SelfRewardRAG: Enhancing Medical Reasoning with Retrieval-Augmented Generation and Self-Evaluation in Large Language Models

Zakaria Hammane*
Faculty of Sciences Tetouan,
Département Informatique,
Data and Intelligent Systems (DIS) Team,
Abdelmalek Essaadi University,
Tetouan, Morocco
zakaria.hammane@etu.uae.ac.ma

Fatima-Ezzahraa Ben-Bouazza
Faculty of Sciences and Technology,
Hassan 1st University, Settat
LaMSN-la Maison des Sciences Numériques, Paris Area
fatimaezzahraa.benbouazza@uhp.ac.ma

Abdelhadi Fennan
Faculty of Sciences and Techniques,
Department of Computer Sciences,
Data and Intelligent Systems (DIS) Team,
Abdelmalek Essaadi University,
Tangier, Morocco

2024 International Conference on Intelligent Systems and Computer Vision (ISCV) | 979-8-3503-5018-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISCV60512.2024.10620139

In this study, we present a pioneering approach known as Retrieval Augmented Generation (RAG), which integrates Large Language Models (LLMs) with dynamic data retrieval to surmount the challenge of knowledge obsolescence, a matter of particular significance in the healthcare domain. This innovative system leverages real-time access to up-to-date clinical records, thereby enabling the generation of precise and informed responses, a notable leap over the conventional limitations faced by LLMs due to their reliance on static datasets. Our methodology embodies the seamless integration of RAG with LLMs to adeptly retrieve pertinent medical information from continuously updated repositories, such as PubMed, and to synthesize this information into accurate responses for medical queries. This advancement marks a considerable enhancement in the application of AI within medical decision-making processes, ensuring that the information provided remains both current and relevant. The effectiveness of our approach is validated through a series of experiments, which demonstrate a significant improvement in the accuracy and timeliness of the AI-generated responses, thereby underscoring its transformative potential for medical AI applications. Furthermore, the foundational principles underlying our system indicate its broader applicability in various other fields confronted with the challenges of rapidly changing knowledge bases. Through this work, we not only address the critical need for real-time information integration in healthcare AI but also establish a paradigm for future AI systems, promoting the incorporation of continuous learning and updating mechanisms to enhance their efficacy and relevance.

***Index Terms*—Retrieval Augmented Generation, Large Language Models, Incorporation, Healthcare**

I. Introduction

LLMs have demonstrated exceptional proficiency in tasks involving the comprehension and production of natural language, thus revolutionizing multiple fields including text completion, translation, and question answering [1][2]. The dynamic integration of external knowledge sources, such as healthcare knowledge graphs, into these models can further enhance their utility in specialized fields. For instance, the HealthPathFinder model leverages neural attention mechanisms within a healthcare knowledge graph to provide personalized health recommendations [3]. However, a significant obstacle that these models encounter is the knowledge cutoff predicament: they rely on unchanging training datasets, causing their knowledge to become obsolete over time, thus limiting their effectiveness in dynamic and rapidly changing environments [2].

Researchers have explored various approaches, including ongoing learning, transfer learning [4], and fine-tuning, to integrate new information into LLMs. These tactics have been investigated in response to specific tasks [5][6][2]. Although these approaches offer some enhancements, they often require regular retraining of the model, which can be resource-intensive and prohibitive, especially in domains characterized by rapidly evolving information.

RAG presents a potential solution to the knowledge cutoff problem. It enables LLMs to acquire and integrate up-to-date and relevant information from external sources immediately before responding to queries [7]. By integrating with a "real-time firehose" of information, such as clinical databases or journal repositories, RAG allows LLMs to access a broad and constantly updated knowledge reservoir, including the latest studies, publications, and research findings.

RAG essentially transforms LLMs from static repositories of information into flexible and responsive systems capable of providing current and evidence-based answers to queries. This capability is particularly crucial in fields like healthcare, where staying updated on the latest research and therapeutic recommendations is essential for providing accurate information and making well-informed decisions [8].

This study offers a comprehensive investigation of the RAG framework and its application in addressing the knowledge cutoff issue in LLMs, particularly in the healthcare domain. We propose an innovative RAG architecture specifically tailored to incorporate dynamic information effectively. We provide a detailed description of its components, integration techniques, and implementation considerations.

We demonstrate the effectiveness and practicality of the proposed RAG framework in enhancing the timeliness and accuracy of LLMs through experimental analysis and a case

979-8-3503-5018-0/24/$31.00 ©2024 IEEE

study in the healthcare sector. Our findings underscore the potential of RAG to revolutionize how LLMs acquire and utilize information, representing a significant advancement in mitigating the limitations of fixed training datasets in rapidly changing fields.

