Kashihara, A. et al. (Eds.) (2024). Proceedings of the 32nd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Developing an LLM-Empowered Agent to Enhance Student Collaborative Learning Through Group Discussion

Sixu AN[a*], Yicong LI[a], Yu YANG[a], Yunsi MA[a], Gary CHENG[b] & Guandong XU[a]
[a]*Centre for Learning, Teaching and Technology*
[b]*Department of Mathematics & Information Technology,*
The Education University of Hong Kong, Hong Kong S.A.R. China
*s1154340@s.eduhk.hk

Abstract: Aiming at improving collaborative learning, the current study builds and tests an LLM-empowered agent to enhance student engagement in group discussion. We introduce a four-module conversational system, providing a user-friendly chat website integrated with an LLM agent, where students can discuss on and learn a specific topic in an online classroom. The LLM agent can continuously monitor the dialogue process and give constructive and reflective responses as a knowledgeable learning peer to engage students in the computer-supported collaborative learning (CSCL) environment. To evaluate the pedagogical performance of the system, three LLMs were tested by prompting. The results showed that LLMs with only prompting were unable to accurately process multi-user dialogue information and lacked pedagogical strategies in their responses.

Keywords: Artificial intelligence, LLM agent, computer-supported collaborative learning, conversational system

1. Introduction

Collaborative learning (CL) is a crucial instructional strategy for enhancing student performance, as evidenced by numerous studies (Johnson et al., 2014; Laal & Ghodsi, 2012). Group discussion, a widely recognized CL method, involves organizing students into groups to discuss specific questions or topics, thereby fostering deeper understanding and critical thinking (Gokhale, 1995). In traditional group discussions, teachers encourage student interaction and review posts to assess understanding, which is time-consuming. With multiple groups discussing simultaneously, a single teacher cannot effectively participate in all, reducing teaching efficiency. Sun et al. (2024) argued that a well-designed computer-supported collaborative learning (CSCL) environment, coupled with effective instructional strategies and pedagogical approaches has beneficial effects on students' academic performance and interaction.

With its rapid development in recent days, artificial intelligence (AI) and large language models (LLMs) show great potential in enhancing the CSCL environment by supporting the student-AI collaboration (Kim et al., 2022). Do et al. (2022) examined various agent communication strategies for facilitating group discussions. However, most of the existing tutor systems for group discussion are limited in scope and lack substantial pedagogical depth. Some tutor systems incorporating VR technology and LLMs were constrained in their functionality, as the agents primarily addressed limited classroom actions such as note-taking, posing questions, and interacting with students (Kim et al., 2024; Liu et al., 2024). To further enhance the capabilities of LLM-based tutor agents, Mao et al. (2024) developed a multi-user discussion assistant designed to increase user engagement in group discussions; however, this system does not emphasize the educational role of the agent.

Building on existing research, this study explores how to harness the capabilities of LLMs to develop an agent that can assist teachers and facilitate group discussions in the CSCL environment. The researchers posit that an effective CL assistant agent should not only respond to students' inquiries but also actively promote student engagement. A prevalent challenge in CSCL environment is the varying levels of student participation. Specifically, some students dominate discussions, while others remain silent, often due to a lack of understanding of the topics being discussed, which diminishes their willingness to participate. Therefore, the present research aims to create a novel educational platform integrated with an LLM agent capable of answering students' questions, offering pedagogically sound responses, and enhancing overall student engagement.

2. Existing Chatbots for CSCL

Many works have tried to build CSCL tools. Intent-based chatbots and retrieval-augmented generation chatbots are two different ways. Intent-based chatbots try to understand user intentions from conversational inputs. Kumar, J. A (Kumar, J. A. 2021) built intent-based chatbots for a team-based design course. In this work, the author used different intent-based chatbots such as a welcome bot, project registration bot, peer-to-peer evaluation bot, etc. to guide students during the discussion process. For retrieval-augmented generation chatbots, chatbots can retrieve from a pre-set educational corpus to find the most possible response. Nguyen, H. D. et al. (Nguyen, et al. 2022) designed an intelligent educational chatbot for information retrieval. Thway, M, et al.(Thway, M., et al. 2024) built a teacher role retrieval-augmented chatbot to answer questions from students but not focus on engagement. However, one drawback of the retrieval-augmented method is that it usually requires that developers build an appropriate educational corpus in advance, which is time-consuming. Although some existing methods can be used to build CSCL, few of them focus on how to improve student engagement. In this work, we built an LLM-empowered agent to enhance student engagement during group discussions.

