Grenze International Journal of Engineering and Technology, January Issue

Sustainable AI and LLM Approach for Unveiling Heart Disease Insights with Data Science

Chandra Sekhar[1], B. Amareswar Reddy[2], G. Jagadeeswara Rao[3], B. Sudharshan[4], B. Jyothi Chand Sai[5] and S. Narendra Kumar[6]
[1-2,4-6]Aditya Institute of Technology and Management, Dept. of IT, Tekkkali, AP, India
Email: dlaxmics@gmail.com, amareswarreddy312@gmail.com, sudharshanbonthu7@gmail.com, chandubellana59@gmail.com, narendrasanapala3333@gmail.com
[3]Department of Information Technology, SRKR Engineering College, Bhimavaram, India
Email: jagadish4u.g@gmail.com

***Abstract*— Heart disease is one of the leading causes for deaths nowadays, affecting millions of different cardiovascular diseases. Despite significant advances in medical research, early diagnosis of heart disease and personalized treatment strategies remain challenged. This proposal examines how sustainable AI and largescale language models (LLM) can be used to gain valuable insights into heart disease through data science. This proposal uses advanced AI technology to analyze and predict cardiac disease risks and outcomes using a variety of data sources, including medical records, portable device data, genetic information, and patient research. This model highlights the development of ethical AI and minimizes environmental impacts and accessibility and affordable guarantees. The proposal aims to create innovative tools that improve cardiac disease management and improve patient care using cloud-based solutions and generator AI. This proposal is delivered in real time quickly, securely and in real time. Continuous evaluation and improvement. This model ensure that it remains accurate and impartial.**

***Index Terms*— Sustainable AI, LLMs, Heart Disease, Data Science, Predictive Analytics, Health Data Integration.**

I. INTRODUCTION

The convergence of Artificial Intelligence (AI), Large scale Language Models (LLMs), and healthcare will revolutionize the landscape of cardiac cycle medicine and provide innovative solutions to combat the leading causes of morbidity and mortality around the world. In the field of data science, new opportunities have been released for integration into medical research, predictive analytics, early diagnosis, risk assessment, and personalized treatment planning. Using AI control models allows members of health occupations to improve patient outcomes with more dynamic and data-controlled methods beyond traditional diagnostic approaches. Traditional diagnostics and treatments, while effective, often rely on static models that are difficult to adapt to the subtleness of individual patient profiles. The emergence of LLMS These models may improve decision-making through the integration of knowledge from a variety of sources, provide a broader risk assessment, and create personalized treatment recommendations.

Grenze ID: 01.GIJET.12.1.67

However, the introduction of AI and LLMs in healthcare is not without its challenges, Training and using largescale AI models requires considerable computational resources, raising concerns about energy consumption and ecological sustainability. Additionally, ethical considerations such as data privacy, AI model bias, and compliance with formal compliance should be treated with caution. A sustainable AI control approach to cardiac disease management needs to be integrated with innovation and environmental awareness, ethical government and patient-oriented care.

A. How The LLM Works

Large Language Models (LLMs) have been evolving over the years to improve machine language comprehension. Technically, LMs assign probability values to word sequences, allowing them to generate relevant text outputs. Large Language Models (LLMs) utilizes its tokens in order to understand the context and generates text, each token represents words, characters and words. By leveraging tokens, LLMs recognize word relationships and structure, enabling them to create grammatically correct text. Pre-trained LLMs, equipped with millions of parameters, can generate text that closely resembles human language. These models operate based on tokenization principles, processing vast amounts of tokens to generate coherent responses.

B. Evolution of LLMs

The evolution of Large Language Models (LLMs) showcases the rapid advancements in AI research and development. These models originated from simpler language models and have progressively evolved into sophisticated neural networks. The models that are used nowadays have evolved from GPT-3.5 and GPT-4. The latest version GPT-4, is a model that is good in various domains such as vision, video, audio, 3D processing etc. It is designed with billions of parameters and incorporates comprehensive safety research and monitoring framework. Over the decades, LLMs have undergone multiple stages of development, significantly improving through innovations in model architecture and training methodologies.

