

Enhancing Healthcare Management: A Case Study of a Medical Chatbot in Egypt

Shimaa Ismail, Lina A.Salem, Noha E.El-Attar and Tarek A.Elshishtawy

Information Systems Dept., Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt

Email: Shimaa.mustafa@fci.bu.edu.eg

Abstract

In hospitals, receptionists are generally responsible for the following functions: greeting and answering inquiries from visitors, providing them with the appropriate information, and ensuring that hospital staff and patients receive timely and professional communication. This paper presents the implementation and deployment of a medical chatbot designed to replace the traditional receptionist role in hospitals where their visitors speak Arabic. The proposed case study here is an Egyptian hospital. The user can ask questions in text and the answers can be text or voice. The presented chatbot utilizes the power of the Generative Pre-Trained Transformers (GPT-4), which represents one of the most powerful large language models (LLMs) available to generate text. This model is merged with prompt engineering capabilities for fine-tuning specific tasks or instructions. This merger has gained traction to enhance model performance and adaptability. The model is integrated with an SQLite database to provide immediate information to patients about doctor availability, examination costs, hospital policies, and more. The chatbot demonstrates a significant potential to streamline hospital operations, improve patient satisfaction, and reduce administrative workload. The evaluation shows a 99% accuracy rate, indicating the high reliability of the system.

Keywords Chatbot, GPT-4, Question Answering Bot, Automatic Hospital Receptionist, LLM, Prompt Engineering

1. Introduction

The word “chatbot” consists of the terms “chat” and “robot.” The word “chatbot” was first used to describe a computer program that uses a text-based dialogue system to mimic human speaking [1]. Chatbots have been used since the 1960s when Weizenbaum [2] developed the very first chatbot system for psychiatric reasons. Followed by summarizing the impact elements and findings on social, psychological, and behavioral aspects of user-chatbot contact. Chatbots could make multiple social interactions with humans by identifying keywords and generating responses according to a set of pre-programmed rules. In the beginning, the meaning of the sentences (the semantics) and the grammatical structure (the syntax) were not understood, but an illusion of understanding was given. Lately, chatbots are changing the patterns of interaction between humans and computers. By assisting users in solving problems that require multi-step solutions as an assistant and support in a specific task. Nevertheless, the extension of chatbots to more domains has led to more complex designs and architectures.

To make chatbots understand human natural language, an artificial intelligence model should be used to facilitate interaction with humans through text or voice. One of the generative models that are frequently used to generate text is the Large Language Model (LLM). LLM is a deep learning algorithm that can perform a variety of natural language processing (NLP) tasks such as text generation, translation, summarization, and question-answering. It is also referred to as neural networks (NNs), which are computing systems inspired by the human brain. These neural networks function via a layered network of nodes, just like neurons do. LLMs are not only capable of teaching artificial intelligence (AI) applications to human languages but they can also be trained to do a wide range of activities, such as understanding protein structures and generating software code [3]. They need to be pre-trained and then fine-tuned to tackle challenges related to text categorization, question answering,

document summarizing, and text generation, just like the human brain. Their ability to solve problems can be used in industries including healthcare, education, research, and entertainment [4], where big language models are used for a range of NLP uses, including chatbots, AI assistants, translation, and more.

A major development in healthcare, the integration of LLMs holds the potential to revolutionize the handling, sharing, and application of medical data. These models, exemplified by state-of-the-art technologies such as OpenAI’s GPT-4, possess the capability to understand and generate human-like text with remarkable accuracy and contextual awareness [5]. Their application in healthcare can lead to improved patient interactions, streamlined operations, and enhanced decision-making processes, thus addressing several critical challenges faced by the industry. Li et al. [5] also emphasized the necessity of maintaining accurate and available datasets for AI applications to function effectively in healthcare. Any errors or outdated information can lead to incorrect responses, potentially affecting patient trust and care quality.

In recent years, the adoption of LLMs in healthcare has garnered considerable attention. LLMs can interpret complex medical terminologies, respond to patient queries, and provide preliminary medical advice with high accuracy. This functionality is particularly beneficial in clinical settings, where efficient communication can significantly impact patient outcomes. Vito [6] highlighted the potential of LLMs to augment healthcare professionals by assisting in patient triage, symptom checking, and administrative tasks, thereby freeing up valuable time for medical staff to focus on critical care.

One of the most compelling advantages of LLMs in healthcare is their multilingual capability. In diverse and multicultural societies, language barriers can significantly hinder access to quality healthcare. LLMs can bridge these gaps by accurately translating and delivering medical

information in multiple languages, ensuring that all patients receive the information they need in a language they understand [7]. The impact of multilingual LLMs in improving patient engagement and inclusivity and highlighting their role in enhancing healthcare accessibility for non-native speakers is underscored, thereby enhancing patient satisfaction. Raiaan et al, [8] underscored the importance of language support in AI applications for global accessibility.

The deployment of a chatbot can lead to substantial cost savings for hospitals by decreasing reliance on human receptionists for routine inquiries. This allows human resources to be reallocated to more complex tasks that require personal interaction. Javaid et al, [9] discussed how AI-driven chatbots can result in significant cost reductions by automating administrative tasks, allowing hospitals to optimize their human resource allocation. Implementing and maintaining an AI-driven chatbot requires significant initial investment and ongoing maintenance, which can be a barrier for some hospitals. Lui et al, [10] pointed out the financial and technical challenges associated with integrating advanced AI solutions into existing healthcare infrastructures.

Moreover, the integration of LLMs into hospital systems represents a leap forward in the efficiency and quality of healthcare services. These models not only provide instantaneous responses to patient queries but also offer personalized interactions based on patient data and history. This capability is crucial to managing patient expectations and delivering tailored healthcare services. By utilizing patient data securely and effectively, LLMs can predict patient needs, schedule follow-ups, and even remind patients about medication adherence, thereby contributing to better health outcomes [11].

Despite the numerous benefits, the integration of LLMs in healthcare also presents challenges that need to be addressed. Concerns regarding data privacy, the accuracy of AI-generated medical advice, and the ethical implications of using AI in healthcare are critical areas requiring ongoing attention. However, advancements in AI technology and stringent regulatory frameworks are continuously improving the reliability and ethical deployment of LLMs. In 2024, researchers emphasized the importance of establishing robust data governance and ethical guidelines to ensure that LLMs are used responsibly and effectively in healthcare settings [12].

Customizing the LLM for a specific task has four patterns: prompt engineering, retrieval-augmented generation (RAG), fine-tuning an LLM, and training your LLM. In this paper, the prompt engineering pattern has been used to customize the aim of the chatbot. Prompt engineering is the art of crafting inputs that guide AI systems to generate texts and predictions. The input of an AI system is called a 'prompt', so basically, prompt engineering means creating the best input to get your desired output back. The presented prompt directs the model's response in two ways: a zero-shot or a few-shot setting, where the model generates responses without further training or with minimal examples. Finally, a well-engineered prompt ensures that the AI understands the

task, which leads to more aligned output with the expectations.

For the healthcare sector, utilizing large language models in the healthcare industry offers a remarkable opportunity to enhance patient care and streamline hospital operations. By automating routine administrative tasks, providing accurate multilingual support, and aiding in medical decision-making, LLMs are poised to become an essential part of modern healthcare systems.

