Human Interaction and Emerging Technologies (IHIET-AI 2024), Vol. 120, 2024, 458–463
https://doi.org/10.54941/ahfe1004597

ReActIn: Infusing Human Feedback Into Intermediate Prompting Steps of Large Language Models

Manuel Delafor[1], Claire Gendron[2], Wangfan Li[2], Cecilia Delgado Solorzano[2], and Carlos Toxtli[2]

[1]Human-AI Empowerment Laboratory, Clemson, SC 29631, USA
[2]Clemson University, Clemson, SC 29631, USA

ABSTRACT

This paper introduces ReActIn, a framework designed to infuse human feedback into the intermediate prompting steps of large language models. The practicality and effectiveness of ReActIn are validated through experiments that apply four established prompting strategies, evaluated both with and without human feedback integration. The proposed architecture's performance is compared against traditional large language models across various tasks using four standard evaluation tests. Our findings reveal that the integration of human feedback has a direct impact on the reasoning, action prompting, and overall decision-making capabilities of the language models. This study underscores the potential of ReActIn to shape a future where sophisticated, context-aware AI systems, empowered by human feedback, can effectively navigate complex real-world scenarios.

Keywords: Large language models, Human computer interaction

INTRODUCTION

Large language models (LLMs), despite their impressive ability to generate human-like text, often struggle with inaccuracies and the generation of non-factual or "hallucinated" information. As their responses tend to prioritize what appears plausible over what is factually correct, the technology frequently delivers inconsistent outputs, which mislead users or obscure essential information. To tackle these issues, we are working on an approach to improve LLM responses' factual and logical accuracy and reduce hallucination phenomena.

Our approach, ReactIn (Reason, Act, and Interact), integrates both AI and human inputs and is divided into two critical steps. The first involves crafting a well structured initial prompt to constrain the LLM's basic response space. The second introduces a human-assisted intervention in the language generation process to evaluate and guide the LLM's outputs. By providing this "check" we attempt to eliminate some of the problems of common average possible outputs based on learned patterns.

To validate our approach, we are conducting our tests against a series of problematic prompts that have proved to be challenging for LLMs.

Our hand-crafted initial explorations show that our methodology leads to improvements in LLM performance, so we are moving towards its operationalization and evaluation. Although at an early stage, our work paves the way towards altering the distribution of pre-generated answers, placing a higher priority on factual correctness and logical consistency rather than mere linguistic plausibility. Ultimately, we aim to improve the credibility and utility of LLMs, instigating a shift in LLMs text-generation that fosters accuracy, reliability, and trustworthiness.

RELATED WORK

The field of artificial intelligence has seen a surge of research focusing on the enhancement of language models through various methods, particularly the integration of human feedback and in-context learning. These studies have laid the groundwork for our research on ReActIn, a novel cognitive architecture that dynamically evolves by incorporating human feedback into its decision-making processes. Previous research has demonstrated the effectiveness of Reinforcement Learning from Human Feedback (RLHF) in aligning models with the human goals (Bai et al., 2022). This approach, which involves training models to improve their comprehension and reasoning abilities, aligns with our goal of integrating human feedback into ReActIn's decision-making processes but from the In-Context learning perspective (Wang et al., 2023).

The idea of training AI agents to think like humans, known as Thought Cloning, has been proposed as a way to improve AI safety and interoperability (Hu and Clune, 2023). This approach not only clones the behaviors of human demonstrators but also their thoughts as they perform these behaviors, providing valuable insights into the potential of human-like reasoning in AI systems, a key aspect of our work on ReActIn. The Socratic Method, which uses a series of probing questions to get to the truth of things, has been shown to reduce hallucination and improve logical reasoning when integrated into LLMs (Chang, 2023). The Chain of Thought prompting technique encourages step-by-step reasoning and mitigates some risks of disorganization in model-generated text (Wei et al., 2022). Additionally, Tree of Thought prompting leverages hierarchical reasoning to break down complex prompts (Yao et al., 2023). Other methods such as the ReAct framework use tools such as web browsing and calculators to validate or identify relevant information outside of the LLM knowledge (Yao et al., 2022).

