Enhanced Retention Performance Modeling for Intelligent Tutoring System

N. Sharada
Research Scholar
Dept. of CS & SE
College of Engineering
Andhra University

M. Shashi
Professor
Dept. of CS & SE
College of Engineering
Andhra University

Xiaolu Xiong
Data Researcher
Worcester Polytechnic Institute
100 Institute Road, Worcester
MA 01609

ABSTRACT

The aim of Personalized Adaptive Scheduling System is not only to maximize learning, but also to minimize the tendency to forget. Empirical studies from psychology and cognitive theory of memory indicate that repetition with gradually increasing spacing interval is necessary to promote optimal long-term retention. While all individuals differ in their capacity to learn and retain information, the existing personalized systems attempt to model student retention performance based on learning capabilities only. The present study aims to investigate the extent to which individual differences in learning and forgetfulness help in predicting retention performance, which is essential for designing personalized retrieval practice schedules for a student. This study was conducted using data from Personalized Adaptive Scheduling System in ASSISTments, an adaptive mathematical tutor. The results illustrate the importance of student learning and forgetfulness features in predicting retention performance of a new skill, the proposed prediction model showed a significant improvement from an R^2 of 0.2975 with an existing baseline model to an R^2 value of 0.3989. In addition the newly identified features are used to predict the retention interval of a student, since ideal personalized retention schedules should be roughly equal to the retention interval of the student. The predictive accuracy of the linear regression model thus obtained was found to be statistically significant with an R^2 value of 0.413.

General Terms

Educational Data Mining, Intelligent Tutoring Systems

Keywords

Retention Performance, Personalized Adaptive Scheduling System, Automatic Reassessment and Relearning System.

1. INTRODUCTION

Intelligent Tutoring System offers assistance to the student while learning new skills and at the same time provides detailed evaluation of the student knowledge to the teacher. Over the past two decades, researchers in the field of Intelligent Tutoring Systems have focused on developing skill diagnosis models to determine in real time whether the student has mastered the skill. The probability that a student learned a skill is estimated based on the pattern of correct and incorrect responses given, through each opportunity to apply that skill. Bayesian Knowledge Tracing (BKT), for modeling student learning has been extensively used in the area of Intelligent Tutoring System [1]. The model computes the probability that a student knows a skill from four types of parameters namely initial knowledge, knowledge transition, probability of slip and guess during learning process. A lot of variations [2, 3, 4, 5] of the standard BKT models have been developed to improve performance. However, all these models typically assume that forgetting does not happen and that students retain knowledge after their initial mastery of the skill.

Y Wang and Heffernan [6] reported that simple modification of Standard Bayesian Knowledge Tracing can't handle forgetting and relearning and hence some other techniques need to be explored. Furthermore the performance observed during real time skill learning, may be mediated by rote learning or hints specific to the learning process rather than being indicative of any substantial knowledge or understanding. Long term retention of a skill is one of the three essential factors of durable and robust learning. This paper investigates the comprehensive mechanisms of predicting long-term retention performance. Interest in this issue has been triggered by two aspects. One is the idea, proposed by Pashler [7], that educational failures often reflect problems in retaining information over time, rather than in acquiring the information in the first place. The second trigger is the counterintuitive finding [8,9] that the number of correct and incorrect responses used in conventional methods for predicting student performance are actually not effective in predicting long-term retention.

Long term retention is often promoted by practicing the skill through repeated testing referred to as retrieval practice. Researchers [10] believe that repeated testing with gradually increasing spacing interval referred to as Expanding Spacing Retrieval Practice ought to promote long-term retention. The idea behind expanding retrieval practice is that students should practice retrieval sooner after learning a new skill and then gradually increase the spacing interval between successive retrieval attempts. Expanding Retrieval Practice is considered as a superior technique for promoting long term retention compared to Equally Spaced Retrieval Practice [11]. Education learning demand the student's capability to recall the knowledge they acquired over a long period of time. This is more important for subjects like mathematics where it is essential to master as well as transfer the knowledge on the pre-requisite skills for learning more complex skills. Although teachers recognize the importance of retrieval practice, the time demands of retesting the old concepts compete against the imperative to regularly introduce new concepts.

