



HCI - Learning Discontinuity (Second Progress Report)

Jie Luo, Naifei Pan, Yiwon Zhang

Agenda

- ❏ Introduction/ Clients Information
- ❏ Data
- ❏ Methods & Preliminary results
- ❏ Next Steps & Roadblocks



Introduction

Advisor:

Zach Branson, Assistant Teaching Professor, CMU

Members:

Name: Yiwen Zhang

Background: MSP

Name: Naifei(Julia) Pan

Background: MSP

Name: Jie Luo

Background: MSP

Client Info

Vincent Aleven

- **Professor and Director of Undergraduate Programs in Human-Computer Interaction Institute, CMU**
- **Co-founder of Carnegie Learning & MathTutor**





Introduction - Background

- This project is tasked with developing a way to detect learning discontinuities within tutor log data to measure effects of out-of-tutor events in Intelligent Tutoring System (ITS)s.
- Research Questions:
 - Do these interventions put students on a different learning trajectory, with respect to the specific skills?
 - How can we measure effect?
 - Do we see struggles before tutor interventions?
- Purposes:
 - Improve Learning with tutor system
 - Improve scientific understanding of learning with ITS and teachers

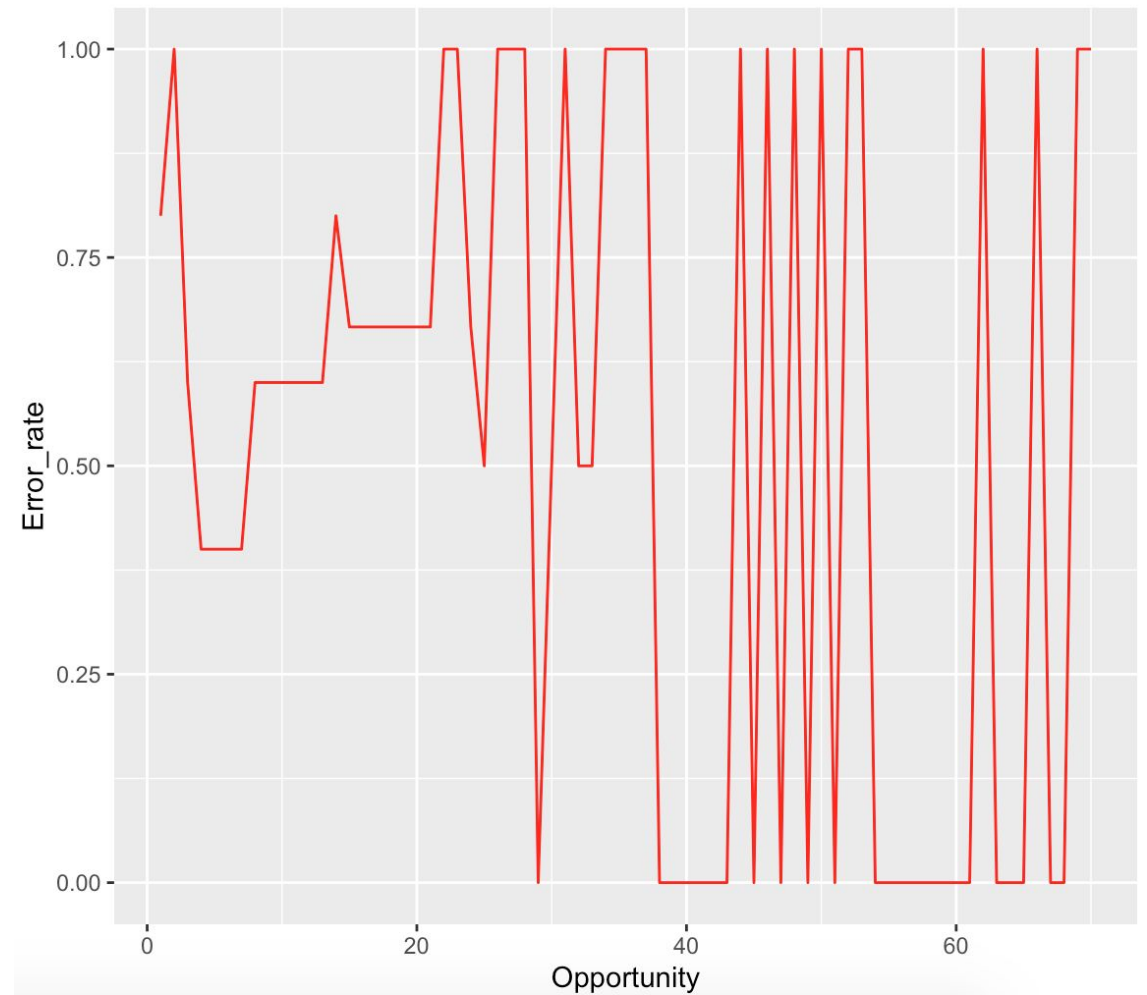


Data

- 2 Datasets: Students transaction dataset (104,550 transactions). Student-Step dataset (195 students).
- Transaction dataset: transaction time, **tutor response**, problem name, relevant KC, student actions, ...
- Student-step dataset: **opportunity**, problem name, relevant KC, Derived from Transaction dataset.
- **KC**: A Knowledge Component is “an acquired unit of cognitive function or structure that can be inferred from performance on a set of related tasks”. We have 7 KCs in our dataset. (Combine variable terms, Compute quotient for constant, etc)
- Each problem can test more than 1 KC.
- **Opportunity**: An opportunity is the first chance on a step for a student to demonstrate whether he or she has learned the associated KC. Opportunity number is therefore a count that increases by one each time the student encounters a step with the listed knowledge component.

