1. Action Selection (for list and descriptions of all features and citations, please refer to this)

a. Motivational Model Features (total 14): The final motivation model integrated individual mastery approach variables and self-efficacy, as motivation encompasses both. Mastery approach signifies sustained engagement and goal-orientation, while self-efficacy reflects perceived confidence in problem-solving. Some key features, in addition to post MAP and post SE scores:

Features	Justification
Total # of sessions per problem (Total # of sessions, normalized by # of problems attempted)	All of these measures should likely reflect student's tendency to persist and sustain efforts to learn, which indicates student's tendency for mastery-oriented (Palmer, 2019; Wolters, 2003)
Session persistence (# of problems completed each session)	
Attempt-to-hint ratio (# of attempts on average before requesting hint)	Help seeking behaviors are related to students' self-efficacy. Students with low efficacy will likely rely on more hints, while those with high efficacy would not. Efficacy beliefs can also change depending on how they perform (accuracy) so the "hint-usage pattern" reflects this. (e.g., Bernacki et al., 2015; Miles & Vela, 2022).
Hint-usage pattern (# of hints that was followed by OK)	

b. Metacognitive Model Features (total 13): The final metacognitive model incorporates features based on Zimmerman's three phases of self-regulated learning: planning, monitoring, and evaluation, as each phase incorporates metacognitive learning processes of learning (Zimmerman, 2001). Some key features, in addition to Post Meta scores:

Features	Justification
Concept building time	Reflects whether students spent time reviewing prior content before attempting problems
Average hints per problem attempted(Mean # of hints requested for each problem)	These features reflect help-seeking behavior (HSB). HSB is also associated with metacognitive skills as it requires students to understand and be aware of what they know and do not (monitoring and evaluating) (Aleven et al., 2006; Sumadyo et al., 2021).
Optimal help-seeking rate (average of time spent on problem + accuracy % + average hint level)	

c. Math Learning Model (total 23): The final math learning model has features that directly reflect student math proficiency, including accuracy, skill development, and time spent on tasks. Some features overlap with motivation and metacognition features, aligning with existing literature that metacognition and motivation both predict math proficiency. Some key features:

Features	Justification
Pre-math baseline score	Initial math proficiency will affect how students perform later

Average workspace time	Study time both actual and perceived are linked to better study outcomes (e.g., Spitzer, 2022; Theobald, 2025)
Overall error rate	Error rates clearly reflects how well the student is performing

2. Model Performance

Model	Metrics	Interpretation
Metacognition	QWK = 0.31	A QWK of 0.31 shows modest agreement, implying the model grasps some metacognitive trends but might miss subtle reflection patterns. Adding text-based features like student self-reports of cognitive processes could boost performance.
Math	$R^2 = 0.88$ RMSE = 0.088 MAE = 0.046	The model explains 88% of post-test math score variance ($R^2 = 0.88$). Unmodeled factors (e.g., prior knowledge, in-class instruction variations) or temporal features (e.g., learning trajectories over time) may account for the remaining error.
Motivation	QWK = 0.42	A QWK of 0.42 suggests modest agreement, yet challenges persist in modeling affective outcomes. Integrating clickstream features from motivational scaffold prompts could improve performance.

a. Reasons why we might be finding such results:
- Small sample size or skewed score distributions may limit accuracy.
- Behavioral features are indirect proxies for internal states, making them harder to model.
- Finer-grained temporal or textual data (e.g., free responses) could improve performance.
- Including pre-test scores may enhance predictions for metacognition and motivation.

3. Impact for Educators and Ed-Tech

For educators: Our results suggest that 1) online learning platforms can offer valuable data to forecast math learning outcomes, serving as an effective method for educators to track student advancement. Therefore, educators should monitor student performance in real time via the platform to detect early signs of struggle or disengagement, enabling timely intervention. Also, 2) pre-test scores are good indicators for later performance, so use them effectively for guidance. Lastly, 3) to better understand motivation and metacognition, which are harder to observe with platform performance data. We recommend combining these behavioral patterns with qualitative measures like student surveys, goal-setting prompts, or self-reflections.

For developers (EdTech): We suggest 1) embedding lightweight formative assessments and reflection prompts directly in the learning platform so that it will reveal metacognitive skills and motivation for both teacher and students. This can be embedded into MATHia’s dashboard. For instance, if a student’s metacognition score is predicted to be low, the system could auto-prompt a reflection question, and if motivation dips, teachers receive an alert and can assign a quick, confidence-boosting mini-challenge. Over time, this closed-loop integration can personalize hint timing, scaffold difficulty, and coach self-regulatory behaviors—ultimately improving both learning outcomes and the learner experience.