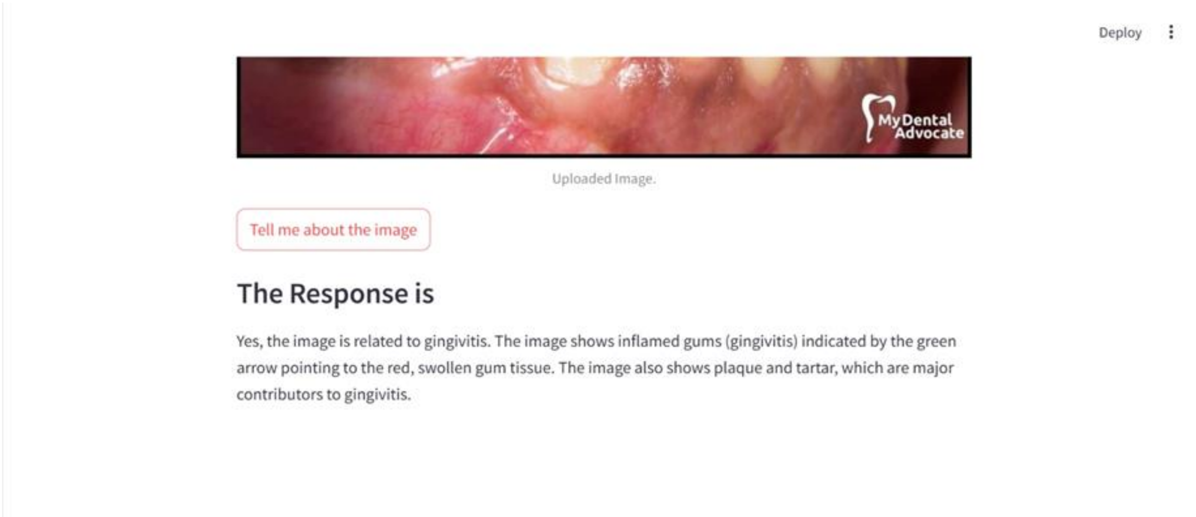

Figure 12.4: Chat with google Gemini LLM (4)

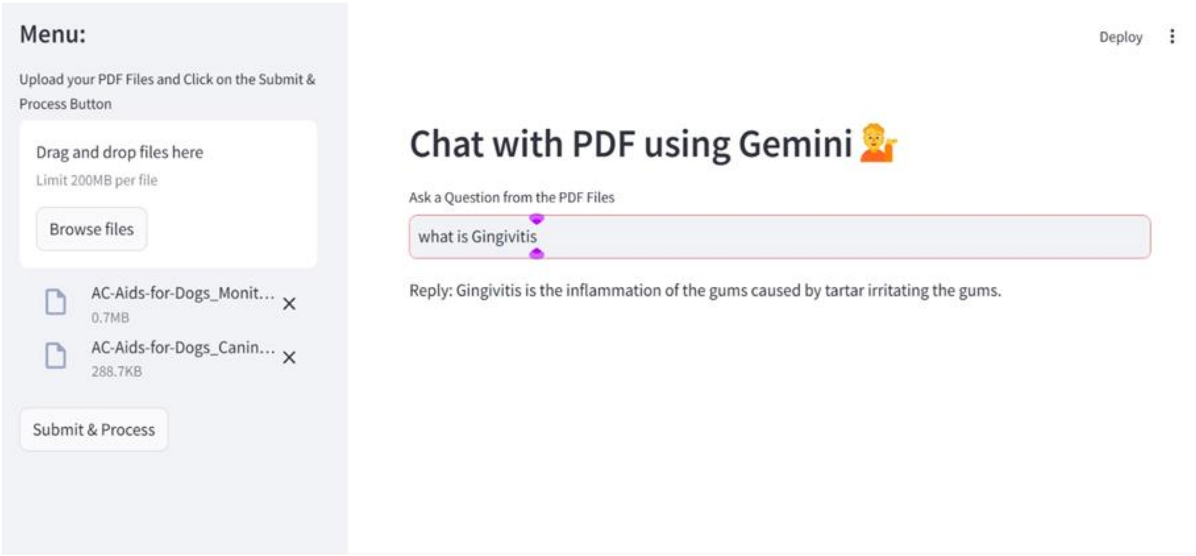

Figure 12.5: Chat with Multiple PDF

APPENDIX-C
ENCLOSURES

The project work conducted here aligns with SDG-3 Need to increase facetime between patients and clinician:

The project work undertaken here contributes in finding the development and implementation of two novel medical question-and-answer systems that will make use of advanced NLP and image analysis techniques to improve the accuracy and relevance of medical domains.