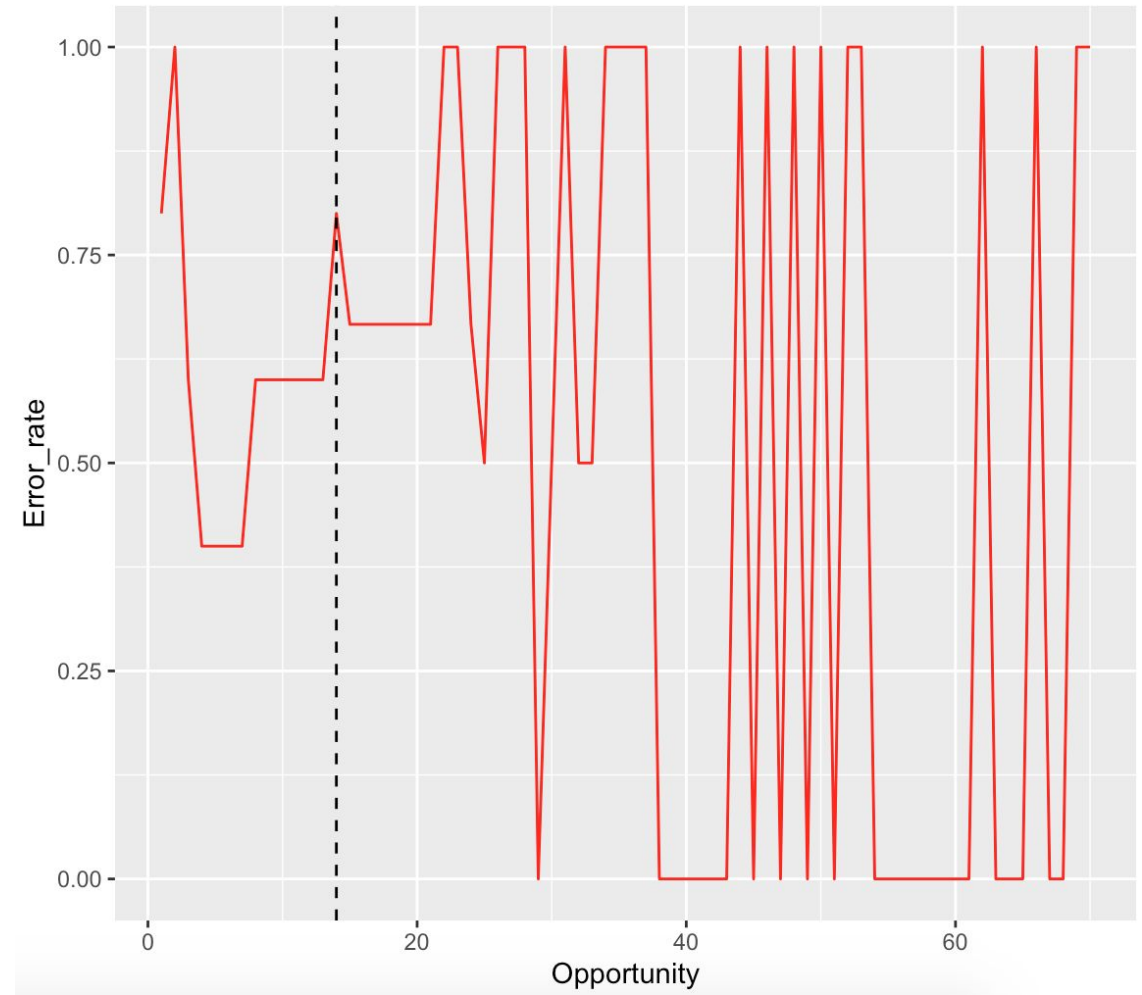
Data

- Incorrect attempt: once a student makes a mistake or asks for a hint in one attempt, we would call it an incorrect attempt
- Error rate: the proportion of incorrect attempts among total attempts