Other relevant approaches include prompt programming, which uses demonstrations to teach LLMs new skills (Reynolds and Mc-Donell, 2021), conversational semantic parsing for iterative querying (Wei et al., 2022), and debate modeling, where two models argue opposing sides of an issue (Liang et al., 2023). Active learning provides human feedback on model outputs to refine predictions (Margatina et al., 2023), while calibrated perturbations help models determine when to abstain from answering (Zhao et al., 2023). While these studies have made significant contributions to the field, our work on ReActIn introduces a novel approach to integrating human feedback into the decision-making processes of AI systems. We demonstrate the efficacy of ReActIn through a series of experiments and provide an in-depth analysis of how the incorporation of human feedback contributes to the continuous evolution and enhancement of the cognitive architecture. Our work has significant implications for the development of more sophisticated, context-aware AI systems that can better understand and respond to complex, real-world scenarios.

SYSTEM DESCRIPTION

The system is a web application with front-end and back-end components. The front-end provides a chatbot-style user interface using HTML, CSS and JavaScript. Conversations are displayed between a user and AI agent. The UI allows users to provide feedback on the intermediate reasoning steps generated by the AI and rewrite them in their own words. The back-end is built using Python and the Flask framework. It delivers the static webpages and implements REST APIs to interface with the OpenAI API. User feedback is collected through the front-end UI and stored in a MySQL database.

To generate responses, the system leverages pre-recorded conversations consisting of questions from the MMLU benchmark dataset. Established prompting techniques like Tree of Thought, Chain of Thought, and Socratic Questioning are used to produce intermediate reasoning steps along with the final answer. The ReActIn framework in the back-end integrates the collected user feedback into the LLM using three approaches:

- N-Shot Learning: User rewritten reasoning steps are provided as demonstrations to the LLM to learn to formulate new reasoning chains.
- Substitution: The original intermediate reasoning messages generated by the AI are replaced with the user's rewritten versions.
- Conversational: User feedback is provided to the LLM following its initial reasoning to encourage iterative refinement of the chain of thought.

The modular architecture allows easy extensibility. New prompting techniques and human feedback integration strategies can be incorporated as needed.

METHODOLOGY

Our methodology utilizes a three-phase approach:

Phase 1 - Prompt Engineering: Prompts are crafted using questions from the MMLU, ARC, HellaSwag, and WinoGrande QA benchmarks benchmarks to constrain the initial response space and encourage factual accuracy over plausibility.

Phase 2 - Human-in-the-Loop: Participants interact with pre-generated responses produced using established prompting techniques. They provide feedback and rewrite intermediate reasoning chains in their own words to steer away from flawed averaged outputs.

Phase 3 - Human Feedback Integration: The collected feedback is integrated into the LLM using the ReActIn framework through three strategies - N-Shot Learning, Substitution, and Conversational.

Finally, ReActIn's performance with and without human feedback is evaluated by testing on the MMLU, ARC, HellaSwag, and WinoGrande QA benchmarks. The optimal combinations of prompting techniques and human feedback integration methods are identified.

This phased approach allows comprehensive analysis of human feedback's impact on the LLM's reasoning and decision-making capabilities. The methodology is designed to enable iterative enhancement of the cognitive architecture. We hypothesize that integrating human input through ReActIn will significantly enhance logical accuracy and reduce hallucination phenomena in LLM outputs.

DISCUSSION

Large language models (LLMs), as they stand today, are prodigious achievements. However, their output accuracy is currently limited by at least two key issues. Firstly, these models are prone to generate "hallucinated" responses—outputs that are plausible sounding, but factually incorrect or nonsensical. Secondly, LLMs often prioritize what "sounds right" over what is factually correct, due to their training on vast text corpora where they learn to mimic human like text patterns without truly comprehending the semantic and factual elements of the context.

Our view is that this is primarily because the models, trained on a probabilistic basis, are not wired to think or reason in a way that humans do. Instead, they often average or blend potential responses based on learned patterns from their training data, a factor that can result in plausible but incorrect answers. These problems create an inconsistent output quality, leading to situations where LLMs provide incomplete or incorrect responses that can mislead users or obscure crucial information.

Our research offers an approach to rectify these issues, aiming to improve the factual accuracy of LLM responses and reduce the incidence of hallucinations. We propose two things, one is a well structured initial prompt, and secondly an intermediate human-assisted step in the language generation process. This step includes human evaluation and guidance, effectively helping LLMs to "think" and ensuring a more accurate and logically sound response generation, instead of simply averaging possible outputs based on learned patterns.