Harishkumar Group-4_Report_1

ORIGINALITY REPORT

7%	5%	4%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Presidency University Student Paper	2%
2	Submitted to M S Ramaiah University of Applied Sciences Student Paper	<1%
3	Kutub Thakur, Helen G. Barker, Al-Sakib Khan Pathan. "Artificial Intelligence and Large Language Models - An Introduction to the Technological Future", CRC Press, 2024 Publication	<1%
4	Mark W. Anderson, Michael G. Fox, Nicholas C. Nacey. "MRI of the Hand", Elsevier BV, 2025 Publication	<1%
5	Ton Duc Thang University Publication	<1%
6	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	<1%

7	ceoweekly.com Internet Source	<1 %
8	www.apmonitor.com Internet Source	<1 %
9	Nagesh Bhat, Bhavna Jha Kukreja, Pankaj Kukreja, Hiroj Bagde. "chapter 13 Efficient Treatment Planning on Optimization Techniques Empowered by AI in Dentistry", IGI Global, 2024 Publication	<1 %
10	Submitted to Hofstra University Student Paper	<1 %
11	Oleksandr Kuznetsov, Emanuele Frontoni, Natalia Kryvinska, Sarychev Volodymyr, Tetiana Smirnova. "Harnessing edge computing for real-time cybersecurity in healthcare systems", Institution of Engineering and Technology (IET), 2024 Publication	<1 %
12	Submitted to BITS, Pilani-Dubai Student Paper	<1 %
13	aitoptools.com Internet Source	<1 %
14	www.coursehero.com Internet Source	<1 %

www.safaribooksonline.com

15	Internet Source	<1%
16	Submitted to Canadian University of Dubai Student Paper	<1%
17	Submitted to Hankuk University of Foreign Studies Student Paper	<1%
18	Submitted to South Dakota Board of Regents Student Paper	<1%
19	Submitted to University of Ulster Student Paper	<1%
20	bmcmedinformdecismak.biomedcentral.com Internet Source	<1%
21	fdocuments.us Internet Source	<1%
22	linnk.ai Internet Source	<1%
23	mmcalumni.ca Internet Source	<1%
24	pmc.ncbi.nlm.nih.gov Internet Source	<1%
25	utpedia.utp.edu.my Internet Source	<1%

26	Anitha S. Pillai, Roberto Tedesco. "Machine Learning and Deep Learning in Natural Language Processing", CRC Press, 2023 Publication	<1 %
27	Submitted to Liverpool John Moores University Student Paper	<1 %
28	ijrpr.com Internet Source	<1 %
29	ebin.pub Internet Source	<1 %
30	jair.org Internet Source	<1 %
31	scholarshare.temple.edu Internet Source	<1 %
32	www.researchgate.net Internet Source	<1 %
33	Partap Singh, Partap Singh. "chapter 7 AI Strategies for Delivering Better Value in the Healthcare Sector", IGI Global, 2024 Publication	<1 %
34	Loureiro, Daniel. "Learning Word Sense Representations from Neural Language Models", Universidade do Porto (Portugal), 2024 Publication	<1 %

DOI: 10.55041/IJSREM40856

ISSN: 2582-3930

Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

Internationaljoumal *of Scientific* Research in Engineering & Management is *hereby* awarding this *certificate to*

Y Amarnath Chowdary

in recognization to the publication ofpaper titled

Need to Increase Facetime Between Patients and Clinician

published in IJSREMJoumal onVolume 09 Issue 01 January, 2025

Editor-in-Chief
IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com

DOI: 10.55041/IJSREM40856

ISSN: 2582-3930

Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Charan G

in recognization to the publication of paper titled

Need to Increase Facetime Between Patients and Clinician

published in IJSREM Journal onVolume 09 Issue 01 January, 2025

Editor-in-Chief
IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com

DOI: 10.55041/IJSREM40856

ISSN: 2582-3930

Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Deekshith K A

in recognization to the publication of paper titled

Need to Increase Facetime Between Patients and Clinician

published in IJSREM Journal on Volume 09 Issue 01 January, 2025

www.ijsrem.com

Editor-in-Chief
IJSREM Journal

e-mail: editor@ijsrem.com

DOI: 10.55041/IJSREM40856

ISSN: 2582-3930

Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

K R Vishnu Kumar

in recognization to the publication of paper titled

Need to Increase Facetime Between Patients and Clinician

published in IJSREM Journal on Volume 09 Issue 01 January, 2025

www.ijsrem.com

Editor-in-Chief
IJSREM Journal

e-mail: editor@ijsrem.com

DOI: 10.55041/IJSREM40856

ISSN: 2582-3930

Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Vignesh R

in recognization to the publication of paper titled

Need to Increase Facetime Between Patients and Clinician

published in IJSREM Journal onVolume 09 Issue 01 January, 2025

Editor-in-Chief
IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com