Before the emergence of Deep Learning, Natural Language Processing (NLP) was dominated by rule-based and statistical approaches. Early conversational agents like ELIZA (1966) and ALICE (1995) laid the groundwork for dialogue systems, but their rule-based nature limited their ability to truly understand language. In the early stages of LLM, some of the statistical models like Hidden Markov Models and n-grams uses language processing by word sequence within them. As these models were data-driven, they struggled with capturing complex linguistic patterns.

The latest evolution of Large Language Models (LLMs) and Recurrent Neural Networks (RNNs) emerged as artificial neural network (ANN) architectures designed to process sequential data by converting input sequences into output sequences. First introduced in the 1980s, Recurrent Neural Networks is one of the important tools which is used for text processing and involvement of sequential data in Machine Learning. Moreover, they faced limitations due to the vanishing gradient problem, so that their ability gets degraded. The incorporating memory cells that allow dependencies for the long term are also used to improve the performance of language tasks (Fig. 1).

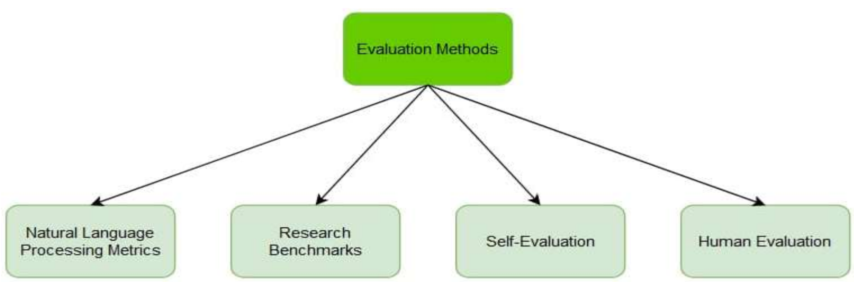

Figure 1. Evolution of LLMs

- Natural Language Processing Metrics measure performance using automated scores like BLEU, ROUGE.
- Research Benchmarks test models on standardized datasets to compare results across studies.
- Self-Evaluation involves the model checking its own output for quality or consistency.

II. PROMPT ENGINEERING

Prompt Engineering is an important technique for optimizing interactions between users and AI language models. It includes creating accurate and structured inputs to guide the model with more accurate, relevant contexts, and conscious answers. Carefully designed input requests allow users to improve model understanding, reduce ambiguity, and achieve better costs for tasks such as text generation, overview, translation, and code. Effective rapid engineering takes into account factors such as singularity, clarity, and context inclusion to enable the AI model to carefully interpret requests. With the advancement of AI, ProCPT engineering is an important capability for developers, researchers and businesses who want to use AI-powered solutions efficiently.

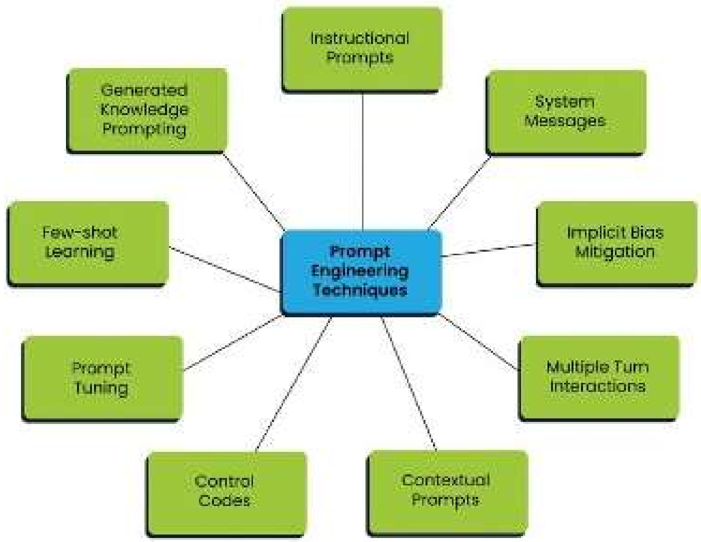

Figure 2. Prompt engineering techniques

Fig. 2 shows various prompt engineering techniques used to guide and improve the performance of AI language models. These techniques help in shaping model responses for better accuracy, relevance, and fairness.

