MedAId

Medical Q&A

MED277 Final Project

Jay Sawant, Maria Tavares & Mayukh Sharma

Table of contents

01 Introduction /Background

02 Datasets & Methods

03 Results

04 Discussion & Conclusion

Introduction/Background

Healthcare needs

Health inequity and health care barriers

Q&A systems

Break down barriers like limited health literacy and language

Objective

Build a multimodal and multilingual medical QA system to process speech, text, and medical images to democratize information

Purpose

Empower patients by providing more accessible and personalized medical information

Datasets

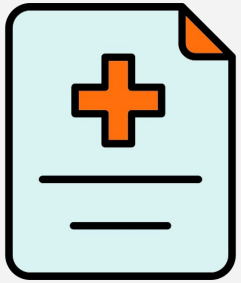
PubmedQA - expert labeled answers to 1000 queries

pubid int32	question string	context sequence	long_answer string	final_decision string
25,429,730	Are group 2 innate lymphoid cells (ILC2s …	{ "contexts": ["Chronic rhinosinusitis (CRS) is a…	As ILC2s are elevated in patients with…	yes
25,433,161	Does vagus nerve contribute to the…	{ "contexts": ["Phosphatidylethanolamine…	Neuronal signals via the hepatic vagus…	yes
25,445,714	Does psammaplin A induce Sirtuin 1-dependent…	{ "contexts": ["Psammaplin A (PsA) is a natural…	PsA significantly inhibited MCF-7/adr…	yes
25,431,941	Is methylation of the FGFR2 gene associated…	{ "contexts": ["This study examined links between DNA…	We identified a novel biologically…	yes

Datasets

MIMIC-IV-Note: Deidentified free-text clinical notes

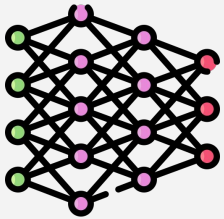
Radiology reports (2.3M)
from 237K patients



Discharge summaries (331K)
from 145K patients

Vector Embedding Models

We researched two different types of Embedding models



BiomedNLP-BiomedBERT-base
(110M params) by Microsoft

Context window: 512 tokens

Trained on abstracts and full-texts
from PubMed and PubMed Central

“Text-embedding-ada-002”
by OpenAI

Context window: 8192 tokens

Trained on Internet perhaps?

Knowledge-base preparation

Long discharge summaries are condensed into shorter summaries

3K (from 331K) discharge summaries

Average length ~ 10K characters per note

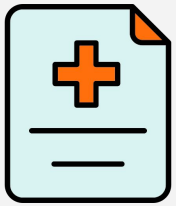
3K (from 331K) shortened discharge summaries

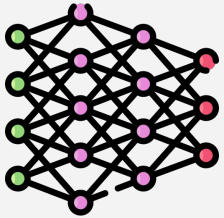
Average length ~ 700 characters per note

