CSC591/791: Assignment 3

September 7, 2017

Due 9/19/2017 8:30am.

In this assignment you will gain familiarity with the BKT student model. You have been provided with a dataset in CSV format representing a series of students working on a multi-KC system. Your task is to implement a system that trains a per-student, per-KC BKT model and to report on their general success. You must also define a multi-KC success function by extending the equation covered in Corbett and Anderson Corbett and Anderson.

Dataset

The dataset has been provided as a CSV File `AssignmentData.csv`. The file contains 36,628 rows & 10 columns:

Student Anonymous student ID.

StepID Numeric id for each unique step.

Correct Binary indicator of whether the step was correct 1 = Yes 0 = No.

KC_1 - KC_27 Per-KC relevance columns. 1 if the Step is relevant to the KC 0 if not.

Model & Analysis

You should implement a basic version of the BKT model for each individual KC using the equations introduced in Lecture 6 and in Baker, Corbett, and Aleven You may *not* use an existing BKT library for the implementation.

You will then train the models on a per-student basis as described in the lecture. You may use static values for the slip, guess, train, and L_0 parameters. You are *not* required to apply expectation maximization or some other form of parameter learning. You should look at the parameter constraints presented in Lecture 7 and/or the required readings for suitable values. For each of the trained models you will also generate prediction scores via the correct performance equation included in Corbett and Anderson.

In addition to the individual models you will also design a multi-expert model that uses the individual per-KC models to predict the likelihood that students will succeed on a multi-KC step. This can be based upon the correct performance equation. This model should be structured to generate a valid prediction on single-KC problems as well. Having implemented the trained models you will then evaluate their *individual* success over all of the steps.

In your written report you should:

- Include a description of the code structure along with a high-level description of the implementation and use.
- Describe how you implemented the per-KC and multi-expert models including citing any sources for your parameter settings.
- Evaluate the relative performance of the per-KC and multi-expert models. Which one performs better? You should base your comparison on some quantitative measure of performance and support it with appropriate visual representations.

As always all code should be readable, well structured, and execute without errors for grading.

Submission

You should submit two separate files:

- Your BKT code for processing the dataset (`YourName_code.zip`).
- Your written report (`YourName_report.pdf`).

References

Baker, Ryan S. J. d., Albert T. Corbett, and Vincent Aleven (2008). "More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing". In: *Intelligent Tutoring Systems: 9th International Conference, ITS 2008, Montreal, Canada, June 23-27, 2008 Proceedings*. Ed. by Beverley P. Woolf et al. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 406–415. ISBN: 978-3-540-69132-7. DOI: 10.1007/978-3-540-69132-7_44. URL: https://doi.org/10.1007/978-3-540-69132-7_44.

Corbett, Albert T. and John R. Anderson (1995). "Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge". In: *User Modeling and User-Adapted Interaction* 4, pp. 253 –278.