Towards Modeling Affecto-Cognitive-Metacognitive Interactions in Machine Learning Education

Aditya Rajmane*
Centre for Educational Technology
Indian Institute of Technology Bombay
Mumbai, India
aditya.rajmane@iitb.ac.in

Ramkumar Rajendran
Centre for Educational Technology
Indian Institute of Technology Bombay
Mumbai, India
ramkumar.rajendran@iitb.ac.in

Kshitij Sharma
Department of Computer Science
Norwegian University of Science and Technology
Trondheim, Norway
kshitij.sharma@ntnu.no

ABSTRACT

This doctoral research investigates the interplay between cognition, affect and metacognition (CAM) in students' engagement with machine learning (ML) tasks in computer-based learning environments. While prior work has examined these CAM processes, limited research has explored how they dynamically interact during complex, data-driven problem-solving. To address this gap, this work proposes the design and development of a tutoring system (ML-Tutor) to teach Machine Learning. This system captures learners' interactions and identifies the affective and metacognitive processes while learning. The research conceptualizes metacognition as embedded within the learning activity and operationalizes it through planning, monitoring, and control processes. A preliminary paper-and-pencil study has informed the design of tasks, prompts, and data collection strategies, demonstrating the feasibility of eliciting multimodal traces of learner activity. This work is ongoing, with current efforts focused on developing the CBLE, refining the analytical framework, and identifying suitable units of analysis for studying temporally unfolding learning processes. Feedback is sought on the conceptual framing, operationalization of metacognition, and methodological approach for analyzing multimodal data.

Keywords

Affective Computing, Cognitive Processes, Metacognition, Multimodal Learning Analytics, Intelligent Tutoring Systems, Computer-based Learning Environment, Personalization

1. INTRODUCTION

Student engagement is widely recognized as a core component of learning outcomes, particularly in complex, open-ended domains that require sustained effort and self regulation. Some perspectives conceptualize engagement as a multidimensional construct encompassing behavioral, cognitive, and affective components, each contributing uniquely to learning processes ([12, 5, 13]). In recent years, this view has been extended to emphasize the dynamic and situated nature of engagement, where learners' involvement fluctuates over time in response to task demands, contextual factors, and internal states ([20]). This shift toward process-oriented understandings of engagement has become especially relevant in computer-based learning environments, where data enables the study of engagement as it unfolds in real time [2].

Affective Engagement

Affective engagement [19] has emerged as a critical dimension of learning, particularly in computer-based learning environments where learners interact with complex, open-ended tasks. Affective states are not merely byproducts of learning but play a constitutive role in influencing attention, persistence, and strategy use [7]. As such, understanding affective engagement is essential for designing learning environments that can support productive learning trajectories. With the integration of machine learning techniques, [17] detecting affect is possible in more authentic environments. Within the context of educational data mining, affect detection has relied on multimodal learning analytics, enabling richer representations of learners' experiences across cognitive and behavioral dimensions [16]. Current affective computing systems largely treat affect as an observable signal to be detected and responded to, with limited consideration of its interaction with cognitive processes during problem solving [17].

Affect, Cognition and Metacognition

The relationship between affect and cognition is central to understanding self-regulated learning [6]. Efklides' Metacognitive and Affective Model of Self-Regulated Learning [8] (MASRL) conceptualizes this interaction by linking cognitive processing, metacognitive monitoring, and affective experiences such as effort, difficulty, and confidence. Within this framework, affect actively shapes learners' judgments, strategy use, and regulation during task engagement. Recent work has further emphasized the importance of capturing these processes through fine-grained, multimodal data,

*corresponding author

Aditya Rajmane, Ramkumar Rajendran, and Kshitij Sharma. Towards Modeling Affecto-Cognitive-Metacognitive Interactions in Machine Learning Education. In Anthony Botelho, Maria Mercedes T. Rodrigo, Adish Singla, Hiroaki Ogata, Hyojeong So, and Young Hoan Cho (eds.) Proceedings of the 19th International Conference on Educational Data Mining, Seoul, Republic of Korea, June, 2026, pp. 892–895. International Educational Data Mining Society (2026).

https://doi.org/10.5281/zenodo.21039859

demonstrating that understanding self-regulated learning requires integrating trace data, self-reports, observable interaction data [11, 9, 10]. There also is a dynamic and temporal nature of affect and metacognitive processes during learning. Studies have shown that learners' emotional states evolve over time and are closely tied to regulation and task performance [15]. However, there is limited research that integrates cognition, metacognition, and affect to provide a holistic understanding of learning processes. One key challenge lies in the multimodal nature of data collection. Affective states are often captured using modalities such as facial expressions, while cognitive and metacognitive processes are typically inferred from interaction data such as clickstreams. Integrating these heterogeneous data sources remains complex.