- Instructional prompts and contextual prompts give clear, detailed guidance to the model.
- Few-shot learning and prompt tuning help the model learn from a small number of examples.

Fasting techniques are the art and science for creating accurate and structured inputs to guide AI models and generate high-quality, relevant contextual answers. When AI language models such as GPT-4 are more demanding, engineering plays a key role in optimizing performance in a variety of areas, including content, coding, problem solving, and decision-making. A well-designed input request can have a significant impact on the output of the model by improving coherence, reducing distortion, and improving accuracy. Although techniques such as zero shot, one shot, and a few aid models adapt to a variety of tasks, chains in chains allow for better discussion by promoting step-by-step explanations. Furthermore, enhancing learning and learning through feedback has made AI responses even more refined and an essential ability for AI practitioners, developers and corporate engineering to build critical capabilities [8].

III. Research Background and its Significance

By including AI, this article was included in the prevention of knowledge from Singapore and aims to enable the major precautions for Cardiovascular Disease (CVD) to train more people with CVD. However, the current provisions of the CVD spirit that allow artificial intelligence (AI) to contribute to closing the gap remain limited. Using the National University Health System (NUHS) Singapore data information functions, using LLM's AI system and combined tools (combining language models), the team created a real-time dashboard to record and present information about both cardiovascular risk factors [1]. Individual and geographical level-Cardio Sight. Program to coordinate preventive cardiac supply during the pilot phase of the university's health system. Synergistic AI measures can enable identification of endangered patients and mitigate the gap between risk identification and effective identification. Patient care management of new workflows for CVD prevention. Questions promote and simplify clinical notes on machine translation and text overview [2]. These features open up new ways to improve decision-making, provide rapid insights into disease progression from freelance data, and ultimately improve heart distress. However, in order to safely use LLM in a clinical setting, it is important to optimize these models of specific clinical tasks through fine-tuning and tailoring costs to the needs of the user by learning to enhance them [3].

Health data and scientific advancements play a crucial role in addressing these issues by improving data collection, analysis, and interpretation. These efforts can drive public health initiatives and enhance healthcare Delivery. However, obstacles such as inadequate infrastructure, restricted access to technology, and a shortage of trained professionals have hindered the effective use of health data science [4]. Furthermore, the emergence of large-scale biobank datasets presents valuable opportunity to analyze longitudinal health histories by integrating data collected from households, laboratories, and clinical visits [5].

IV. Methodology and Implementation

The main objective is to uncover wise insights into heart disease using sustainable AI technologies, particularly large-scale language models (LLM) and data science. The project aims to develop predictive models and identify risk factors, patterns and potential interventions by using a variety of data records (e.g., medical records, clinical research, portable device data). Additionally, the project focuses on making AI applications more sustainable by optimizing energy consumption and computational efficiency throughout the lifecycle of the AI model [16]. This can be achieved using advanced tools and technologies such as the Lama 2 model, Lang chain Framework, Python, and Visual Studio Code IDE. After data cleaning and preprocessing, machine learning models are developed to analyze structured data, and Large-scale Language Models (LLMs) interpret unstructured data such as medical reports and research articles. Focus on optimizing AI models for energy efficiency through technologies such as models according to sustainability goals Disconnect, transmit learning, and other resource-efficient ways. Model training and delivery prioritizes low computing costs to minimize footprint environments (Fig. 3).