To illustrate the usefulness and significant advantages of LLMs in a real-world healthcare setting, this article examines the deployment and development of a medical chatbot utilizing GPT-4 at an Egyptian hospital. Based on a review of the literature, conversations with patients, their families, receptionists, and patient coordinators, as well as knowledge of the Egyptian healthcare system, several crucial regulations have been established.

- **Usability and Availability:** In the beginning, when the user looks at the interface of the chatbot, it should be simple and suitable for everyone. It must be usable by patients, their attendants, doctors, and patient coordinators. It should be available 24/7.
- **Timeliness:** Prompt responses are important as delays often lead patients to seek information elsewhere. The main model rule is obtaining patient satisfaction. So, the responses of the model should be fast and accurate at the same time. Therefore, the next requirement is accuracy.
- **Accuracy:** The chatbot's capability to answer all types of questions that a receptionist answers is a fundamental requirement. So, using a suitable LLM might offer accurate information. It might not be tailored to specific healthcare systems or hospital-specific nuances. However, it can be justified to ensure accurate and contextual responses.
- **Security and Privacy:** If large language models are not adequately maintained or monitored, they pose significant security threats. They can create spam, take part in phishing schemes, and leak people's personal information. Users with malicious intent can reprogram AI to their ideologies or biases, which helps disseminate false information. Therefore, the query answers of patients or users shouldn't display all data records in the hospital database. They need specific authorizations for the available data.
- **Hallucinations Prevention:** When a LLM generates an output that is false or inconsistent with the user's purpose, it's called a hallucination. Saying it is human, that it feels emotions, or that it is in love with the user are a few examples. Large language models cannot fully grasp human meaning because they predict the next syntactically right word or phrase. Occasionally, the outcome is what's known as a "hallucination." The presented chatbot should prevent hallucination results for users.

The rest of the paper is organized as follows: Section 2 provides a literature review of relevant studies about the application of LLM in different fields. The model requirements and the integrations for building the proposed medical chatbot are discussed in section 3. Section 4 presents the gathered dataset and reports the experimental results. Lastly, the paper concludes by summarizing the findings and discussing potential avenues for future enhancements to the model.

2. Literature Review

LLMs were created expressly to understand and generate human language [13]. LLMs have made a substantial contribution to the field of artificial intelligence and have been used in a variety of contexts, including research, education, communication, entertainment, healthcare, content creation, and article composing [13;14]. The development of neural network-based approaches in the field of NLP is credited with giving rise to LLMs. Artificial neural networks (ANNs) were first proposed by Warren McCulloch and Walter Pitts in the 1940s [15]. Later, in the 1950s and 1960s, the first language model was launched, which developed the early neural networks and rule-based models. In addition, language processing was made easier by their use of precisely defined linguistic features and rules [8]. After transformers were introduced, scaling-up LLM models were created by increasing the number of parameters and training data in later years [14]. In 2018, OpenAI made a substantial contribution to the advancement of LLMs (a transformer-based architecture) called generative pre-trained transformers (GPT). GPT is a popular foundational model whose numbered iterations are improvements on their predecessors, as shown in Figure. 1. It can be fine-tuned to perform specific tasks downstream.

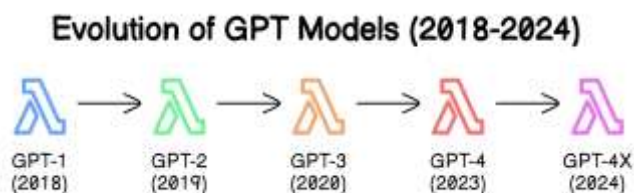


Fig. (1) Evolution of GPT Models (2018-2024)

GPT-1 needs simple fine-tuning and supervised training for use in downstream tasks. As Radford et al, [16] introduced a framework that uses discriminative fine-tuning and GPT-1 to provide robust natural language understanding using a single task-agnostic model. They improved the state-of-the-art datasets to successfully transfer and solve discriminative tasks like question answering, semantic similarity assessment, entailment determination, and text classification [17]. Based on GPT-1, GPT-2 had additional improvements such as adding more training data, enhancing the generalization ability, making small adjustments to the model structure, and training the requirement for supervised fine-tuning. After that, GPT-3 extended the parameter space and the data scale in 2020. In zero-shot and few-shot conditions, the model performs exceptionally well and can handle downstream tasks without the need for fine-tuning.

GPT-3 has produced great results on a variety of new tasks, including arithmetic addition, news article generation, language interpretation, and code writing [18], based on the multi-task generalization skills of GPT-2. The cost, model behavior, and privacy issues are some of the main issues in integrating this language model into the healthcare delivery process [19]. There are typically a lot of daily visits to clinics and hospitals. For each patient, an ideal system might submit several queries to produce relevant results. The implementation of GPT-3 may incur significant operating costs overall. In the healthcare industry, patient interactions are held to much higher standards. Therefore, it is necessary to assess the benefits and dangers before implementation. Lastly, the convention provides strong protection for medical records. However, asking GPT-3 for assistance would require sharing these data with OpenAI. To safeguard patient privacy, OpenAI and healthcare providers must adhere to the same set of regulations [20]. Handling sensitive patient information necessitates robust data privacy and security measures. There are inherent risks of data breaches and misuse of personal health information. Xu and Wang [21] discussed the critical importance of implementing stringent data privacy protocols to protect patient information in AI-driven healthcare solutions.

Lim and Schmäzle [22] investigated the use of LLMs like GPT-3 for generating personalized health messages. They compared the effectiveness of AI-generated messages with human-written ones, providing insights into the strengths and limitations of LLMs in health communication. They used prompt engineering to generate awareness messages about vitamins compared to the most retweeted human-generated messages. They showed that AI-generated messages ranked higher in message quality and clarity across both samples. Their findings ensured high accuracy in information retrieval and response generation, minimizing human errors and inconsistencies, and demonstrating how these technologies can enhance communication between healthcare providers and patients.

According to OpenAI, GPT-4 has much more knowledge and is capable of addressing problems. GPT-4 is said to be able to understand eight times as many words as the previous version, be harder to trick, comprehend visuals, and be less likely to reply to requests that are denied [23]. Karanikolas et al, [24] presented a hybrid model between LLM and Natural Language Understanding (NLU) / Natural Language Generation (NLG) systems. The presented model could filter irrelevant text. In addition, it could produce semantically correct information based on recorded events. The corrected information was converted into syntactically correct and relatively testable sentences with the help of NLG.

