Predicting Student Aptitude Using Performance History

Anthony F. Botelho
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609-2280
abotelho@wpi.edu

Seth A. Adjei
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609-2280
saadjei@wpi.edu

Hao Wan
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609-2280
hale@wpi.edu

Neil T. Heffernan
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609-2280
nth@wpi.edu

ABSTRACT

Many tutoring systems currently in use provide a wealth of information pertaining to student learning over long periods of time. Providing meaningful representations of student performance can indicate levels of knowledge and understanding that can alert instructors to potential struggling students in order to provide aid where it is needed; it is the goal of many researchers to even provide such indication preemptively in order to intervene before students become frustrated when attempting new skills. The goal of this work is to utilize student performance history to provide a means of quantizing student aptitude, defined here as the speed at which a student learns, and then using this measurement to predict the speed at which each student will learn the next skill before beginning. Observing a dataset of 21 skills, we compare two methods of predicting aptitude to majority class predictions at the skill level. Our results illustrate how our proposed methods exhibit different strengths in predicting student aptitude when compared to majority class, and may be used to direct attention to a struggling student before attempting a new skill.

Keywords

Aptitude, Student Knowledge, Intelligent Tutoring Systems

1. INTRODUCTION

Many instructors rely on intelligent tutoring systems (ITS) as a means of extending student learning outside the classroom. Many such systems, such as the ASSISTments system used in this work, provide a wealth of student performance data that is often underutilized. While many systems have focused on and have shown success in predicting next problem correctness, such information is only useful to instructors in a short time-span as students are completing assignments. Furthermore, many of these models rely on latent variables that leas to problems of identifiability [1] when attempting to draw conclusions of student knowledge.

The purpose of this work is to observe and predict student learning rates, referenced throughout this paper as aptitude; this value is expressed as a metric in terms of completion speed (cs), or the number of problems a student needs to complete the assignment (described further in the next section). Such a measure of aptitude in prerequisite skills has shown to be successful in predicting initial knowledge, represented as correctness, on a subsequent skill [2], illustrating that the two concepts are related, but from that work, it is unclear as to whether student aptitude is transitive across skills. In this work, therefore, we strive to answer the following research questions:

1. Do students exhibit similar degrees of aptitude across skills?
2. Are changes in student aptitude across skills predictable?
3. Can a student's aptitude in previous skills be used to construct a reliable prediction of completion speed in a new skill before it is begun?

2. METHODOLOGY

The dataset[1] used in this work is comprised of real-world data from PLACEments test data reported from the ASSISTments tutoring system. Data pertaining to 21 unique observable skills was extracted. Here, we define a skill as observable if it contains data from more than 10 unique students, and no less than half of the students must have completed the skill. ASSISTments defines skill completion in terms of 3 consecutive correct answers.

We used a simple binning method implemented in similar research [2][3] to place students into one of five categories based on completion speed in order to represent different levels of aptitude. As aptitude is an independent concept of domain knowledge, a student's entire recorded performance history, regardless of the prerequisite structure, was used to categorize each student. Observing each student's performance over several skills, we used a moving average of student completion rates of each skill ordered from oldest

[1]The original raw dataset can be found at the following link: http://bit.ly/1DVbHdB.

to most recent. Equation 1 displays the formula for this method. For our implementation, we used a value of 0.3 for alpha.

$$A_t = ((1 - \alpha) * A_{t-1}) + (\alpha * V_t) \tag{1}$$

Table 1: The ranges of completion speed represented by each bin with corresponding the quantized aptitude value.

Bin Number	Completion Speed(cs)	Quantized Value
1	$3 \leq$ cs ≤ 4	1
2	$4 <$ cs < 8	0.75
3	$8 \leq$ cs	0.5
4	DNF, pcor $\geq .667$	0.25
5	DNF, pcor $< .667$	0

Once an average completion speed, in terms of number of problems needed to reach three sequential correct responses, each student is placed in the corresponding bin described in Table 1. Bins 4 and 5 contain students that did not finish (DNF) at least one previous skill, and are instead split based on the average percent correctness (pcor) across all previous skills. The quantized values are chosen arbitrarily to discretize the learning rate that is intended to be represented by each bin.

