Altriva: Empowering wellness by blending tradition and innovation for a healthier you.

Mohith Sai Yandra
dept. of Data Science
University of North Texas
Texas, USA
Yandramohithsai@gmail.com

Sai Manikanta Kasireddy
dept. of Data Science
University of North Texas
Texas, USA
saikasireddy1999@gmail.com

Sri Sai Durga Katreddi
dept. of Data Science
University of North Texas
Texas, USA
katreddisrisaidurga@gmail.com

Saketh Kaveti
dept. of Data Science
University of North Texas
Texas, USA
Sakethchintu4u@gmail.com

***Abstract*—In the era of increasing health awareness, and demand for personalized care, integrating a diverse form of medical system has become essential. This study introduces Altriva, a chatbot designed to suggest alternative medicines and provide personalized recommendations, educational content, and symptom analysis for users based on its extensive knowledge of Ayurveda, homeopathy, and allopathy.**

Accessing timely medical care is difficult, with long appointment waits, often forcing people to delay treatment for minor health concerns. To address this problem, We developed a chatbot offering natural and home remedies to users. The chatbot is developed using LLama 3.1 and Langchain, facilitating efficient natural language processing and integrating diverse medical information. Interacting with Altriva, users will be able to receive responses suggesting treatments from Ayurveda, allopathy, and homeopathy.

The traditional medical system has notable disadvantages, particularly regarding the painful nature of diagnosis and treatment. Ayurveda in contrast offers a holistic approach that emphasizes the natural remedies that help alleviate symptoms with minimal discomfort. The knowledge base this Altriva gives responses from is a large set of research papers, which are published by authorized healthcare professionals, to ensure safety and give prompt responses.

***Index Terms*—Altriva, Chatbot, Alternative Medicine, Personalized Recommendations, Ayurveda, Homeopathy, Allopathy**

I. Introduction

In recent times, AI in healthcare emerged up and boomed into prominence because it is viewed as improving patient care and promoting accessibility to health information. Equipped with NLP and machine learning algorithms, chatbots are turning out to be a powerful tool for providing real-time health advice and information. Altriva is one of those innovations centered on alternative medicine practices, including the two holistic systems of Ayurveda and homeopathy that delve into treatment and prevention on an individual basis.

Ayurveda, with its roots in ancient Indian tradition, allows for personalized health care by integrating lifestyle, diet, and environmental factors into treatment. Homeopathy is a system of medicine based on the principle that highly diluted substances can treat patients, possibly by stimulating the body's natural healing processes. Access to dependable guidance through these alternative medicine systems is sometimes quite inadequate due to the lack of qualified professionals on the one hand and knowledge specific to the context on the other. Altriva is an AI-driven chatbot that will be proposed in the subsequent sections of this research, providing personalized recommendations from both the Ayurvedic and homeopathic perspectives for user inputs regarding specific health symptoms or concerns.

The key contributions of this work are the assessment of Altriva's efficacy in providing accurate, relevant, and user-friendly responses compared to other medical chatbots. Altriva is a fine-tuned large language model for Ayurveda and homeopathy where embeddings are learned to utilize large datasets for better response quality. This work describes the design and implementation of Altriva, from its underpinning architecture to data collection and evaluation metrics.

This study adds to the growing literature related to AI in healthcare and it shows how AI-enabled chatbots have the potential to be informative assistants in both traditional and alternative medicine practices while fostering a more holistic mode towards patients. These results will provide valuable insight into how this AI technology has been effective in increasing access to Ayurvedic and homeopathic health advice as a method of developing an integrated healthcare model.

II. Related Work

The integration of AI chatbots into healthcare is proliferating. Altriva is built upon various foundational works which are incorporated using advanced natural language processing (NLP) techniques along with domain specific training focusing on retrieval strategies while delivering the reliable responses in the alternative medicine. There are various researches which are exploring the effectiveness, limitations and applications of AI in the medical context. One significant contribution in this field is the study "Chatbots in healthcare" [9] which produced a chatbot that illustrates the potential of these systems supporting healthcare professionals along with patients. As per this paper, the privacy concerns and accuracy remain while the chatbots can reduce costs and improve accessibility to care. This research enhanced the context of regulatory healthcare by [4]. Kim developed a QA-RAG model for regulatory compliance in pharma sector. They combined Gen AI with retrieval mechanisms which demonstrates a potential RAG setup in managing specialized, regulated content effectively. Altriva is built on this concept while optimized to retrieve alternative medicines while retrieving reliable content from Ayurvedic, and Homeopathic sources. While QA-RAG focused on precision, and regulatory adherence, Altriva will present sourcing to medically reliable information while presenting it to user in friendly language. They developed a model which combines Gen AI with retrieval techniques, which showed better accuracy in retrieving relevant information. This is mainly useful for domain specific retrieval system where accuracy and relevance of the response plays a major role. This method uses a user query and a hypothetical answer suggesting a promising framework for retrieving reliable information. This aligns Altriva's aim of presenting accurate responses based on verified sources.

