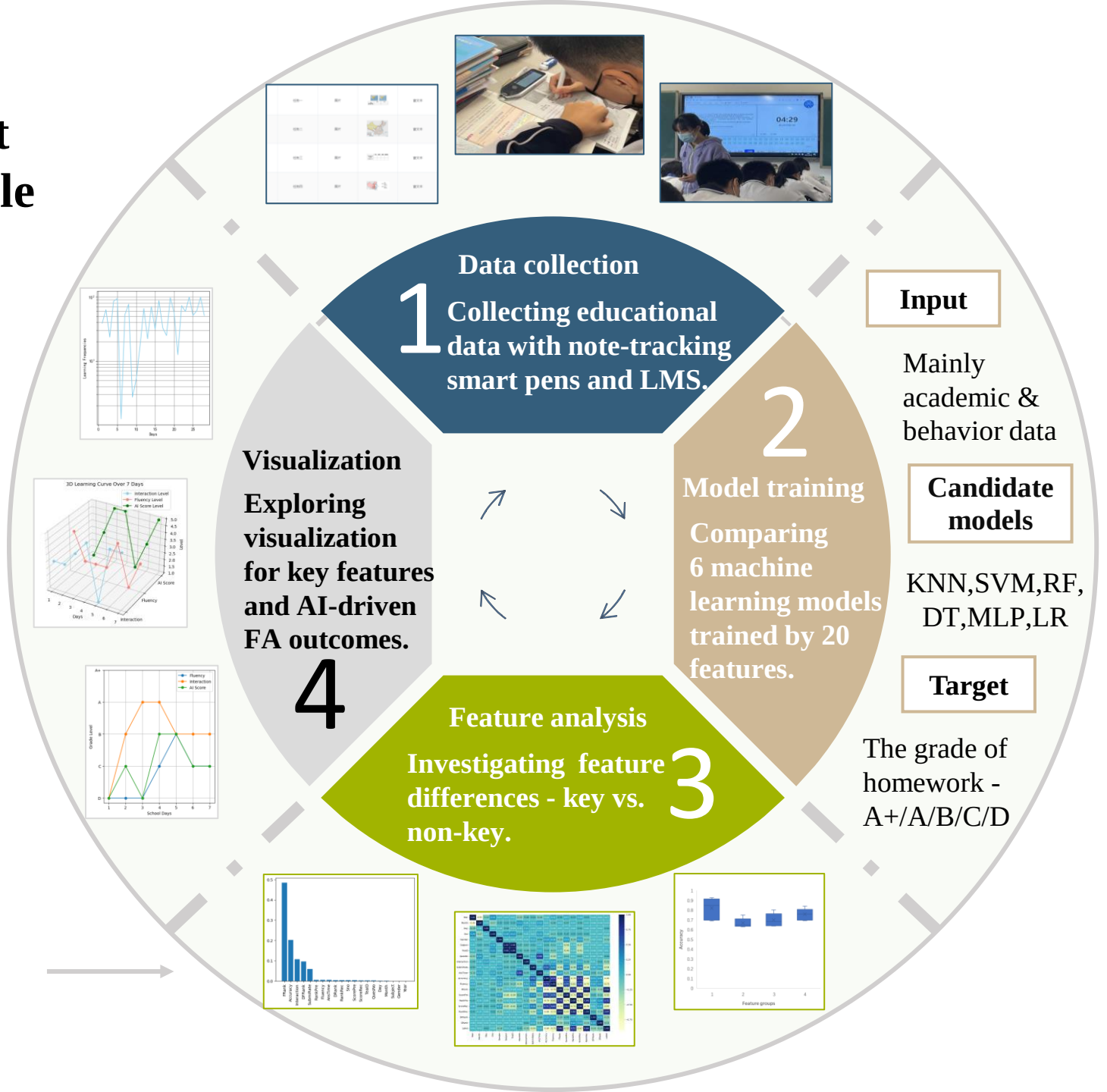
Investigating AI-driven formative assessment with note-tracking smart pens - An exploratory study in middle school

RQ1: How accurate is an AI-driven FA model for middle school students integrating PT-related features, past academic performance, learner and lesson information, and LMS-recorded data?

RQ2: What are the pivotal features in model training, and do these key features exhibit statistically significant differences among other features?

RQ3: How can the outcomes of AI-driven FA be effectively depicted in a timeline chart that adheres to the principles of PT theory?



Methodological framework of AI-driven formative assessment

Instruments for data collection

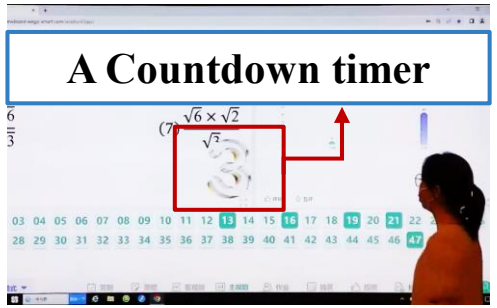
The screen on the smart pen shows in -class tasks.

