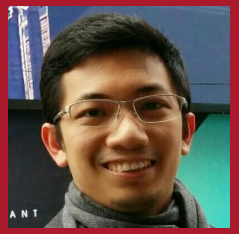
Building Models to Predict Hint-or-Attempt Actions of Students


Francisco Castro


Seth Adjei


Tyler Colombo


Neil Heffernan

The 8th International Conference on Educational Data Mining
Madrid, Spain
26-29 June 2015

Motivation

A great deal of EDM research focus on modeling student performance

- Bayesian Knowledge Tracing
- Performance Factors Analysis

A lot on affect (Baker's BROMP Protocol)[1]

[1] http://www.columbia.edu/~rsb2162/bromp.html

Motivation

Should we know if the student is "confident" enough to attempt a problem, without asking for help?

The Impact of Incorporating Student Confidence Items into an Intelligent Tutor: A Randomized Controlled Trial[2]

- self report on confidence might hurt students or be unreliable

[2] Charles Lang, Neil Heffernan, Korinn Ostrow, and Yutao Wang

Motivation

Understanding student behavior is crucial

- Better tutoring practices
- Improved content selection for ITSs
- Identify low-performing students

Research Questions

1. How do we determine when students will ask for help when using an ITS?

2. What information may be useful for developing models that forecast students' need for assistance?

Methods

- Used information on problem attempts and help (hint) requests to predict first action on the next problem

- Tabling methods for generating predictions[3]

[3] Wang, Q.Y., Kehrer, P., Pardos, Z. and Heffernan, N. Response Tabling – A simple and practical complement to Knowledge Tracing. KDD 2011 Workshop: Knowledge Discovery in Educational Data.

Dataset

ASSISTments

- Online tutoring system maintained at WPI
- www.assistments.org
- Data spans 5 months within the 2012-2013 school year
- A total of 599,368 log entries by 14,658 students across 589 problem sets
- Data is at http://bit.ly/1KaEsJO

Experimental Models

1. **Attempt/Hint Count** (**AHC**) Model
   - Number of attempts and hints used

2. **Hint History** (**HH**) Model
   - History of hint request as first action in preceding questions

Example: AHC Prediction

Attempts Taken	Hints Taken			
	0 / 3	1 / 3	2 / 3	3 / 3
1	0.0211	0.1001	0.2213	0.4025
2	0.0261	0.0558	0.0747	0.1105
3	0.0237	0.0447	0.0737	0.0916
4	0.0363	0.0287	0.0743	0.0949
5	0.0132	0.0263	0.0857	0.0912

Student	A_C	H_C	H_T	FANP
92677	1	0	3	0.0211
92680	2	3	3	0.1105

Experimental Models

*. **Baseline** (**BL**) Model

- No gold standard for first-course-of-action prediction
- Hint instances on students' second action