[3] OpenAI privacy policy provides some guidelines on how ethical data handling should be done as it is critical in AI applications within domain specific applications like healthcare. The openAI privacy standards has been prioritized while implementing Altriva to process the queries securely and without risk of unauthorized data exposure. This helps in safeguarding user trust and aligns with ethical requirements.

[10] Krizhevsky has introduced a significantly advanced IR in healthcare applications which uses CNN based model and AlexNet that set a milestone in Computer vision to achieve state-of-art accuracy on ImageNet dataset. This work's influence extends into healthcare applications, while inspiring deep learning architectures. Similar kind of deep learning techniques are used in Altriva like utilizing embeddings, and attention mechanisms that helps to understand complex medical terminologies across Ayurveda, and Homeopathy. This creates a bridge between complex inputs and accurate , context-aware responses.

Another notable model is ChatDoctor [6]. This model finetunes Meta's LLaMA on a dataset which have patient-doctor interactions. This helps improve response accuracy in a medical context. While effective, the drawback lies where the data privacy and validation requirements concerned. Altriva will benefit from the same kind of finetuning that helps in enhancing the performance in the allopathic, ayurvedic, and homeopathic context. This addresses specific knowledge gaps with medically reliable datasets. While this lacks in privacy, a PMC-LLaMA fine-tuned model on 4.8 million medical papers was developed [7]. There are medical QA performance benchmarks like PubMedQA, which showcase how the domain-specific data can refine chatbot capabilities. Altriva's reliance on research papers and other verified medical sources for training aligns with PMC-LLaMA's approach, enhancing its potential to deliver reliable, context-specific medical advice.

The author, [8] developed the LLM-AMT system. This integrates medical textbooks within a RAG framework. The models such as Hybrid Textbook Retriever and Knowledge Self-Refiner are used to reduce hallucinations and risks, which helps in enhancing the accuracy of the medical responses. This can be used to ensure high accuracy and trust in the advice it provides.

While developing Altriva, there are several foundational works like Mohler [2] emphasizing the importance of quality reporting in the complex healthcare systems and proposed the CReDECI list which helps to enhance transparency and also applicability in the medical research publications. Altriva followed this approach to sourcing from verified and high quality healthcare documents, while ensuring that the users receive reliable information in response to the health enquiries. Additionally, Chen [5] introduced vectorized context, which advanced the field of retrieval Augmented Generation(RAG). This improved the relevance of open-domain question-answering enhancing response accuracy. This application helped Altriva reduce hallucinations in alternativemedicine context. [9] Moreover, Božić mentioned the role of healthcare chatbots in reducing clinician workloads, while highlighting their potential to support both healthcare providers and patients. This aligns closely with the Altriva's objective to serve as an AI-driven assistant, which is accessible particularly in regions where qualified professionals and knowledge of alternate medicines may be scarce. SynaergyAltriva helps to alleviate the pressure on practitioners, inturn fostering patient autonomy in managing minor health concerns.

III. Methodology

A. Overview

This study established a process for building a RAG-based chatbot model that addresses specific user needs like, 24/7 availability of medicinal data, and instant and reliable answers to questions specific to domains ayurveda, homeopathy, and allopathy. This RAG-based approach reduces hallucinations, by enabling accurate retrieval of information, and by ensuring responses that are trustworthy and relevant. So, in industries where accuracy is critical like healthcare, this chatbot will be ideal. Users always want responses considering past interactions or inquiries making the conversation feel less robotic. We maintained conversational context and retrieved personalized information, Altriva can provide nuanced answers that were built on previous chats. Regardless of the topic's complexity, users can always seek expert insights over a range of topics by helping reduce the user to search for correct data in multiple sources. Altriva also cites the specific source, where the data originates, ensuring transparency and fostering trust on it. Altriva can retrieve targeted information from extensive verified sources for getting detailed answers quickly to facilitate learning.