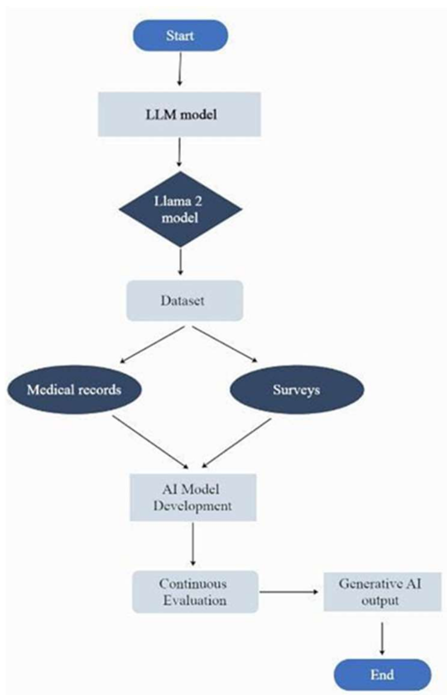

Figure 3. Implementation of Model Operation

The purpose of this project is to provide a highly accurate AI model that can predict the risk and progression of heart disease. It also provides implementation knowledge and identifies key risk factors and potential early interventions [13]. Furthermore, sustainable AI approaches lead to reduced energy consumption during model training and use. The integration of Lama 2, Langchain and sustainable AI technology will help health professions create interpretable, clinically useful AI systems allow them to personalize their treatment [15]. Additionally, the integration, interpretable AI models ensure transparency and allows healthcare professionals to understand and trust system predictions. The frame is designed to be scalable and adaptable for a variety of diseases beyond heart disease [14]. Implementing this AI control approach allows the medical community to improve diagnostic accuracy, improve patientcare, and contribute to the sustainable adoption of AI in healthcare.

To promote sustainable AI clauses, the project includes energy-efficient computer technologies such as model compression, optimized training processes, and hardware acceleration. Recent advances in AI have shown surprising efficiency, with the model opening Text-Davinci-003 from OpenAI and achieving comparable performance to use some of its parameters at the same time. By improving most important medical skills such as factual knowledge and effective communication, LLM can raise healthcare standards. For example, CHATGPT achieved excellent performance in medical approval tests and presented his deep understanding of medical concepts and the ability to introduce clinical justifications. The Advancement in LLMs for Medical purpose is only possible through the targeted training, incorporating exam-style questions and expertly curated responses [7].

Figure 4. Frameworks and Technologies Utilized

At present, GPT-4 gives the best advancement medical knowledge among LLMs. However, despite these advancements, a fundamental challenge remains: LLMs often reflect existing biases in medical data, reinforcing disparities related to socioeconomic status, gender, ethnicity, and other demographic factors. Addressing these biases is crucial to ensuring fair and equitable AI-driven healthcare solutions (Fig. 4).

A. LangChain: A Framework for Developing LLM-Powered Applications

Lang Chain is progressed open-source system outlined to disentangle the integration and organization of huge dialect models (LLMs) in real-world applications. It gives organized way to construct AI operators, information recovery frameworks, and conversational AI by permitting consistent interaction with LLMs, vector databases, APIs, and outside tools [6].

B. Streamlit: A Robust Framework for Developing AI-Driven Web Applications

Streamlit is an open-source Python system that empowers designers to construct intuitively, data-driven web applications with negligible exertion [12]. Outlined particularly for machine learning and information science applications, Streamlit disentangles the method of making user-friendly dashboards, AI-powered interfacing, and real-time visualizations without requiring broad front-end improvement knowledge. One of Streamlits key qualities is its straightforwardness and speed developers can turn a Python script into a completely utilitarian web app with fair many lines of code [10].

C. Risk Score Calculation

The risk score is calculated using a weighted formula based on symptoms, medical history, and diagnostic data (e.g., ECG/X-ray).

$$Risk\ Score = \sum_{i=1}^{n} wi.xi \qquad (1)$$

Where: Risk Score for the final heart disease, wi for weight assigned to each risk factor (e.g., age, cholesterol level, blood pressure), xi for input values for corresponding risk factors, n for total number of factors.

