

A Statistical Approach to Student Modeling

Rahul Yedida, Ankush Kumar, Vima Rai, Srihari Joshi
Under the guidance of Ms. Kundhavai K. R.

Abstract

Educational data mining is a research area that has been gaining increasing interest in recent years. The subject deals with student modeling and test score predictions to advance education methodologies and develop reliable student models. Student modeling is primarily used in Intelligent Tutoring Systems (ITS), which are software that teach new concepts to students. Bayesian Knowledge Tracing (BKT) and various extensions to it have become increasingly popular choices for this task.

This project takes a statistical approach to the problem using various Hidden Markov Models (HMMs) and their extensions to understand what factors are important for student learning and interpret the resulting model to extract information about how students learn best. We will use the 2009-10 ASSISTment data, which has 401,756 records of student responses to problem and associated data. We will implement several underlying models proposed, from the early Bayesian Knowledge Tracing model, to more recently proposed models, such as the Intervention-BKT and the KAT models. We anticipate to get Root Mean Squared Error (RMSE) and Mean Absolute Errors (MAE) at levels similar to those obtained in literature.

1 Introduction

In recent years, an increasing focus of research in data mining has been focused on education. Problems related to this domain include predicting student performance, individualizing student models, and clustering [1].

Knowledge tracing is a technique of modeling a student's mastery of various concepts in coursework. Bayesian Knowledge Tracing (BKT) [2] is the most popular approach used in Intelligent Tutoring System (ITS) research [3]. Formally, the BKT is a two-state Hidden Markov Model (HMM), with

two states, *learned* and *unlearned*. These states correspond to whether or not a student has learned a particular concept.

Chen Lin and Min Chi proposed an extension to the BKT model, called the Intervention-BKT (I-BKT) [4] that incorporates student response time. Fundamentally, the I-BKT model is an Input Output Hidden Markov Model (IOHMM) [5]. Their model directly incorporated response time into the Bayesian framework to model student knowledge.

Recently, an increasing interest in the effect of the student’s affective, or emotional, states has driven research to investigate including this in the student model successfully. [6] [7]. These models use tools that discern the student’s affective states by recording the facial expressions throughout the duration of the session.

In this project, we build on top of the ASSISTments data [1], an ITS. We first compare the original BKT model with the I-BKT model. We then add affect to the model and compare performance of all the models. The ITS with affect will record the student’s facial expression and group it into one among a set of categories defined in the ASSISTments data.

2 Literature Review

Several extensions to the BKT have been proposed. Pardos and Heffernan [8] incorporated problem difficulty in the model and showed an improvement in performance on the ASSISTments dataset. Yudelson et al [9] demonstrated that parameterizing student learning speed improves the model as well.

Lin and Chi [4] added student response time to the BKT model, and called the resulting model, the Intervention-BKT (I-BKT) model. The I-BKT model is an Input Output Hidden Markov Model, and their paper showed that it outperformed the classic BKT model in predicting next-step performance.

Several Deep Neural Network (DNN) models have also been proposed for this task. Piech et al [10] showed that Recurrent Neural Networks (RNNs) show a major improvement in Area Under ROC Curve (AUC) over BKT and Item Response Theory (IRT) models. Lin and Chi [11] compared various RNN and Long Short-Term Memory (LSTM) models, and also incorporated target replication, a technique proposed by Lipton et al [12]. They showed that in general, RNN and LSTM models outperformed BKT models. However, deep neural networks (DNNs) have several well-known downsides. Perhaps the most notable of these is that DNNs are difficult to interpret as compared to simpler statistical models, and it can be impossible to determine why a trained DNN made a certain decision. For this reason, we prefer to use

the BKT model and its extensions instead of neural networks for building the ITS, since the formal definition of an HMM can explain the student model significantly better.

Research efforts have also been focused on incorporating student’s emotional state, or affect, into student models. For example, Schultz and Arroyo [7] combined the BKT model with the Hidden Markov Model-Item Response Theory (HMM-IRT) model proposed by Johns and Woolf [13], to create the Knowledge and Affect Tracing (KAT) model. The KAT model includes transition probabilities between *engaged* and *disengaged* states, which are affect states introduced in the HMM-IRT paper. Spaulding, Gordon, and Breazeal [6] instead used several affective states, including confused, bored, and smiling using a commercial affect-analysis tool called Affdex.

While the original BKT models suffers from some limitations, we believe that the extensions proposed provide a balance between good performance and ease of interpretability. For this reason, we use BKT and its extended models as the underlying engine for our ITS.

References

- [1] “Assistments 2009-10 skill-builder data.” <https://sites.google.com/site/assistmentsdata/home/assistment-2009-2010-data/skill-builder-data-2009-2010>.
- [2] A. T. Corbett and J. R. Anderson, “Knowledge tracing: Modeling the acquisition of procedural knowledge,” *User modeling and user-adapted interaction*, vol. 4, no. 4, pp. 253–278, 1994.
- [3] R. S. d Baker, A. T. Corbett, and V. Aleven, “More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing,” in *Intelligent Tutoring Systems*, pp. 406–415, Springer, 2008.
- [4] C. Lin and M. Chi, “Intervention-bkt: incorporating instructional interventions into bayesian knowledge tracing,” in *International Conference on Intelligent Tutoring Systems*, pp. 208–218, Springer, 2016.
- [5] S. Chiappa and S. Bengio, “Hmm and iohmm modeling of eeg rhythms for asynchronous bci systems,” tech. rep., IDIAP, 2003.
- [6] S. Spaulding, G. Gordon, and C. Breazeal, “Affect-aware student models for robot tutors,” in *Proceedings of the 2016 International Conference on*

- Autonomous Agents & Multiagent Systems*, pp. 864–872, International Foundation for Autonomous Agents and Multiagent Systems, 2016.
- [7] S. Schultz and I. Arroyo, “Tracing knowledge and engagement in parallel in an intelligent tutoring system,” in *Educational Data Mining 2014*, 2014.
 - [8] Z. A. Pardos and N. T. Heffernan, “Kt-idem: introducing item difficulty to the knowledge tracing model,” in *International Conference on User Modeling, Adaptation, and Personalization*, pp. 243–254, Springer, 2011.
 - [9] M. V. Yudelson, K. R. Koedinger, and G. J. Gordon, “Individualized bayesian knowledge tracing models,” in *International Conference on Artificial Intelligence in Education*, pp. 171–180, Springer, 2013.
 - [10] C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. J. Guibas, and J. Sohl-Dickstein, “Deep knowledge tracing,” in *Advances in Neural Information Processing Systems*, pp. 505–513, 2015.
 - [11] C. Lin and M. Chi, “A comparisons of bkt, rnn and lstm for learning gain prediction,” in *International Conference on Artificial Intelligence in Education*, pp. 536–539, Springer, 2017.
 - [12] Z. C. Lipton, D. C. Kale, C. Elkan, and R. Wetzell, “Learning to diagnose with lstm recurrent neural networks,” *arXiv preprint arXiv:1511.03677*, 2015.
 - [13] J. Johns and B. Woolf, “A dynamic mixture model to detect student motivation and proficiency,” 2006.