SymptoCare: A Speech Recognition and Translation System for Medical Communication

Talha Ahmed, Syed Muqeem Mahmood, Faizan Ali, Maryam Shakeel, Nehal Ahmed Shaikh, Ubaid Ur Rehman

May 6, 2024

Background

Language barriers in healthcare are a significant challenge that can lead to misunderstandings, misdiagnosis, and even incorrect treatment plans. This is particularly problematic in multicultural societies or regions where patients and healthcare providers may need to share a common language. Traditional translation methods, such as human interpreters or basic translation software, often fall short. They can be slow, leading to delays in urgent care situations. They can also be inaccurate, especially when dealing with complex medical terminology, potentially leading to severe consequences for patient care.

Introduction

Effective patient-provider communication is vital, especially with diverse linguistic backgrounds, posing risks for refugees, immigrants, and underserved areas. Our proposed solution is an SMTS (Speech-to-Machine Translation System), integrating speech recognition, machine translation, and text-to-speech technologies to bridge language gaps in healthcare. Development involves database selection, LLM fine-tuning, and RAG pipeline utilization, enabling real-time access to medical advice and information.

Methodology

The system receives a user's query as an audio input (Figure 1). The query is transcribed into English via STT module. A pre-trained LLM is employed to fetch medically accurate diagnosis as responses to patients symptoms. Additionally, this LLM is fine-tuned on medical documents. The English text response is translated into user-specific target language and then converted to audio using TTS module. The final output is played for the patient allowing them to receive an expert diagnosis.
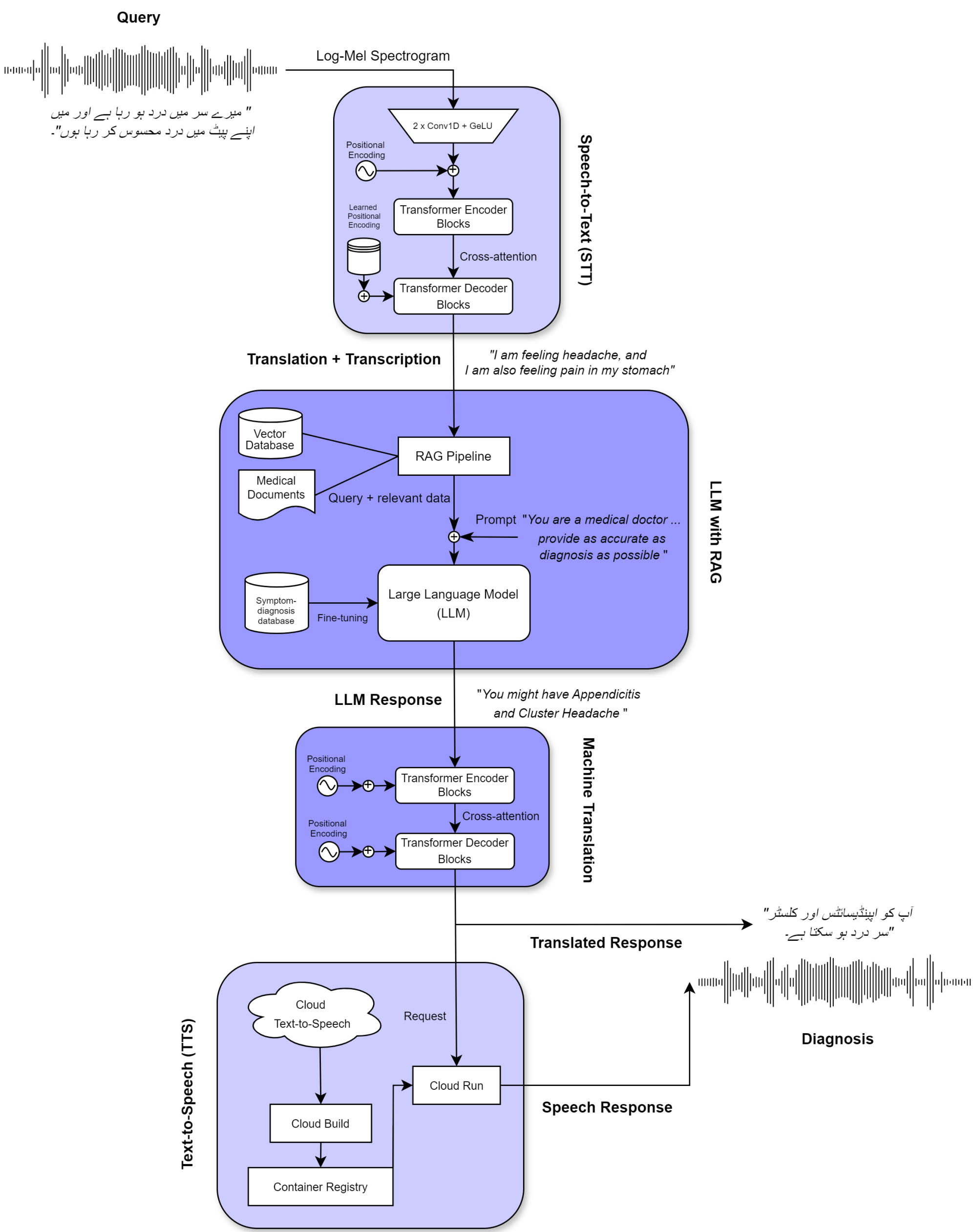

Figure 1. Proposed System Architecture

Tools & Techniques

To achieve our objective of generating expert diagnosis, we have utilized state-of-the-art tools and methods currently employed in the industry. An exhaustive list of is provided in Table 1.

Table 1. Tools & Techniques for Medical Diagnosis

Category	Tools & Techniques
Vector Database	Chroma DB
Large Language Models	Mistral (7b bnb 4bit & 8x7b-instruct-v0.1), Nous (hermes2)
Datasets	MeDAL, Disease-Symptom Dataset, Disease Symptoms & Patients Profile DatasetShaip, MS^2
Fine-tuning LLMs	Replicate, GradientAI, Unsloth
RAG Pipelines	LlamaIndex
Prompt Engineering	Role Assignment
Speech-to-Text (STT)	FasterWhisper
Machine Translation	MarianMT
Text-to-Speech (TTS)	gTTS
Deployment	Gradio

Conclusion

We successfully built and deployed our proposed SMTS breaking linguistic barriers in healthcare. We integrated cost-effective solutions like Chroma for vector databases and LLMs such as LLaMA and Mistral. MeDAL dataset improved medical abbreviation disambiguation. FasterWhisper enhanced transcription efficiency, MarianMT ensured accurate and swift machine translation, and GTTS provided simple text-to-speech conversion. Gradio Audio addressed real-time recording challenges. Future plans include further LLM fine-tuning and multi-language expansion.

SBA School of
Science & Engineering

SBASSE
Department of Computer Science

SBASSE
Department of Electrical Engineering