3. Methodology

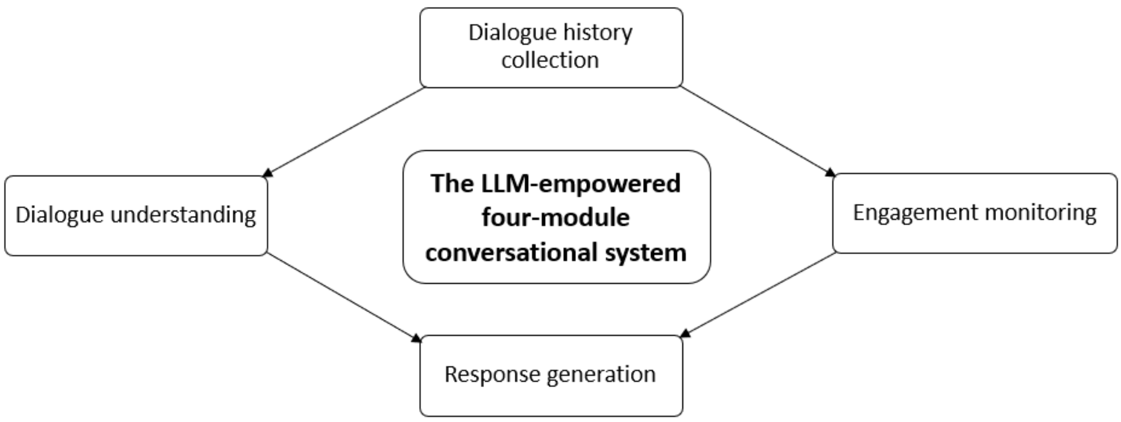


Figure 1. The LLM-empowered four-module conversational system

In this research, we developed an LLM-empowered Computer-Supported Collaborative Learning (CSCL) platform, structured as a multi-person conversational system with an integrated discussion website interface. Students who join the discussion room can access the platform and log in as individual users. An LLM agent is embedded within the discussion room to provide pedagogically appropriate responses, guide the discussion, and monitor students' engagement levels. The system is divided into four modules: dialogue history collection, dialogue understanding, engagement monitoring, and response generation (Figure 1). Additionally, extensive tutor dialogue data will be utilized to train the LLM, ensuring the generation of more professional responses for instructional and pedagogical guidance. An appropriate database will also be selected to store user dialogues, enabling the agent to process multi-user interactions with greater accuracy.

In Dialogue history collection, dialogue information is stored in the Mysql database with user ID, speak content, and speak time. When students do not send more information within 1 minute, the dialogue understanding and engagement monitoring module starts to work. The period between these two modules starting to work and the agent giving a response is called the detection period. In the dialogue understanding module, LLM uses spoken content extracted from the database and then generates an appropriate response. In the engagement monitoring module, the user ID information is extracted and the speech frequency of each speaker is calculated. The speaker whose speech frequency is 0 will be identified as inactive frequency during the detection period. The identified inactive student name from the engagement monitoring module will be sent to the response generation module. Then, the LLM agent will call inactive students to express their ideas more.

To use the system, students log in and enter the discussion room interface. Once all participants are ready, the agent initiates the discussion by posting an introductory message about the topic. Students can then type their contributions directly into the website interface, sharing their ideas. Throughout the discussion, the agent continuously collects and processes the dialogue to monitor each student's level of engagement. If the agent identifies inactive students, it will tag them in the chat and post corresponding messages to encourage participation. Additionally, if the agent detects differing opinions among students on a particular question, it will analyze the dialogue history and generate pedagogical responses to guide the discussion and help students arrive at the correct answers.

The LLM agent was implemented through extensive prompting, and several LLM models, including Qwen, Ernie-Speed-128k, and Ernie-Lite-8K, were tested to evaluate their effectiveness in achieving learning goals during group discussions (Bai et al., 2023; Baidu, 2024). The prompt used is illustrated in Figure 1, specifying the discussion topic and the agent's tasks. This prompt is used in the Response generation module for prompt-based agent build. This Existing research indicates that an agent speaking in a teacher-like manner can decrease students' willingness to participate (Nguyen, 2023). Therefore, the agent was designed to act as a learning peer, participating in the discussion alongside the students.