> Addressing the challenge of integrating the multimodalities can enable a more comprehensive modeling of learner behavior. Such integrated models have the potential to better capture the dynamics of learning and support the design of more effective, personalized scaffolding strategies in computer-based learning environments.

2. OUR WORK

Our work is situated in the context of machine learning (ML) education for school students. The focus is on developing a foundational understanding of core ML concepts such as classification, features, and similarity-based reasoning through accessible, real-world analogies and interactive activities. The learning tasks are designed to progressively move from intuitive understanding to structured reasoning. For instance, students engage with classification problems using everyday categories (e.g., fruit, objects) before transitioning to simplified representations of algorithms such as K-Nearest Neighbors (KNN). These tasks are intentionally designed to surface cognitive challenges and uncertainty, creating opportunities to observe how learners respond to difficulty in emerging domains such as ML.

The following research goals (RG's) were formulated:

1. To investigate how pedagogical approaches for teaching classification, feature identification, and similarity-based reasoning can be translated into a CBLE.
2. To design and develop a CBLE that embeds opportunities for eliciting and capturing learners' cognitive, metacognitive, and affective processes during task engagement.

Paper-Pencil Based Intervention

A preliminary paper-and-pencil study was conducted to study the feasibility of creating a CBLE aligning with RG 2. The study involved school students in the age group of 11 - 14, engaging with introductory machine learning (ML) concepts through structured activities focused on classification and similarity-based reasoning.

Pedagogical Design and Methodology

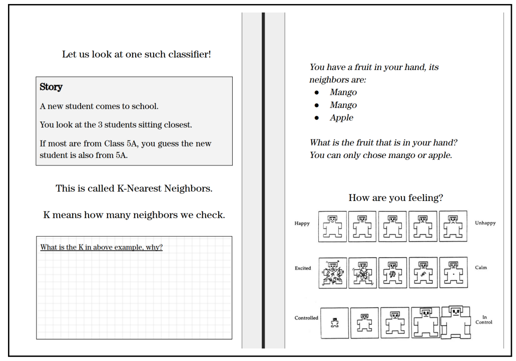

Figure 1: Excerpts from Paper-and-Pencil Intervention

The instructional content was adapted from the CBSE Artificial Intelligence curriculum for grades 6-8, ensuring alignment with age-appropriate learning objectives and prior exposure. Interactive tools such as Teachable Machine [4] were incorporated to enable students to train and test simple models, making abstract processes visible and manipulable. Students then engaged in structured activities like solving classification problems using tabular patient data, requiring them to identify relevant features, compare instances, and make decisions using nearest-neighbor logic. These activities were followed with reflective prompts (e.g., "How are you feeling?") and self-report checklists in the end to encourage monitoring of effort, difficulty, and persistence.

The affective self-report items were adapted from established instruments, the Self-Assessment Manikin (SAM Manikin) and Semantic Differential Scales [3], which have been widely used to capture affective responses in learners of comparable age groups (11 to 14 years). The study was conducted with eight ($N = 8$) participants aged 11 to 14. This design intentionally integrates cognitive engagement with opportunities for metacognitive reflection and affective expression, while also serving as a preliminary exploration of how such task structures and measures may be adapted for implementation within a computer-based learning environment.

Data Collection Strategy

Prior to data collection, formal ethics approval was obtained from the Institute's Ethics and Review Board. Informed consent was obtained from guardians, and assent was obtained from participating students. A multi-source data collection approach was adopted to capture learners' cognitive and affective experiences during task engagement. Data sources included students' responses to structured worksheets (Ref. 1) and embedded self-report measures capturing affective states such as effort, difficulty, and persistence. The instructional material was organized into discrete activities, with affective self-reports using the Self-Assessment Manikin (SAM) [3] administered at the end of each activity to capture affect at multiple points during the learning process 3 In addition to self-reports, video recordings of participants were collected to enable analysis of behavioral and physiological signals. These recordings allow for the extraction of heart rate using remote photoplethysmography (rPPG), a non-contact method for estimating physiological signals from facial videos [21, 18, 1]. To support temporal align-

Figure 2: Student Participant

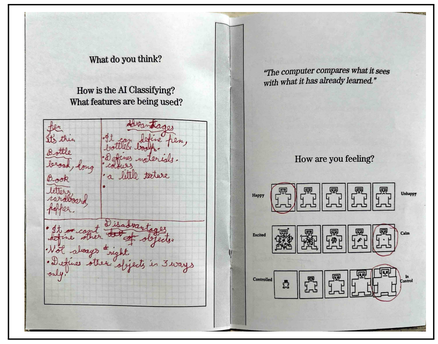

Figure 3: Student Work

ment, the teacher maintained timestamps corresponding to each activity, enabling synchronization of video, self-report, and task data. This multimodal and temporally structured design provides a foundation for examining how affective states, cognitive processes, and physiological indicators co-occur during different phases of task engagement.