IV. System Architecture

Altriva is composed of five main components, as shown in Figure 1, each specifically designed to collaborate effectively in delivering a seamless, reliable, high-quality user experience. This architecture forms the backbone of the system, by ensuring optimal performance.

A. Dataset

The sample dataset Altriva uses includes research papers from Google Scholar published by authorized professionals. This dataset will be compiled from credible sources, including treatment protocols, patient case studies, and scholarly articles. This dataset will form the knowledge base for Altriva's recommendations. As the data given in research papers is considered structured, to facilitate efficient handling and parsing of PDF files, such as research papers, `PyPDFLoader` is utilized. This can efficiently extract content from each page, thus converting it into a retrievable format. Splitting the documents into chunks will improve retrieval efficiency as the retriever focuses on contextually relevant portions of text, providing quality outputs.

As each chunk captures focused content this approach will enable more accurate matching of text thus improving relevance in response generation. The documents are split into manageable text chunks with a slight overlap to maintain continuity between sections to maintain and balance the quality of the data retrieved and efficiently compute it.

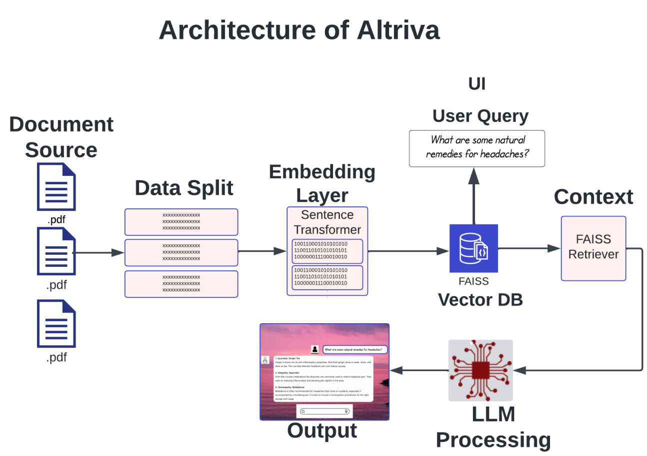

Fig. 1. Altriva architecture showcasing the workflow from document ingestion to personalized response generation through FAISS-based retrieval and LLM processing.

B. Embedding Model Selection

HuggingFaceEmbeddings with the sentence-transformers/all-mpnet-base-v2 model is selected for this specific task as the model is known for its semantic understanding, making it ideal for generating contextually rich embeddings of the text data. It generates semantic-embeddings that capture contextual meaningwhich is ideal for medical question-answering. The accuracy of this model in representing sentence-level embeddings made it an obvious choice for question-answering tasks. Choosing embeddings to be CPU, as in real world, the response time is less critical compared to how much impact this application has on the lives of people making it cost-effective and more accessible to the people, enabling Altriva to run on widely available hardware.

C. Embedding Storage

Facebook AI Similarity Search is chosen to be ideal for retrieving embeddings in a large dataset, as it provides fast and scalable similarity search and clustering. The quick retrieval based on vector similarity will help in optimizing the search by indexing. It is a highly optimized vector database for handling data on large scales and performs well with high-dimensional embeddings, by supporting fast and accurate similarity-based information retrieval. FAISS's efficiency in handling high-dimensional embeddings ensures that the Altriva provides accurate responses.

Embeddings are stored in FAISS as dense vectors called chunks. By using optimized indexing techniques, like IVF (inverted file), quantization enhances search accuracy and efficiency. Finding the nearest embeddings to the query vector, FAISS performs similarity-based retrieval by finding the nearest embedding to the query vector by addressing the challenge that arises when retrieving semantically similar documents, which is essential for accurately answering questions in Altriva. To ensure focused document retrieval, the vector store is converted to a retriever to match queries and embeddings narrowing down relevant text chunks in response to queries. Vector normalization and relevance thresholds are used to ensure accurate and meaningful retrieval. It tunes similar thresholds, where irrelevant documents are filtered out, thus improving the quality of responses.

D. LLM Integration

Ollama LLM Model (llama3.1) is chosen to implement Altriva, considering its high performance, in terms of generating accurate and context-aware responses. Altriva is an expert in understanding and generating human languages for seamless customer interactions. It enhances the RAG setup by providing contextually relevant answers built on retrieved documents.

RetrievalQA combines LLM response generation with document retrieval which enables the Altriva to retrieve relevant content and generate precise answers ensuring the enhancement of the Altriva's capacity for accurate, knowledge-based responses. To create coherent input, the chain type "stuff" is used for Altriva, as it ensures that only the most relevant content is passed to the LLM.