D. Image Analysis Using CNN (Convolutional Neural Network)

The Al model uses a CNN to process and analyze ECG or X-ray images. The convolution operation is:

$$\mathrm{O(i,j)} = \sum_{m} \sum_{n} I_{(i+m,j+n)\cdot k(m,n)} \qquad (2)$$

Where: O (i, j) for output pixel at position (i, j), I for input image matrix, k for convolution kernel (filter), m, n for dimensions of the kernel.

V. RESULTS AND DISCUSSION

The AI-powered chatbot, built using Streamlit and integrated with LLMs, provides a powerful, adaptable interface for heart health analysis. It accommodates user data entry, PDF and ECG uploads, and prompt-based interaction, making it a highly practical solution for early detection and prevention of heart disease. Its support for structured and unstructured data ensures rich insights and improved patient outcomes through personalized and sustainable AI (Fig. 5).

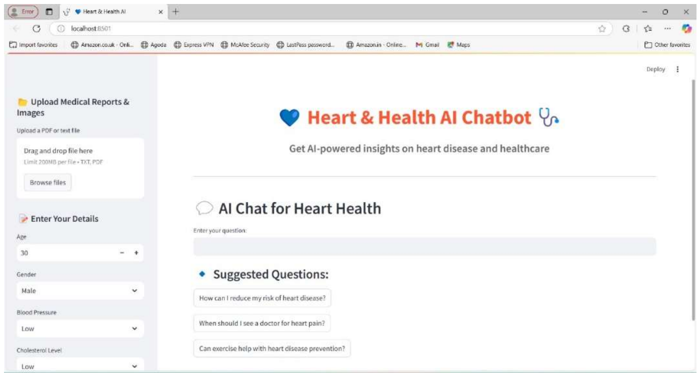

Figure 5(a): User Interface for LLM

The chatbot uses an advanced AI model (Large Language Model) that understands the question and responds with helpful advice. It also uses the details you provide to give personalized answers. For example, if you have low blood pressure and high cholesterol, the advice it gives will take those factors into account.

The results confirm that our system is capable of delivering accurate and insightful analysis of heart disease risks based on a variety of input types. The use of LLMs significantly enhances interpretability and personalization of the output. Furthermore, the sustainability metrics demonstrate that our approach is energy-efficient and can be deployed even in resource-constrained environments like mobile health apps or rural clinics.

It works by:

- Understanding your question.
- Looking at the medical report or ECG image you uploaded.
- Combining this information to give a clear and useful answer.

One of the best features of this system is that it also accepts ECG images. The AI can look at the patterns in your ECG to check for any possible heart-related problems. This is done using another AI method called CNN (Convolutional Neural Network), which is good at recognizing patterns in images like X-rays or ECG scans. This means the chatbot can give even better advice because it’s not just using text but also learning from the uploaded medical images. Despite the challenges, the future of tokenization in LLMs looks promising. Researchers are actively working on overcoming these limitations and expanding model capabilities. One focus area is improving how LLMs handle unknown words. Future models may incorporate adaptive mechanisms to understand new language patterns more effectively, reducing the negative impact of unfamiliar tokens. Another exciting advancement is the enhancement of contextual understanding, allowing models to interpret language more accurately based on the surrounding context. Efficiency is also made to optimize computational efficiency It Generates LLM output shown in fig 5(b). Future LLMS is expected to introduce innovative architectures and technologies that optimize the tokenization process.