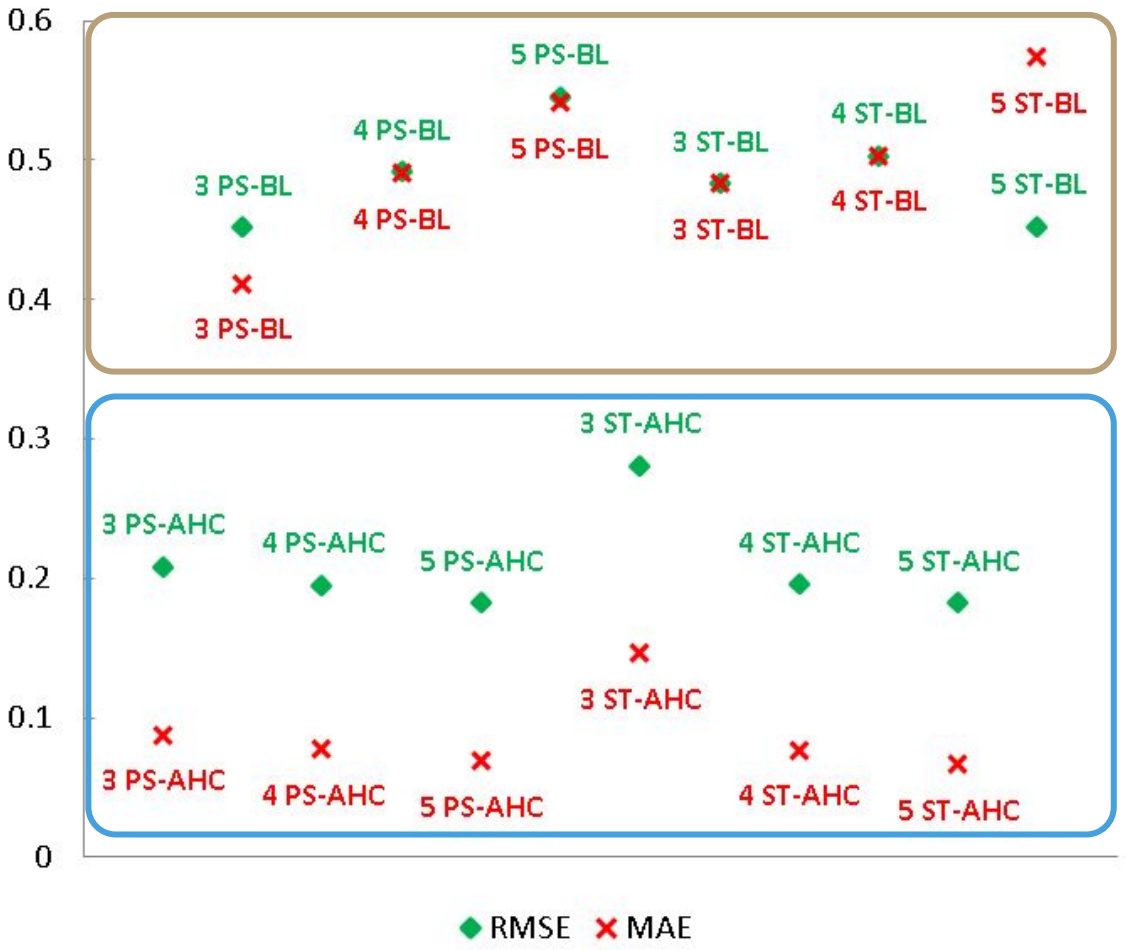
Analysis

- Problem set and student level analysis
- Training, testing: 5-fold cross-validation

Problem entries used:

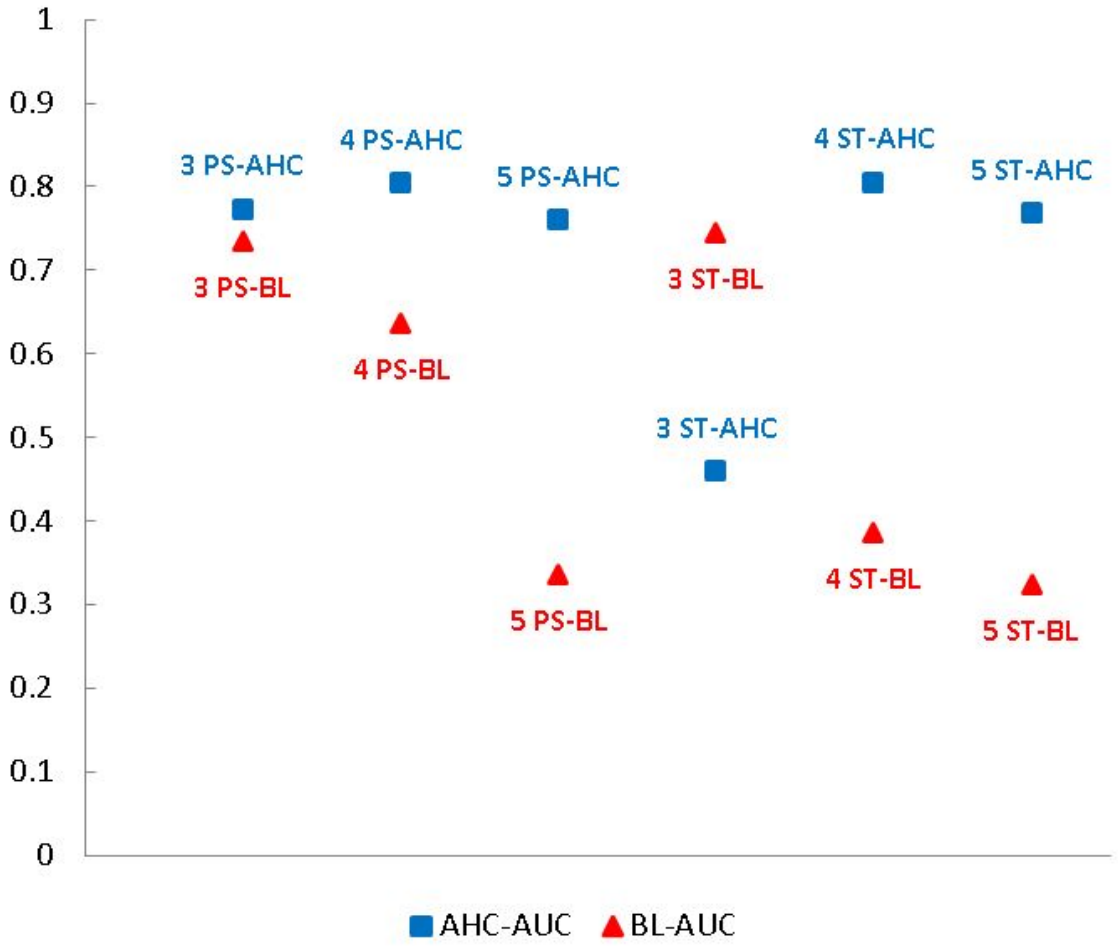
- **AHC**: Problems with 3, 4, 5 available hints
- **HH**: Problems with 3, 4 prior responses per student