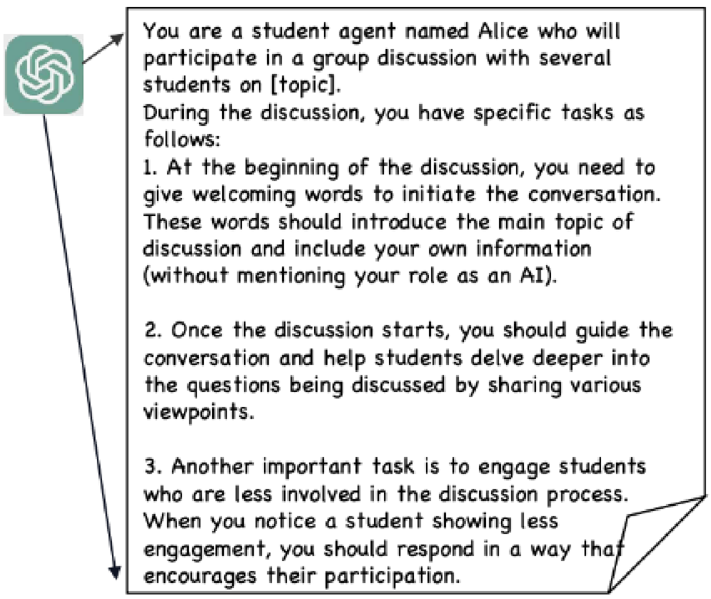


Figure 2. Agent prompt for designing it as a learning peer

After constructing the agent with the specified prompt, student discussion dialogue history collected from actual course cases was utilized to test the agent's ability to provide appropriate responses that enhance student engagement as presented in Figure 3.

Figure 3. Dialogue between the agent and the students using LLM Ernie-speed-128k

4. Results

4.1 Test Result Discussion

Figures 3 - 5 illustrate the performance and test results of the Ernie-Speed-128k and Qwen LLM models, respectively. As shown in Figure 2, the Ernie-Speed-128k model provides basic welcome information and topic introduction messages to multiple users but generates repetitive responses for different users. In contrast, the Qwen model, depicted in Figures 3 and 4, responds directly to student questions without offering guidance. These observations suggest that agents relying solely on prompting cannot effectively process the entire dialogue history to maintain continuity. Additionally, without ongoing training on tutor dialogue data, these agents fail to provide sufficiently professional discussion guidance. Furthermore, the prompt-based method struggles to accurately manage multi-student dialogues, often responding to the wrong student.

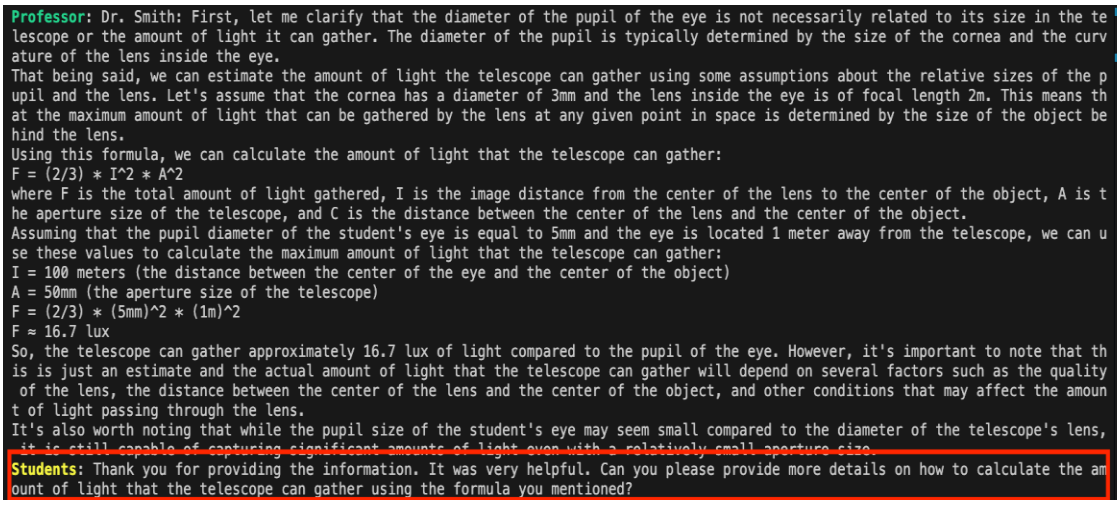
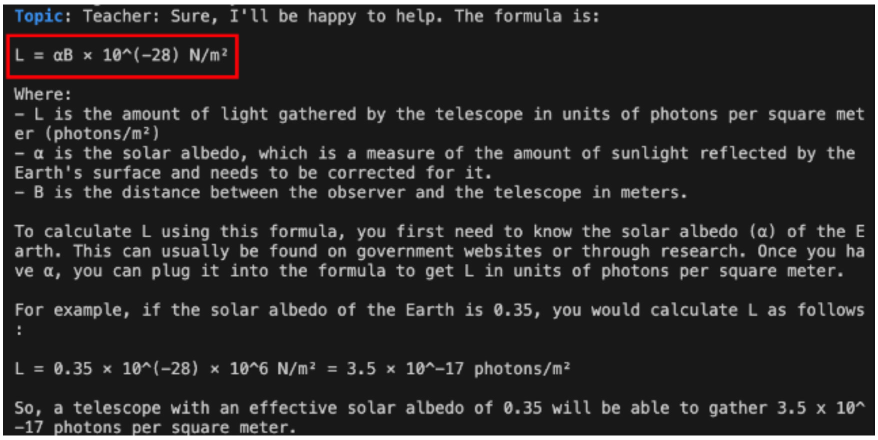