The paper is structured as follows: Section II presents a comprehensive review of related work, establishing the context for our study by discussing previous advancements and identifying gaps in the application of RAG and LLMs within the healthcare sector. In Section III, we introduce our novel approach, SelfRewardRAG, detailing its motivation, methodology, and implementation specifics, including the integration of dynamic information retrieval with LLMs for enhanced medical reasoning. Section IV presents a series of experiments designed to validate our approach, comparing SelfRewardRAG's performance against state-of-the-art models across various medical question-answering benchmarks. Additionally, this section discusses the implications of our findings, highlighting the model's strengths and areas for improvement. Section V concludes the paper by summarizing our contributions and outlining future research directions, emphasizing the potential of SelfRewardRAG to transform AI applications in healthcare by addressing the critical challenge of knowledge obsolescence.

II. Related Works

RAG has shown considerable potential in the medical domain for addressing the challenge of limited knowledge and improving decision-making processes. Several studies have explored the use of RAG in healthcare settings, particularly in clinical decision support and medical question-answering systems.

[9] introduced a clinical decision support system that utilizes RAG to provide personalized therapy suggestions in real-time, incorporating the latest clinical evidence. The technology integrates extensive LLMs with live clinical data streams and dynamically retrieves relevant medical literature to assist healthcare practitioners in making well-informed decisions. By leveraging RAG, the system can incorporate current information into therapy suggestions, thereby improving patient outcomes and reducing medical errors.

[10] developed a real-time medical question-answering system using the RAG model. They utilized it to retrieve and integrate information from continuously updated medical literature databases. This technology allows doctors to inquire about specific medical conditions or treatments and generates evidence-based responses by extracting relevant excerpts from medical literature. By utilizing RAG, the system provides doctors with access to up-to-date research findings and clinical guidelines, facilitating evidence-based decision-making at the point of care.

While the use of RAG in the medical field offers significant advantages, it also presents challenges and limitations. Ongoing research and development efforts focus on ensuring the reliability and accuracy of retrieved information, safeguarding patient privacy and confidentiality, and integrating RAG systems into existing clinical workflows. Additionally, further research is needed to explore the scalability and applicability of RAG across various medical specialties and languages.

RAG approaches have gained prominence in automated radiology report generation [11]. Our methodology combines multimodal embeddings from pre-trained vision language models and utilizes generic domain generative models such as OpenAI's text-DaVinci-003, gpt-3.5-turbo, and gpt-4 to generate reports. This integration has yielded promising results, improving clinical metrics and providing increased flexibility in diverse clinical environments.

Furthermore, in [12], the authors enhance the field by introducing an LLM-RAG pipeline specifically tailored for healthcare, with a focus on preoperative medicine. The model demonstrates the feasibility and benefits of integrating RAG with LLMs in medical contexts by achieving faster answer generation and improved accuracy.

In summary, RAG holds the potential to enhance decision-making processes and patient care in the medical profession by facilitating the integration of real-time knowledge. Future research efforts should prioritize addressing challenges related to data quality, privacy, and integration to fully realize the potential of RAG in healthcare settings.

III. Proposed Approach

A. Motivation

The SelfRewardRAG method was developed to address the complex issue of medical reasoning in artificial intelligence. This requires a system capable of processing natural language queries and accessing and evaluating data from specialized databases like PubMed. RAG is combined with LLMs like GPT-3.5 to enhance the precision of medical question-answering systems. This integration aims to merge advanced NLP capabilities with the vast knowledge found in medical literature.

This methodology employs a multi-layered approach where each layer contributes uniquely to the reasoning process. Initially, the LLMs, adept at understanding complex language, convert medical inquiries into structured queries. To ensure that the data used for generating responses is supported by scientific evidence, these queries are employed to retrieve relevant publications from PubMed, utilizing the capabilities of the PubMed API and the Python library PyMed. The PubMed API, provided by the National Center for Biotechnology Information (NCBI), offers programmatic access to the extensive repository of biomedical literature contained within the PubMed database. It enables users to search and retrieve bibliographic information using various search terms and parameters, making it an invaluable resource for automated literature review processes.

PyMed is a Python library that serves as an interface to the PubMed API, simplifying the process of querying the database and managing the retrieved data. Through PyMed, our system sends structured queries to the PubMed API, which then returns a list of relevant publications based on the search criteria. These publications include articles, reviews, and other scholarly documents that are pertinent to the medical questions at hand.

Different GPT-based agents meticulously generate the following responses, each analyzing different sections of the retrieved data. This redundancy prevents information from being overlooked and broadens the range of perspectives considered while devising solutions.

The self-evaluation layer is crucial as it establishes an internal quality control mechanism. Another LLM reviews the answers to identify any inconsistencies or inaccuracies, ensuring that the final product is both informative and logically coherent.

