

# Knowledge Tracing - Continued

Machine Learning for Behavioral Data

April 1, 2025

# Today's Topic

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
<b>7</b>	<b>Time Series Prediction</b>
8	Time Series Prediction

## Supervised learning on time series:

- Probabilistic graphical models
- GLMMs
- Neural networks: LSTM, GRU, etc.

# Getting ready for today's lecture...

- **If not done yet:** clone the repository containing the Jupyter notebook and data for today's lecture into your Noto workspace
- SpeakUp room for today's lecture:

<https://go.epfl.ch/speakup-mlbd2025>



# Agenda

- Short quiz about the past...
- Learning Curves
- Alternative approaches to knowledge tracing
- Daylight Saving quiz!

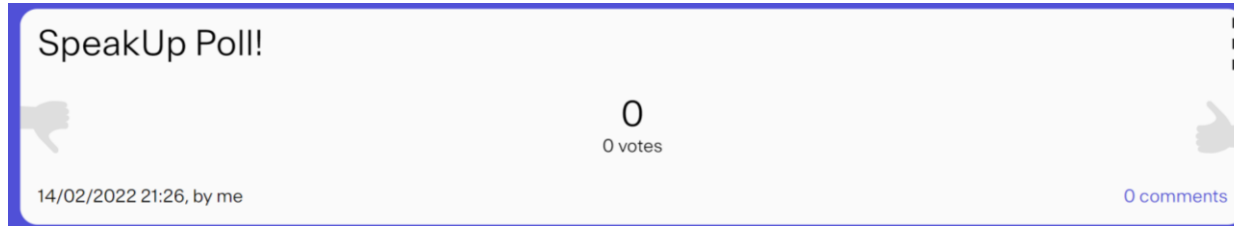
*Win Swiss chocolate!*



# Short quiz about the past...

BKT does account for students guessing the correct answer.

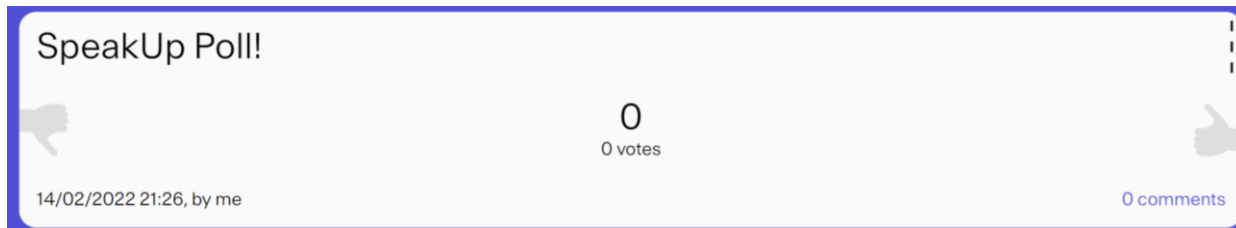
- a) True
- b) False



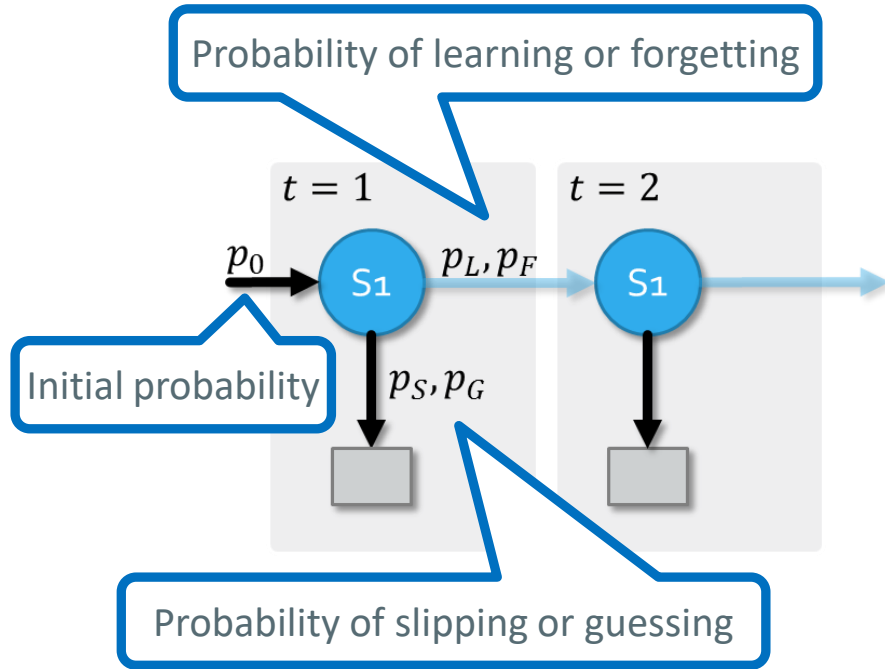
# Short quiz about the past...