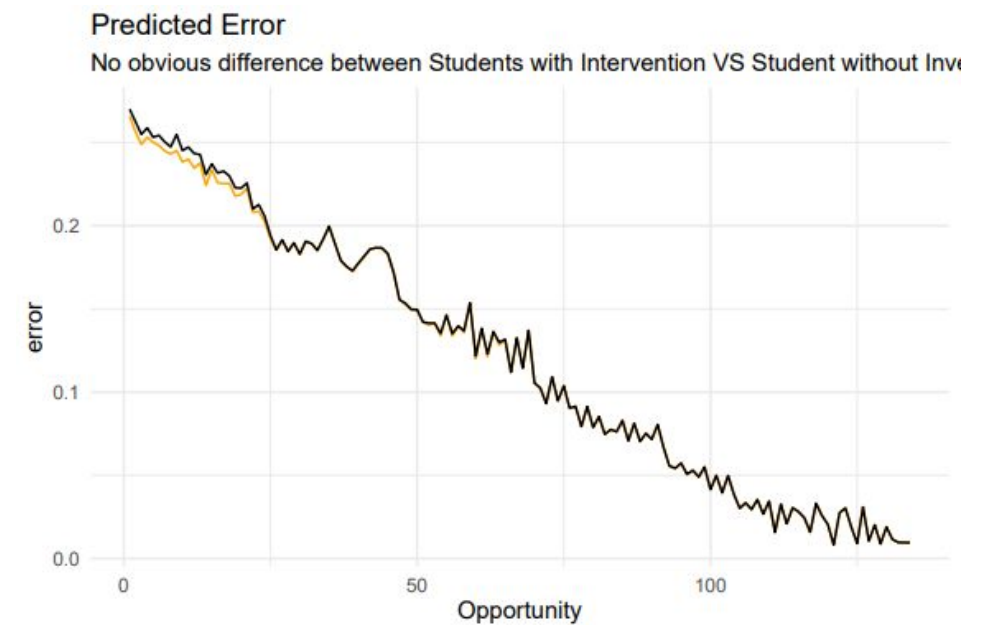
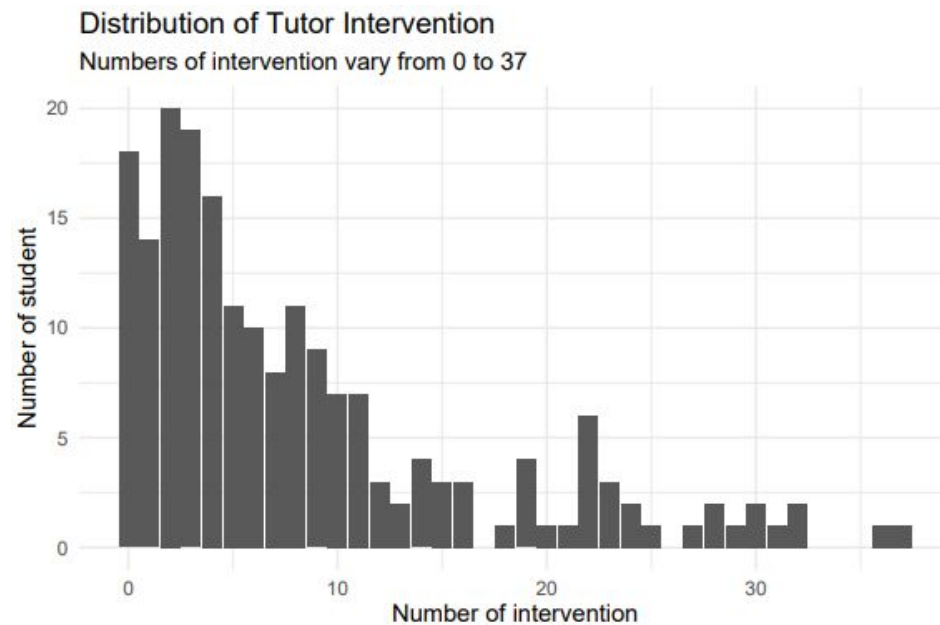


Data

- Tutor intervention: only available in Transaction dataset (observations listed by transaction time), manually combined with Student-step dataset
- Tutor intervention time: the earliest exposure to the tutor for a specific student. Here we assume that a tutor would influence all KCs at the same time.

