Machine Teacher: Reinforcement Learning for Teaching

Anyi Rao
Email: anyirao@outlook.com

Abstract: We seek to use machine learning to teach or give suggestion to teachers. They are able to automatically identify which items to include in a set of curriculum, and how to adaptively select these items, in order to maximize student performance on some specified set of learning objectives.
Keywords: Reinforcement Learning, Deep Learning, Teaching

1 Introduction

It's often the case that novice teachers fail their job even if they get full prepared and hold good teaching materials and detailed teaching plan. After the class, maybe only a few of students are able to get master of what the teachers taught. When the teachers look at the learning objectives list (i.e. a list of capabilities the student should have after taking the course) they are very upset and confused

Figure 1: Teacher Confused

The reason is mainly attribute to that they don't know how to adjust their teaching stuff according to their course learning objectives. It makes an enormous difference to the effectiveness and efficiency in which students learn, when changing the content (slides, videos, quizzes, etc) and the way in which material is taught.

Is there good solution or can we help these poor teachers if we resort to machines?

2 Related Work

To finish these tasks, we need to first know how to evaluate our solution. A natural way to measure the course effectiveness is to evaluate how well students perform after class, i.e. what is the specified input set of learning objectives they reach?

Existing approaches to addressing this often involve researchers performing extensive data analysis to identify places where the existing models or curriculum can be revised in order to improve student learning [4]. Another common approach is to use standard experimental design to compare multiple ways to teach the material, but such work typically requires a very large number of subjects to explore many conditions.

We don't need to fully quantify the effectiveness of all possible approaches, in trying to find the best way to teach some material. Recent machine learning approaches such as using multi-armed bandits optimization to identify the best of a finite set of teaching conditions, or treating the problem as function optimization using Bayesian optimization (BO) to find the maxima of a latent function that describes how adjusting the (continuous) parameters that prescribe how to teach impacts student outcomes in cognitive task learning [2].

3 Model

3.1 Bayesian Optimization

Our goal is to automatically identify which items to include in a set of curriculum, and how to adaptively select these items to maximize student performance on some specified set of learning objectives. In many situations teachers desire both effective and efficient teaching approaches. Therefore, we need a combination of lots of measures for a particular teaching way.

We can view this general challenge as an instance of finding the parameter values (i.e. how to choose a teaching items) that maximize a function value (student performance on learning objectives with the specified parameters). The bad news is that there are a huge number of possible parameter values to try.

Bayesian optimization is a popular machine learning approach for finding the optimal value of a function when evaluating the function at a particular set of parameters is costly, and we wish to minimize the number of parameters tried to find the optimal function value.

This has practical use in education, where each function evaluation corresponds to providing one real student with a particular instructional curriculum or approach to selecting that curriculum, and measuring the resulting outcome value.

3.2 Bayesian Knowledge Tracing

Bayesian Knowledge Tracing (BKT) is a popular model of student learning [1]. We are able to use a simple approach for mastery teaching while adding a threshold parameter. This corresponds to a practical use: when the BKT estimate of the probability that the student has mastered the skill exceeds the threshold, we can stop teaching the skill and move to the nest one.

However, the greatest problem is we don't have student data. It isn't clear how to set the parameters of a BKT mastery teaching policy. Even given data, existing approaches tend to use MLM (maximum likelihood methods) to fit the BKT parameters and then hand set the threshold parameter.

BKT only approximately models student learning, and it may be beneficial to skip setting parameters and to initialize parameters and directly optimize all parameters. For each skill, we directly find the BKT policy π that maximizes the function value f(π), where p_π, s is the normalized post test score for this given skill, p_π is the normalized post test score over all skills (with 0.7 as desired passing score), and l_π, s is the number of practice problems given to the student for this skill (subscript π means following policy π). This objective encourages policies where the student quickly does well broadly and on this skill.

$$f(\pi) = \frac{p_{\pi,s} + I(p_\pi \geq T)}{\sqrt[2]{l_{\pi,s} + 1}} \tag{1}$$

3.3 Current situation

Up to now we need to use our machine teachers (trained models) one by one student since the data is continuously coming and the machine learning process is a kind of reinforcement learning.

We use BO to efficiently search over the space of each skill mastery policies. When a new student starts to use our machine teacher (i.e. our trained models), we first use BO with the popular expected improvement acquisition function to identify a good next set of policy parameters to try for each skill. These parameters define a mastery policy that is used to teach a student and determine when to halt and present the student with the post test. Given the student's results on the post test, we compute the above $f(\pi)$ value that is used by BO to select a new set of parameters π to try for the next student.

4 Conclusion

BO is able to quickly find good parameters, and can be used as a tool to help teachers identify misalignment between their present stuff to their desired learning objectives. Teachers can adjust their teaching things according to our machine teachers.

5 Future Work

The state-of-art group (R Antonova, et al.) verified that Bayesian Optimization can quickly find good parameters by using an experiment with a histogram tutoring system. However, it's not easy to define objective function. Although BKT is a popular student learning model, it may not present true desired intent. And I think model-free learning may be an alternative method. It's more difficult while relies less on the complex student learning process.

And it's ideal to use TensorFlow to do this job since TensorFlow's powerful visualization and model parallelism. These will be very helpful because we want to apply our model to more than one student, and visualization is always better than tedious code.

Although there is still a long distance between this work to developing a excellent self-optimizing tutoring system, it means there is a huge treasure in front of us and will be of great fun if I can work with your Machine Learning Group using TensorFlow.

References

[1] A. Corbett and J. Anderson. 1994. Knowledge Tracing: Modeling the acquisition of procedural knowledge. User Model User Inter 4 (4): 253-278.

[2] R. Lindsay, M. Mozer, W. Huggins and H. Pashler. 2013. Optimizing Instructional Policies. Proc of Neural Info Proc Sys 2778-2786.

[3] Antonova R, Runde J, Lee M H, et al. Automatically Learning to Teach to the Learning Objectives[C]// Proceedings of the Third (2016) ACM Conference on Learning @ Scale. ACM, 2016:317-320.

[4] K.Koedinger, J. Stamper, E. McLaughlin and T. Nixon. 2013. Using data-driven discovery of better student models to improve student learning. Proc of AI Educ 421-430.