GPT-4 is used in different fields, such as tourism. Carvalho and Ivanov [25] explored the potential of ChatGPT in enhancing customer experiences. The research highlighted various applications, such as generating personalized travel recommendations, translating content for international audiences, and automating tasks like booking and itinerary planning. The authors also addressed the benefits, including improved customer satisfaction and operational efficiency, alongside risks like bias and misinformation. GPT-4 can be used in the medical education field. Ghorashi et al, [26]

generated a chatbot for medical students to help them with learning and research. Chatbots have the potential to improve students' understanding, retention, and application of medical information in real time because of their capacity to summarize, and clarify difficult concepts, automate the development of memory aids, and function as an interactive instructor and point-of-care medical reference. While there are many benefits to using AI-powered chatbots in medical education, students need to use these tools as aids rather than as a replacement for more traditional methods. Chatbots ought to be designed to cite evidence-based medical sources and generate accurate, reliable information that complies with ethical norms, scientific writing requirements, and medical science standards. LLM can improve the research process for academics, which becomes increasingly evident as technology develops. Antu et al, [27] explored a tool to analyze academic literature reviews using OpenAI's ChatGPT. They tried to make these literature reviews more efficient, covering additional topics and offering writing scaffolding for this type of academic writing. Baber et al, [28] provided a comprehensive bibliometric review of the literature on ChatGPT, highlighting its inception and the rapid expansion of research in this domain. Their work lays a foundational understanding of the technological advancements and academic interest surrounding LLMs, tracing the development and proliferation of ChatGPT-related studies. They also mentioned the limitations of the available datasets with the predefined prompts, which reflect the limited responses to the queries. In addition, they noted the constraints of AI models in dealing with nuanced and complex medical scenarios, underscoring the need for human oversight.

Product-specific guidance (PSG) development and assessment, which serves as the foundation for recommendations for fasting and fed bioequivalence studies to direct the pharmaceutical industry in developing generic drug products, depends critically on the food effect summarization from the New Drug Application (NDA). It takes a lot of effort to manually summarize the effects of food from lengthy medication application assessment paperwork. Therefore, there is a need to develop automated ways to generate food effect summaries. The recent large language models (LLMs) like ChatGPT and GPT-4 have shown great promise for enhancing automated text summarization efficacy. Shi et al, [29] presented a straightforward and powerful method—iterative prompting—that allows humans to improve the quality of food effect summaries. They conducted comprehensive evaluations for food effect summaries utilizing 100 NDA review documents. Using multi-turn interaction to improve the quality of the summaries that are produced.

Rathore et al, [30] described a novel two-stage approach called Self-Supervised Prompting (SSP) for using in-context learning over very large language models (GPT-4X). Their extensive tests conducted on three natural language processing tasks and eleven low-resource languages from three language groups demonstrate significantly better performance than published baselines, establishing a new state of the art in the field.

The use of large language models (LLMs) in healthcare has evolved significantly over the past few years, marking a pivotal transformation in how medical information is managed and disseminated. This section delves into the historical progression and key contributions of LLMs in healthcare, showcasing seminal papers and notable authors who have driven this field forward.

2.1 LLM in Diagnosis Diseases

Xu et al, [31] conducted an exploratory case study on LLMs' capability on mental health reasoning tasks, illustrating the promising capability of certain models such as GPT-4. They summarized their findings into a set of action guidelines for potential methods to enhance LLMs' capability for mental health tasks. They also emphasized the important limitations before achieving deploy ability in real-world mental health settings, such as known racial and gender bias. Also, Hua et al, [32] highlighted the benefits of chatbots in providing accessible mental health care and improving patient outcomes through timely interventions. GPT-4 provided immediate support for mental health issues, offering resources and guidance based on patient input. Although chatbots can provide valuable mental health support, they may not fully replace the nuanced care provided by human therapists. There are concerns about the chatbot's ability to handle severe mental health crises effectively. They highlight the limitations of relying solely on AI for mental health care, stressing the need for human intervention in critical cases.

Deep integration of AI models such as ChatGPT and GPT-4 into medical situations is expected to enhance medical quality and efficiency while freeing up clinicians' time to engage in patient-centered care and personalized health management. Law et al, [33] presented an artificial intelligence chatbot that holds a lot of promise for the bariatric healthcare industry. They might be crucial in anticipating and addressing issues associated with weight reduction and obesity. They also addressed the medical, legal, ethical, data security, privacy, and liability issues emerging from medical errors produced by ChatGPT/GPT-4, though, given its potential limits.

Yang et al, [34] developed from scratch a large clinical language model to be evaluated on five clinical NLP tasks, including clinical concept extraction, medical relation extraction, semantic textual similarity, natural language inference (NLI), and medical question answering (MQA). They built a medical AI system using NLP powered by GPT-3 to process and interpret electronic health records (EHRs). The findings of this model increased the accuracy and ability of applying medical AI systems to improve healthcare delivery. Li et al, [35] discussed in their review the importance of analyzing electronic health records and electronic medical records (EMR). These records contain helpful information to assist oncologists in computer-aided diagnosis and decision-making. Finally, they determined the existing NLP application constraints that support clinical activities. The deployment of AI in healthcare raises significant ethical and legal questions, including accountability for AI-generated errors and transparency in decision-making processes. Vaishya et al, [36] discussed the ethical considerations and potential legal ramifications of

integrating AI into healthcare, highlighting the need for clear regulatory frameworks.

2.2 LLM in Management in Hospitals

In the healthcare sector, Jameel et al, [37] focused on the development and evaluation of a chatbot aimed at assisting patients in finding suitable doctors. The study demonstrates the chatbot's potential to improve user satisfaction and highlights the challenges in chatbot performance and accuracy. Similarly, Zhang and Zheng [38] reviewed the benefits and challenges of medical chatbots, emphasizing their role in future healthcare delivery. Dammavalam et al, [39] discussed the implementation and evaluation of a chatbot-based system for managing hospital operations. This study demonstrates the potential improvements in efficiency and user satisfaction, while also addressing the technological and practical challenges involved.

LLM is also used as a pre-consultation robot that will help doctors collect the symptoms, signs, medical history, and other information of patients in a standardized and comprehensive way through intelligently guided consultation to optimize the diagnosis and treatment process [40]. Mittal et al, [41] presented a bot engine that integrated several machine-learning approaches with natural language processing (NLP) algorithms. This bot managed user requests and provided quick access to local hospital information, trying to solve the unreliable and non-transparent data generated from the online hospital website.

Zhang and Zheng [42] analyzed the implementation of medical chatbots and virtual assistants, emphasizing their utility in enhancing patient care and administrative efficiency. By automating routine queries related to doctor schedules, examination costs, and hospital navigation, the chatbot ensures that patients receive timely and accurate information, reducing the workload on hospital staff. Their work highlights the early applications of LLMs in healthcare, setting the stage for more advanced implementations. They analyze four kinds of medical chatbots, such as pre-consultation, medical institution customer service, mental health consultation, and elderly care companions. Their results proved that there is a synergistic effect when clinicians and medical chatbots work together, producing better results than either alone. Thereby reducing wait times and improving overall service delivery.

Rajkumar et al, [43] created a user-friendly, intelligent chatbot for medical guidance and support. They presented a chatbot that worked in two different modes: online mode and offline mode, based on patient preferences. A real person will chat with the patient in online mode. Otherwise, in offline mode, the patient will ask the chatbot. If the chatbot fails to answer the patient, the patient will be escalated to the online mode. Large language models can significantly improve patient engagement through conversational AI, offering personalized interactions that encourage patients to participate actively in their health management. Wen et al, [44] discussed the power of conversational AI in enhancing patient engagement and satisfaction on digital health platforms.