Automatic Reassessment and Relearning System a newly developed module in ASSISTments automatically assigns default retention tests at a spacing interval of 7-14-28-56 days after the student's initial mastery of the skill. ASSISTments is a non-profit web-based tutoring system for 4th through 10th grade mathematics. In Skill Building Problem sets, the students are given skill specific problems to solve until they demonstrate mastery of the skill by three consecutive correct responses. In tutoring systems, the mastery speed represents the total number of problems required to achieve the skill mastery status. The slower mastery speed refers to the more

number of attempts required to attain mastery status. The mastery speed represent two important aspects of student while learning a skill, first his/her aptitude to learn a new skill and second how well the student can retain the previous skill knowledge and apply that for new skill. Such a measure of mastery speed has shown to be successful in predicting the long term retention. Furthermore, it has been identified that students with slower mastery speeds had exhibited decline in the retention performance as the retention test intervals lengthened [12].

The expanded spacing time interval tailored to the needs of the student based on their mastery speeds was proposed to enhance retention performance. A Consequently improved version of ARRS known as Personalized Adaptive Scheduling System (PASS) [13] has been developed to schedule retention tests at various levels on 7-14-28-56 days based on the students individual mastery speed. When a student give correct responses he will be promoted to next level however, if he/she gives incorrect responses he will be scheduled for a lower level retention interval. Furthermore Wang et al., [14] performed student retention analysis and studied whether forgetting varies from student to student. All of this effort that has been expended in the quest to find the best predictors for planning retention interval schedules have yielded a varied degree of success.

2. PROBLEM STATEMENT

Educational failures at various levels are often attributed to the inability to retain the skills and knowledge that were once mastered due to the lack of properly planned spaced retrieval practices. The desire to find the best predictors for grouping students to schedule optimal expanding retention intervals is important for a number of reasons. This help the students in improving their retention performance, and thereby achieve robust learning, irrespective of the differences in their speed of learning. Additionally this guides the researchers in the Educational Research Community to model students' knowledge retention more accurately and thereby enhance retention performance prediction in Intelligent Tutoring System.

The mastery speed was shown to be successful in predicting student long term retention performance. Nevertheless, features such as individual differences in retention abilities among students, effectiveness of teaching learning process in the class, and skill characteristics were not modeled in the previous work. The goal of this paper was to further extend individualization and include student specific learning and forgetfulness features while building regression models for scheduling retention tests.

In this work, therefore, we seek to answer the following research questions:

1. Can a student's Individual retention attribute be used to construct a reliable prediction model of retention performance for a new skill?

2. Do class level features, such as class mastery speed which capture teaching learning process experienced by the student help in accurate prediction of retention performance?

3. Do students exhibit similar degree of retention across skills?

4. Can we precisely predict the retention interval of a student in order to design personalized repeating spacing interval retention tests?

3. METHODOLOGY

3.1 Dataset

The dataset used in this work is comprised of real-world data from Personalized Adaptive Scheduling System, a new feature in ASSISTments web based tutoring system. PASS enables the Automatic Reassessment and Relearning System to design personalized retention tests for each skill mastered, with gradually increasing spacing intervals, to help students relearn forgotten skills. The default reassessment setting in ARRS is 7-14-28-56 days, and this indicates each skill requires four levels of testing. The PASS alters the first level retention test schedule based on the individual mastery speed of the student ranging from 1-7 days. When a student passes the first level test he/she is promoted to a 14-day test, and from there onwards the interval schedules are same as in ARRS. When a student fails a retention test, a relearning assignment is assigned to the student and a reassessment schedule is continued from the previous level. The dataset derived from first level retention test performance from PASS is used for modeling retention test performance in the current research to test the stated hypothesis. The resulting dataset records the learning experience of 12238 unique students while mastering 154 mathematical skills described in 185904 data rows up to middle school level. Table 1 shows the sample of the dataset that was used for this study.

The studies from psychological and cognitive theories of memory, indicate that personalization of spaced retrieval schedules require estimates of an individual's retention interval. A good heuristic is to aim for having the length of the retrieval interval to be equal to the retention span of the student. However, existing adaptive-scheduling algorithms [14] base their predictions on observations from only that individual's knowledge level (mastery speed) of a particular skill. However the proposed approach is fundamentally data driven, and perform statistical inference to predict the retention performance of individuals by considering class features, skill features and individual student features in addition to the basic features like mastery speed, skill-id and problem difficulty. Hence we formed a hypothesis that the models retention performance prediction can be enhanced by incorporating individual differences in forgetfulness apart from the ease of learning.

To test the stated hypothesis, skill level features, class level features and student level features were constructed from the PASS dataset. These features are:

The skill level features represent skill nature information:
(1) Sk_id: the unique identifier for each skill, by modeling Sk_id as a factor the effect of each skill can be estimated.
(2) Sk_ms: measures the skill mastery speed and is computed from sum of the mastery speed of all the students by total number of students who mastered the skill.
(3) Sk_rt: measures the skill retention rate and is calculated from the number of correctly answered retention tests by the total number of retention tests taken for a specific skill.