Mixed-effect models are useful when the samples in the data set are uncorrelated.

- a) True
- b) False



# Last Week: Bayesian Knowledge Tracing



- Predict  $p(o_{i_{s_1}, t} | o_0, \dots, o_{t-1})$ , the probability that the student will solve task  $i_{s_1}$  correctly at time step  $t$
- Predict  $p(s_{1,t} | o_0, \dots, o_{t-1})$ , the probability that the student has mastered skill  $s_1$  at time step  $t$

# Assumptions behind BKT

- Knowledge can be divided into different skills
  - Definition of skills is accurate/detailed enough
  - Each task corresponds to a single skill (original)
  - There is **no** connection between the skills
  - Mastery can be achieved through practice
  - There is no forgetting:  $p_F = 0$  (original)
-



# Today

- **Learning Curves**
  - Alternative Models for Knowledge Tracing
    - AFM
    - PFA
-

# Today's Use Case

- ASSISTments is a free tool for assigning and assessing math problems and homework
  - All math problems (tasks/items) are associated to a specific skill/knowledge component
  - 4,151 middle-school students
  - 525,534 observations
-

# Tracing Knowledge – why is it useful?

- Is the student learning?
  - Measure what the student *knows* at a specific time  $t$
  - More specifically: knowledge of the student about relevant knowledge components (skills)

➡ Choose the next appropriate activity

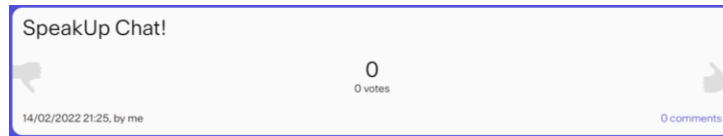
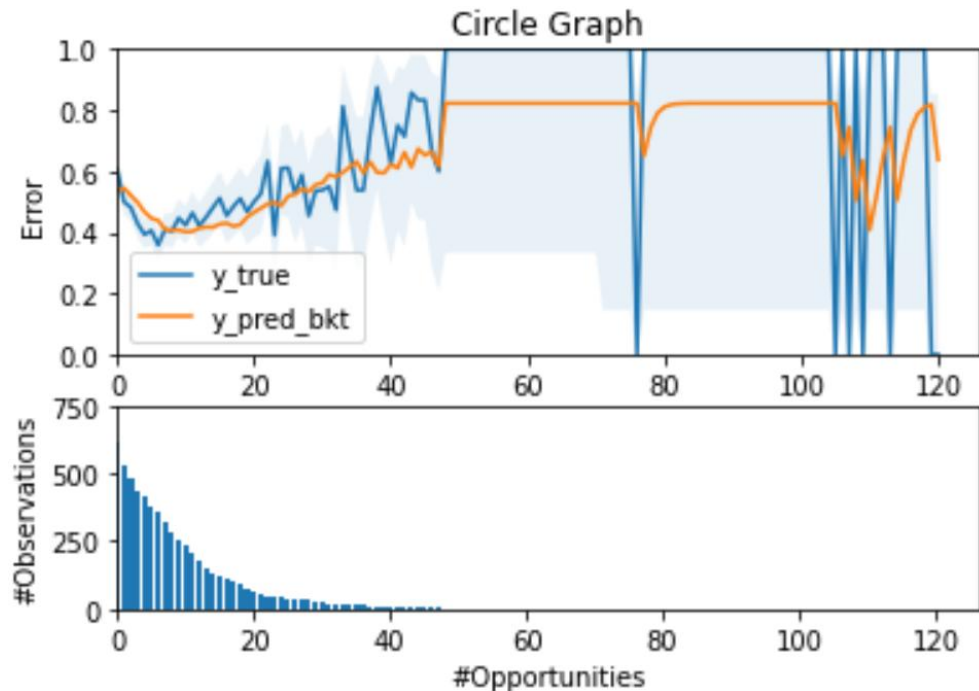
➡ Know which activities support learning

---

# Building a learning curve for skills

Student	Opportunity	y_true	y_pred
0	0	0	0.3
0	1	0	0.5
0	2	1	0.7
0	3	1	0.9
1	0	0	0.3
1	1	1	0.5
2	0	0	0.3
2	1	1	0.5
2	2	1	0.7
3	0	1	0.3
3	1	0	0.7
3	2	1	0.5
3	3	1	0.9