To validate our proposed solution, we will conduct tests against a set of problematic prompts that are typically challenging for LLMs. By reducing hallucination and improving logical and factual accuracy, we hope to help make LLMs more useful and trustworthy for a range of applications, from drafting reports to providing real-time assistance in various professional contexts.

Although we are at an early stage and much work remains to be done, the initial results from our research provide a promising foundation for improving the credibility and usefulness of LLMs. We believe our proposal will instigate a shift in LLM development, putting us on a path towards creating AI models that generate reliable, accurate, and trustworthy responses.

CONCLUSION

Our work-in-progress, ReActIn, seeks to mitigate the challenges faced by large language models (LLMs) in generating reliable, accurate, and logically sound outputs. Grounded in a dual-pronged approach, ReActIn strategically integrates human feedback within the language generation process, thus leading to a significant enhancement in the overall quality of LLM responses. Initial results suggest that a well-structured prompt followed by intermediate human assistance can effectively constrain the LLM's response space, encouraging it to prioritize factual correctness and logical consistency. This human-AI synergy brings a more conscious thought process to the language models, shifting them away from simply generating plausible but often misleading outputs based on learned patterns. Consequently, the incidences of "hallucination" or the generation of non-factual information are expected to be reduced substantially. While we have yet to reach the stage of presenting concrete findings, the promise shown by ReActIn offers a solid foundation for continued investigation. Notwithstanding the limitations of the current stage of research, the framework is revealing a potential pathway for substantial improvements in the performance of LLMs.

Looking ahead, we envision a future where our efforts in incorporating human feedback into AI systems will render LLMs that are not just linguistically adept, but also contextually aware and factually accurate. The ultimate goal is to refine the utility, reliability, and credibility of LLMs, thereby instigating a shift in LLMs text-generation that fosters accuracy, reliability, and trustworthiness. The implications of ReActIn are far-reaching. If successful, our approach could redefine how AI interacts with and responds to human input, thereby fundamentally transforming LLMs' role across a myriad of applications. Future work will focus on the continued optimization of the ReActIn framework, further testing, and more extensive validation with larger and more diverse problematic prompt sets.

In conclusion, despite being a work in progress, the ReActIn framework's potential to enhance the effectiveness of large language models is apparent. This work underlines the pivotal role of human feedback in the intermediate steps of the language generation process, a factor that has been largely unexplored till now. The promising start to this research venture offers hope for a future where AI's ability to navigate complex real-world scenarios is significantly improved, thereby maximizing its utility and reliability. We look forward to sharing our results and further developments in subsequent stages of this research.

REFERENCES

Edward Y Chang. 2023. Prompting large language models with the socratic method. In 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC). IEEE, 0351–0360.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. Advances in Neural Information Processing Systems 35 (2022), 24824–24837.

Katerina Margatina, Timo Schick, Nikolaos Aletras, and Jane Dwivedi-Yu. 2023. Active Learning Principles for In-Context Learning with Large Language Models. arXiv preprint arXiv:2305.14264 (2023).

Laria Reynolds and Kyle McDonell. 2021. Prompt programming for large language models: Beyond the few-shot paradigm. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. 1–7.

Liang Wang, Nan Yang, and Furu Wei. 2023. Learning to Retrieve In-Context Examples for Large Language Models. arXiv preprint arXiv:2307.07164 (2023).

Shengran Hu and Jeff Clune. 2023. Thought Cloning: Learning to Think while Acting by Imitating Human Thinking. arXiv preprint arXiv:2306.00323 (2023).

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of thoughts: Deliberate problem solving with large language models. arXiv preprint arXiv:2305.10601 (2023).

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. 2022. React: Synergizing reasoning and acting in language models. arXiv preprint arXiv:2210.03629 (2022).

Theodore Zhao, Mu Wei, J Samuel Preston, and Hoifung Poon. 2023. Automatic Calibration and Error Correction for Large Language Models via Pareto Optimal Self-Supervision. arXiv preprint arXiv:2306.16564 (2023).

Tian Liang, Zhiwei He, Wenxiang Jiao, Xing Wang, Yan Wang, Rui Wang, Yujiu Yang, Zhaopeng Tu, and Shuming Shi. 2023. Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate. arXiv preprint arXiv:2305.19118 (2023).

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. 2022. Training a helpful and harmless assistant with reinforcement learning from human feedback. arXiv preprint arXiv:2204.05862 (2022).