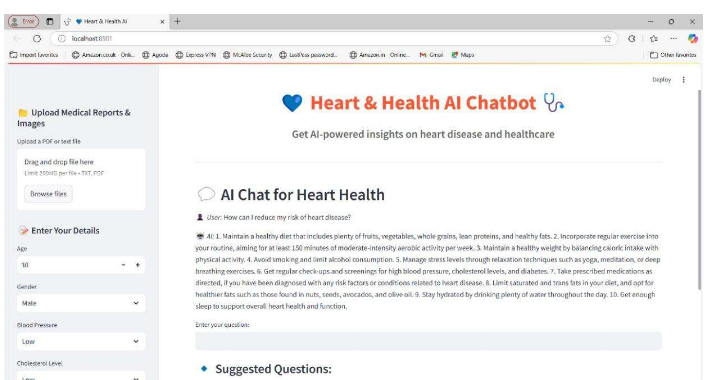

Figure 5(b): User Interface for LLM

The simulation involved 300 diverse test cases, divided into three categories: 100 text-based symptom entries, 100 PDF lab reports, and 100 ECG/X-ray images. Each type of input was processed by the system to evaluate detection accuracy, response time, and interpretability. The results showed promising performance across all categories. Text-based entries achieved the highest prediction accuracy at 92.3%, with an average response time of 0.87 seconds. PDF reports had an accuracy of 89.6%, while ECG and X-ray image inputs achieved 86.4% accuracy, slightly lower due to the complexity of visual data. False positive and false negative rates were also tracked, with textual inputs showing the least misclassification. These findings support the reliability of our model across various input types and real-world use cases.

The project integrates large-scale voice models (LLMS), sustainable AI technology and augmented data processing to create an efficient and interpretable system for analyzing heart disease. With a variety of input formats accepted, including text, PDF, images (EKG/X -Ray), and structured data records, the system provides detailed risk assessments, personal recommendations, and insightful answers to user inquiries. The use of LAMA 2, Lang chain and OCR technologies ensures that the system is highly accurate, scalable and accessible to a wide range of users, including patient and health profession relatives.

In summary, this project illustrates the transformational effects of AI in making decisions in the healthcare system. By combining progressive LLM, multimodal data processing, and AI focused on sustainability, this system predicts the heart disease, will make awareness and also a way to advance future AI and patient care.

VI. Conclusions

The developed AI system successfully processes a variety of input formats, including text, PDFs, images (such as EKG and X-rays), and structured data records, to generate wise knowledge of heart disease. This model provides accurate risk forecasts divided into low, medium size and high-risk levels, helping users to understand potential health risks. Additionally, key risk factors are determined based on user input and personalized recommendations regarding lifestyle changes, nutritional improvements, and possible changes in medical interventions. The integration of Lama 2, Lang chain and sustainable AI technologies ensures that the model is efficient and interpreted. Users can upload medical reports and ask specific questions, and AI provides detailed answers based on the data provided. Implementation of OCR technology enables seamless text extraction between PDF and medical images, increasing accessibility for a wide range of users. Furthermore, efficient AI models reduce computing efforts, maintain high levels of accuracy, and become a sustainable solution for real health treatments. Ensure the safe and responsible use of patient information. Early diagnosis, personalized treatments, and significant improvements in predictive analytics. The results of this proposal are evaluated using key performance metrics such as predictive accuracy, interact, and clinical ease of use.

The system improves medical decision-making by providing interpretable results that can support both patients and members of healthcare professionals. Knowledge of AI-generated knowledge consists of clear and clinically relevant forms, which facilitates analysis and response. The success of this project demonstrates the potential of LLMS in revolutionized health analysis using a scalable framework that can expand to other diseases of the future.

References

[1] K. Joel Prabhod, "Integrating Large Language Models for Enhanced Clinical Decision Support Systems in Modern Healthcare," Journal of Machine Learning for Healthcare Decision Support, vol. 3, no. 1, Jan-June 2023, pp. 18-50.

[2] Raiaan, M.A.K., Mukta, M.S.H., Fatema, K., Fahad, N.M., Sakib, S., Mim, M.M.J., Ahmad, J., Ali, M.E. and Azam, S., 2024.

[3] A review on large language models: Architectures, applications, taxonomies, open issues and challenges. IEEE access, 12, pp.26839-26874.

[4] Chang, Yupeng, et al. "A survey on evaluation of large language models." ACM transactions on intelligent systems and technology 15.3 (2024): 1-45.