The resolution layer, acting as a resolver agent, further refines the responses using feedback from the self-evaluation. This layer is responsible for selecting the optimal answer and improving it to adhere to the highest standards of medical precision and reliability.

The method is driven by the necessity for precision and trustworthiness in medical AI applications, where the stakes are high. By employing a systematic and iterative approach to answer generation and refinement, SelfRewardRAG aims to bridge the gap between human expertise and AI capabilities, providing a tool that could potentially support medical professionals in their decision-making processes.

B. Methodology and Implementation

SelfRewardRAG, as depicted in Fig. 1, utilizes three established techniques in rapid engineering: chain-of-thought, self-reflection/self-rewarding, and resolver.

The first approach, emphasized in SelfRewardRAG, harnesses the emerging chain-of-thought (CoT) reasoning capabilities of LLMs to enhance both the predictive performance and the explainability of models applied to sophisticated tasks such as medical reasoning. This methodology capitalizes on the advances in LLMs, such as GPT-3.5, which are optimized for dialogue, allowing for more robust performance through contextually rich, zero-shot prompting styles. The essence of this approach is to explore how these CoT prompting styles, discovered in earlier model generations, can be generalized and enhanced for new model versions and datasets focusing on medical and scientific domains [13].

The second approach introduces Reflexion, a novel reinforcement learning framework that diverges from traditional methodologies by employing verbal feedback instead of traditional model weight updates. Reflexion enhances the learning capacity of language agents by allowing them to internally reflect on and verbalize the feedback received from their interactions with the environment. This reflection is then stored as episodic memory, providing a concrete direction for future actions and improvements. Essentially, it employs a method akin to human learning processes, encouraging iterative reflection on past mistakes to establish a more effective strategy for future attempts. This method stands out for its versatility in handling various feedback types and its potential to significantly improve task performance across a spectrum of tasks, including sequential decision-making and language reasoning [14].

Lastly, the Dialog-Enabled Resolving Agents (DERA) framework represents a groundbreaking approach to enhancing the output of LLMs, particularly in critical areas such as healthcare. DERA employs a conversational mechanism, leveraging the improved conversational abilities of GPT-4, to facilitate communication between two specialized agents: the researcher and the decider. The researcher identifies key information and problem areas, while the decider integrates this information to make informed judgments. This dialogue-based method facilitates iterative improvement of the model's output, ensuring higher accuracy and completeness crucial for medical applications. DERA's design promotes a more targeted and effective refinement process by assigning each agent a specific role, allowing for more nuanced and quality outputs [15].

SelfRewardRAG is an advanced system that combines GPT-3.5 with the ability to retrieve information from PubMed articles. It also incorporates chain-of-thought reasoning, self-reflection/self-rewarding mechanisms, and a resolver component.

The proposed approach combines the superior natural language processing (NLP) capabilities of LLMs with the retrieval effectiveness of domain-specific databases, like PubMed. This methodology is based on the coordination of multiple cutting-edge computing techniques, such as resolver protocols, self-reflection/self-rewarding mechanisms, chain-of-thought reasoning, and retrieval-augmented creation. The combination of these techniques is designed to handle the complexities and significant needs of medical reasoning assignments.

The SelfRewardRAG technique, aiming to enhance medical reasoning by incorporating retrieval-augmented generation and big language models, can be summarized as follows:

Question Processing: When the SelfRewardRAG pipeline initiates, medical questions are processed by an optimized version of GPT-3.5 capable of comprehending and analyzing intricate clinical question structures. Through this process, questions in natural language are converted into structured queries specifically designed to efficiently query medical databases.

Retrieval Mechanism: The primary function of the SelfRewardRAG system is its ability to search and retrieve relevant articles from PubMed. Structured queries serve as a connection point between the question's purpose and the database's indexed articles, facilitating the retrieval of the most pertinent information necessary to create well-informed responses.

Answer Generation: SelfRewardRAG employs three separate GPT-based agents in a manner akin to a multi-agent system. Each agent specializes in generating responses based on a distinct part of the obtained documents. This three-part structure ensures a thorough and exhaustive analysis of the data, resulting in comprehensive answers covering a wide spectrum of existing literature.

Sequential Justification: Each agent systematically analyzes the material, ensuring that the responses are thorough, precise, and based on a comprehensive assessment of the pertinent literature.

Self-Evaluation Layer: The answers generated by the agents undergo rigorous scrutiny by an additional LLM, which performs self-evaluation to identify and rectify any logical

fallacies or errors. This introspective mechanism acts as a safeguard, reinforcing the accuracy and dependability of the generated responses.

