From Students Struggles to Teachers Insights: A Dashboard of Student-AI Interactions in ML Practical Assignments

Boyun Zhang
boyunzhang@tudelft.nl
Delft University of Technology
The Netherlands

Ilinca Rențea
I.E.I.Rentea@tudelft.nl
Delft University of Technology
The Netherlands

Gosia Migut
m.a.migut@tudelft.nl
Delft University of Technology
The Netherlands

Abstract

When students work on self-paced, unsupervised practical assignments in Machine Learning courses, teachers have little visibility into the difficulties students encounter. In our Machine Learning course, students completed a practical assignment using JELAI, an AI-supported programming environment. The recorded student-AI interactions were classified into pedagogically meaningful categories and visualized in a teacher-facing dashboard. An initial evaluation showed that teachers found the dashboard useful, while also identifying areas for improvement. We describe our experience and discuss the potential and challenges of using student-AI interactions to improve teaching practice.

CCS Concepts

• **Social and professional topics** → **Computing education**; • **Applied computing** → **Education**.

Keywords

teacher-facing dashboard, students-AI interactions, interactions classification, Machine Learning education

ACM Reference Format:
Boyun Zhang, Ilinca Rențea, and Gosia Migut. 2026. From Students Struggles to Teachers Insights: A Dashboard of Student-AI Interactions in ML Practical Assignments. In *Proceedings of the 31st ACM Conference on Innovation and Technology in Computer Science Education V.2 (ITiCSE 2026), July 10–15, 2026, Madrid, Spain.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3803401.3811981

1 Introduction

The growing adoption of AI-supported programming environments in computing education has created new opportunities for both students and teachers. Tools that provide personalized on-demand feedback allow students to work through practical assignments at their own pace, reducing the dependence on instructor availability. However, this independence comes at a cost: teachers lose visibility into the difficulties students encounter, making it harder to adapt instruction to student needs [5].

Learning analytics research has long argued that capturing and surfacing student behavior data can meaningfully support teaching [2]. Teacher-facing dashboards, in particular, have been proposed as a practical mechanism for translating raw interaction data into actionable insights, helping instructors identify struggling students, common misconceptions, and gaps in understanding [6]. Yet most existing dashboards focus on behavioral metrics such as time-on-task or submission counts, rather than the semantic content of student questions and struggles.

The rise of AI tutoring and programming assistants introduces a new and largely untapped data source: the questions students ask AI while working [4, 5]. Analyzing these interactions offers a window into student thinking that traditional log data cannot provide. Early work on student-AI interaction analysis has shown that such data can reveal patterns in conceptual confusion and help-seeking behavior [1], although its application to teacher-facing tools remains underexplored.

In this paper, we build on these strands of research by collecting and categorizing student-AI interactions from a Machine Learning course and presenting them through a teacher-facing dashboard. We describe our design process, the proposed pedagogical classification scheme, and the teachers' initial responses to the resulting tool.

2 System Design

To visualize student-AI interactions in a teacher-facing dashboard, we follow the pipeline shown in Figure 1. Students complete the ML assignment in JELAI [7], a web-based environment that integrates a Jupyter notebook interface with an AI chatbot to provide on-demand Python programming assistance. When a student submits a question, the chatbot returns a response while the question is automatically classified into predefined pedagogical categories. The labeled questions and responses are stored in a SQLite database and retrieved for visualization in the teacher-facing dashboard.

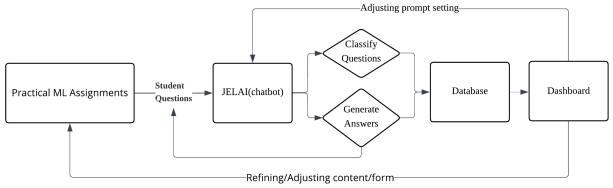

Figure 1: Overview of the proposed pipeline, including the ML assignment, student questions submitted through JELAI chatbot, question classification and visualization of the results in a dashboard.

2.1 Data Collection

We collected data during a pilot study in a second-year undergraduate Machine Learning course at Delft University of Technology,

ITiCSE 2026, Madrid, Spain

ACM ISBN 979-8-4007-2633-0/2026/07
https://doi.org/10.1145/3803401.3811981

The Netherlands. The course is mandatory for Computer Science and Engineering students and focuses on understanding and implementing core machine learning algorithms. For this study, students were invited to complete an optional Neural Networks assignment using JELAI [7] as their programming environment. While working on the assignment, they could ask the chatbot questions, and all interactions were logged for analysis with students' informed consent. A total of 13 students participated, generating 99 student-AI interactions covering assignment clarification, code debugging, and explanations of machine learning concepts.

2.2 Data Processing

We aim to categorize student-AI interactions into five pedagogically motivated categories [8]: *Code Implementation*, *Concept*, *Error Debugging*, *Contextual*, and *Other*. To that end, we fine-tuned a RoBERTa-base model on StudyChat [3], a publicly available dataset of student dialogues following the same labeling scheme. We then manually labeled our collected interactions and used them as a held-out test set. The classifier achieved 68% accuracy, with most errors involving the *Contextual* category, often confused with *Concept* and *Error Debugging*. This likely reflects the ambiguity of student questions, which frequently combine multiple concerns, suggesting that multi-label classification may better capture their complexity.