A prototype chatbot application was designed for patients to use while in the hospital, yielding more accurate patient experience (PX) data, but it was also an opportunity for staff

to respond in real-time. They applied different tools, such as ChatGPT, to recognize relevant PX complaints from natural language data and improve patient experience in real-time [45]. This study underscores the importance of stakeholder involvement in designing effective and user-friendly AI tools. Another system incorporated ChatGPT into clinical decision support systems [46]. They tried to improve patient outcomes, streamline healthcare services, and support healthcare providers with evidence-based decisions. By maintaining current knowledge of artificial intelligence developments and investigating their use in healthcare environments. The benefits of this application are optimized treatment protocols, improved patient care, reducing the cognitive load on healthcare professionals, and allowing them to focus more on patient care and less on administrative tasks.

3. The Proposed Medical Chatbot

The deployment of a medical chatbot to replace the traditional receptionist role involves two main stages: data collection and data integration. The main architecture of the presented chatbot is displayed in Figure. 2. The local database, prompt, and utilized LLM presents the key components of the proposed chatbot.

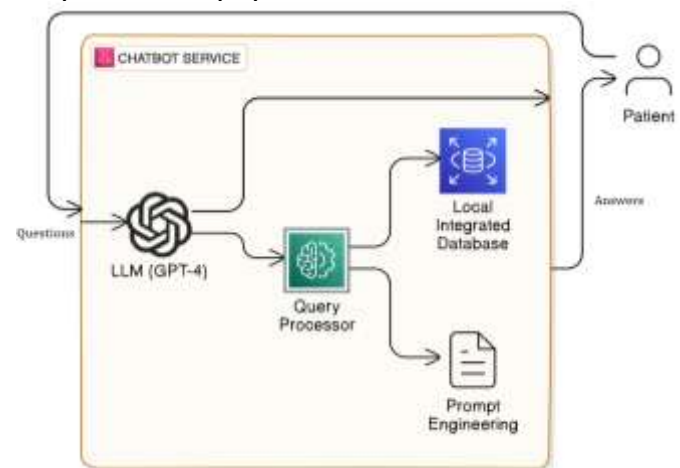


Fig. (2) The Main Architecture of the Proposed Medical Chatbot

3.1 Data Collection and Preparation

The first step is to gather comprehensive data from a real hospital in Egypt. The data is collected by conducting interviews with hospital staff using multiple sources to verify its correctness with the hospital administrators. These interviews showed us the types of questions that the patient needs to answer and the structure of the local database that will be built. This database included detailed information about:

- **Doctors:** Names, working hours, examination costs, job title, years of experience.
- **Hospital Sections:** Names of sections, and their locations within the hospital.
- **Assistants:** Names, Contact numbers, office locations.
- **Room Prices:** Costs associated with different types of rooms.

- **Hospital Policies:** General policies and procedures that patients need to be aware of.

The collected data is stored in two partitions as shown in Figure. 2.

1)

Local Database that includes data about doctors, assistants, and hospital sections. SQLite 3 is the utilized type of database which is considered the most used database engine in the world. It is a C-language library that implements a small, fast, self-contained, high-reliability, full-featured, SQL database engine.

2)

Prompt Engineering file (prompt) includes the metadata about the database in addition to the hospital policies and room prices. The policies are also added to the prompt to enhance patient satisfaction by generating accurate responses. The data is entered manually to ensure the completeness and reliability of the collected data.

3.2 Data Integration

There are several steps to apply the data integration:

3.2.1 The Framework Selection

The next step was to design the system and select the appropriate framework for launching the chatbot:

- **The LangChain framework** was chosen for its robust support for building applications that integrate language models with databases. LangChain launched LangChain Hub as a home for uploading, browsing, pulling, and managing prompts, chains, and agents for large language models (LLMs). LangChain Hub has become *the* go-to place for developers to discover new use cases and polished prompts. The LangChain hub has LangSmith access, which provides full read and write permissions for the existing prompts or uploading the proposed prompt by logging in. The proposed prompt is designed and uploaded to LangChain Hub with LangSmith access.
- **Google Colab Notebook** is used for implementing the proposed chatbot. It is a free cloud service that allows you to write and execute Python code in your browser. It provides access to free computing resources, including Graphics processing units (GPUs) and tensor processing units (TPUs), without any setup being required [47].

- The utilized open AI Model is (**GPT-4-turbo**) which is considered the latest GPT-4 model with vision capabilities. Unlike any of the earlier models, GPT-4 is a huge multimodal model that can tackle challenging issues with higher accuracy. It can accept text or image inputs and output text.

3.2.2 Chatbot Implementation

The main contribution of this paper is how to insert the data in the prompt. Crafting well-structured prompts is crucial for guiding the chatbot to generate accurate and relevant responses. The prompt was designed to provide the chatbot with all the necessary context and instructions, ensuring that it could manage a wide range of queries effectively. In addition, it prevents the users from hallucination by answering questions if it knows otherwise asking the user to rewrite his question in another way. The unknown questions are saved in a log file. This file is reviewed by the manager and the programmer to enhance this problem. This problem can be solved by adding additional information /rules in the local database or the prompt. Table 1 displays some rules that are given to the prompt to know what is it and its usage.

Table 1: An example of the prompt's rules

You are a powerful expert SQLite agent for an Arabic hospital.

You must be precise in writing SQL and make sure it returns the correct answer.

Your job is to extract the right information from the database and if you don't have information don't try to make up and say 'I don't have this information'.

You are given all the access to the database to which you are connected.

You will make a conversation with the user in Classical Arabic language and be polite to the user.

As indicated in Table 1, The chatbot was designed to evaluate user input and determine its responses. Patients' input text may include greetings or information requests. The answer will create and run an SQL query if it needs information from the local database. Otherwise, as illustrated in Fig. 3, it reacts directly in accordance with the prompt's instructions.