RMSE/MAE Results: AHC vs BL



Note: **PS** = Problem set **ST** = Student
Numbers = no. of available hints

AUC Results: AHC vs BL

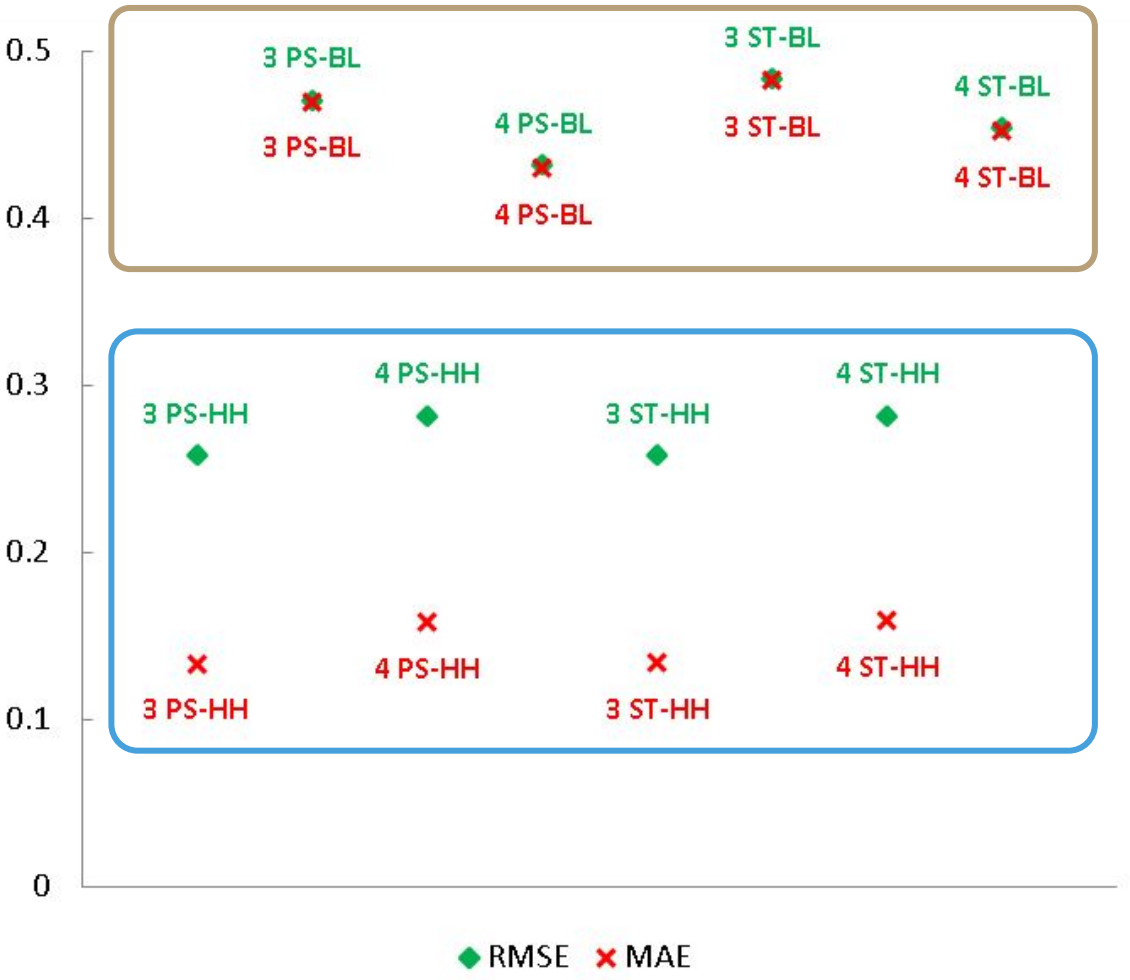


Note: **PS** = Problem set **ST** = Student
Numbers = no. of available hints

Results Summary: AHC model