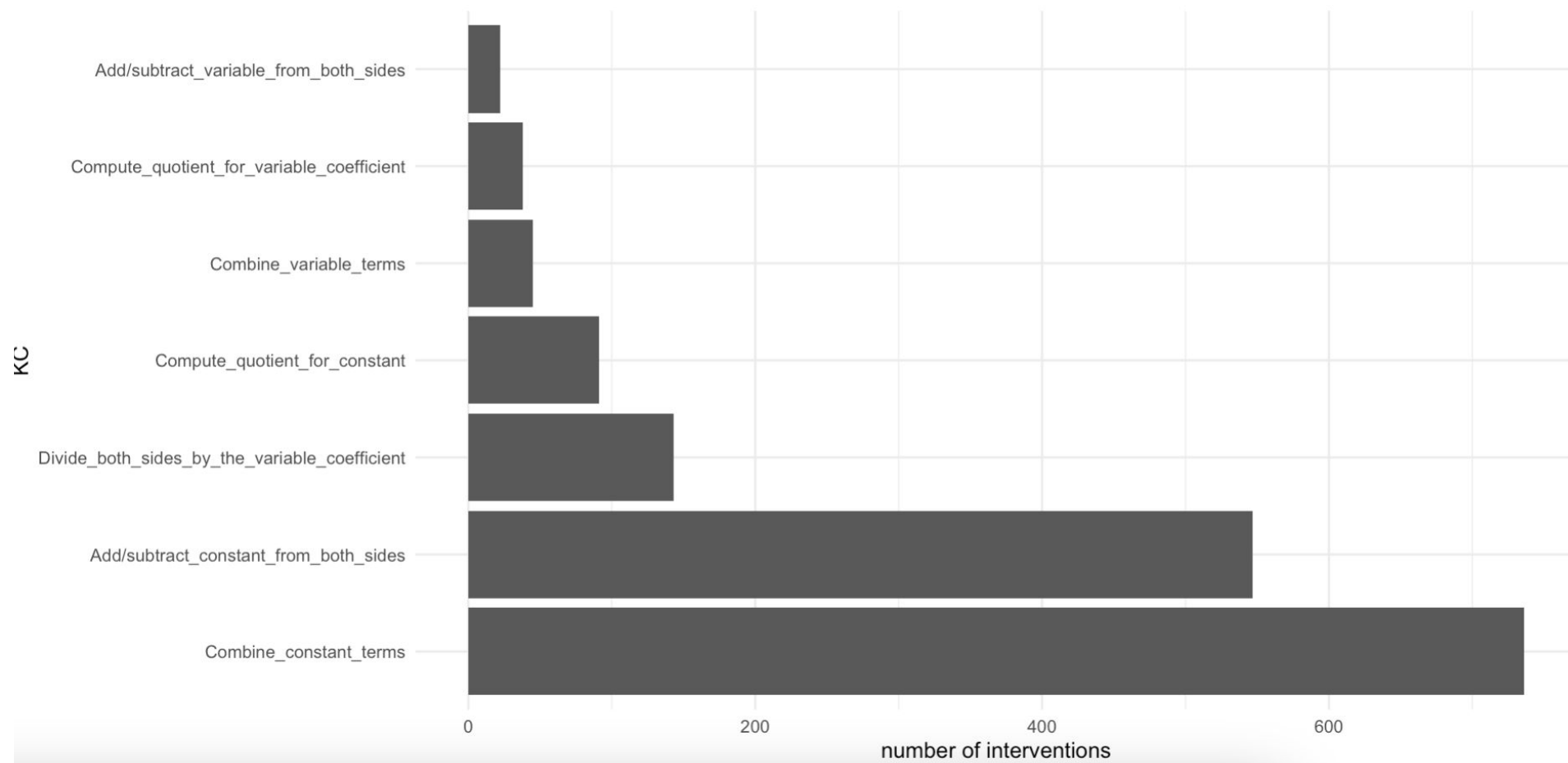


Data

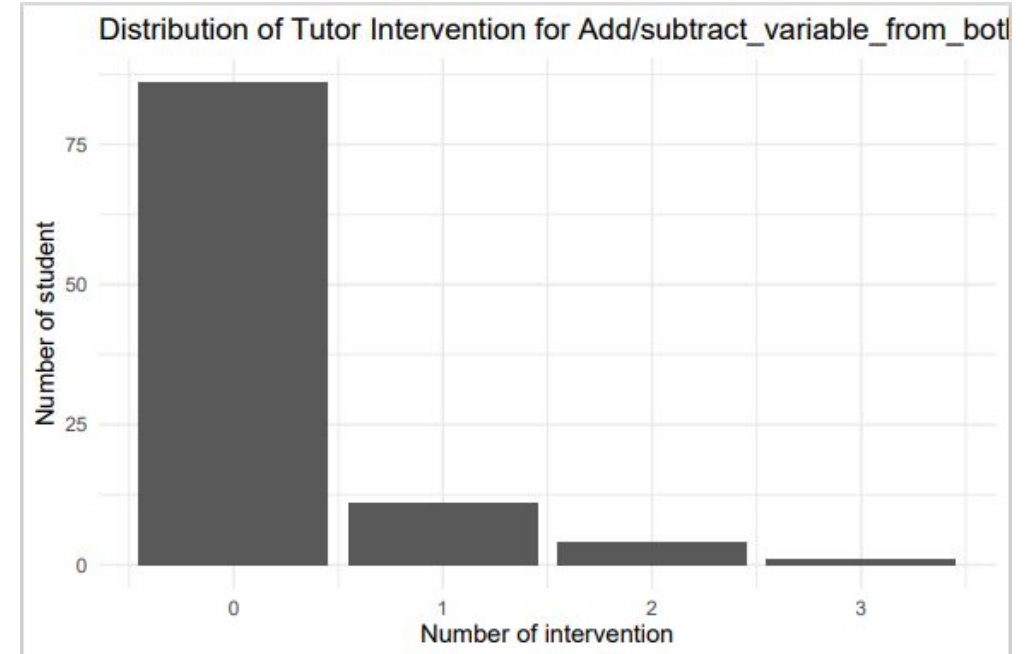
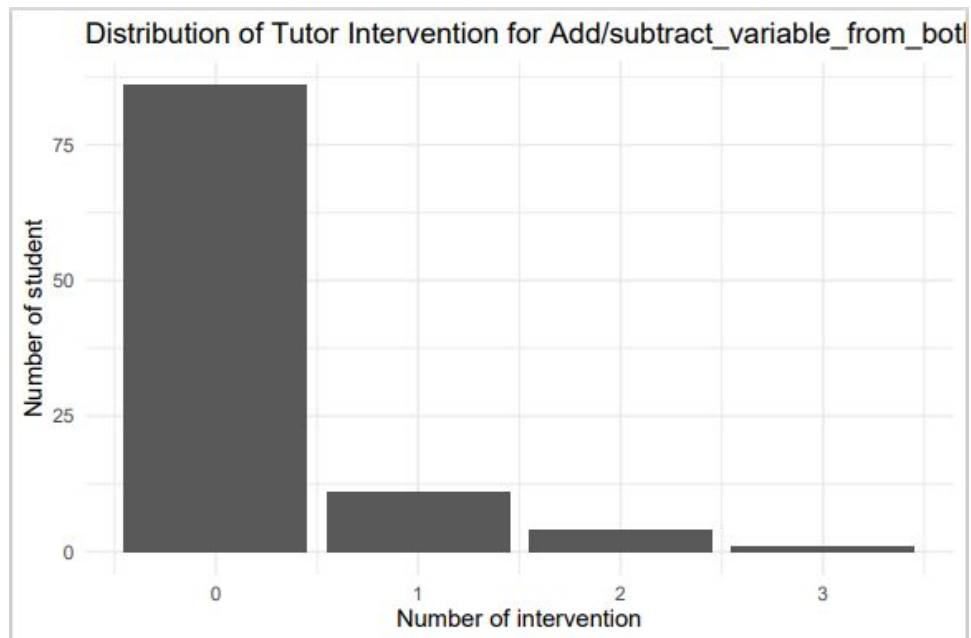


