Reddy, D. Majumdar, M. (Eds.) (2025). Proceedings of the 1st International Conference on Learning Evidence and Analytics. Asia-Pacific Society for Computers in Education

Student Behavioral Analysis in Problem-based Learning: Problem Identification

Deepti REDDY[a*], **Rwitajit MAJUMDAR**[b]

[a] *Mukesh Patel School of Technology Management & Engineering, SVKM's NMIMS, Mumbai, India'*

[b] *Research and Education Institute for Semiconductors and Informatics, Kumamoto University, Japan*

Abstract: This study explores the implementation of Problem-Based Learning (PBL) in a Distributed Computing course for third-year engineering students. Conducted in an online learning environment, the study introduced students to the fundamentals of PBL, including its steps and characteristics. Students were tasked with developing their problem statements, supported by reading materials and ChatGPT. The quality of these problem statements was used to classify students as low, medium, or high performers. Log data was analyzed to examine the interaction patterns of students across these performance levels.

Keywords: PBL, Problem identification, Quality of problem, Behavior patterns.

1. Introduction

Problem-based learning (PBL) is a pedagogical approach that enables students to learn while actively engaging with meaningful problems. Students are allowed to solve problems in a collaborative setting and form self-directed learning habits through practice and reflection. It can be viewed as an iterative process made up of, first, a problem analysis phase, second, a solution design phase, and third, an implementation phase. A tutor acts as a facilitator and scaffold for students' learning, particularly in the problem analysis and solution design phases to facilitate students' inquiry paths as they make sense of their ideas through discussion and sharing (Yew, 2016).

Learning analytics has been widely used to study student behaviors and their impact on learning outcomes in PBL environments (Saleh, 2022; Sáiz-Manzanares, 2021). However, little attention has been given to analyzing how students behave when selecting and defining their problem statements for PBL, particularly in the context of generative AI. We have conducted a study in which students interacted in an online micro-learning platform that shared resources to understand the characteristics of the problem. Students were then tasked with defining their problem statements. To support this process, ChatGPT was integrated into the activity, allowing students to seek feedback and ask questions. The expected outcome was for students to formulate an ill-structured, open-ended problem rooted in a real-life scenario.

The students were categorized as high, medium, or low performers based on the quality of the problem statements they defined. Their interactions within the LMS were logged and analyzed to identify the behavioral patterns exhibited by students in each category. This study addresses the following research question: What behavioral patterns do students exhibit when using ChatGPT to define their problem statements for PBL?

2. Related work

PBL has been widely adopted in engineering education to foster skills such as problem-solving, collaboration, and self-directed learning. Studies have emphasized that well-crafted problems are central to successful PBL, as they stimulate critical reasoning, promote teamwork, and allow application of knowledge (Yew, 2016).

A study (Saleh, 2022) examined how PBL can support the design and interpretation of learning analytics in a collaborative game-based learning environment. The principal component analysis (PCA) was used to describe the patterns in game interaction data, and clustered students based on the PCA. Results showed that the groups with higher

collaborative sense-making demonstrated better learning outcomes. Another study by Sáiz-Manzanares (2021) explored the application of educational data mining techniques to monitor student learning, with a primary focus on identifying patterns in student behavior using LMS log data. The study was conducted in the context of Problem-Based Learning (PBL) and self-regulated learning. Clustering algorithms were employed to group students based on similar interaction patterns. The results indicated that, in PBL settings, students engaged more actively with LMS resources during the initial and middle phases of the learning process. Interaction typically decreased in the final phase, as students shifted their focus to preparing project presentations.

The existing studies have established the important characteristics of the problem in PBL, which lead to effective learning outcomes and use of learning analytics to study student behaviors and their effect on learning outcomes. In this paper, we aim to study student behaviors while framing their problem statements while communicating with chatbots in the era of LLMs.

3. Intervention

The activity was created in LA-ReflecT platform (Majumdar et al. 2023) to help students understand PBL and the characteristics of the problem. The reading resources in terms of what is PBL, steps in PBL, characteristics of problems in PBL with examples were given in the LA-ReflecT. The students were asked to read and identify their problem statement and justify whether their problem satisfies the given characteristics of PBL. The students were encouraged to interact with Chat-GPT, which was integrated within the activity, to seek feedback on their problem, self-evaluate, and modify if necessary.

4. Method

70 participants (58% male and 42% female) were purposively sampled for this study who participated in problem-based learning in the Distributed Computing course. All students are pursuing a four-year Computer Science and Engineering degree in a private engineering college located in India, and were enrolled in the 15-week course Distributed Computing course in their third year.

The students were given access to the LA-ReflecT activity in the lab. They individually interacted with the system for approximately one hour. The interactions, such as clicks and student responses and interactions with ChatGPT, were logged in the system.

The problem statements entered by the students were scored by one of the authors on a scale of 1-3. The highest score (3) is given when the problem addresses a real-life scenario and is ill-structured. The example of the student problem statement that scored high is *“Booking tickets for popular concerts is a challenging task due to high demand during peak reservation periods. Many users face difficulties such as system hangs, slow processing, and even website crashes. The system becomes unresponsive, causing frustration among users, leading to an inefficient and unsatisfactory experience.*” The medium score (2) is given if the problem is not based on a real-life scenario or is not ill-structured. The lowest score (1) is given if the problem is not based on a real-world scenario and is well-structured. The student's problem statement that was scored low is “*Blockchain-Based Notary System for File Authentication*”.

