Artificial Agents that Learn to Teach

Rheeya Uppaal, Hira Dhamyal, Kadie Clancy, Janet Garcia, Akanksha Malhotra
Mentors: Emma Brunskill, Rika Antonova, Joe Rune & Dexter Lee

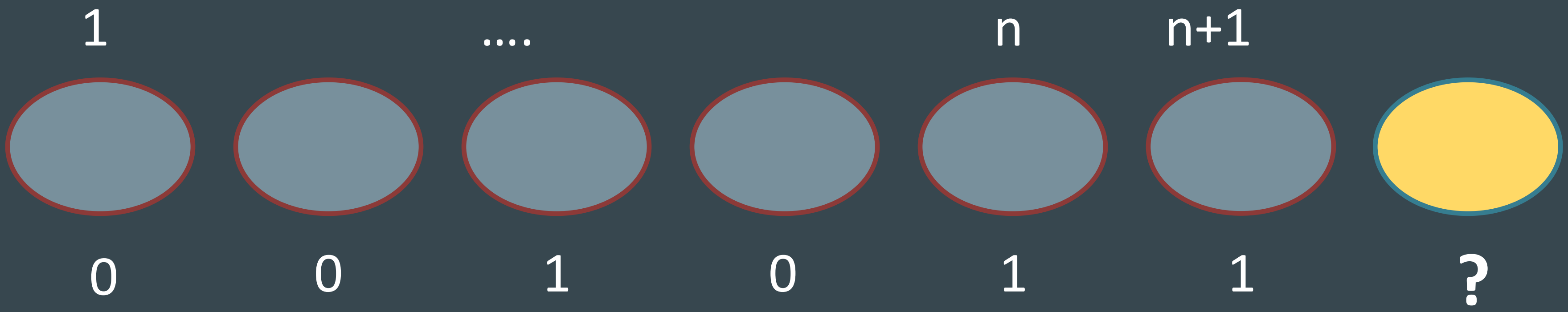
Problem

- Modeling student learning has direct implications for how to best teach students.
- Investigate how best to model student learning of histograms in a statistics intelligent tutor
  - Hidden Markov Model
    - Bayesian Knowledge Tracing (BKT)
  - Logistic Regression
    - Additive Factor Model (AFM)
    - Performance Factor Model (PFM)

Setting

Student gets a series of problems about a skill

Get to see if get each problem correct or not



Chronological response sequence for student *Y*

[0 = Incorrect response 1 = Correct response]

Approach: Logistic Regression

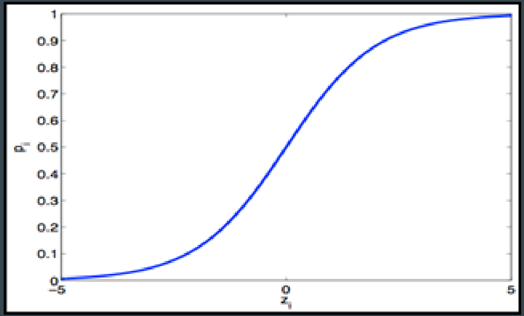
$p_i = P(Y_i = 1 \mid \text{features of student and step for example i})$

P(correct)

adjustable weights

features of student & step

$$P(Q_i = 1) = 1 \Big/ 1 + e^{-(\alpha_0 + \sum_{j=1:d} \alpha_j X_{ij})}$$



Additive Factor Model (AFM)

Essentially

-One student parameter

-One parameter that scales number of problems a student answers

-One parameter for skills difficulty

Performance Factor Model (PFM)