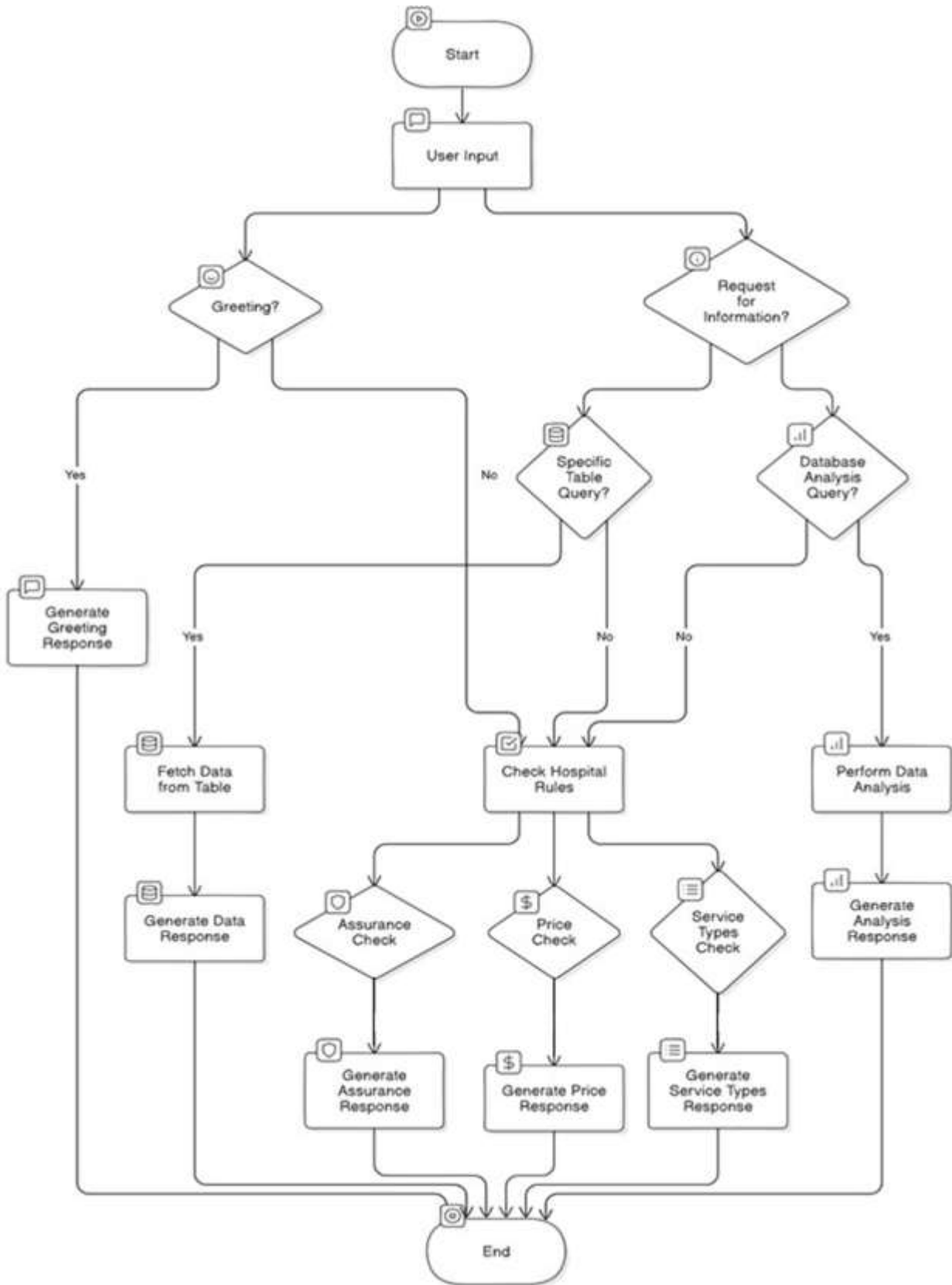


Fig. (3) The hospital reception chatbot flowchart

For each session for a new patient, the model saves the history of the conversation with limited queries to predict the new answers. This means if the patient asks about a specific department assistant and then asks to obtain his phone, the model will give the correct answers because it saves the history of each chat till finished.

3.2.3 Integration with the Local Database

Integrating the chatbot with the hospital's local SQLite database was a critical step. This integration involved:

- **Database Connectivity:** Establishing a secure and reliable connection between the chatbot and the SQLite database.
- **Query Execution:** Developing mechanisms for the chatbot to dynamically construct and execute SQL queries based on user inputs. This allows the chatbot to retrieve specific information about doctors, services, and policies from the database.
- **Error Handling:** Implementing robust error handling procedures to manage situations where the database might be unavailable or where the query returns no results. By entering erroneous and ambiguous queries and recording the responses or handling the unclear inputs. Such as if the user asks "Is there a surgery department in the hospital", "هل يوجد قسم للجراحة في المستشفى", then the chatbot answers "There are four departments for surgery what do you need?", "نعم يوجد في جراحة المستشفيات أقسام للجراحة وهي كالتالي جراحة المخ والأعصاب جراحة العظام, جراحة التجميل, القلب والأوعية الدموية". Thereby the chatbot can handle unclear or ambiguous questions.

3.2.4 Text-to-Speech responses

Speech is a general way of communication between human beings. The presented model offers a feature of displaying the input query answers as a speech using Google API text-to-speech. Texting and voice are becoming a major bot interface. Soon there will be a high demand for speech-enabled chatbots. Finally, the speaker is used to deliver the audio response output, which is played using the Python module's play sound method.

4. Evaluations

In this section, the dataset that was collected is explained. In addition, the used techniques for evaluating the presented chatbot performance are outlined. By displaying a comparison between the collected data and answers of the proposed chatbot.

4.1 Dataset Description

The collected dataset from the patients and hospital staff is about a hundred questions with their answers. Table 2 displays a sample of these questions.

Table 2: A sample of questions collected from the hospital

Questions in English	Questions in Arabic
May I know the location of the Psychiatric and Neurological Department clinic?	ممكن أعرف مكان عيادة قسم الأمراض النفسية والعصبية؟
I want to know the phone number of the orthopedic department assistant.	عايز أعرف رقم تليفون مساعد قسم العظام؟
May I know the dental clinic timings?	ممكن اعرف مواعيد عيادة الأسنان؟
I need an Otolaryngology consultant and how can I contact him?	محتاج استشاري أنف وأذن وحنجرة وكيفية التواصل معه؟
Does the hospital have an MRI?	هل يوجد بالمستشفى اشعة رنين؟
For me to do a test or an x-ray, must a doctor from the hospital write to me or is it normal from outside?	عشان اعمل تحليل او اعمل اشعة لازم دكتور من المستشفى يكون كاتبهالي ولا عادي من برا؟
Does the examination price differ depending on the insurance companies?	هل يختلف سعر الكشف تبع شركات التأمين؟

4.2 Evaluation and Testing

The chatbot's performance was evaluated through two primary methods:

4.2.1 Manual Evaluation:

Evaluators posed a diverse set of queries to the chatbot, encompassing all areas of hospital operations such as doctor schedules, service details, room pricing, and hospital policies. This comprehensive testing ensured that the chatbot was evaluated across a wide spectrum of potential user questions. Then, each response generated by the chatbot was meticulously cross-checked against the corresponding entries in the database by the evaluators. A 99.9% accuracy was obtained from their manual evaluation. This step was crucial in verifying the chatbot's ability to retrieve and relay accurate information.

4.2.2 Automatic Evaluation:

An evaluation dataset was created, comprising a set of model-generated answers and their corresponding correct answers stored in an Excel file. Table 3 reviews the examples of the answers to the questions written in Table 2 in sequence.

Table 3: A comparison between the collected data from hospital staff and the responses of the presented chatbot for all questions in Table 2