1) Retrieval and Embedding Process: To improve response accuracy, Altriva uses similarity-based retrieval with FAISS. When a user query q is submitted, the system computes the similarity S between q and each document chunk d_i in the FAISS index based on equation (1). The similarity score is calculated as follows:

$$S(q, d_i) = \frac{\sum_{k=1}^{n} q_k \cdot d_{i,k}}{\sqrt{\sum_{k=1}^{n} q_k^2} \cdot \sqrt{\sum_{k=1}^{n} d_{i,k}^2}} \tag{1}$$

where:

- q_k and $d_{i,k}$ are the components of the query and document vectors.
- n is the dimensionality of the embeddings.

E. User Query Handling and response generation

The query processing loop allows in continuous interaction, which lets users to input multiple queries, ensuring the responsiveness of the chatbot and its flexibility ensuring user engagement as shown in Figure 2. Handling edge cases (queries with no relevant information) the system is configured to recognize the low similarity returning an informative message if no relevant information is found, preventing irrelevant responses to the user.

Altriva retrieves the most relevant content, then passes it through LLM for contextual understanding, and outputs and synthesized response to the user, by maintaining clarity and relevance, taking user feedback or query logging will help Altriva to refine the responses, and adjusting the retrieval and similarity thresholds based on observed patterns in user interactions.

V. RESULTS

The following table summarises the responses generated by Altriva to sample queries which showcases its ability to provide context by condition identification, treatment options, lifestyle modifications, and advice.

The table I and II displays a detailed summary of Altriva's response for GRED and diabetes related question covering both conventional and herbal treatment options along with recommending practical lifestyle changes. For instance, Altriva responds saying the use of antacids and H-2 receptor blockers for the query related to GRED for immediate relief, emphasizing the importance of consulting a professional healthcare provider for severe symptoms. It also provides lifestyle adjustments like dietary changes and bed elevation to manage symptoms effectively demonstrating Altriva's ability to merge evidence-based advice with holistic approaches.

By comparing these two cases, we can observe that Altriva's adaptability across different medical conditions along with reliable and condition-specific responses are ensured.

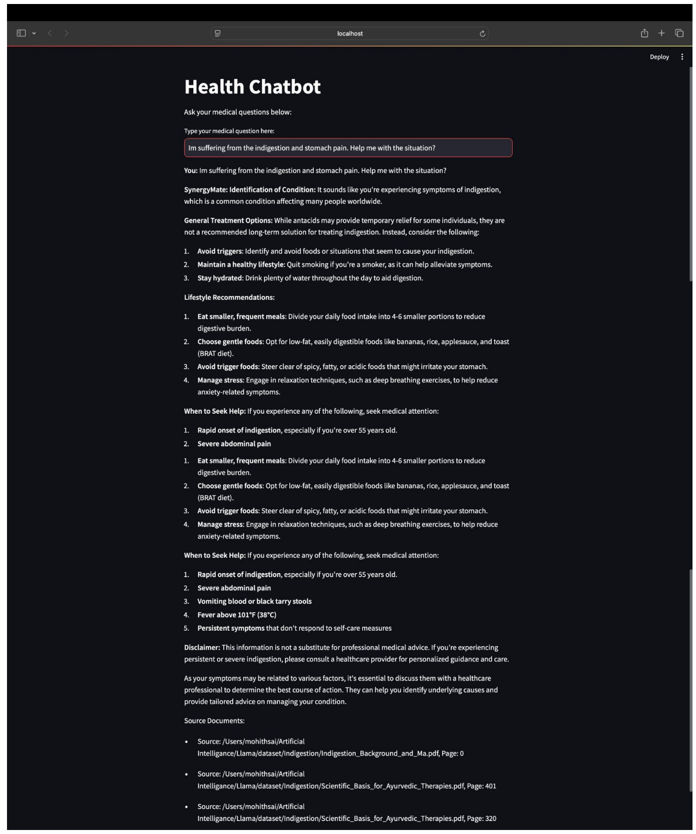

Fig. 2. Sample Responses from Altriva AI-Powered Medical Assistant This figure showcases Altriva's interactive interface, highlighting its ability to deliver context-aware and user-friendly responses across Ayurveda, Homeopathy, and Allopathy domains. The screenshots illustrate the chatbot's retrieval of reliable medical advice based on user queries, demonstrating its practical application in alternative healthcare guidance.