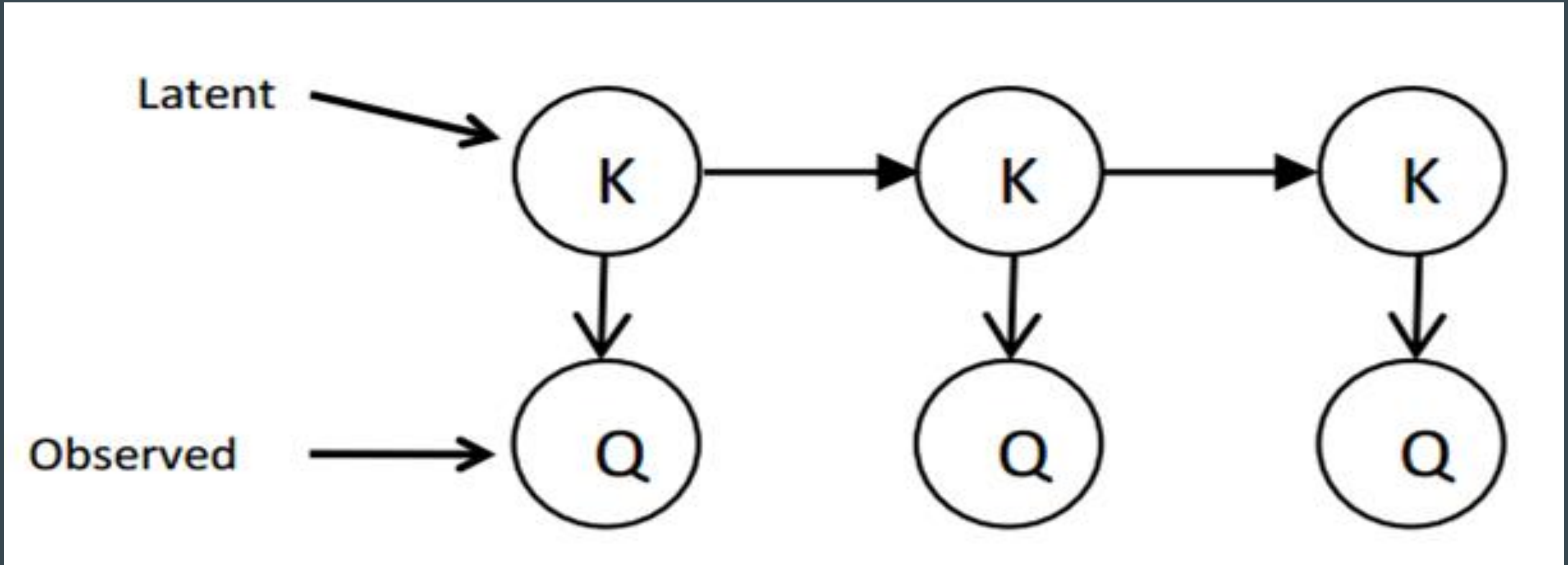
Essentially

-One student parameter

-One parameter for the number of correct answers and another parameter for the number of incorrect answers

-One parameter for skills difficulty

Bayesian Knowledge Tracing (BKT)



Bayesian Knowledge Tracing(BKT)

Ran the model on the whole data with 8 different skills

Made all the skills same

Divided one skill (y axis) into 2 different skills

- y axis had the lowest P(T)- probability to learn at each step
- Make the division of problem set in a better way

Student Model Evaluation

Using learned model parameters, predict how new students perform

How do you determine how good the predictions are?

Using Root Mean Squared Error

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(Q_i - P(Q_i)\right)^2}$$

Results and Findings for BKT

Original RMSE (for entire data set) = 0.486

On dividing the data set by skills,

Centre: 0.438	X Axis: 0.457	Y Axis: 0.487	Histogram: 0.458
Shape: 0.423	Spread: 0.446	H to D: 0.481	D to H: 0.459

On making all problems have the same skill,

RMSE drops to 0.473

On splitting the skill Y axis into ‘Y axis assess’ and ‘Y axis assign’

RMSE increases to 0.486

Results and Findings for Logistic Regression Models

Holding Out Some Data Per Student

PFM RMSE = 0.48

AFM RMSE = 0.48

- PFM variants
  - x-axis: 0.49
  - y-axis: 0.50
  - center: 0.42
  - spread: 0.45
  - h to d: 0.48
  - d to h: 0.49
  - shape: 0.45
  - histogram: 0.48
- AFM variants
  - x-axis: 0.50
  - y-axis: 0.52
  - center: 0.42
  - spread: 0.45
  - h to d: 0.49
  - d to h: 0.49
  - shape: 0.45
  - histogram: 0.48

Results and Findings for Logistic Regression

Holding Out Some Students

PFM RMSE = 0.50

- PFM variants
  - x-axis: 0.48
  - y-axis: 0.50
  - center: 0.53
  - spread: 0.52
  - h to d: 0.52
  - d to h: 0.52
  - shape: 0.48
  - histogram: 0.52

Surprizing Observations

Comparing BKT and AFM:

BKT:	RMSE = 0.486	(holding out students)
PFM:	RMSE = 0.48	(holding out items)
AFM:	RMSE = 0.4796	(holding out items)

OurCS
ML and Education

Emma Brunskill

With Rika Antonova, Joe Runde and Dexter Lee