Figure 4. Qwen test results



```
Topic: Teacher: Sure, I'll be happy to help. The formula is:
L = αB × 10^(-28) N/m²

Where:
- L is the amount of light gathered by the telescope in units of photons per square met
er (photons/m²)
- α is the solar albedo, which is a measure of the amount of sunlight reflected by the
Earth's surface and needs to be corrected for it.
- B is the distance between the observer and the telescope in meters.

To calculate L using this formula, you first need to know the solar albedo (α) of the E
arth. This can usually be found on government websites or through research. Once you ha
ve α, you can plug it into the formula to get L in units of photons per square meter.

For example, if the solar albedo of the Earth is 0.35, you would calculate L as follows
:

L = 0.35 × 10^(-28) × 10^6 N/m² = 3.5 × 10^-17 photons/m²

So, a telescope with an effective solar albedo of 0.35 will be able to gather 3.5 x 10^
-17 photons per square meter.
```

Figure 5. Qwen response results

4.2 User Interface of the LLM-empowered Chatbot System

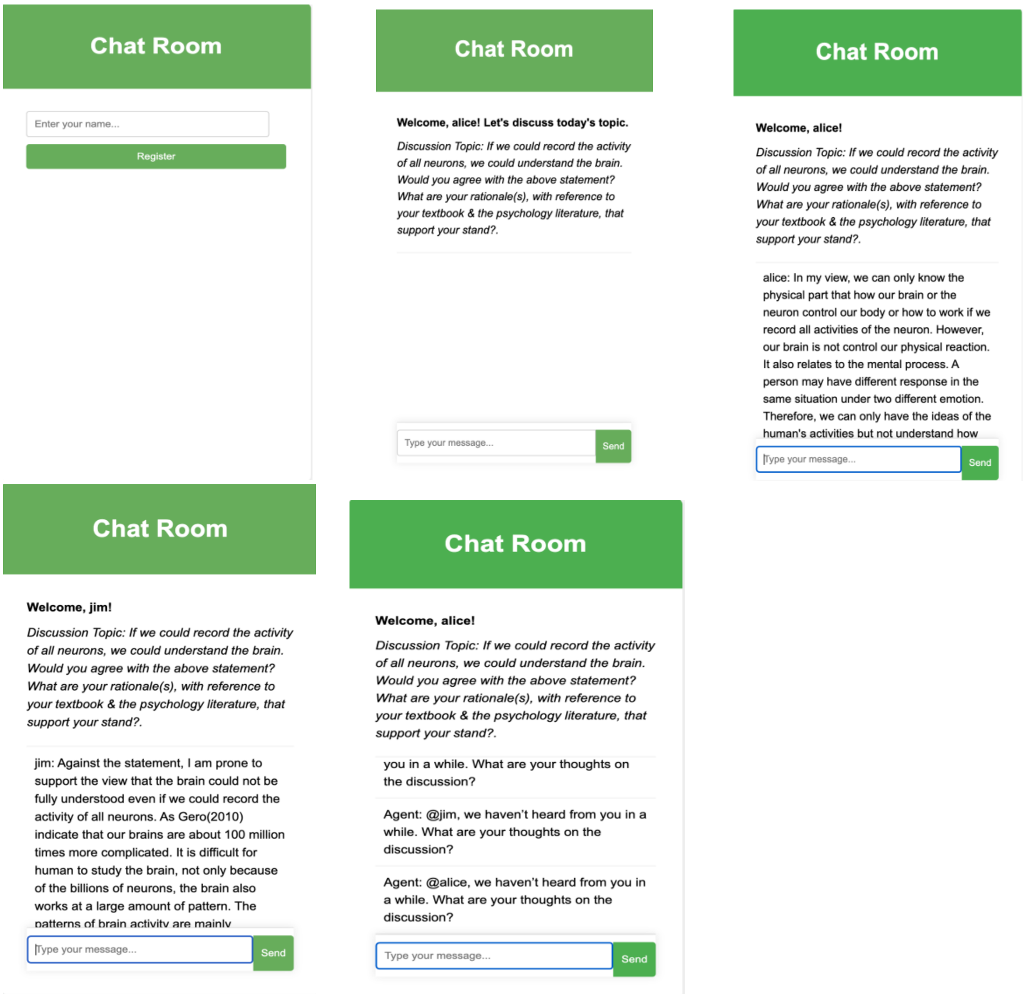


Figure 6. Proposed system

The user chat room interface is illustrated in Figure 6. The welcome message and discussion topic were shown at the top of the website. The agent will give a response based on the user's dialogues. In the chatroom, users can see messages from others and when they become inactive, the agent will @ the user to talk more.

5. Conclusion

In this work, we develop a group discussion chatroom with an LLM-empowered agent to enhance student engagement. We found that general-purpose LLMs with prompting techniques are unable to process multi-user dialogue information accurately, and the generated responses failed to accommodate pedagogical strategies. In the future, we will further redesign engagement monitoring and promotion algorithms, carry out system evaluation with students, and collect feedback for further model improvement.

Acknowledgment

This work thanks for the support from EdUHK-ECNU Joint AI Grant Scheme.

References

Bai, J., Bai, S., Chu, Y., Cui, Z., Dang, K., Deng, X., ... & Zhu, T. (2023). Qwen technical report. *arXiv preprint arXiv:2309.16609*.

Baidu Intelligent Cloud Documentation. 2024. ERNIESpeed-128K - Qianfan Large Model Platform. Retrieved on 19 August 2024 from https://cloud.baidu.com/doc/WENXINWORKSHOP/s/6ltgkzya5.

Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational psychology review*, *19*, 15-29.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.