Knowledge-base preparation

3K shortened discharge summaries

Avg chars ~ 700





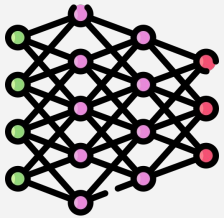
BiomedBERT model

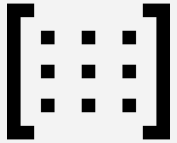
768 dimensional vector embeddings

Used for text-based retrieval

50K (out of 2.3M) Radiology reports

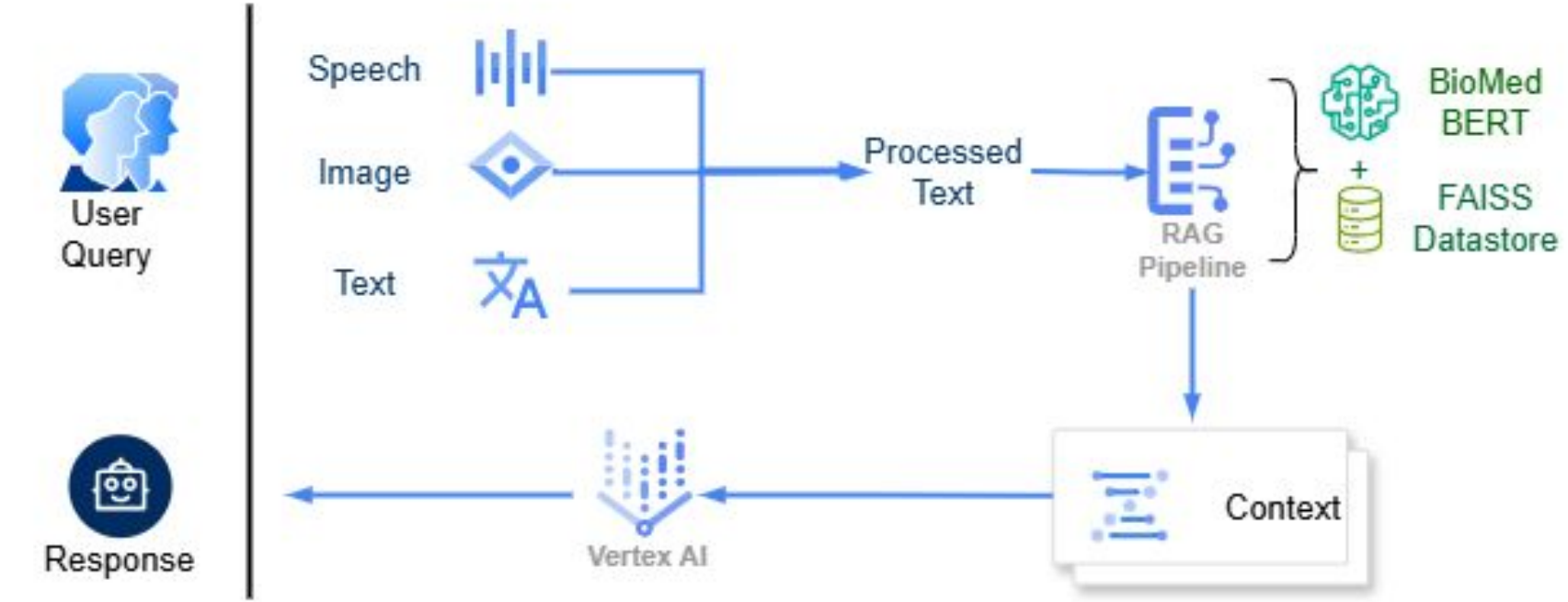
Avg chars ~ 1K per report





Used for image-based retrieval

Architecture: MedAid



Prompt Structure

"""
Based on the following information, provide a clear and accurate response to the last question in the conversation.