- Predicted error rate: estimated by AFM model

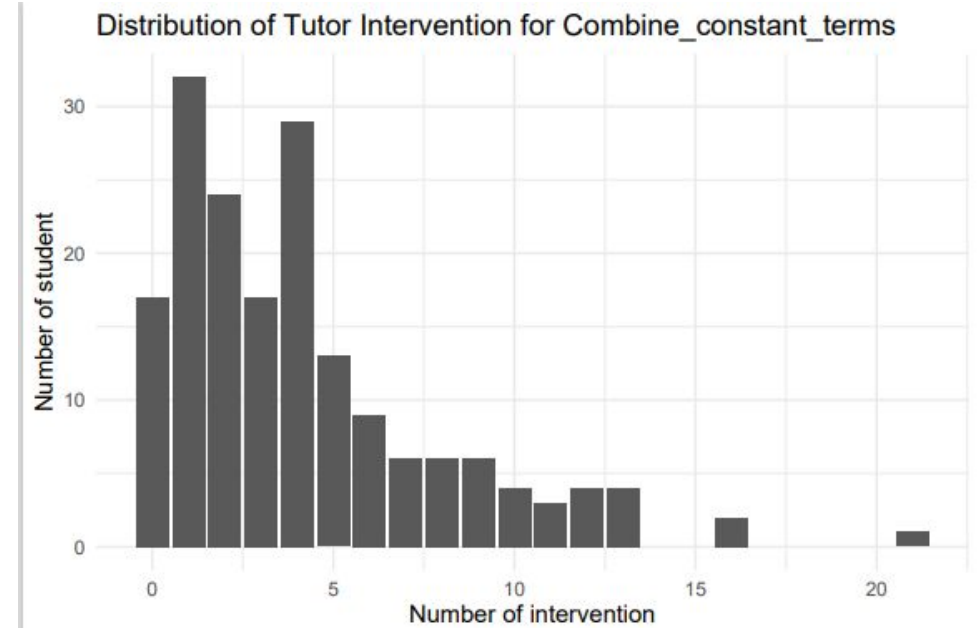
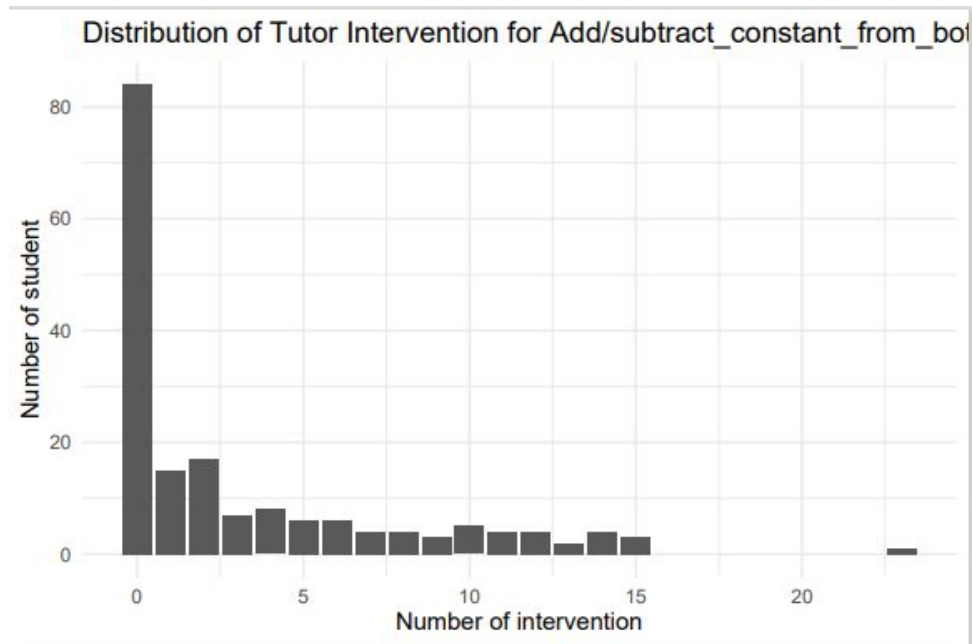
Data