# What could this curve indicate?



# Your Turn – Learning Curves

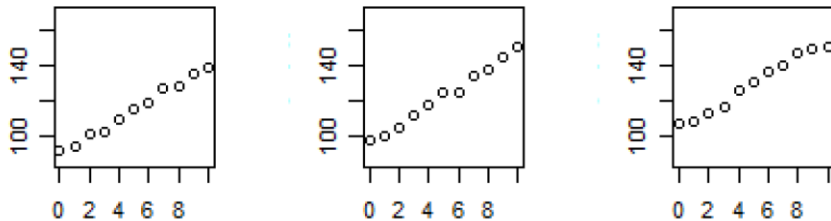
- In the student notebook, you have:
    - BKT model trained on all skills and students
    - List of available skills
    - Function for plotting learning curves and student numbers for a specific skill
  - Your task:
    - Pick 1-2 skills, generate the learning curves for them, and interpret them
    - Send us your plots and interpretations
-

# Today

- Learning Curves
  - **Alternative Models for Knowledge Tracing**
    - AFM
    - PFA
-

# Generalized Linear **Mixed Effects** Models revisited

- Example: strength gain by weight training
  - Each person has individual starting strength



$$y_n = \beta_0 + u_n + \beta_1 x_{n,1}$$

“Fixed” Effects

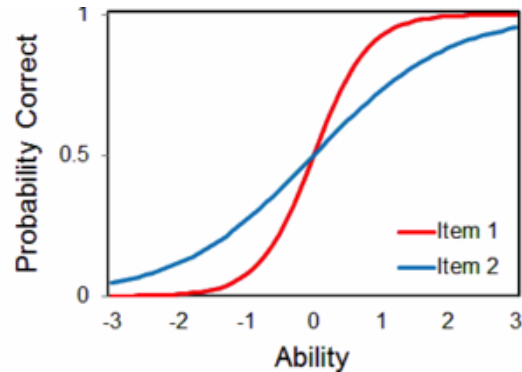
“Random” Effect



# Rasch Model

$$\log\left(\frac{p_{i,n}}{1 - p_{i,n}}\right) = \theta_n - b_i$$

Probability that student  $n$  will solve item  $i$  correctly.



$\theta_n$ : student ability

$b_i$ : difficulty of item  $i$

# Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

Probability that student  $n$  will solve task  $i$  correctly.

# Additive Factors Model (AFM)

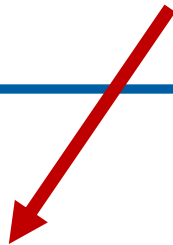
$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

# Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$




Student proficiency

# Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$



Student proficiency

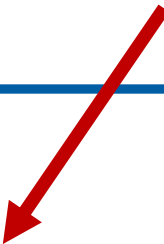


$q_{ik} = 1$ , if item  $i$  uses skill  $k$


# Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$



Student proficiency



Difficulty of  
skill  $k$



$q_{ik} = 1$ , if item  $i$  uses skill  $k$

# Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

Student proficiency

Difficulty of  
skill  $k$

Number of practice  
opportunities  
student  $n$  had at  
skill  $k$

$q_{ik} = 1$ , if item  $i$  uses skill  $k$

Learning rate  
at skill  $k$

# AFM - Assumptions

- Students may initially know more or less
  - Students learn at the same rate
  - Some skills are more likely to initially be known
  - Some skills are easier to learn than others
  - Students learn with each practice opportunity
  - Each item belongs to one or more skills
-



# Performance Factors Analysis (PFA)

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

# Performance Factors Analysis (PFA)

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot s_{n,k} + \rho_k \cdot f_{n,k})$$

Number of prior  
**successes** student  
 $n$  had at skill  $k$

Number of prior  
**failures** student  $n$   
had at skill  $k$

# PFA - Assumptions

- Students may initially know more or less
  - Students learn at the same rate
  - Some skills are more likely to initially be known
  - Some skills are easier to learn than others
  - Students learning rate differs for correct and wrong practice opportunities
  - Each item belongs to one or more skills
-

# AFM/PFA in action...

## ➡ Jupyter Notebook

Cheat sheet for mixed effect models:

<https://go.epfl.ch/mlbd-mixed-effects>

# Your Turn: Comparing Models

- We have evaluated AFM, PFA, and BKT on a subset of six skills. Your task:
    - Visualize the overall RMSE and AUC of the models such that it can easily be compared
    - Discuss the obtained results
-

# Summary

- Learning Curves
  - Alternative Models for Knowledge Tracing:
    - AFM
    - PFA
-

# Daylight Saving Quiz – Join us on Kahoot!



[www.kahoot.it](https://www.kahoot.it)

Enter the game pin!



*Win some Swiss chocolate!*

---