Resolution and Improvement: The feedback received from the self-evaluation layer is incorporated, and another LLM is used to select the most optimal answer, further refining it to ensure it meets the highest criteria of precision and dependability.

Output: The system provides a revised solution based on reliable medical literature and evaluated for coherence and logical validity.

The SelfRewardRAG framework represents a sophisticated implementation of AI in the medical domain, aiming to augment human expertise with a meticulous and iterative computational reasoning process. It signifies a significant stride in the ongoing endeavor to harmonize artificial intelligence with the nuanced realm of medical inquiry.

C. Prompting example

The three prompts provided in figure 1 serve as examples of how tasks can be structured for language models to handle complex academic research tasks, especially in the medical and scientific domains. Here's a brief explanation of each prompt based on the context provided by the associated papers: **Prompt 0**: "You are a medical research assistant capable of formulating precise search queries for the PubMed database based on a given medical question. Your task is to extract the essential keywords from the question that would lead to the most relevant articles addressing the specific medical issue at hand. Once you have identified these keywords, you are to combine them using logical operators to create an effective search string. This string should be capable of returning articles that could potentially provide insights or answers to the medical question posed. Your response should list only the keywords identified and the final search string crafted, without including any extraneous information or discussion."

Tasks the model with acting as a medical research assistant, where it must extract relevant keywords from a medical question and construct a precise PubMed search query. This requires an understanding of medical terminologies and the essence of complex questions.

prompt 1: "Assume the role of a medical research assistant with expertise in analyzing complex medical queries. You have a profound understanding of various medical disciplines and are adept at extracting relevant keywords in English from complex medical scenarios to aid in literature searches. Upon receiving a detailed description of a medical query along with specific challenges encountered within that query, your skill involves crafting search queries using the extracted keywords to find the most relevant scholarly articles from PubMed that address those challenges. You always ensure that the articles are relevant to the medical query's context. Follow these steps, listing only the keywords and the crafted search queries in your response:

1. Analyze the text to understand the medical query's context, objectives, and the nature of the challenges encountered.

2. Identify the main medical concepts present in the text. These concepts are the core principles, theories, or terminologies relevant to the query and the challenges described.

3. Extract the most relevant keywords from the text and express them in English, avoiding acronyms and writing out the full words. These terms are crucial for conducting a scholarly article search aimed at addressing the challenges within the query's context. Consider the medical query's context and the challenges to ensure the keywords are highly relevant and specific.

4. Ensure that the keywords you extract facilitate searches that are specific enough to narrow down search results to scholarly articles or resources that directly address the challenges within the medical query's framework.

5. The output must always be expressed in English. This is the only output you return: Using the extracted keywords, craft the most efficient search queries to find relevant articles on PubMed that solve the medical query's challenges, combining these keywords with logical operators (AND, OR). Each query you craft must include keywords from both the medical query's context and its challenges. All queries aim to find articles that address the query's challenges in the specific medical context. Then, craft broader search queries in case the initial ones yield no results. Next, craft search queries using synonyms of the keywords. Finally, explore adjacent fields or broader topics, look into similar technologies, research challenges separately, and include related technologies or keywords. All queries must not include acronyms; write out the full words. List all the crafted search queries in English."

Inspired by the study on chain-of-thought reasoning[13], asks the model to not only understand a scientific query but also identify and use scientific concepts to create relevant PubMed search queries. This prompt tests the model's capability in handling complex information and extracting actionable insights.

Prompt 2: "You are a researcher tasked with investigating the response options provided. List the flaws and faulty logic of each answer option. Let's work this out in a step by step way to be sure we have all the errors."

Aligned with the approach described in "Reflexion: An Autonomous Agent with Dynamic Mentor and Self-Reflection"[14], this prompt focuses on the model characterized as a researcher. Here, the task revolves around the critical examination of multiple answer options by identifying errors and faulty logic. This function simulates a critical peer review process, where the model scrutinizes the reasoning presented in various responses to highlight inaccuracies or fallacies, demonstrating an advanced level of understanding and critical analysis.

Prompt 3: "You are a resolver tasked with 1) finding which of the answer options the researcher thought was best, 2) improving that answer, and 3) Printing the improved answer in full. Let's work this out in a step by step way to be sure we have the right answer."

Based on the DERA[15] approach, assigns the role of a "resolver" to the model, where it must improve upon an identified best answer. This showcases advanced problem-solving through dialogue and iteration.