{preliminary_findings (if image is present)}

{medical_history_section}

{retrieved_contexts_section}

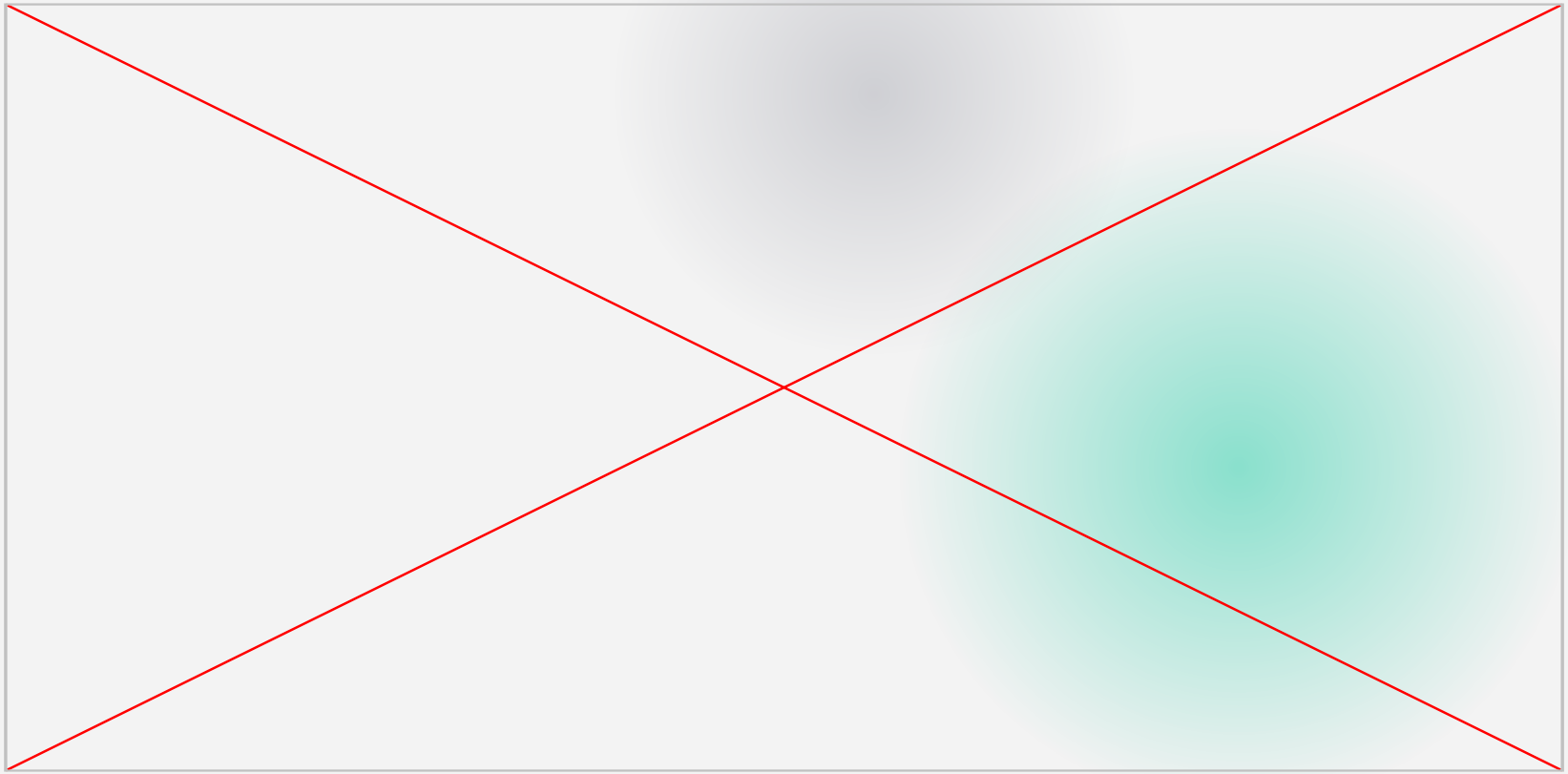
Conversation History:
user: query 1
assistant: Response 1
user: query 2
"""

Results (RAG-based responses)

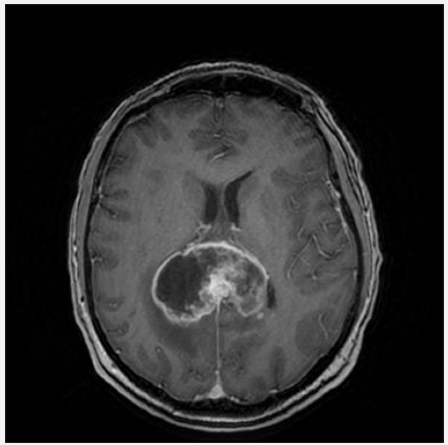
PubMedQA Dataset (expert labeled)	PubMedBERT embeddings	OpenAI embeddings
Rogue1	0.2758	**0.3213**
Rogue2	0.0774	**0.0939**
RogueL	0.1848	**0.2061**
BLEU	0.0256	**0.0294**

LLM: Gemini-1.5-flash Temperature 0.1 Number of retrievals(K) = 5

MedAId Demo: X-ray image and speech



MedAId Multimodal Example: MRI



Ground truth diagnosis:
Glioblastoma
(subtype of Gliomas)