2.1 Experiments

Our first prediction method, referenced as Same Bin Prediction (SBP) in our results section, simply uses the average completion speed of each student's performance history to determine in which bin to place each student. The method then simply uses that bin's quantized value as a prediction for the new skill. Both the SBP and majority class are then compared to each student's actual completion speed, expressed as a quantized bin value, to determine both error rates.

Our second experiment attempts to make predictions again using each student's performance history, but by also taking into account changes in aptitude across skills. Our first experiment assumes that most students will exhibit the same level of aptitude in a new skill as in previous skills. This experiment takes into account the realization that differences in skill difficulty may cause fluctuations in our aptitude measurements. Our second method, referenced as Transitioning Bin Prediction (TBP) in our results section, builds off of the previous SBP prediction by calculating an offset transition value. For example, if half the students in bin 1 (value = 1) remained in that bin for the new skill, while half transitioned to bin 2 (value = 0.75), an offset value of -0.125 would be applied to all predictions of bin 1. A negative offset indicates that many students required more opportunities to complete than normal, while a positive offset indicates the reverse. The prediction is normalized to a value between 0 and 1 to make full use of our quantized values

3. RESULTS AND CONCLUSIONS

Table 2 contains the RMSE results of each prediction method divided by each bin of the new skill. The success of the majority class predictions extends across higher aptitude students, while the TBP method provides the most accurate predictions over students in the lower aptitude bins.

Table 2: Average RMSE of the skill level analysis divided by bin.

Bin of New Skill	Majority Class	SBP	TBP
1	0.230	0.498	0.358
2	0.120	0.356	0.170
3	0.284	0.362	0.205
4	0.307	0.526	0.251
5	0.571	0.659	0.497

Table 3: Percent correctness at the skill level divided by bin.

Bin of New Skill	Majority Class	SBP	TBP
1	0.709	0.479	0.500
2	0.280	0.245	0.268
3	0.102	0.251	0.200
4	0	0.029	0.129
5	0	0.041	0.333

Each method described in this work exhibited different strengths, including the simple majority class predictions. It is often for the benefit of both teachers and students that a model represent meaningful information beyond the provision of predictive accuracy. The SBP method, for example, while not excelling in any one category, illustrates tendencies of aptitude mobility. Such methods may act as a means of better understanding and developing course structure and skill relationships.

The fact that the proposed prediction methods fail to outperform majority class overall suggests that using all performance history is not by itself a strong predictor of future performance, and is instead dependent to some degree on skill-based attributes. This work ignores prerequisite skill hierarchies available in many tutoring systems and MOOCs, using all previous performance history. Using prerequisite data may lead to stronger predictions, or at the very least provide indications of strong and weak skill relationships. Knowing more information about such skill relationships could provide better indications of when performance history is most useful as a predictor.

4. ACKNOWLEDGMENTS

We acknowledge funding for ASSISTments from the NSF (1316736, 1252297, 1109483, 1031398, 0742503, 1440753), the U.S. Dept. of Ed. GAANN (P200A120238), ONR's "STEM Grand Challenges," and IES (R305A120125, R305C100024).

5. REFERENCES

[1] J. E. Beck and K. min Chang. Identifiability: A fundamental problem of student modeling. In *User Modeling 2007*, pages 137–146. Springer Berlin Heidelberg, 2007.

[2] A. Botelho, H. Wan, and N. Heffernan. The prediction of student first response using prerequisite skills. In *Learning at Scale*, 2015.

[3] X. Xiong, S. Li, and J. E. Beck. Will you get it right next week: Predict delayed performance in enhanced its mastery cycle. In *FLAIRS Conference*, 2013.