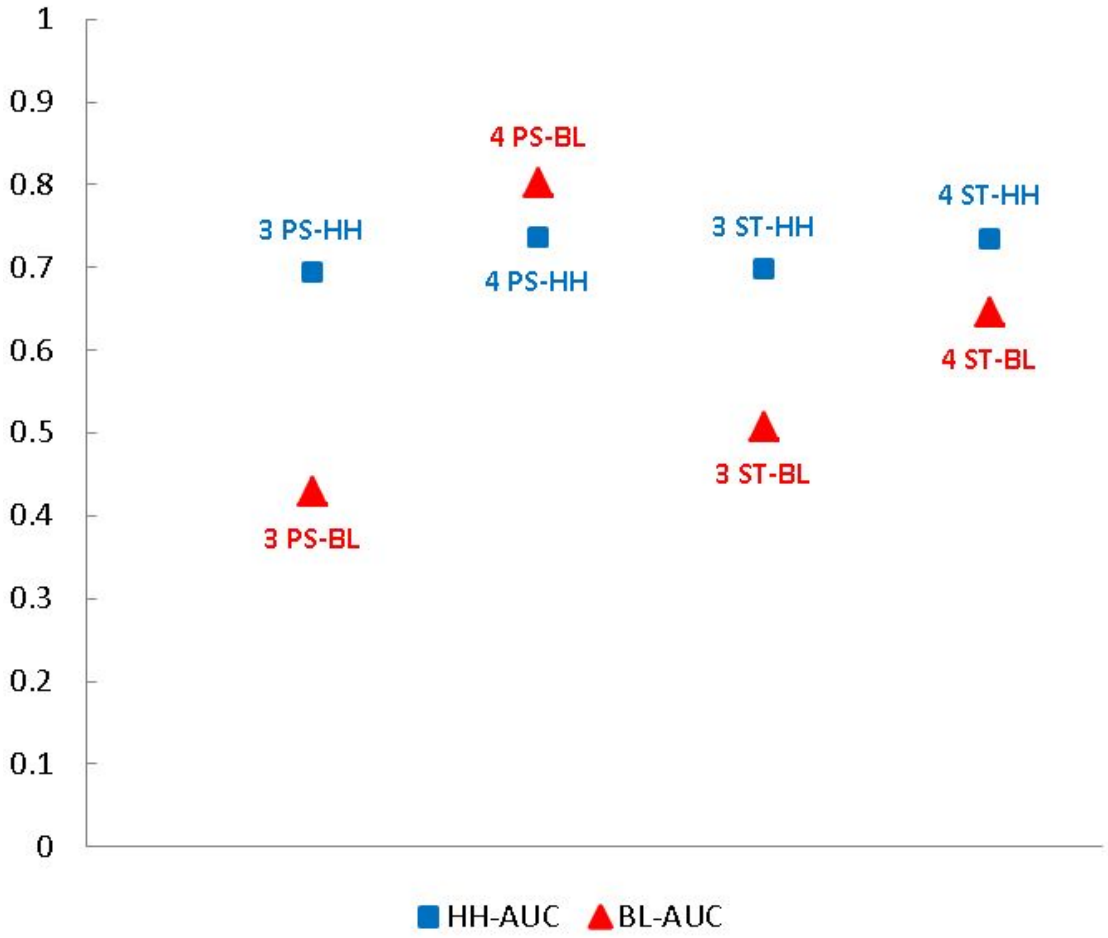
- AHC predictive performance in all metrics is fairly consistent
- Model is fairly generalizable across problems with varying number of hints
- For student level analysis, model performs well provided there is a high number of opportunities to ask for help