Table 1. Sample dataset

St_id	Sk_id	Sk_ms	Sk_rt	St_ms	St_rt	Cl_ms	Cl_rt
220427	190	4.8	0.7	5	0.7	5.8	0.8
110541	209	3.6	0.9	3	1	6.8	0.8

The class level features represent the effectiveness of teaching – learning process experienced by the students in the class: (1) Cl_id: Represent the unique identifier for each class. By modeling Cl_id as a factor, we are estimating an overall effect of the classroom and the material being taught. (2) Cl_ms: measures the mastery speed of the class and is calculated from the sum of mastery speeds of all students by the number of students who reached mastery status in the class. (3) Cl_rt: measures the class retention rate and is calculated from the number of correctly answered retention tests by the total number of retention tests taken in the class.

The student level features represent student information: (1) St_id: Represent unique identifier for each student and by modeling St_id as factor, the individual differences in student learning and retention can be modeled. (2) St_ms: measures the overall mastery speed of student and is computed from the sum of mastery speeds of all the skills by the number of skills mastered by the student. (3) St_rt: measures overall retention performance of a student and is calculated as the total number of correct responses in a retention test by total number of retention tests taken by the student.

3.2 Regression Models

To investigate how student-level features, class-level features and skill attributes could impact prediction of student retention test performance, initially a base model is developed with fundamental features relating to mastery speed and the retention test problem easiness. The mastery speed is an estimate of student knowledge and is obtained from the number of attempts the student requires to reach mastery status for a skill. The impact of student-level, class-level and skill features are investigated separately, by developing three different models besides the base model. The model with skill features explore the impact of skill difficulty level and the nature of the skill attributes on the retention performance. The skill difficulty is captured by the skill mastery speed and skill nature is indicated by retention performance. The class feature model capture the effectiveness of the teaching-learning process experienced by the class which is represented by a class-id, class-mastery speed and class-retention. The student-level model represent student learning ability, retention ability through student mastery speed, student retention performance features.

4. EXPERIMENTAL RESULTS

In order to address the stated first three research questions, the linear regression models are fitted with data obtained from PASS after the first level retention test. The purpose of these models is to illustrate the effectiveness of using skill-level features, class-level features and student-level features in order to predict the retention performance of a student for a new skill. Table 2 provides the results for each of these models with the prediction performance being measured in terms of R^2 on the dataset.

Table2: Prediction performance of different models

Model	R^2
Base Model + Student Features	0.3989
Base Model + Class Features	0.3326
Base Model + Skill Nature	0.2976
Base Model	0.2975

The experimental results indicate that the base regression model with student specific speed of learning feature gives only a marginal model improvement. At the same time, improved models were obtained both for student features and class features. The prediction model constructed using student level features achieved a maximum of R^2 and hence clearly indicates that the variations in student capabilities for learning and retention are helpful in predicting retention performance for a new skill. The results also demonstrate that retention performance of a student is also a function of class characteristics represented by class learning and retention history. Furthermore, it is also noted that, there is no significant improvement in prediction of retention performance by extending the base model with features related to skill nature. This implies that the features considered in the base model have captured the skill characteristics.

Personalization of student retention schedule demands identifying the retention interval of the student. The ideal repeating spacing interval should be roughly equal to the retention interval of the student. The attempt to retrieve sooner after initial learning of a skill, ensures a high level of success, hence it is crucial to accurately estimate the first retrieval practice interval. In the previous works the retention interval or delay days after which the retention test has to be conducted was identified based on the student mastery speed using simple lookup table. In the current work, the models constructed with student and class features, for predicting retention performance outperformed the base model, hence these features are included for modeling delay days.

Linear Regression model is constructed with the data obtained from the first level retention test of PASS to address the fourth research question. The dataset obtained from PASS module comprises of an attribute "correct" which indicates whether a student was successful in retention test conducted after an estimated number of delay days once he masters the skill or not. The success in retention test endorses the estimation made for delay days reflecting the length of duration he can retain the mastered skill. However being unsuccessful in retention test disapproves the estimation made for delay days leading to unknown delay days for such case. Hence the rows corresponding to successful retention tests are included for the purpose of estimating the number of delay days.

The predictive accuracy of the linear regression model constructed for estimating delay days was found to be statistically significant with an R^2 value of 0.4136, p-value of 2.2e-16. The improvements in the model fit suggest that PASS should incorporate student features of learning and forgetfulness, class features as they have the potential to enhance the precision of predicting delay days.