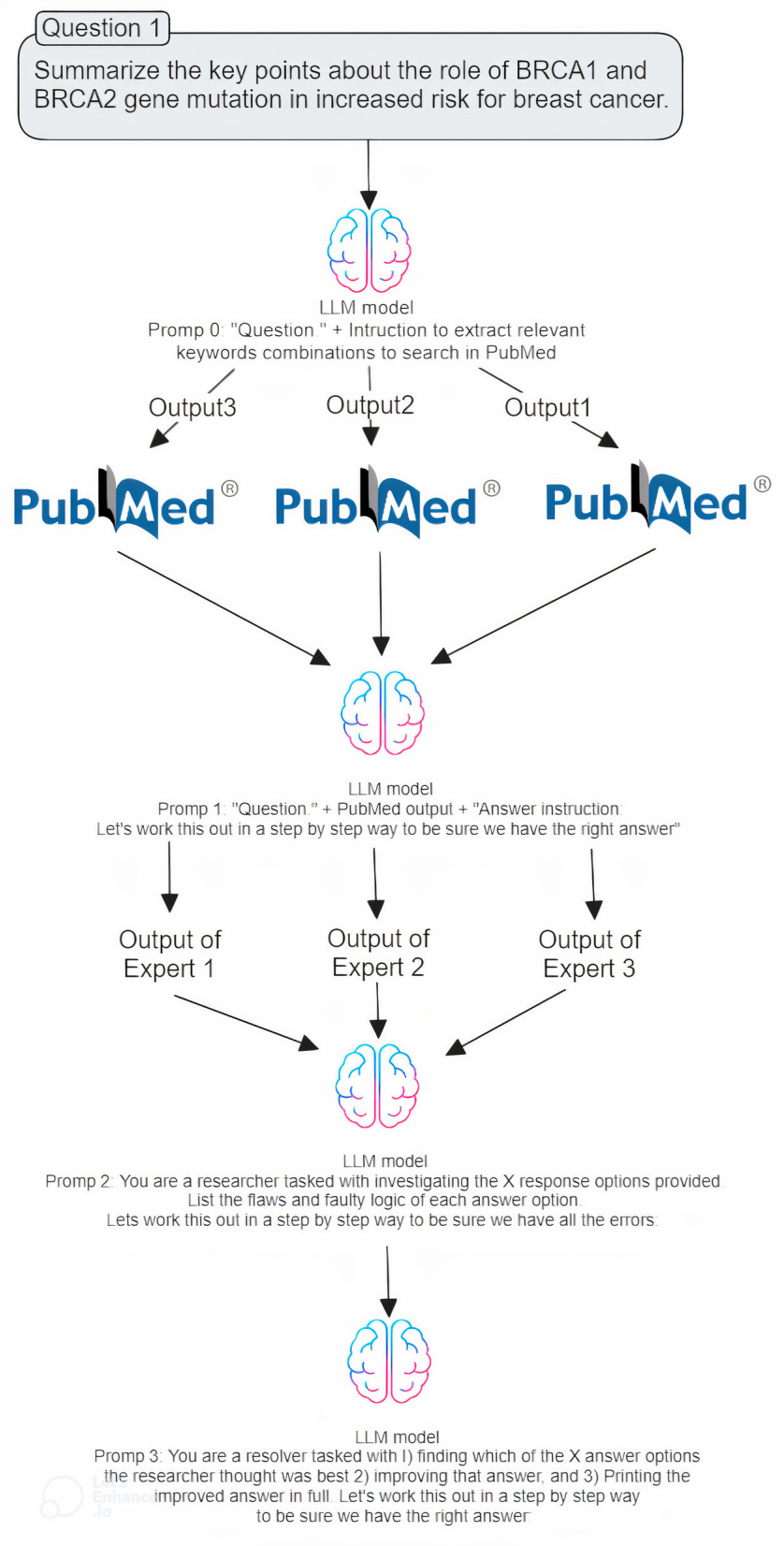

Fig. 1. Key aspects of Healthcare 4.0.

TABLE I
PERFORMANCE ON MEDICAL QUESTION-ANSWERING BENCHMARKS

Model Name	Size	Accuracy (%)			Reference Paper
		PubMedQA	MedQA-USMLE	BioASQ	
GPT-3.5 + SelfRewardRAG (**Ours**)	NA	**81.1**	50.0	95.0	-
GPT-4 (Medprompt)	NA	82.0	51.0	96.0	[16]
Med-PaLM 2	NA	81.8	49.0	95.5	[17]
MEDITRON	70B	81.6	48.0	95.0	[18]
Palmyra-Med	40B	81.1	47.0	94.5	[19]
GPT-4-base	NA	80.4	46.0	94.0	[20]
GPT-3.5 + Z-Code++	175B	79.6	45.0	93.5	[21]
Flan-PaLM (3-shot)	540B	79.0	44.0	93.0	[22]
Codex (5-shot)	175B	78.2	43.0	92.5	[23]
Galactica*	120B	77.6	44.4	94.3	[24]
GPT-4	NA	75.2	42.0	92.0	[20]
DRAGON	360M	73.4	47.5	96.4	[25]
PMC-LLaMA	7B	73.4	40.0	91.5	[26]
BioLinkBERT (large)	340M	72.2	45.1	94.9	[25]
BioLinkBERT (base)	110M	70.2	40.0	91.0	[25]
BioBERT (multi-phase)	110M	68.1	39.0	90.0	[27]
PubMedBERT	110M	55.8	38.1	87.5	[28]