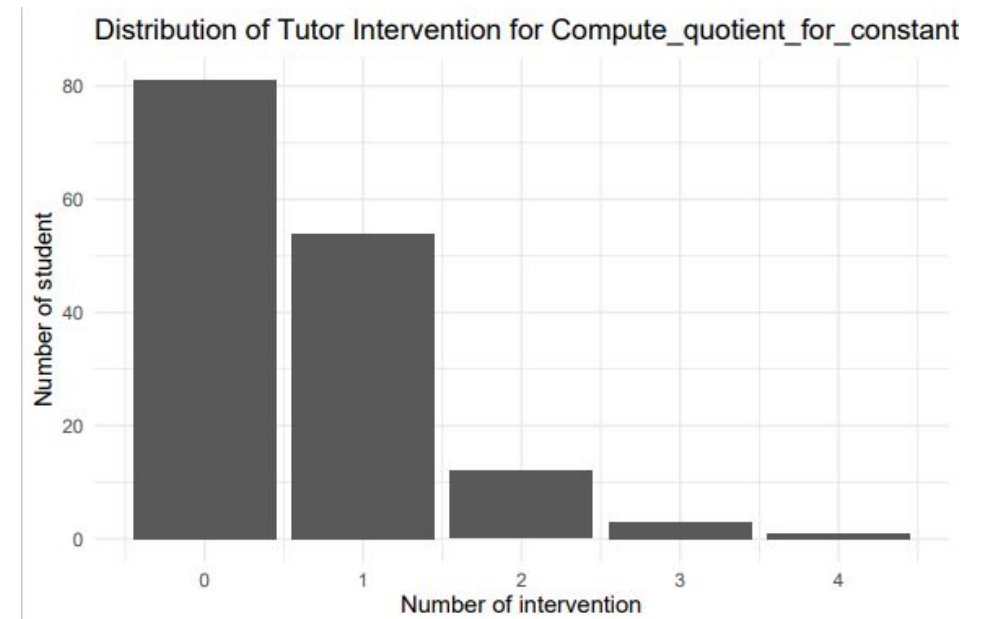
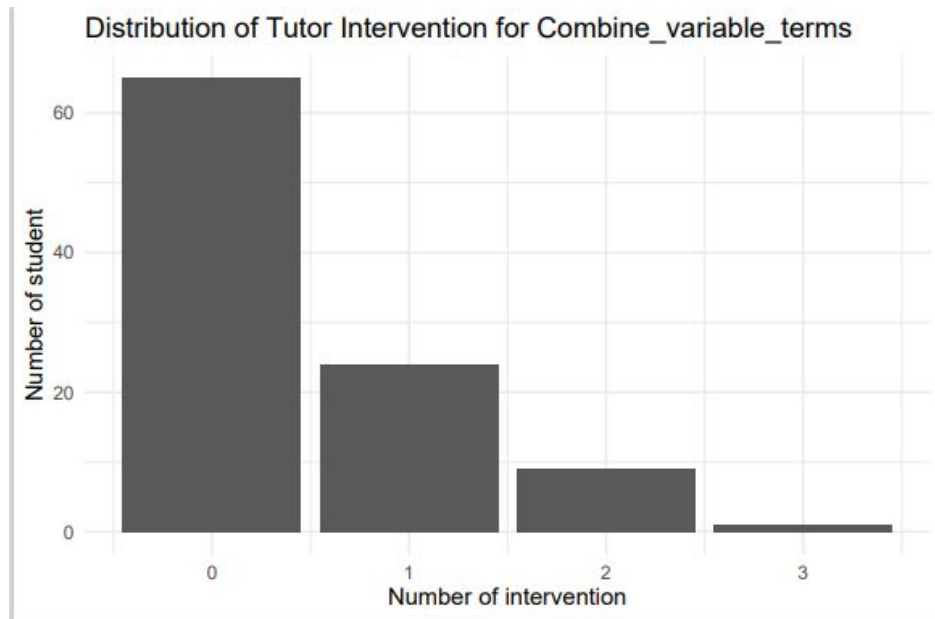
Data



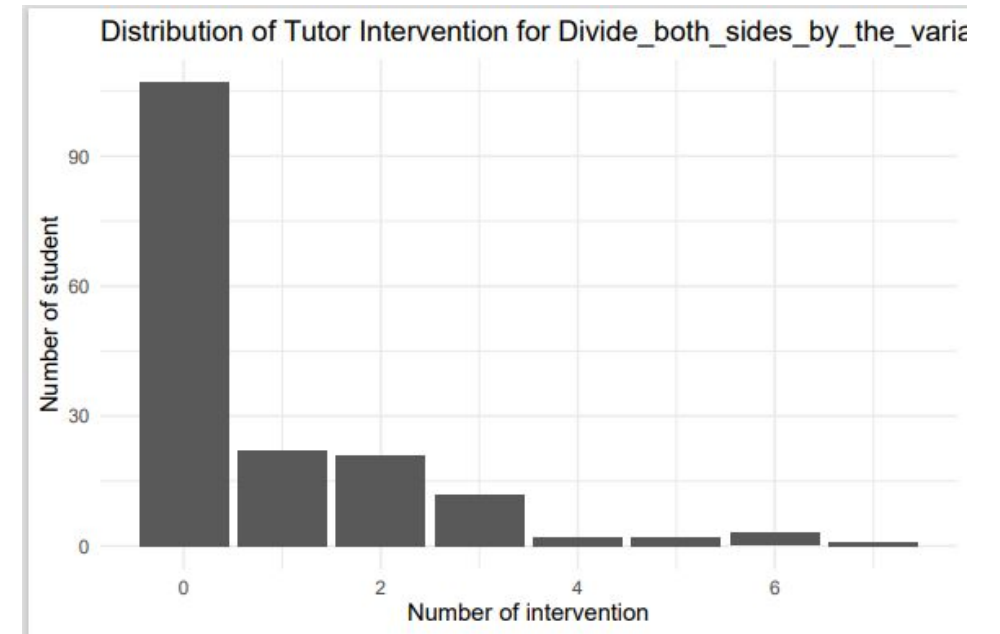
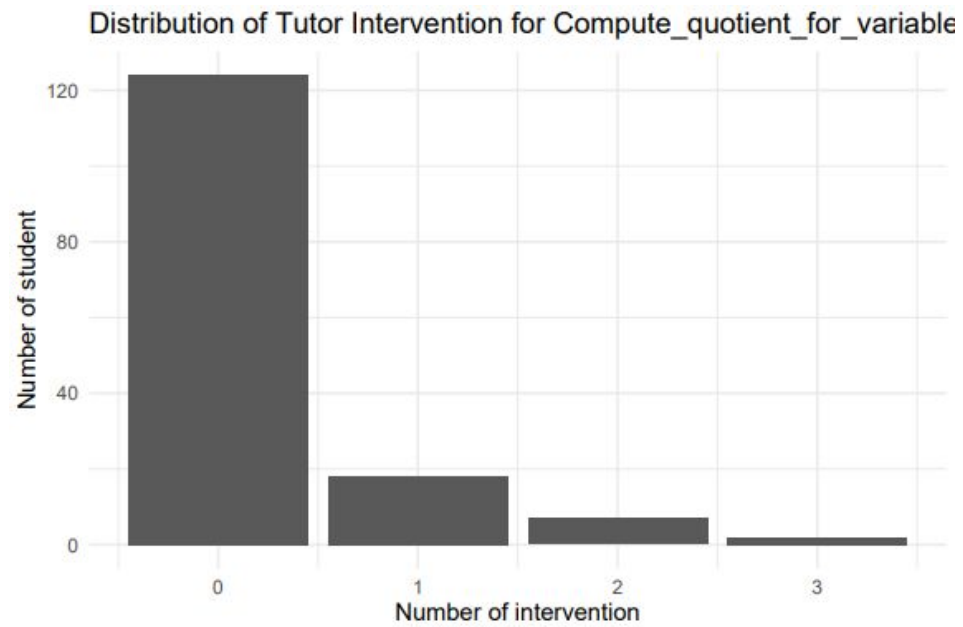
Data