Since outliers in the dataset could skew the results, data transformation is applied to minimize the effects of outliers in the dataset. The Skill Builder problem sets in ASSISTments have a default daily limit of 10 problems which the students can attempt in a day. Hence the mastery speed values greater than 10 in the dataset are transformed to 10 to fairly account student performance. The impact of this transformation is not significant, since more than 95% of the dataset has mastery speeds less than 10.

5. CONCLUSIONS

This study made three contributions. First, this paper provides a methodological evidence that the student's individual learning and retention are useful new features that can be modelled to predict retention performance with reasonable accuracy in Personalized Adaptive Scheduling Systems. The model thus developed represented meaningful information beyond with the provision of predictive accuracy. Second, this paper explored and identified that class level features are also effective in modeling retention. The class features are considered for analysis to account for the teaching learning process experienced by the students in the class. The design of Personalized Retention Schedules demands estimation of a student's individual memory strength for a particular skill. The ideal reassessment spacing interval should be roughly equal to the retention interval of the student. The third contribution of this paper is modeling student's retention interval by adopting newly identified features such as student features, class features apart from the mastery speed of the student. These findings emphasize the necessity for having the length of retention interval to be equal to the retention span of the student, while designing retention test schedules

6. FUTURE WORK

This research work proposes the usage of the average learning and retention performance of the mastered skills as a basis for modeling and predicting the retention performance of a new skill. It may be beneficial to consider other aspects such as hint usage, response time and knowledge estimation in terms of number of corrects in the retention tests conducted at various levels. Precise predictions may be achieved through the use of prerequisite skill structures and inter-skill relationships. Furthermore, the ability to predict retention performance may benefit from other models like Bayesian Knowledge Tracing. Also exploring the interference of the previously mastered similar skills on the retention performance of the new skills is an interesting future direction.

7. REFERENCES

[1] Corbett, A.T. and Anderson, J.R., 1994. Knowledge tracing: Modeling the acquisition of procedural knowledge. User modeling and user-adapted interaction, 4(4), pp.253-278.

[2] Pardos, Z. and Heffernan, N., 2010, June. Navigating the parameter space of Bayesian Knowledge Tracing models: Visualizations of the convergence of the Expectation Maximization algorithm. In Educational Data Mining 2010.

[3] Lee, J.I. and Brunskill, E., 2012. The Impact on Individualizing Student Models on Necessary Practice Opportunities. International Educational Data Mining Society.

[4] d Baker, R.S., Corbett, A.T. and Aleven, V., 2008, June. More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In International Conference on Intelligent Tutoring Systems (pp. 406-415). Springer Berlin Heidelberg.

[5] Yudelson, M.V., Koedinger, K.R. and Gordon, G.J., 2013, July. Individualized bayesian knowledge tracing models. In International Conference on Artificial Intelligence in Education (pp. 171-180). Springer Berlin Heidelberg.

[6] Wang, Y. and Heffernan, N., 2010, June. Towards modeling forgetting and relearning in ITS: preliminary analysis of ARRS data. In Educational Data Mining 2011.

[7] Pashler, H., Rohrer, D., Cepeda, N.J. and Carpenter, S.K., 2007. Enhancing learning and retarding forgetting: Choices and consequences. Psychonomic bulletin & review, 14(2), pp.187-193.

[8] Wang, Y. and Beck, J., 2012, June. Incorporating Factors Influencing Knowledge Retention into a Student Model. In Educational Data Mining 2012.

[9] Xiong, X., Li, S. and Beck, J.E., 2013, May. Will You Get It Right Next Week: Predict Delayed Performance in Enhanced ITS Mastery Cycle. In FLAIRS Conference.

[10] Landauer, T.K. and Bjork, R.A., 1978. Optimum rehearsal patterns and name learning. Practical aspects of memory, 1, pp.625-632.

[11] Kang, S.H., Lindsey, R.V., Mozer, M.C. and Pashler, H., 2014. Retrieval practice over the long term: Should spacing be expanding or equal-interval? Psychonomic bulletin & review, 21(6), pp.1544-1550.

[12] Xiong, X. and Beck, J.E., 2014, June. A study of exploring different schedules of spacing and retrieval interval on mathematics skills in ITS environment. In International Conference on Intelligent Tutoring Systems (pp. 504-509). Springer International Publishing.

[13] Xiong, X., Wang, Y. and Beck, J.B., 2015, March. Improving students' long-term retention performance: a study on personalized retention schedules. In Proceedings of the Fifth International Conference on Learning Analytics and Knowledge (pp. 325-329). ACM.

[14] Wang, Yutao, and Joseph E. Beck., 2012 Using Student Modeling to Estimate Student Knowledge Retention. International Educational Data Mining Society.