Preliminary Observations

While no formal analysis has been conducted, the collected multimodal data provides an initial basis for exploring relationships between cognitive processes, affective experiences, and physiological indicators such as heart rate (rPPG), informing the design of a subsequent computer-based learning environment.

3. CBLE AND EMBEDDED METACOGNITION

The next phase of our research involves the design of an adaptive computer-based learning environment (CBLE) that embeds metacognitive processes within task interaction. Rather than treating metacognition as an external construct measured through post-hoc reports or think-aloud protocols, this work conceptualizes metacognition as an activity that is enacted during learning.

In this work, metacognition is operationalized through three core processes: planning, monitoring, and control [14]. The CBLE will incorporate explicit support for planning by prompting learners to articulate their approach before attempting a task. For example, *learners may be asked to identify which features are relevant for a classification problem or to select a strategy for solving it.* These prompts are designed to provide observable indicators of how learners prepare to engage with machine learning tasks. Such interactions allow planning to be captured as part of the learning process rather than inferred retrospectively. Monitoring processes will be embedded through periodic prompts and checkpoints during task engagement. *These include self-reports of perceived difficulty, effort, and confidence, as well as affective indicators captured through instruments such as the Self-Assessment Manikin. Artifacts for monitoring such as an interactive clock and progress bar will also be available.* By integrating these prompts within task flow, the system captures learners' moment-to-moment awareness of their cognitive and affective states. Control processes will be inferred from learners' interactions with the system, including actions such as *revisiting problems, modifying responses, switching strategies, and persisting after incorrect attempts.* These behaviors provide observable evidence of how learners regulate their learning in response to perceived difficulty or feedback. By capturing these actions alongside planning and monitoring data, the system enables the examination of how metacognitive regulation unfolds during task engagement.

These processes will be integrated within the CBLE consisting of a domain module, a pedagogical module, and a learner module. The domain module will structure machine learning tasks such as classification and, in future iterations, clustering. The pedagogical module will orchestrate prompts and scaffolds that elicit planning, monitoring, and control. Content will be in the form of text, videos, and interactive animations. The assessment will include MCQ's, MSQ's and descriptive questions. The learner module will capture multimodal data, including task performance, self-reports, and physiological signals, enabling the construction of temporally aligned traces of learner activity. Together, this design provides a foundation for studying how metacognitive processes are enacted and how they relate to cognitive and affective dynamics during learning.

4. FEEDBACK SOUGHT

a. Theoretical–Operational Alignment: The adequacy of the proposed conceptualization of affect, cognition, and metacognition, particularly the alignment between the MASRL inspired cognitive–affective loops and their operationalization through multimodal indicators. Whether the mapping from latent constructs (e.g., perceived difficulty, confidence, regulatory control) to observable features is sufficiently well-specified and theoretically defensible.

b. Unit of Analysis and Temporal Segmentation: The validity and analytical utility of defining these affecto-metacognitive episodes as the primary unit of analysis in computer-based learning environments, including the proposed criteria for segmenting continuous interaction data into episodes and the implications of segmentation choices for modeling temporal dynamics and cognitive–affective coupling.

5. CONCLUSION

This work outlines a research trajectory for studying the dynamic interplay between affect, cognition, and metacognition in ML learning environments through the design of an adaptive tutoring system. The preliminary study informs this direction, with future work focusing on refining the analytical framework and leveraging multimodal data from CBLEs to examine learning processes at a finer temporal level.

6. ACKNOWLEDGMENTS

The work is part of Enhancing AI Education between India and Norway (INNOAIED) project (INCP2-2024/10367). The authors would like to thank the support provided by members of the Learning Analytics Lab at IIT Bombay.

7. REFERENCES

[1] Y. Benezeth, D. Krishnamoorthy, D. J. Botina Monsalve, K. Nakamura, R. Gomez, and J. Mitéran. Video-based heart rate estimation from challenging scenarios using synthetic video generation. *Biomedical Signal Processing and Control*, 96:106598, Oct. 2024.