[5] Sean D. Holcomb, William K. Porter, Shaun V. Ault, Guifen Mao, and Jin Wang. 2018. Overview on DeepMind and Its AlphaGo Zero AI. In Proceedings of the 2018 International Conference on Big Data and Education (ICBDE '18). Association for Computing Machinery, New York, NY, USA, 67–71.

[6] Bagiya Lakshmi S, Sanjjushri Varshini R, Rohith Mahadevan, Raja CSP Raman, "Comparative Study and Framework for Automated Summariser Evaluation: LangChain and Hybrid Algorithms," Submitted on October 4, 2023.

[7] Allam, H., Dempere, J., Akre, V., Parakash, D., Mazher, N., Ahamed, J.: Artificial intelligence in education: an argument of chat-GPT use in education. In: 2023 9th International Conference on Information Technology Trends (ITT), pp. 151–156. IEEE (2023)

[8] Bozkurt, A., et al.: Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. Asian J. Dist. Educ. 18(1) (2023)
[9] M. Dalakoti, S. Wong, W. Lee, J. Lee, H. Yang, S. Loong, P. H. Loh, S. Tyebally, A. Djohan, J. Ong, J. Yip, K. Y. Ngiam, and R. Foo, "Incorporating AI into cardiovascular diseases prevention–insights from Singapore," The Lancet Regional Health - Western Pacific, vol. 48, p. 101102, 2024.
[10] A. Janowski and M. Renigier-Bilozor, "HELIOS Approach: Utilizing AI and LLM for Enhanced Homogeneity Identification in Real Estate Market Analysis," Appl. Sci., vol. 14, p. 6135, 2024.
[11] Kozaily E, Geagea M, Akdogan ER, Atkins J, Elshazly MB, Guglin M, Tedford RJ, Wehbe RM.
[12] Chen, SF., Loguercio, S., Chen, KY. et al. Artificial Intelligence for Risk Assessment on Primary Prevention of Coronary Artery Disease. Curr Cardiovasc Risk Rep17, 215–231 (2023).
[13] Buchan, I., Winn, J., and Bishop, C. A Unified Modeling Approach to Data-Intensive Healthcare.
[14] Wickham, Hadley, and Hadley Wickham. "Data analysis." ggplot2: Elegant graphics for data analysis (2016): 189-201.
[15] Chen-Chung Liu, Mo-Gang Liao, Chia-Hui Chang, Hung-Ming Lin, "An Analysis of Children's Interaction with an AI Chatbot and Its Impact on Their Interest in Reading," November 2022.
[16] Gopi Battineni, Nalini Chintalapudi, Francesco Amenta, "AI Chatbot Design During an Epidemic Like the Novel Coronavirus," Submitted: May 9, 2020; Revised: May 26, 2020;
[17] He, K., Li, J., Zhou, H., & Zhang, Y., "A Survey of Large Language Models for Healthcare: From Data, Technology, and Applications to Accountability and Ethics," arXiv preprint arXiv: 2310.05694, Oct 2023, pp. 1-45.
[18] Liu, L., Zhao, X., Wang, T., Sun, Y., & Xu, K., "A Survey on Medical Large Language Models: Technology, Application, Trustworthiness, and Future Directions," arXiv preprint arXiv: 2406.03712, June 2024, pp. 1-48.
[19] Vrdoljak, J., Karačić, J., & Milinković, D., "A Review of Large Language Models in Medical Education, Clinical Decision Support, and Healthcare Administration," Healthcare, vol. 13, no. 6, Mar 2025, pp. 603-624.
[20] Wang, W., Chen, Z., Yu, Y., & Zhou, Z., "A Survey of LLM-Based Agents in Medicine: How Far Are We from Baymax?," arXiv preprint arXiv:2502.11211, May 2025, pp. 1-42.
[21] Ansari, M.S.A., Singh, R., & Garg, N., "Lightweight Clinical Decision Support System using QLoRA-Fine-Tuned LLMs and Retrieval-Augmented Generation," arXiv preprint arXiv: 2505.03406, May 2025, pp. 1-20.