Collected Answers	Chatbot Answers
<p>هتلاقبها في الدور الثالث مبنى أ عيادة رقم ٣ يمين الأسانسير عندك أي حاجة ثاني أو حاجة ثانية عايز تعرفها، أنا جاهز أساعدك.</p> <p>You'll meet her on the third floor at clinic number 3 right Asancer. You have any further inquiries or needs you need to know, I'm ready to help you.</p> <p>اتفضل يا فندم، تَقْدَر تتواصل على الرقم ده: ٠١٠٨٢٤٥١٢٤١. لو عندك أي سؤال ثاني أو حاجة ثانية عايز تعرفها، أنا جاهز أجابوك.</p> <p>Here you go, then. You can dial this number at 01082451241. If you have any more questions or needs you know, I'm ready to answer.</p> <p>دي مواعيد دكاترة قسم الأسنان: - دكتور طارق محدث (استشاري) مواعيده من ٨:٠٠ صباحاً لحد ٦:٠٠ مساءً - دكتور منه سليمان (أخصائي) مواعيده من ٩:٠٠ صباحاً لحد ٥:٠٠ مساءً دكتور أحمد وائل (مقيم) ودكتور حمدي فتحي (مقيم) مواعيدهم من ١٠:٠٠ صباحاً لحد ٧:٠٠ مساءً لو عاوز تحجز ميعاد مع أي دكتور منهم ممكن أديك معلومات المساعد الخاص بالقسم للتواصل معاه.</p> <p>Dr. Tariq Meddhat (adviser) scheduled from 8:00 a.m. to 6:00 p.m. Dr. Minh Sulaiman (specialist) scheduled from 9:00 a.m. to 5:00 p.m. Dr. Ahmed Wail (resident) and Dr. Hamdi Fathi (resident) scheduled from 10:00 a.m. to 7:00 p.m. if you want to book an appointment with any of them, may I direct the Department's assistant's information to communicate with him.</p> <p>دكتور أحمد الفقي(استشاري) مواعيده من ٨:٠٠ صباحاً لحد ٦:٠٠ مساءً لو عاوز تحجز ميعاد مع أي دكتور منهم ممكن أديك معلومات المساعد الخاص بالقسم للتواصل معاه.</p> <p>Dr. Ahmed Al-Faki (adviser) is scheduled from 8:00 a.m. to 6:00 p.m. if you want to make an appointment with any of them, I can bring you the information of the department's assistant to communicate with him.</p> <p>نعم يا فندم يوجد عمل اشعة رنين بالمستشفى. لو محتاج حجز الأشعة ممكن أديك معلومات سكرتارية قسم الأشعة للتواصل معاه.</p> <p>Yes, sir. There's an MRI at the hospital. If there's a need for X-rays, I can get you information from the X-ray department to get in touch with her.</p> <p>لا، مش ضروري يكون الدكتور من المستشفى. سياستنا تسمح بإجراء التحاليل والأشعة للمرضى، حتى لو كان الدكتور من خارج المستشفى. يعني تقدر تبجي تعمل التحاليل أو الأشعة اللي محتاجها من غير مشاكل. لو عايز تعرف تفاصيل أكثر أو تحجز موعد، قولني وأنا تحت أمرك. ولو احتجت أي سؤال ثاني اتفضل أنا تحت أمرك</p> <p>No, it's not necessary to have the doctor from the hospital. Our policy allows patients to be tested and X-rayed, even if the doctor is out of the hospital. It means Tigi can do the tests or the X-rays she needs without problems. If you want to know more details or make an appointment, just say it when I'm at your disposal. And if you need any more questions, go ahead. I'm all yours.</p> <p>نعم يا فندم، بيختلف سعر الكشف على حسب التأمين أو زيارة جديدة. لو حضرتك بتشتغل معانا في الشركة فمفيش مصاريف تدفعها. لو من أسر العاملين هتدفع ١٠% من قيمة الكشف، ولو زيارة جديدة بتتحمل كل قيمة الكشف المحددة من الدكتور المختص. لو عندك أي استفسار ثاني أو حابب تحجز ميعاد كشف، أنا جاهز أساعدك.</p> <p>Yeah, we're sorry, but it's different on insurance or a new visit. If I come to work with you at the company, there's no expense to pay. If a worker's family pays 10% of the disclosure value, and if a new visit takes on all the disclosure value specified by the competent doctor. If you have any further questions or love to book an appointment, I'm ready to help you.</p>	<p>هتلاقبها في الدور الثالث مبنى أ عيادة رقم ٣ يمين الأسانسير عندك أي حاجة ثانية عايز تعرفها، أنا جاهز أساعدك.</p> <p>You'll meet her on the third floor... a clinic building, number three... right of the Asancer... you have any other need you know, I'm ready to help you.</p> <p>اتفضل حضرتك تَقْدَر تتواصل على الرقم ده: ٠١٠٨٢٤٥١٢٤١. لو عندك أي سؤال ثاني أنا جاهز أجابوك.</p> <p>Your Honor, you may contact this number at 01082451241. If you have any more questions, I'm ready to answer you.</p> <p>مواعيد دكاترة قسم الأسنان: - دكتور طارق محدث (استشاري) مواعيده من ٨:٠٠ صباحاً لحد ٦:٠٠ مساءً - دكتور منه سليمان (أخصائي) مواعيده من ٩:٠٠ صباحاً لحد ٥:٠٠ مساءً دكتور أحمد وائل (مقيم) ودكتور حمدي فتحي (مقيم) مواعيدهم من ١٠:٠٠ صباحاً لحد ٧:٠٠ مساءً لو عاوز تحجز ميعاد مع أي دكتور منهم ممكن أزودك معلومات المساعد الخاص بالقسم للتواصل معاه.</p> <p>Dental doctor's appointments: Dr. Tariq Meddhat (adviser) scheduled from 8:00 a.m. to 6:00 p.m. - Dr. Minh Sulaiman (specialist) scheduled from 9:00 a.m. to 5:00 p.m. Dr. Ahmed Wail (resident) and Dr. Hamdi Fathi (resident) scheduled from 10:00 a.m. to 7:00 p.m. if you want to book an appointment with any of them, please provide your assistant information for the department to communicate.</p> <p>دكتور أحمد الفقي(استشاري) مواعيده من ٨:٠٠ صباحاً لحد ٦:٠٠ مساءً لو عاوز تحجز ميعاد مع أي دكتور منهم ممكن أديك معلومات المساعد الخاص بالقسم للتواصل معاه.</p> <p>Dr. Ahmed Al-Faki (adviser) is scheduled from 8:00 a.m. to 6:00 p.m. if you want to make an appointment with any of them, I can bring you the information of the department's assistant for communication.</p> <p>يوجد بالمستشفى عمل اشعة رنين ولو عايز رقم مساعد قسم الأشعة أنا مستعد أجابوك</p> <p>There's an MRI at the hospital, and if you want the number of the X-ray assistant, I'm ready to answer you.</p> <p>لا. سياستنا تسمح بإجراء التحاليل والأشعة للمرضى، حتى لو كان الدكتور من خارج المستشفى. تقدر حضرتك تعمل التحاليل أو الأشعة اللي محتاجها. لو عايز تعرف تفاصيل أكثر أو تحجز موعد، أنا موجود</p> <p>No. Our policy allows patients to be tested and X-rayed, even if the doctor is out of the hospital. Your Honor, you can do the tests or the X-rays you need. If you want to know more details or make an appointment, I'm here.</p> <p>يا فندم بيختلف سعر الكشف على حسب التأمين أو زيارة جديدة. لو حضرتك بتشتغل معانا في الشركة فمفيش مصاريف تدفعها. لو من أسر العاملين هتدفع ١٠% من قيمة الكشف، ولو زيارة جديدة بتتحمل كل قيمة الكشف المحددة من الدكتور المختص. لو عندك أي استفسار ثاني أو حابب تحجز ميعاد كشف، أما جاهز اعطيك المعلومات المطلوبه.</p> <p>We regret that the rate of disclosure varies according to insurance or a new visit. If I come to work with you at the company, there's no expense to pay. If a worker's family pays 10% of the disclosure value, and if a new visit takes on all the disclosure value specified by the competent doctor. If you have any further questions or love to book an appointment, I'll give you the information you need.</p>

Table 3 assures that the chatbot's answers are very accurate according to the collected information. The generated Excel file is fed into an available large language model to measure the similarity between the two answers. Key aspects of the automated evaluation included: Precision and Recall Metrics: These metrics were used to measure the correctness (precision) and completeness (recall) of the chatbot's responses.