IV. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents a comprehensive evaluation of our SelfRewardRAG model, comparing its performance across several benchmarks in the medical question-answering domain. The goal is to showcase the model's capabilities and position it relative to the current state-of-the-art models. We utilized benchmarks such as PubMedQA[27], MedQA-USMLE[29], and BioASQ[30], as detailed in Table I, to assess the model's proficiency in understanding and responding to complex medical questions derived from a wide range of biomedical literature and clinical scenarios.

A. Benchmark Analysis

- **PubMedQA**: This benchmark assesses the model's comprehension of biomedical literature, requiring answers to questions based on abstracts from PubMed articles. With an accuracy of 81.1%, SelfRewardRAG demonstrates robust capabilities in parsing and interpreting scientific texts, which is fundamental for any AI in the medical field.
- **MedQA-USMLE**: Designed to mimic the United States Medical Licensing Examination, this benchmark evaluates clinical reasoning and decision-making skills. Although SelfRewardRAG's accuracy of 50.0% shows there is room for improvement, it also highlights the model's potential in clinical reasoning, which is essential for further development.
- **BioASQ**: This challenge involves various types of questions, including yes/no, factoid, list-based, and summary. Achieving an accuracy of 95.0%, SelfRewardRAG excels at synthesizing and applying diverse biomedical knowledge, thereby setting a new benchmark for AI in this area.

B. Comparative Performance

The SelfRewardRAG model's performance, as summarized in Table I, showcases its competitive edge against other leading models in the medical question-answering domain. Our model maintains high accuracy across different benchmarks, critical for assessing its real-world applicability in medical settings.

Its nuanced understanding of biomedical literature, evidenced by an 81.1% accuracy on PubMedQA, is on par with leading models like GPT-4 (Medprompt) and Med-PaLM 2. However, it stands out for its comprehensive performance across various types of medical questions, highlighting its versatility.

Moreover, the exceptional 95.0% accuracy on the BioASQ benchmark signifies the model's ability to understand and effectively synthesize a wide range of biomedical information. This performance is noteworthy compared to other models that may focus on broader NLP tasks but fall short in meeting domain-specific requirements.

The model's architecture, which integrates RAG with advanced NLP capabilities and is enhanced by self-reflection/self-reward mechanisms, contributes to its robust performance. This innovative approach leverages the emergent reasoning abilities of LLMs and ensures the relevance and accuracy of the retrieved information. The iterative self-evaluation process further refines the model's outputs, ensuring they meet the high standards required in the medical field.

Overall, the SelfRewardRAG's performance reflects not only a technical achievement but also a significant advancement in AI's application to healthcare. By approaching the cognitive processes of medical experts, the model promises to enhance the quality of information available to healthcare professionals, potentially leading to better patient outcomes.

C. Model's Practical Applicability

The SelfRewardRAG model's real-world applicability is evident from its strong performance across the discussed benchmarks. It presents a promising solution to the critical need for accurate and up-to-date medical information in clinical settings. Outperforming generalist models like GPT-4-base and PubMedBERT in medical-specific tasks reaffirms the potential of specialized AI systems in healthcare. By mimicking the cognitive processes of medical experts, the model ensures the relevance and accuracy of its provided information.

D. Future Directions and Conclusion

Our experimental evaluation validates the efficacy of the SelfRewardRAG framework in enhancing medical reasoning through AI. It opens avenues for further refinement and development of AI systems capable of supporting medical professionals with high precision and reliability. The model's success in these benchmarks indicates that continued enhancements in knowledge integration and reasoning methodologies will be crucial for the future of AI in healthcare.

As we continue to develop AI technologies for medicine, the SelfRewardRAG model exemplifies the potential of such systems to significantly improve our capabilities in understanding and addressing complex medical queries. Its strategic combination of RAG, LLMs, and iterative refinement processes represents a significant step forward in applying AI within specialized knowledge domains.

V. Conclusion

In this research, we introduced SelfRewardRAG, a novel framework that integrates LLMs with RAG to enhance medical reasoning. This method transcends the limitations of traditional LLMs, which are often constrained by static datasets. By dynamically retrieving current health information and synthesizing this data into precise responses to medical inquiries, SelfRewardRAG marks a significant advance in the application of AI in healthcare.

