Udacity capstone proposal

Anoop Aroor

December 2016

1 Domain Background

Adaptive learning systems in education are interactive teaching devices that can adapt to the learning needs of students. The system can adopt the presentation of the materials (for example the sequence of questions to be asked) based on their understanding of the topic measured from their response to questions, task or student background. When costs of education keeps rising, adaptive learning systems promise to offer a scalable yet customized mode of instruction thus reducing cost while simultaneously improving the quality of instruction. It could be used in a classroom setting or in large online education platforms like Coursera, Edx, Udacity or Khan Academy.

2 Problem statement

Knowledge Tracing is a problem where the goal is for a machine to model the student's knowledge about a subject as they work through the coursework so that the machine can make predictions on the students performance on future assignments.

More formally given a sequence of student interactions $x_0, x_1, .., x_t$, where each x_i is a record of an interaction at time t, one has to predict aspects of the next interaction x_{t+1}. Each interaction is defined as follows $x_t = (s, q, r)$, where s is a student, q is the question the student is attempting to answer, r is the correctness of the student response to that question. The goal of this project is to predict the likelihood of a student correctly answering a particular question at time $t + 1$, given a sequence of student interactions upto t.

This is a popular problem in the field of adaptive learning systems and a recent paper compared three different approaches and compared their performance to generate a benchmark performance[2].

3 Datasets

ASSISTment 2010 dataset and Bridge to Algebra 2006-2007 dataset from KDD cup 2010 are the two public datasets that will be used for this project. ASSISTment dataset, after pre-processing, consistes of 346,740 interactions for 4,097 users on 26,684 items arising from 815 templates and 112 skills. The overall percent correct was 64.54.

https://sites.google.com/site/assistmentsdata/home/assistment-2009-2010-data/skill-builder-data-2009-2010

http://www.kdd.org/kdd-cup/view/kdd-cup-2010-student-performance-evaluation/Data

4 Solution statement

Idea 1: In a recent attempt called deep knowledge tracing [1], a LSTM with a single hidden layer was trained predict the student response. They claimed that the model was better than the state of the art. However it turned out as described by the another paper, that LSTM model being the state of the art was a not correct and was due to a bug in data cleaning[2]. Here my goal is to give the LSTM approach another attempt but train a multi layer LSTM instead of a single layer.

Idea 2: Attempt to use probabilistic graphical models to model the structure and learning behavior of the student. The use of template models could be explored.

5 Benchmark Model

Currently a model based on Item response theory called HIRT (hierarchical item response theory) is the state of the art in prediction for student responses[2].

6 Evaluation metrics

The standard way to evaluate the model is to split the data into training and testing populations. Each model is trained on the training population and the model parameters that are not student-level (item parameters for IRT-based models, weights for neural networks are frozen. Then for each time less than t in each testing student's history, we train the student-level parameters in the model on the interactions of the student history and allow it to compute the probability that the t'th response is correct. This process mirrors the practical task that must be completed by an interactive teaching system.

Accuracy of the model on the test data is computed as the percent of responses in which the correctness coincides with the probability being greater than 0.5. AUC is the Area Under the ROC Curve of the probability of correctness for each response. The benchmark AUC for ASSISTment dataset is 0.7292 by HIRT model.

References

[1] Chris Piech et al. "Deep knowledge tracing". In: *Advances in Neural Information Processing Systems*. 2015, pp. 505–513.

[2] Kevin H Wilson et al. "Back to the Basics: Bayesian extensions of IRT outperform neural networks for proficiency estimation". In: *arXiv preprint arXiv:1604.02336* (2016).