[2] M. Bond, O. Viberg, and N. Bergdahl. The current state of using learning analytics to measure and support K-12 student engagement: A scoping review. In *LAK23: 13th International Learning Analytics and Knowledge Conference*, LAK2023, pages 240–249, New York, NY, USA, Mar. 2023. Association for Computing Machinery.

[3] M. M. Bradley and P. J. Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, Mar. 1994.

[4] M. Carney, B. Webster, I. Alvarado, K. Phillips, N. Howell, J. Griffith, J. Jongejan, A. Pitaru, and A. Chen. Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8, Honolulu HI USA, Apr. 2020. ACM.

[5] S. L. Christenson, A. L. Reschly, and C. Wylie, editors. *Handbook of Research on Student Engagement*. Springer US, Boston, MA, 2012.

[6] S. de Mooij, J. Lämsä, L. Lim, O. Aksela, S. Athavale, I. Bistolfi, F. Jin, T. Li, R. Azevedo, M. Bannert, D. Gašević, S. Järvelä, and I. Molenaar. A Systematic Review of Self-Regulated Learning through Integration of Multimodal Data and Artificial Intelligence. *Educational Psychology Review*, 37(2):54, June 2025.

[7] S. K. D'mello and J. Kory. A Review and Meta-Analysis of Multimodal Affect Detection Systems. *ACM Computing Surveys*, 47(3):1–36, Apr. 2015.

[8] A. Efklides. Affect, Epistemic Emotions, Metacognition, and Self-Regulated Learning. *Teachers College Record: The Voice of Scholarship in Education*, 119(13):1–22, Dec. 2017.

[9] Y. Fan, L. Lim, J. van der Graaf, J. Kilgour, M. Raković, J. Moore, I. Molenaar, M. Bannert, and D. Gašević. Improving the measurement of self-regulated learning using multi-channel data. *Metacognition and Learning*, 17(3):1025–1055, Dec. 2022.

[10] Y. Fan, M. Rakovic, J. van der Graaf, L. Lim, S. Singh, J. Moore, I. Molenaar, M. Bannert, and D. Gašević. Towards a fuller picture: Triangulation and integration of the measurement of self-regulated learning based on trace and think aloud data. *Journal of Computer Assisted Learning*, 39(4):1303–1324, 2023.

[11] Y. Fan, J. van der Graaf, L. Lim, M. Raković, S. Singh, J. Kilgour, J. Moore, I. Molenaar, M. Bannert, and D. Gašević. Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, 17(3):949–987, Dec. 2022.

[12] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris. School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, 74(1):59–109, 2004.

[13] S. Heikkinen, M. Saqr, J. Malmberg, and M. Tedre. A longitudinal study of interplay between student engagement and self-regulation. *International Journal of Educational Technology in Higher Education*, 22(1):21, Apr. 2025.

[14] J. E. Jacobs and S. G. Paris. Children's Metacognition About Reading: issues in Definition, Measurement, and Instruction. *Educational Psychologist*, 22(3-4):255–278, June 1987. _eprint: https://doi.org/10.1080/00461520.1987.9653052.

[15] S. P. Lajoie, J. Zheng, S. Li, A. Jarrell, and M. Gube. Examining the interplay of affect and self regulation in the context of clinical reasoning. *Learning and Instruction*, 72:101219, Apr. 2021.

[16] K. Mangaroska, K. Sharma, D. Gasevic, and M. Giannakos. Multimodal Learning Analytics to Inform Learning Design: Lessons Learned from Computing Education. *Journal of Learning Analytics*, 7(3):79–97, Dec. 2020.

[17] L. Mathur, M. Mataric, and L.-P. Morency. Expanding the Role of Affective Phenomena in Multimodal Interaction Research. In *Proceedings of the 25th International Conference on Multimodal Interaction*, ICMI '23, pages 253–260, New York, NY, USA, Oct. 2023. Association for Computing Machinery.

[18] D. McDuff, S. Gontarek, and R. Picard. Remote measurement of cognitive stress via heart rate variability. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, 2014:2957–2960, 2014.

[19] R. W. Picard. *Affective computing*. MIT Press, Cambridge, Mass., 1. paperback ed edition, 2000.

[20] G. M. Sinatra, B. C. Heddy, and D. Lombardi. The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist*, 50(1):1–13, Jan. 2015. _eprint: https://doi.org/10.1080/00461520.2014.1002924.

[21] W. Verkruysse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 16(26):21434–21445, Dec. 2008.