Kim, H., Han, B., Kim, J., Lubis, M. F. S., Kim, G. J., & Hwang, J. I. (2024, May). Engaged and Affective Virtual Agents: Their Impact on Social Presence, Trustworthiness, and Decision-Making in the Group Discussion. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-17).

Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, *27*(5), 6069-6104.

Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia-social and behavioral sciences*, *31*, 486-490.

Liu, Z., Zhu, Z., Zhu, L., Jiang, E., Hu, X., Peppler, K. A., & Ramani, K. (2024, May). ClassMeta: Designing Interactive Virtual Classmate to Promote VR Classroom Participation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-17).

Mao, M., Ting, P., Xiang, Y., Xu, M., Chen, J., & Lin, J. (2024). Multi-User Chat Assistant (MUCA): a Framework Using LLMs to Facilitate Group Conversations. *arXiv preprint arXiv:2401.04883*.

Nguyen, H. (2023). Role design considerations of conversational agents to facilitate discussion and systems thinking. *Computers & Education*, *192*, 104661.

Radmehr, B., Singla, A., & Käser, T. (2024). Towards Generalizable Agents in Text-Based Educational Environments: A Study of Integrating RL with LLMs. *arXiv preprint arXiv:2404.18978*.

Sun, D., Looi, C-K., Yang, Y., & Jia, F. (2024). Exploring Students' Learning Performance in Computer-Supported Collaborative Learning Environment During and After Pandemic: Cognition and Interaction. *British Journal of Educational Technology.* https://doi.org/10.1111/bjet.13492

Kumar, J. A. (2021). Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course. International journal of educational technology in higher education, 18(1), 65.

Xie, T., Liu, R., Chen, Y., & Liu, G. (2021). MOCA: A motivational online conversational agent for improving student engagement in collaborative learning. IEEE Transactions on Learning Technologies, 14(5), 653-664.

Mahafdah, R. F., Bouallegue, S., & Bouallegue, R. (2024, April). Applying Artificial Intelligence in the E-Learning Field. In International Conference on Advanced Information Networking and Applications (pp. 392-403). Cham: Springer Nature Switzerland.

Nguyen, H. D., Tran, T. V., Pham, X. T., Huynh, A. T., Pham, V. T., & Nguyen, D. (2022). Design an intelligent educational chatbot for information retrieval based on integrated knowledge bases. IAENG International Journal of Computer Science, 49(2), 531-541.

Thway, M., Recatala-Gomez, J., Lim, F. S., Hippalgaonkar, K., & Ng, L. W. (2024). Harnessing GenAI for Higher Education: A Study of a Retrieval Augmented Generation Chatbot's Impact on Human Learning. arXiv preprint arXiv:2406.07796.