Our experimental results across multiple benchmarks, including PubMedQA, MedQA-USMLE, and BioASQ, demonstrate the effectiveness of SelfRewardRAG in delivering accurate and timely answers to complex medical questions. The model not only competes with but in some instances surpasses, existing state-of-the-art models, showcasing its potential to revolutionize medical AI applications by ensuring the information remains up-to-date and relevant.

Despite these promising outcomes, there are several limitations to our approach that require consideration. Firstly, the success of SelfRewardRAG heavily relies on the quality and accessibility of external databases like PubMed. Limitations in these databases, such as coverage gaps or inaccuracies, could adversely affect the model's ability to retrieve pertinent information and generate accurate responses.

Secondly, the computational complexity and resource demands of running SelfRewardRAG, due to its dependence on advanced LLMs and dynamic data retrieval, might pose challenges for its widespread adoption, especially in resource-limited settings.

Thirdly, although the model is proficient in synthesizing medical information, there is a continuous challenge in ensuring the reliability and accuracy of AI-generated advice in healthcare settings. The risk of misinterpretation or excessive reliance on AI-generated responses without sufficient human oversight is a significant concern.

Addressing these limitations presents several avenues for future research. Improving the robustness of data retrieval mechanisms to ensure access to high-quality, comprehensive databases is essential. Moreover, enhancing the computational efficiency of SelfRewardRAG could promote its broader implementation across various healthcare environments. Additionally, further research into incorporating human oversight mechanisms and ethical considerations into the AI decision-making process is vital to ensure that such systems can be safely and effectively integrated into clinical workflows.

References

[1] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[2] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," 2020.

[3] Z. Hamane, A. Samih, and A. Fennan, "Healthpathfinder: Navigating the healthcare knowledge graph with neural attention for personalized health recommendations," in *Innovations in Smart Cities Applications Volume 7*, ser. Lecture Notes in Networks and Systems, vol. 906. Proceedings of the International Conference on Smart City Applications: SCA 2023, 2024, pp. 429–446.

[4] F.-E. Ben-Bouazza, Y. Bennani, G. Cabanes, and A. Touzani, "Unsupervised collaborative learning based on optimal transport theory," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 698–719, 2021.

[5] J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, and D. Amodei, "Scaling laws for neural language models," 2020.

[6] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," 2023.

[7] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. tau Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-augmented generation for knowledge-intensive nlp tasks," 2021.

[8] J.-B. Lamy, V. Ebrahiminia, C. Riou, B. Seroussi, J. Bouaud, C. Simon, S. Dubois, A. Butti, G. Simon, M. Favre, H. Falcoff, and A. Venot, "How to translate therapeutic recommendations in clinical practice guidelines into rules for critiquing physician prescriptions? methods and application to five guidelines," *BMC Medical Informatics and Decision Making*, vol. 10, no. 1, p. 31, 2010. [Online]. Available: https://doi.org/10.1186/1472-6947-10-31

[9] A. Rao and J. Palma, "Clinical decision support in the neonatal icu," in *Seminars in Fetal and Neonatal Medicine*, vol. 27, no. 5. Elsevier, 2022, p. 101332.

[10] Y. Guo, W. Qiu, G. Leroy, S. Wang, and T. Cohen, "Cells: A parallel corpus for biomedical lay language generation," *arXiv preprint arXiv:2211.03818*, 2022.