Overall Accuracy Calculation: The overall accuracy rate of the chatbot was calculated based on the evaluation results. The chatbot achieved an impressive accuracy rate of 99%, indicating its high reliability and effectiveness in providing accurate information. Finally, the presented chatbot reduces the workload on hospital staff and minimizes operational costs by automating the receptionist role. In addition, it provides patient satisfaction by providing immediate and accurate responses to their queries. Last, the chatbot is available 24/7, ensuring that patients have access to information at any time.

5. Conclusion and Future Works

The implementation and deployment of a medical chatbot as a receptionist in Egypt highlights the transformative potential of LLMs in healthcare management systems. The integration of LLM with the hospital database with fine-tuning using prompt engineering achieves all the requirements needed to obtain secure, immediate, and accurate information. It also achieves reliability and enhances operational efficiency. Despite its success, the chatbot's implementation also revealed certain limitations, such as the need for continuous updates to the database, potential challenges in handling complex medical queries, and preventing patients from hallucinations. Addressing these limitations will be crucial for the broader adoption and success of such systems in healthcare settings. Finally, the presented model achieved a 99% accuracy rate. This high accuracy rate underscores the reliability of the system in replacing traditional receptionist roles, thereby streamlining hospital operations and enhancing patient experience.

While the current version of the chatbot focuses on providing information, future enhancements could include:

- **Appointment Booking and Cancellation:** Enabling the chatbot to schedule and cancel appointments, further streamlining hospital operations.
- **Medical Advice:** Incorporating functionalities to provide general medical advice based on patient symptoms, under strict ethical guidelines and within the scope of safe practice.
- **Integration with EHR Systems:** Connecting the chatbot to Electronic Health Records (EHR) systems to provide more personalized information and support to patients.
- **Integration with Medical knowledge databases:** Integrating more advanced medical knowledge databases or incorporating collaboration with medical professionals can help the chatbot handle more complex queries effectively.
- **Prompt engineering self-learning:** Implementing mechanisms for continuous learning from new interactions and regular updates to the database will

help keep the chatbot's information current and accurate.

- **Enhancing Data Privacy:** Strengthening data privacy and security measures will ensure that patient information is protected. This includes encryption, access controls, and regular security audits.
- **Measuring patient satisfaction:** After the patient obtains his answers, the chatbot can display a screen for measuring his satisfaction from 0 to five for the generated answers. That will ensure the high accuracy of the presented chatbot.
- **Establishing transparent guidelines:** Establishing transparent guidelines for the chatbot's decision-making process and maintaining a clear chain of accountability for its responses will enhance trust and reliability.

References

- [1] D. Zumstein and S. Hundertmark, "chatbots—an interactive technology for personalized communication, transactions and services," *iadis International Journal on WWW/Internet*, vol. 15, no. 1, 2017.
- [2] J. Weizenbaum, "ELIZA—a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 26, no. 1, pp. 23–28, 1983.
- [3] L. Zhuo, Z. Chi, M. Xu, H. Huang, H. Zheng, C. He, X.L. Mao, and W. Zhang, "Protlm: An interleaved protein-language llm with protein-as-word pre-training," *arXiv preprint arXiv:2403.07920*, 2024.
- [4] S.K. Dam, C.S. Hong, Y. Qiao, and C. Zhang, "A Complete Survey on LLM-based AI Chatbots," *arXiv preprint arXiv:2406.16937*, 2024.
- [5] J. Li, A. Dada, B. Puladi, J. Kleesiek, and J. Egger, "ChatGPT in healthcare: A taxonomy and systematic review," *Journal of Biomedical Informatics*, vol. 129, p. 103826, 2024.
- [6] G. De Vito, "Assessing healthcare software built using IoT and LLM technologies," in *Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering*, pp. 476–481, June 2024.
- [7] R. Yang, T.F. Tan, W. Lu, A.J. Thirunavukarasu, D.S.W. Ting, and N. Liu, "Large language models in health care: Development, applications, and challenges," *Health Care Science*, vol. 2, no. 4, pp. 255–263, 2023.
- [8] M.A.K. Raiaan, M.S.H. Mukta, K. Fatema, N.M. Fahad, S. Sakib, M.M.J. Mim, J. Ahmad, M.E. Ali, and S. Azam, "A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges," *Artificial Intelligence Review*, 2024.
- [9] M. Javaid, A. Haleem, and R.P. Singh, "ChatGPT for healthcare services: An emerging stage for an innovative perspective," *Journal of Medical Systems*, vol. 47, no. 5, p. 102, 2023.
- [10] Z. Liu, L. Zhang, Z. Wu, X. Yu, C. Cao, H. Dai, N. Liu, J. Liu, W. Liu, Q. Li, D. Shen, X. Li, D. Zhu, and T. Liu, "Surviving ChatGPT in healthcare," *Computers in Biology and Medicine*, vol. 141, p. 105316, 2024.
- [11] S. Mukherjee, P. Gamble, M.S. Ausin, N. Kant, K. Aggarwal, N. Manjunath, D. Datta, Z. Liu, J. Ding, S. Busacca, and C. Bianco, "Polaris: A Safety-focused LLM Constellation Architecture for Healthcare," *arXiv preprint arXiv:2403.13313*, 2024.
- [12] T. Olorunsogo, A.O. Adeniyi, C.A. Okolo, and O. Babawarun, "Ethical considerations in AI-enhanced medical decision support systems: A review," *World*