Data



Data



Methods

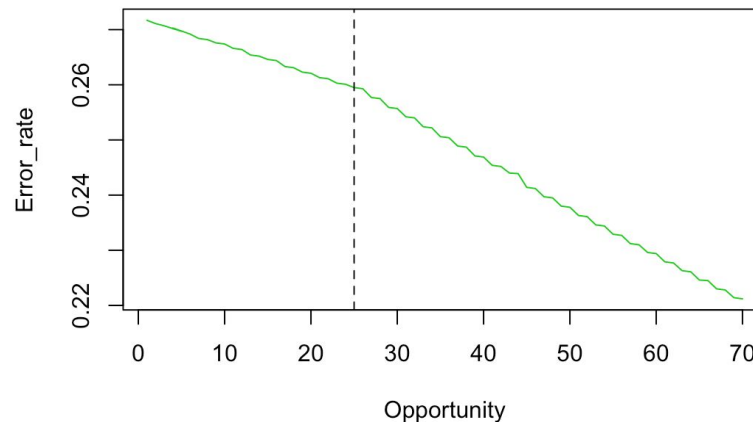
AFM model: logistic regression for predicting the success of the next step

Additive Factors Model (AFM)

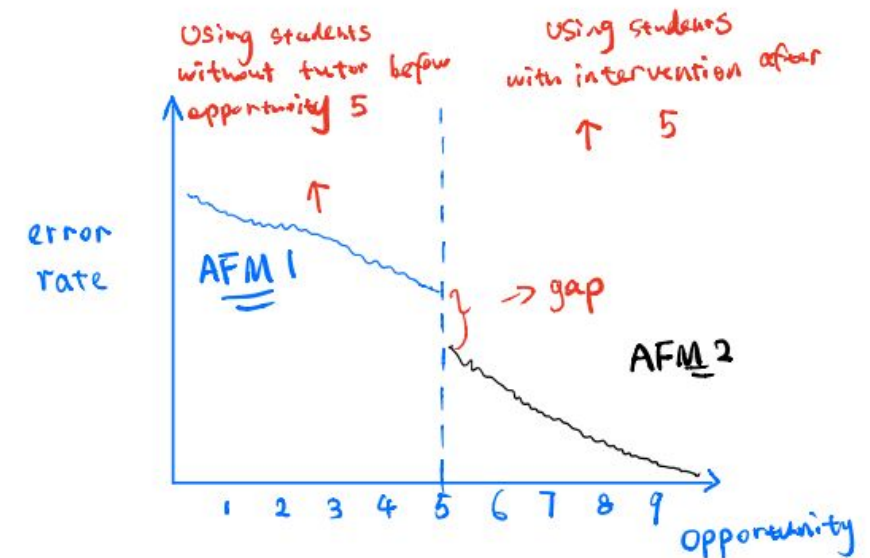
Log likelihood that student gets step correct = Student's initial proficiency + Ease of the KC + How much student learned on prior opportunities for this KC

$$\ln \frac{p_{ij}}{1-p_{ij}} = \theta_i + \sum_k \beta_k Q_{kj} + \sum_k Q_{kj} (\gamma_k N_{ik})$$