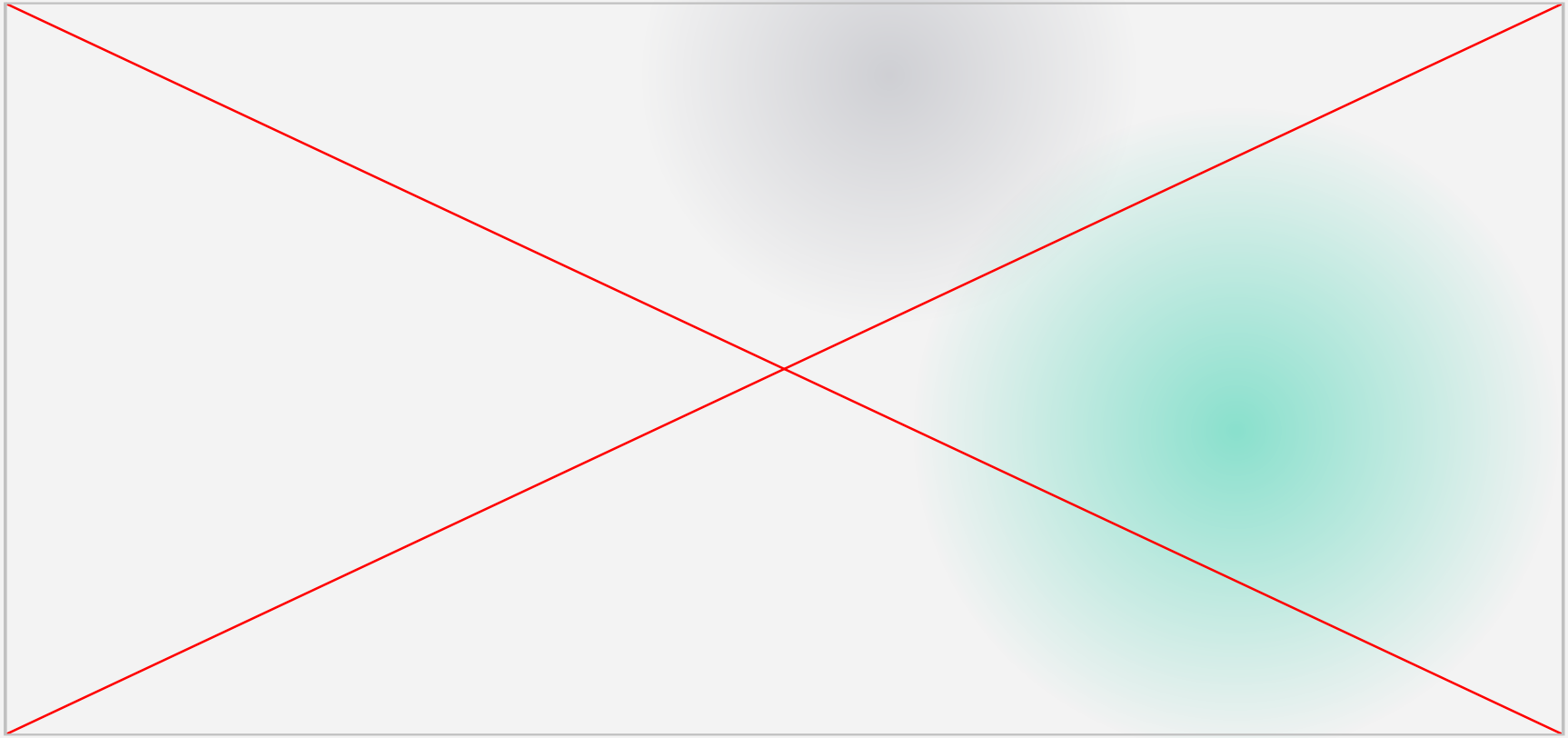
What abnormalities do you see in this brain MRI?