TABLE I
ALTRIVA RESPONSE SUMMARY - GERD

Aspect	**Description**
Condition Identification	Describes GERD and hyperacidity symptoms.
Treatment Options	1. Antacids (e.g., sodium bicarbonate, calcium). 2. H-2 receptor blockers (e.g., cimetidine). 3. Herbal remedies (e.g., psyllium, turmeric).
Lifestyle Recommendations	Avoid trigger foods, elevate bed, wear loose clothing.
Advice	Consult professionals for severe symptoms.
Disclaimer	Not a substitute for medical advice.
Source Documents	1. Scientific_Basis_for_Ayurvedic_Therapies.pdf, p. 349 2. Encyclopedia_of_Medicine.pdf, pp. 321, 231

TABLE II
ALTRIVA RESPONSE SUMMARY - DIABETES

Aspect	Description
Condition Identification	Type 2 diabetes indicators.
Treatment Options	1. Sulfonylurea for insulin. 2. Biguanides for liver glucose production. 3. Thiazolidinediones for insulin sensitivity.
Lifestyle Recommendations	Regular physical activity, balanced diet, stress management.
Advice	Consult healthcare provider for severe symptoms.
Disclaimer	For general guidance only.
Source Documents	1.DIABETES_AND_ANTI-DIABETIC_HERBAL_FORMULATIONS.pdf, pp. 4, 10 2. Encyclopedia_of_Medicine.pdf, p. 276

VI. FUTURE WORK

One important extension this project includes, integrating the ability for the users to attach their health records, generate personalized response using advanced document processing techniques, and addressing the security concerns using NVIDIA Nemo-Guardrails.

[1] Zhang, conducted a comprehensive survey on deep learning models in sentiment analysis. This covers architectures like CNN, RNN, LSTM to understand language sentiment and context. This research shows the importance of nuanced understanding in NLP tasks. Specially when responding the human health related queries, which can be used in Altriva to use LLM that is fine-tuned to discern sentiment and context integrating audio for natural conversation with chatbot, while ensuring that the responses are empathetic, medically appropriate and adjusted to the user's emotional tone.

Future integration of other alternative medicines like popular search engines for better retrieval of latest and comprehensive data. We can establish feedback loops with healthcare professionals for real-time validation can be done through video access. Enabling image input for personal assessment for disease understanding like skin, posture chronic conditions.

A multi-lingual Cultural sensitivity layer can be added to ensure wide range of audience can access this. Advanced machine learning techniques can be used for sentiment analysis on user queries. A predictive health calender which can predict potential future health needs based on user health history and seasonal trends, using NLP, emotion detection for mental well being based on the inputs can increase the capability of Altriva by providing truly personalized forward-thinking healthcare experiences.

REFERENCES

[1] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1253, 2018.

[2] R. Möhler, G. Bartoszek, and G. Meyer, "Quality of reporting of complex healthcare interventions and applicability of the CReDECI list-a survey of publications indexed in PubMed," *BMC Medical Research Methodology*, vol. 13, pp. 1-8, 2013.

[3] "Privacy policy," OpenAI. [Online]. Available: https://openai.com/policies/privacy-policy. [Accessed: 31-Mar-2023].

[4] J. Kim and M. Min, "From RAG to QA-RAG: Integrating Generative AI for pharmaceutical Regulatory Compliance Process," *arXiv*, vol. 2402.01717v1 [cs.CL], 26-Jan-2024. Available: https://arxiv.org/pdf/2402.01717.

[5] Z. Chen, X. Wang, Y. Jiang, P. Xie, F. Huang, and K. Tu, "Improving Retrieval Augmented Open-Domain Question-Answering with Vectorized Contexts," *arXiv preprint*, arXiv:2404.02022, 2024.

[6] Y. Li, Z. Li, K. Zhang, R. Dan, S. Jiang, and Y. Zhang, "Chatdoctor: A medical chat model fine-tuned on a large language model meta-ai (LLaMA) using medical domain knowledge," *Cureus*, vol. 15, no. 6, 2023.

[7] C. Wu, X. Zhang, Y. Zhang, Y. Wang, and W. Xie, "PMC-LLaMA: Further finetuning LLaMA on medical papers," *arXiv preprint*, arXiv:2304.14454, vol. 2, no. 5, p. 6, 2023.

[8] Y. Wang, M. Xueguang, and C. Wenhu, "Augmenting black-box LLMs with medical textbooks for clinical question answering," *arXiv*, arXiv:2309.02233, 2023.

[9] V. Božić, "Chatbots in healthcare," 2023.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012.