2.3 Data Presentation

The dashboard prototype was developed using the *Plotly Dash* and embedded into the JELAI as a JupyterHub service. It primarily relies on visualizations to present analysis results. The dashboard contains three main tabs. The *Course Overview* tab, shown in Figure 2, provides aggregated insights across assignments, including category distributions and question counts. The *Assignment Overview* tab displays a selected assignment's question distribution, rule-based comparison with the course average, student activity timeline, keyword cloud, question list. The *Student Overview* tab shows individual statistics, including total questions, category distribution, comparison with class average, and timeline of student questions. A fourth tab allows teachers to adjust chatbot prompt settings, such as system instructions, learning objectives, and assignment descriptions. These tabs address 18 predefined design questions identified with educators at the course, assignment, and student levels. The design prioritizes at-a-glance readability and enables teachers to identify assignment-level difficulty patterns and students who may need additional support.

3 Expert Review

To examine whether the dashboard prototype effectively helps teachers understand the learning difficulties students encounter in machine learning assignments, we conducted an expert review with 7 participants.

The results of the review indicate that participants generally perceived the interaction-based analytics and dashboard as useful to support their understanding of students' learning difficulties in machine learning assignments. They also highlighted several limitations and areas in which the dashboard could be improved. They expressed a desire for in-depth analysis, such as question clustering, pattern of student questioning, and relating interactions to assignment performance.

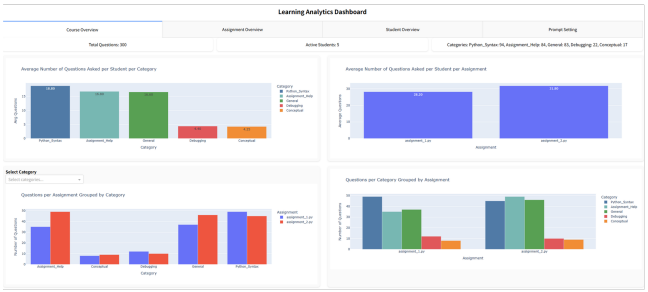

Figure 2: Example screenshot of the *Course overview* tab in the teacher-facing dashboard.

4 Future Work

A key direction for future work is to collect larger-scale interaction data over a full course. This would enable a more robust analysis of interaction patterns, better training and validation of classification models, and exploration of whether interactions can inform timely instructional interventions. Multi-label classification should also be explored to better capture the complexity of student questions. Embedding the dashboard in a real course over a longer period would allow teachers to integrate analytics into their daily teaching practice. Combining interaction data with teacher feedback could reveal how the dashboard influences instructional decisions, such as adapting lectures or lab support, moving beyond perceived usefulness toward measurable impact. Replicating the study across multiple machine learning courses would further support the generalization of the findings.

References

[1] Paul Denny, Viraj Kumar, and Nasser Giacaman. 2023. Conversing with Copilot: Exploring Prompt Engineering for Solving CS1 Problems Using Natural Language. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1* (Toronto ON, Canada). 1136–1142.

[2] Xu Du, Juan Yang, Brett E Shelton, Jui-Long Hung, and Mingyan Zhang. 2021. A systematic meta-review and analysis of learning analytics research. *Behaviour & information technology* 40, 1 (2021), 49–62.

[3] Fareya Ikram Hunter McNichols and Andrew Lan. 2025. The StudyChat Dataset: Student Dialogues With ChatGPT in an Artificial Intelligence Course. arXiv:2503.07928 [cs.AI] https://arxiv.org/abs/2503.07928

[4] Minsun Kim, SeonGyeom Kim, Suyoun Lee, Yoosang Yoon, Junho Myung, Haneul Yoo, Hyungseung Lim, Jieun Han, Yoonsu Kim, So-Yeon Ahn, Juho Kim, Alice Oh, Hwajung Hong, and Tak Yeon Lee. 2024. Designing Prompt Analytics Dashboards to Analyze Student-ChatGPT Interactions in EFL Writing. *CoRR* (2024). arXiv:2405.19691

[5] Sam Lau, Kianoosh Boroojeni, Harry Keeling, and Jenn Marroquin. 2025. Barriers that Programming Instructors Face While Performing Emergency Pedagogical Design to Shape Student-AI Interactions with Generative AI Tools. arXiv:2510.09492

[6] Inge Molenaar and Carolien AN Knoop-van Campen. 2018. How teachers make dashboard information actionable. *IEEE Transactions on Learning Technologies* 12, 3 (2018), 347–355.

[7] Manuel Valle Torre, Thom van der Velden, Marcus Specht, and Catharine Oertel. 2025. JELAI: Integrating AI andnbsp;Learning Analytics innbsp;Jupyter Notebooks. In *Artificial Intelligence in Education: 26th International Conference, AIED 2025, Palermo, Italy, July 22–26, 2025, Proceedings, Part VI*. 68–75.

[8] Sarah Heckman Zhikai Gao, Collin Lynch and Tiffany Barnes. 2021. "Automatically classifying student help requests: a multi-year analysis". In *Proceedings of The 14th International Conference on Educational Data Mining*. 81–92.