The axial T2-weighted MRI shows a posterior fossa mass, likely a cerebellar lesion.

What are the potential diagnoses based on these findings?

Potential diagnoses include:
Meningioma
Metastatic tumor
Gliomas
Hemangioblastoma
Ependymoma
Abscess

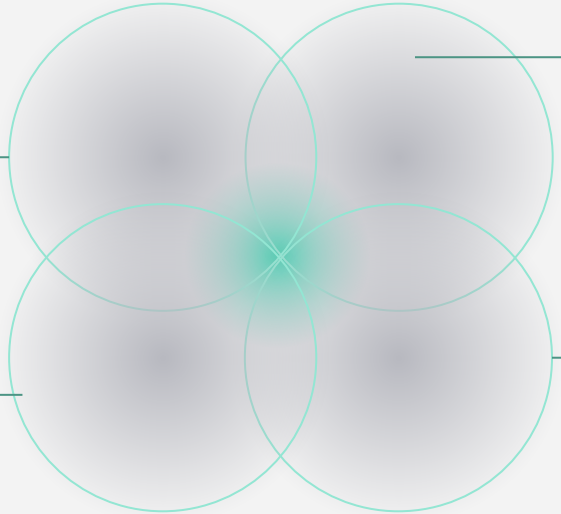
MedAId Demo: multilingual



Grading: metrics

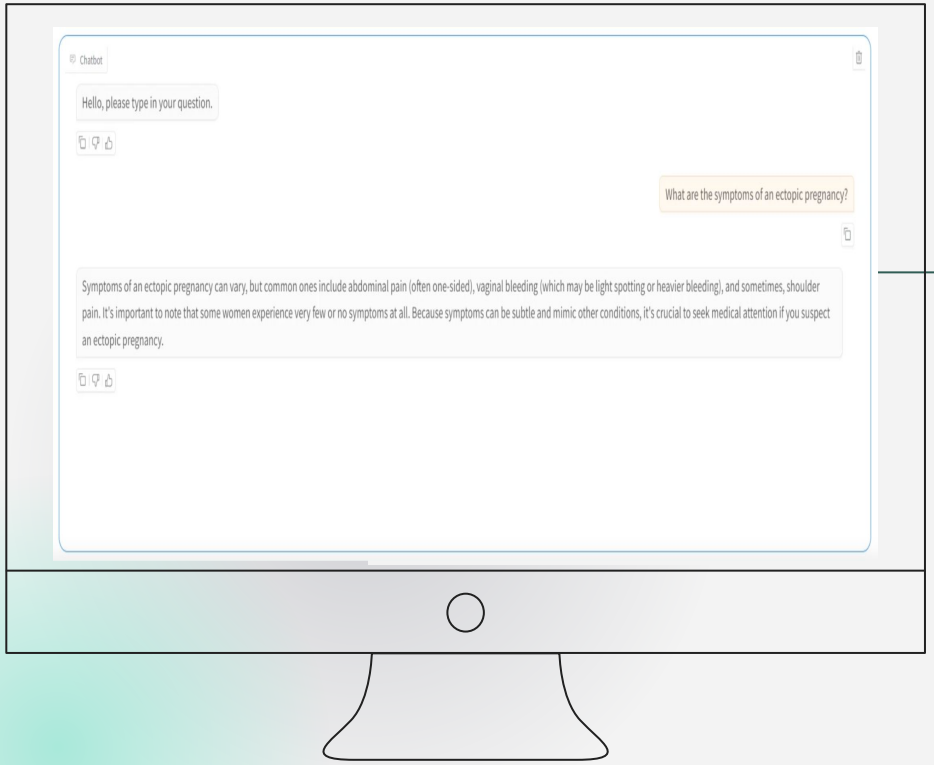
Medical accuracy

Clarity and Readability



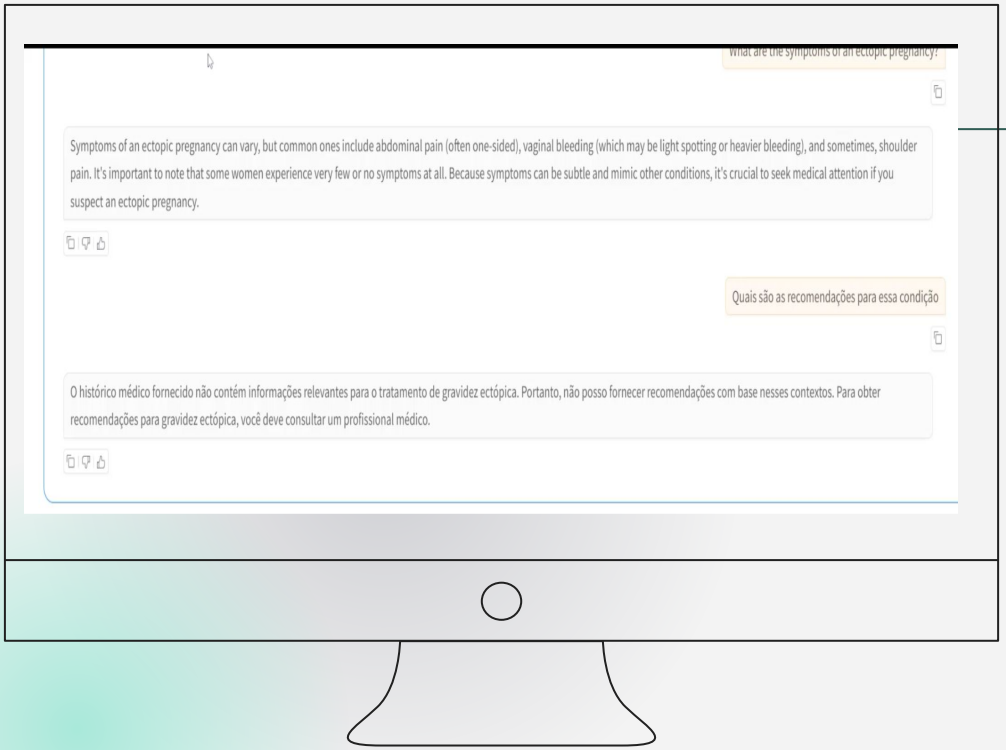
Comprehensiveness

Ethics and Tone



Results:Grading(EN)

Medical professional 1:	**9**
Medical professional 2:	**9.12**
Medical professional 3:	**8.75**
LLM:	**9.25**



Results:Grading(PT)

Medical professional 1: **8**
Medical professional 2: **7.5**
Medical professional 3: **7.2**
LLM: **8.5**

Discussion

By integrating multimodal and multilingual inputs and addressing accessibility barriers, this system demonstrates the potential to enhance both patient outcomes and clinical efficiency

Strengths

- Multimodal capabilities (text, speech, images) enhance flexibility and usability
- Bridges accessibility gaps for linguistically diverse groups
- Supports patients with tailored, context-rich responses

Limitations

- Budget constraints
- Compute constraints

Future Work

1. **Enhanced Embeddings**: Leverage advanced embeddings like OpenAI embeddings for improved performance and relevance.
2. **Vision-Language Models (VLMs)**: Incorporate medical imaging-specific VLMs (e.g., LlavaMed) for superior preliminary findings and retrieval accuracy.
3. **Comprehensive Use of MIMIC-IV**: Utilize the full potential of MIMIC-IV with extended context capabilities offered by OpenAI embeddings.
4. **Expanded Knowledge Base**: Augment the MIMIC-IV dataset with additional high-quality datasets to enhance model understanding.
5. **User Experience Enhancement**: Focus on refining the interface for a more intuitive and seamless user interaction.

Summary/ Conclusion

Our multimodal and multilingual QA system represents a significant step toward improving healthcare accessibility and decision-making by combining speech, text, and medical imaging into an intuitive and reliable tool

Thank you!
Questions?