RMSE/MAE Results: HH vs BL



Note: **PS** = Problem set **ST** = Student
Numbers = no. of prior problems

AUC Results: HH vs BL



Note: **PS** = Problem set **ST** = Student
Numbers = no. of prior problems

Results Summary: HH model

- HH predictive performance in all metrics is fairly consistent
- Model is fairly generalizable across unseen skills and unseen students, as well as across the number of first action history points

Research Questions Answered

RQ1:

How do we determine when students will ask for help when using an ITS?

- Building models that use students' hint usage and attempt counts produce fairly reliable models that seem to generalize to unseen student and unseen problems

Research Questions Answered

RQ2:

What information may be useful for developing models that forecast students' need for assistance?

- Previous Hint and Attempt Usage
- Attempt and hint history models

Contribution

- Experimental results suggest students' help request behavior can be predicted from data descriptive of student action information

- Starting initiative in using action information to build up future studies

Future Work

- Student action patterns
- Leverage other information:

  e.g. Student response times, skill difficulty
- Models' performance with other datasets

Questions?

Results: AHC vs BL

PS	3 AHC	3 BL	4 AHC	4 BL	5 AHC	5 BL
RMSE	0.2075	0.4506	0.1942	0.4910	0.1813	0.5445
MAE	0.0866	0.4104	0.0763	0.4899	0.0677	0.5403
ST	**3 AHC**	**3 BL**	**4 AHC**	**4 BL**	**5 AHC**	**5 BL**
RMSE	0.2799	0.4826	0.1945	0.5023	0.1811	0.4514
MAE	0.1452	0.4821	0.0758	0.5022	0.0653	0.5729

PS	3 AHC	3 BL	4 AHC	4 BL	5 AHC	5 BL
AUC	0.7737	0.7332	0.8043	0.6338	0.7602	0.3338
ST	**3 AHC**	**3 BL**	**4 AHC**	**4 BL**	**5 AHC**	**5 BL**
AUC	0.4599	0.7419	0.8056	0.3841	0.7689	0.3223

Note: **PS** = Problem set **ST** = Student
Numbers = no. of available hints

Results: HH vs BL

PS	3 HH	3 BL	4 HH	4 BL
RMSE	0.2574	0.4697	0.2809	0.4307
MAE	0.1327	0.4687	0.1572	0.4291
ST	**3 HH**	**3 BL**	**4 HH**	**4 BL**
RMSE	0.2573	0.4821	0.2808	0.4528
MAE	0.1328	0.4810	0.1580	0.4513

PS	3 HH	3 BL	4 HH	4 BL
AUC	0.6936	0.4298	0.7357	0.8026
ST	**3 HH**	**3 BL**	**4 HH**	**4 BL**
AUC	0.6989	0.5071	0.7355	0.6458

Note: **PS** = Problem set **ST** = Student
Numbers = no. of prior problems

Example: HH Prediction Table

	Previous 3 First Action Hints / Attempts			
	0H / 3A	1H / 2A	2H / 1A	3H / 0A
# Attempt	111017	17219	3330	683
# Hint	5859	3254	1833	1663
% Hint	0.0501	0.1589	0.3550	0.7089

Example: BL Prediction

Problem entries	Hint Count: 2nd Action	Hint % (BL)	Attempt %
2200	852	0.3872	0.6127