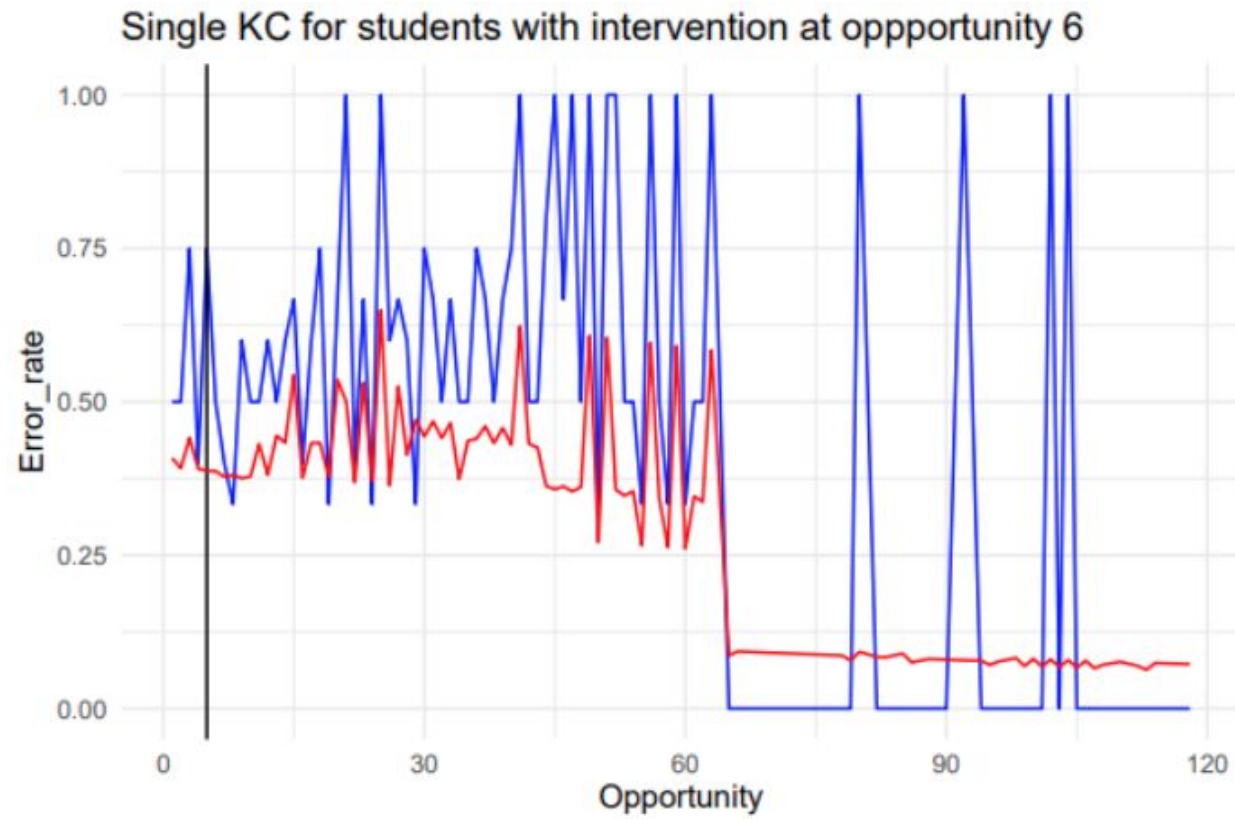
Example:
learning rate
1 Student,
1 KC



Fitting 2 AFM models (ideally with different slope and intercept) KC by KC
Start from KC where sample can be easily separated



Result- preliminary



- 7 students
- Single KC
- intervention happened at opportunity 6
- 1 AFM model

Challenges:

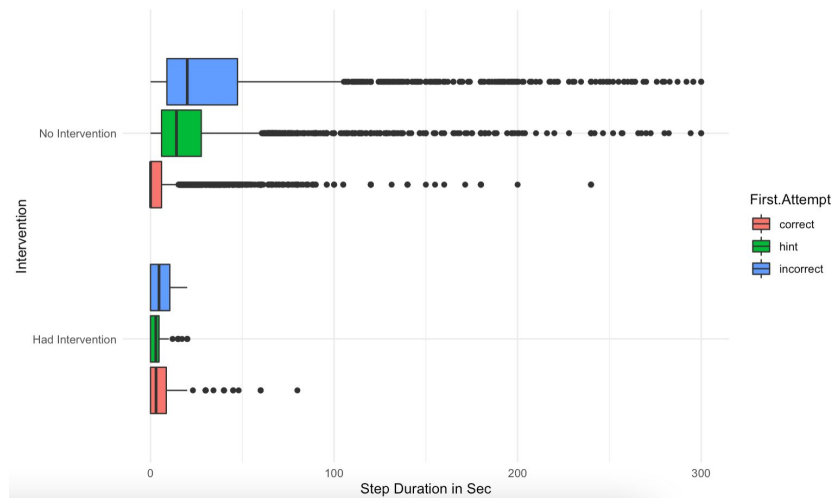
- Sampling observation
- Not every students has the same opportunity(Reasons for the drop at opportunity 65)

Next Steps

1. Intervention

a. **Happens to the struggling students at their struggling time?**

- Do students who never receive tutor help have lower error rate?
- Do we see long step duration before intervention?





Next Steps

1. Intervention
 - a. Happens at the struggle time?
 - b. **Accelerate learning rate?**
 - i. Fit two AFM models
 1. For pre-tutor data: fit a AFM
 2. For post-tutor data: fit another AFM (If intervention happens at opportunity M, then opportunity M+1 will be treat as opportunity 1)
 3. Compare the two AFM model (jump?)

Next Steps

2. Improve AFM model (adding intervention to the model)

Fit PFA (Performance Factors Analysis) model

AFM:

$$m(i, j \in KCs, \boxed{n}) = \alpha_i + \sum_{j \in KCs} (\beta_j + \boxed{\gamma_j n_{ij}})$$

PFA:

$$m(i, j \in KCs, \boxed{s, f}) = \sum_{j \in KCs} (\beta_j + \boxed{\gamma_j s_{ij} + \rho_j f_{ij}})$$



Q&A



Contacts

Project Members:

Jie Luo <jieluo@andrew.cmu.edu>

Naifei Pan <npan@andrew.cmu.edu>

Yiwen Zhang <yiwenzh3@andrew.cmu.edu>

Faculty Advisor:

Zach Branson <zach@stat.cmu.edu>