- Journal of Advanced Engineering Technology and Sciences, vol. 11, no. 1, pp. 329–336, 2024.
- [13] E. Kasneci, K. Seßler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, S. Günnemann, E. Hüllermeier, and S. Krusche, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, p. 102274, 2023.
 - [14] M.U. Hadi, R. Qureshi, A. Shah, M. Irfan, A. Zafar, M.B. Shaikh, N. Akhtar, J. Wu, and S. Mirjalili, "A survey on large language models: Applications, challenges, limitations, and practical usage," *Authorea Preprints*, 2023.
 - [15] M. Kardum, "Rudolf Carnap – the grandfather of artificial neural networks: The influence of Carnap's philosophy on Walter Pitts," *Guide to Deep Learning Basics: Logical, Historical and Philosophical Perspectives*, pp. 55–66, 2020.
 - [16] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," 2018.
 - [17] M. Zhang and J. Li, "A commentary of GPT-3 in MIT Technology Review 2021," *Fundamental Research*, vol. 1, no. 6, pp. 831–833, 2021.
 - [18] J.G. Meyer, R.J. Urbanowicz, P.C. Martin, K. O'Connor, R. Li, P.C. Peng, T.J. Bright, N. Tatonetti, K.J. Won, G. Gonzalez-Hernandez, and J.H. Moore, "ChatGPT and large language models in academia: opportunities and challenges," *BioData Mining*, vol. 16, no. 1, p. 20, 2023.
 - [19] M. Zong and B. Krishnamachari, "A survey on GPT-3," *arXiv preprint arXiv:2212.00857*, 2022.
 - [20] E. Sezgin, J. Sirrianni, and S.L. Linwood, "Operationalizing and implementing pretrained, large artificial intelligence linguistic models in the US health care system: outlook of generative pretrained transformer 3 (GPT-3) as a service model," *JMIR Medical Informatics*, vol. 10, no. 2, p. e32875, 2022.
 - [21] R. Xu and Z. Wang, "Generative artificial intelligence in healthcare from the perspective of digital media: Applications, opportunities and challenges," *Digital Health*, vol. 1, p. 20552076211069021, 2024.
 - [22] S. Lim and R. Schmalzle, "Artificial intelligence for health message generation: an empirical study using a large language model (LLM) and prompt engineering," *Health Communication*, vol. 38, no. 5, pp. 531–540, 2023.
 - [23] E. Waisberg, J. Ong, M. Masalkhi, S.A. Kamran, N. Zaman, P. Sarker, A.G. Lee, and A. Tavakkoli, "GPT-4: a new era of artificial intelligence in medicine," *Irish Journal of Medical Science (1971)*, vol. 192, no. 6, pp. 3197–3200, 2023.
 - [24] N. Karanikolas, E. Manga, N. Samaridi, E. Tousidou, and M. Vassilakopoulos, "Large language models versus natural language understanding and generation," in *Proceedings of the 27th Pan-Hellenic Conference on Progress in Computing and Informatics*, pp. 278–290, November 2023.
 - [25] I. Carvalho and S. Ivanov, "ChatGPT for tourism: applications, benefits and risks," *Tourism Review*, vol. 79, no. 2, pp. 290–303, 2024.
 - [26] N. Ghorashi, A. Ismail, P. Ghosh, A. Sidawy, and R. Javan, "AI-powered chatbots in medical education: potential applications and implications," *Cureus*, vol. 15, no. 8, 2023.
 - [27] S.A. Antu, H. Chen, and C.K. Richards, "Using LLM (Large Language Model) to Improve Efficiency in Literature Review for Undergraduate Research," *LLM@ AIED*, pp. 8–16, 2023.
 - [28] H. Baber, K. Nair, R. Gupta, and K. Gurjar, "The beginning of ChatGPT – a systematic and bibliometric review of literature," *Journal of Artificial Intelligence in Medicine*, vol. 115, p. 102452, 2023.
 - [29] Y. Shi, P. Ren, J. Wang, B. Han, T. ValizadehAslani, F. Agbavor, Y. Zhang, M. Hu, L. Zhao, and H. Liang, "Leveraging GPT-4 for food effect summarization to enhance product-specific guidance development via iterative prompting," *Journal of Biomedical Informatics*, vol. 148, p. 104533, 2023.
 - [30] V. Rathore, A. Deb, A. Chandresh, and P. Singla, "SSP: Self-Supervised Prompting for Cross-Lingual Transfer to Low-Resource Languages using Large Language Models," *arXiv preprint arXiv:2406.18880*, 2024.
 - [31] X. Xu, B. Yao, Y. Dong, S. Gabriel, H. Yu, J. Hender, M. Ghassemi, A.K. Dey, and D. Wang, "Mental-llm: Leveraging large language models for mental health prediction via online text data," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 1, pp. 1–32, 2024.
 - [32] Y. Hua, F. Liu, K. Yang, Z. Li, Y.H. Sheu, P. Zhou, L.V. Moran, S. Ananiadou, and A. Beam, "Large language models in mental health care: a scoping review," *arXiv preprint arXiv:2401.02984*, 2024.
 - [33] S. Law, B. Oldfield, and W. Yang, "ChatGPT/GPT-4 (large language models): Opportunities and challenges of perspective in bariatric healthcare professionals," *Obesity Surgery*, 2024.
 - [34] X. Yang, A. Chen, N. PourNejatian, H.C. Shin, K.E. Smith, C. Parisien, C. Compas, C. Martin, A.B. Costa, M.G. Flores, and Y. Zhang, "A large language model for electronic health records," *NPJ Digital Medicine*, vol. 5, no. 1, p. 194, 2022.
 - [35] C. Li, Y. Zhang, Y. Weng, B. Wang, and Z. Li, "Natural language processing applications for computer-aided diagnosis in oncology," *Diagnostics*, vol. 13, no. 2, p. 286, 2023.
 - [36] R. Vaishya, A. Misra, and A. Vaish, "ChatGPT: Is this version good for healthcare and research?" *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 18, no. 3, pp. 207–210, 2023.
 - [37] U. Jameel, A. Anwar, and H. Khan, "Doctor recommendation chatbot: A research study: Doctor recommendation chatbot," *Journal of Applied Artificial Intelligence*, vol. 2, no. 1, pp. 1–8, 2021.
 - [38] H. Zhang and J. Zheng, "The application analysis of medical chatbots and virtual assistant," *Frontiers in Social Science and Technology*, vol. 3, pp. 11–16, 2021.
 - [39] S.R. Dammavalam, N. Chandana, T.R. Rao, A. Lahari, and B. Aparna, "AI based chatbot for hospital management system," in *2022 3rd International Conference on Computing, Analytics and Networks (ICAN)*, IEEE, pp. 1–5, November 2022.
 - [40] D. Madhu, C.N. Jain, E. Sebastain, S. Shaji, and A. Ajayakumar, "A novel approach for medical assistance using trained chatbot," in *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, IEEE, pp. 243–246, March 2017.
 - [41] M. Mittal, G. Battineni, D. Singh, B. Tech, T.S. Nagarwal, and P. Yadav, "Web-based chatbot in the management of local hospital information," 2021.
 - [42] H. Zhang and J. Zheng, "The Application Analysis of Medical Chatbots and Virtual Assistant," *Journal of Medical Internet Research*, vol. 25, no. 11, p. e23541, 2023.
 - [43] K. Rajkumar, T. Ragupathi, and S. Karthikeyan, "Intelligent Chatbot for Hospital Recommendation System," in *2024 2nd International Conference on Disruptive Technologies (ICDT)*, IEEE, pp. 664–668, March 2024.
 - [44] B. Wen, R. Norel, I. Liu, T. Stannenbeck, F. Zulkernine, & H. Chen "Leveraging Large Language Models for Patient Engagement: The Power of

- Conversational AI in Digital Health”. *arXiv preprint arXiv:2406.13659*, (2024).
- [45] X. Wang, S.M. Abubaker, G.T. Babalola, and S.T. Jesso, “Co-Designing an AI Chatbot to Improve Patient Experience in the Hospital: A human-centered design case study of a collaboration between a hospital, a university, and ChatGPT,” *International Journal of Medical Informatics*, vol. 159, p. 104669, 2024.
- [46] J. Miao, C. Thongprayoon, T. Fülöp, and W. Cheungpasitporn, “Enhancing clinical decision-making: Optimizing ChatGPT’s performance in hypertension care,” *The Journal of Clinical Hypertension*, vol. 26, no. 5, p. 588, 2024.
- [47] V. Sharma, G.K. Gupta, & M. Gupta, “performance benchmarking of gpu and tpu on google colaboratory for convolutional neural network,” In *Applications of Artificial Intelligence in Engineering: Proceedings of First Global Conference on Artificial Intelligence and Applications (GCAIA 2020)* (pp. 639-646). Springer Singapore. 2021.