[11] M. Ranjit, G. Ganapathy, R. Manuel, and T. Ganu, "Retrieval augmented chest x-ray report generation using openai gpt models," 2023.
[12] Y. Ke, L. Jin, K. Elangovan, H. R. Abdullah, N. Liu, A. T. H. Sia, C. R. Soh, J. Y. M. Tung, J. C. L. Ong, and D. S. W. Ting, "Development and testing of retrieval augmented generation in large language models – a case study report," 2024.
[13] K. Hebenstreit, R. Praas, L. P. Kiesewetter, and M. Samwald, "An automatically discovered chain-of-thought prompt generalizes to novel models and datasets," *arXiv preprint arXiv:2305.02897*, 2023, submitted on 4 May 2023 (v1), last revised 3 Aug 2023 (this version, v2).
[14] N. Shinn, F. Cassano, E. Berman, A. Gopinath, K. Narasimhan, and S. Yao, "Reflexion: Language agents with verbal reinforcement learning," *arXiv preprint arXiv:2303.11366*, 2023, submitted on 20 Mar 2023 (v1), last revised 10 Oct 2023 (this version, v4).
[15] V. Nair, E. Schumacher, G. Tso, and A. Kannan, "Dera: Enhancing large language model completions with dialog-enabled resolving agents," *arXiv preprint arXiv:2303.17071*, 2023, submitted on 30 Mar 2023.
[16] H. Nori, Y. T. Lee, S. Zhang, D. Carignan, R. Edgar, N. Fusi, N. King, J. Larson, Y. Li, W. Liu, R. Luo, S. M. McKinney, R. O. Ness, H. Poon, T. Qin, N. Usuyama, C. White, and E. Horvitz, "Can generalist foundation models outcompete special-purpose tuning? case study in medicine," 2023.
[17] K. Singhal, T. Tu, J. Gottweis, R. Sayres, E. Wulczyn, L. Hou, K. Clark, S. Pfohl, H. Cole-Lewis, D. Neal, M. Schaekermann, A. Wang, M. Amin, S. Lachgar, P. Mansfield, S. Prakash, B. Green, E. Dominowska, B. A. y Arcas, N. Tomasev, Y. Liu, R. Wong, C. Semturs, S. S. Mahdavi, J. Barral, D. Webster, G. S. Corrado, Y. Matias, S. Azizi, A. Karthikesalingam, and V. Natarajan, "Towards expert-level medical question answering with large language models," 2023.
[18] Z. Chen, A. H. Cano, A. Romanou, A. Bonnet, K. Matoba, F. Salvi, M. Pagliardini, S. Fan, A. Köpf, A. Mohtashami, A. Sallinen, A. Sakhaeirad, V. Swamy, I. Krawczuk, D. Bayazit, A. Marmet, S. Montariol, M.-A. Hartley, M. Jaggi, and A. Bosselut, "Meditron-70b: Scaling medical pretraining for large language models," 2023.
[19] K. Kamble and W. Alshikh, "Palmyra-med: Instruction-based fine-tuning of llms enhancing medical domain performance," 2023.
[20] H. Nori, N. King, S. M. McKinney, D. Carignan, and E. Horvitz, "Capabilities of gpt-4 on medical challenge problems," 2023.
[21] P. He, B. Peng, L. Lu, S. Wang, J. Mei, Y. Liu, R. Xu, H. H. Awadalla, Y. Shi, C. Zhu, W. Xiong, M. Zeng, J. Gao, and X. Huang, "Z-code++: A pre-trained language model optimized for abstractive summarization," 2023.
[22] K. Singhal, S. Azizi, T. Tu, S. S. Mahdavi, J. Wei, H. W. Chung, N. Scales, A. Tanwani, H. Cole-Lewis, S. Pfohl, P. Payne, M. Seneviratne, P. Gamble, C. Kelly, N. Scharli, A. Chowdhery, P. Mansfield, B. A. y Arcas, D. Webster, G. S. Corrado, Y. Matias, K. Chou, J. Gottweis, N. Tomasev, Y. Liu, A. Rajkomar, J. Barral, C. Semturs, A. Karthikesalingam, and V. Natarajan, "Large language models encode clinical knowledge," 2022.
[23] V. Liévin, C. E. Hother, A. G. Motzfeldt, and O. Winther, "Can large language models reason about medical questions?" 2023.
[24] R. Taylor, M. Kardas, G. Cucurull, T. Scialom, A. Hartshorn, E. Saravia, A. Poulton, V. Kerkez, and R. Stojnic, "Galactica: A large language model for science," 2022.
[25] M. Yasunaga, J. Leskovec, and P. Liang, "Linkbert: Pretraining language models with document links," 2022.
[26] C. Wu, W. Lin, X. Zhang, Y. Zhang, Y. Wang, and W. Xie, "Pmc-llama: Towards building open-source language models for medicine," 2023.
[27] Q. Jin, B. Dhingra, Z. Liu, W. W. Cohen, and X. Lu, "Pubmedqa: A dataset for biomedical research question answering," 2019.
[28] Y. Gu, R. Tinn, H. Cheng, M. Lucas, N. Usuyama, X. Liu, T. Naumann, J. Gao, and H. Poon, "Domain-specific language model pretraining for biomedical natural language processing," *ACM Transactions on Computing for Healthcare*, vol. 3, no. 1, p. 1–23, Oct. 2021. [Online]. Available: http://dx.doi.org/10.1145/3458754
[29] D. Jin, E. Pan, N. Oufattole, W. Weng, H. Fang, and P. Szolovits, "What disease does this patient have? A large-scale open domain question answering dataset from medical exams," *CoRR*, vol. abs/2009.13081, 2020. [Online]. Available: https://arxiv.org/abs/2009.13081
[30] G. Tsatsaronis, G. Balikas, P. Malakasiotis *et al.*, "An overview of the bioasq large-scale biomedical semantic indexing and question answering competition," *BMC Bioinformatics*, vol. 16, no. 1, p. 138, 2015. [Online]. Available: https://doi.org/10.1186/s12859-015-0564-6