Details of smart pens

The LMS monitor

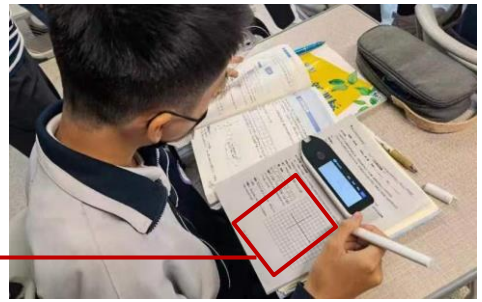
Function1: For students to receive, view, and submit in-class tasks

Students participate by touching the engagement button on smart pens.



The LMS monitor

Function2: For students to engage in "snatch-and-answer" activities



The function description of note-tracking smart pens

Function3: For synchronized tracking of students' handwritten notes

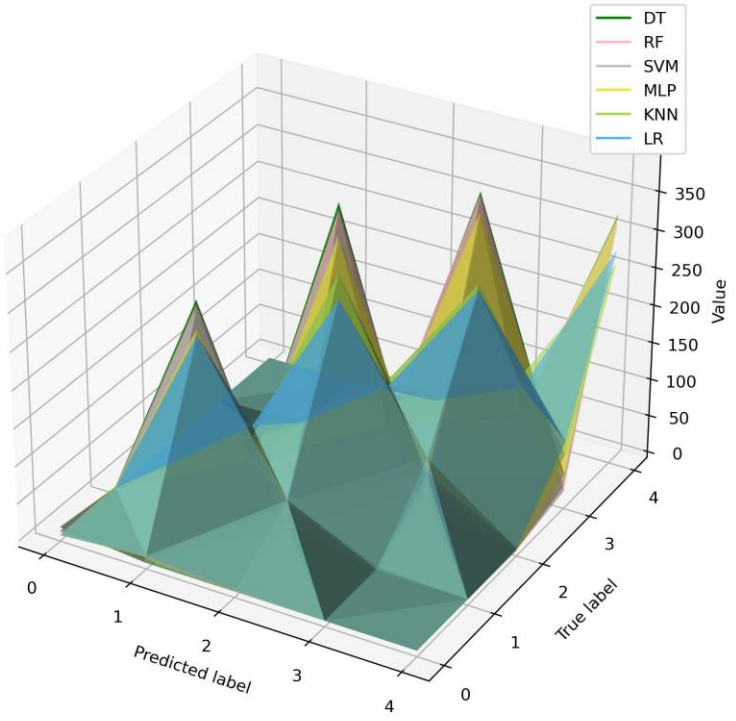
Results of AI-driven formative assessment

RQ1 – Accuracy, recall, f1 score

The most robust model was DT:

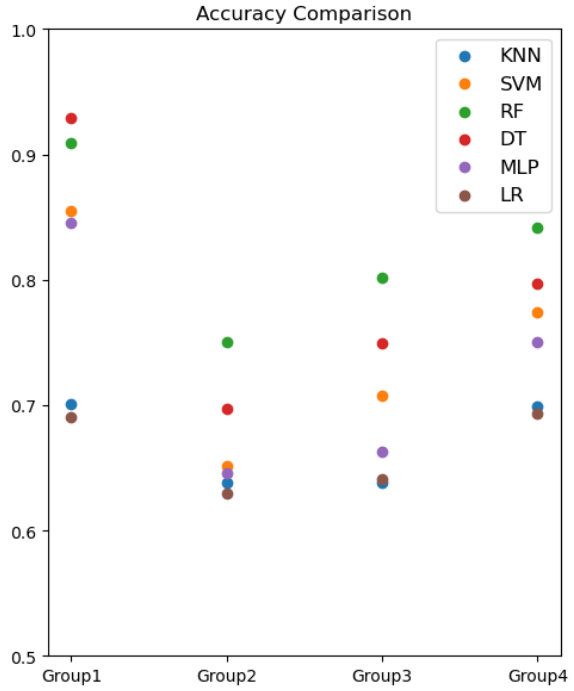
- accuracy 92.862%,
- recall 94.447%,
- f1 score 92.857%.

The DT model exhibited the highest true positive counts for all label types.



RQ2 – Key feature types

- The key feature types fell into learning achievement and engagement data.
- ANOVA and Tukey HSD post hoc test confirmed the significant difference between key and non-key features.



RQ3 – Formative assessment chart

The FA chart, based on the AI-driven FA model, uses metrics:

- fluency for learning achievement,
- interaction reflecting engagement status,
- AI-driven FA grade for thorough evaluation.