Out of the total 70 students, 27 were grouped as high performers, 34 as medium performers, and 9 as low performers. The log data of each group was analyzed to find the interaction behaviors exhibited, as shown in Figure 1. The graphs show the frequency and percentage of actions performed by students in each step and in transitions.

5. Results and Discussion

High and medium performers accessed the reading material 70% and 71% of the time, respectively, compared to 66% for low performers, indicating that low performers engaged with the material less frequently. Overall interaction patterns showed that the medium performers exhibited non-linear patterns, frequently moving back and forth between *enter your problem statement*, *self-evaluate*, *justify*, and *interact with ChatGPT*. In contrast, high

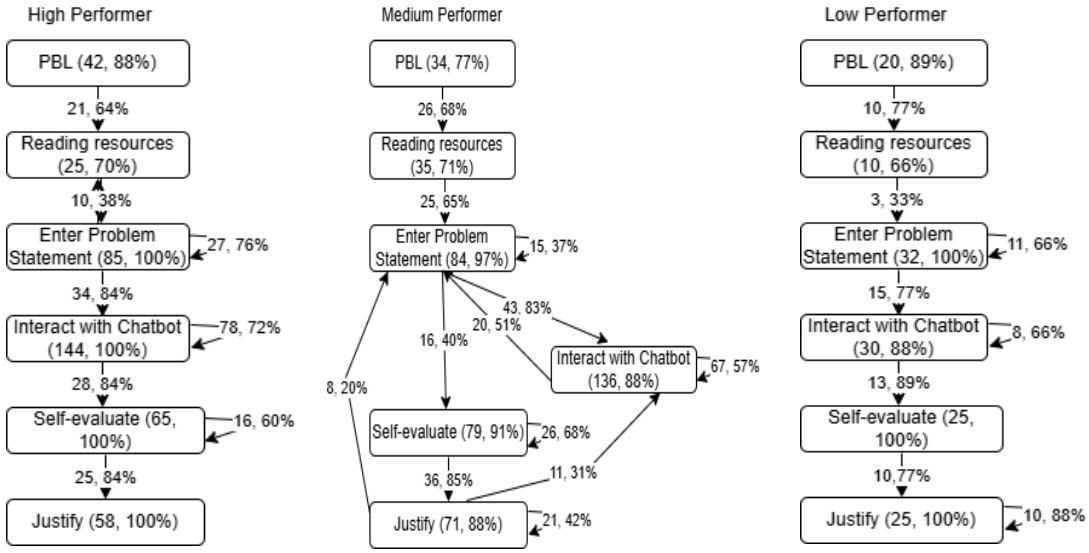

Figure 1. Interaction patterns of low, medium, and high performers.

performers followed a more structured approach, revisiting *problem entry*, *ChatGPT interaction*, and *self-evaluation*. Percentage of ChatGPT interactions were highest among high performers (100%) and lower for the other two groups (88%). Low performers rarely revisited the *self-evaluation* step, unlike medium (68%) and high performers (60%).

These patterns suggest that low and medium performers exhibited limited help-seeking behavior, as reflected in their minimal use of ChatGPT. Medium performers also showed signs of confusion, frequently revisiting previous steps and making multiple attempts even after completing the justification step. For instance, after the *Justify* step, 20% returned to *Enter Problem Statement* and 31% returned to *Interact with Chatbot*. In contrast, high performers demonstrated more strategic help-seeking, engaging with ChatGPT and self-evaluation before finalizing their justification. Once justified, they rarely revised their responses, indicating greater confidence in their solutions. Overall, low performers lacked metacognitive behaviors, neither revisiting self-evaluation nor revising their problem statements. High performers engaged in active revision, improving their responses based on ChatGPT feedback and self-assessment. Medium performers demonstrated trial-and-error behaviors with limited ChatGPT interaction.

6. Conclusion

This paper examines student behavior during problem identification in PBL, focusing on their interactions with core PBL elements, such as access to reading material, and their use of chatbot dialogues (ChatGPT) to refine understanding. The analysis revealed that low performers tended to avoid help-seeking and were primarily task-focused, medium performers showed signs of confusion and indecision, while high performers demonstrated strategic help-seeking and confidence. In future work, we aim to develop predictive models to identify low and medium performers based on their online interactions and offer timely, targeted interventions.

References

Majumdar, R., Prasad, P., Kadam, K., Gatare, K., Warriem, J.M. (2023). LA-ReflecT: A Platform Facilitating Micro-learning and Its Multimodal Learning Analytics. In *Lecture Notes in Computer Science*, vol 14200. Springer, Cham. https://doi.org/10.1007/978-3-031-42682-7_69

Sáiz-Manzanares, M. C., Rodríguez-Díez, J. J., Díez-Pastor, J. F., Rodríguez-Arribas, S., Marticorena-Sánchez, R., & Ji, Y. P. (2021). Monitoring of student learning in learning management systems: An application of educational data mining techniques. *Applied Sciences*, *11*(6), 2677.

Saleh, A., Phillips, T. M., Hmelo-Silver, C. E., Glazewski, K. D., Mott, B. W., & Lester, J. C. (2022). A learning analytics approach towards understanding collaborative inquiry in a problem-based learning environment. *British Journal of Educational Technology*, *53*(5), 1321-1342.

Yew, E. H., & Goh, K. (2016). Problem-based learning: An overview of its process and impact on learning. Health professions education, 2(2), 75-79.