Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/37/20230477

Investigation of recent advances in exercise-enhanced knowledge tracing models

Yilin Zheng

Computer Science and Engineering Department, School of Engineering, University of Connecticut, Storrs, CT, 06901, United States

yilin.zheng@uconn.edu

Abstract. Exercising plays a significant role in Knowledge Tracing (KT), however, the majority of KT models struggle to effectively extract the abundant information embedded within students' exercise histories, leading to a prevalent issue of information erosion. Hence, it is vital for exploring methods to mine the data from the exercises to better forecast students' future performance. This paper is a review that aims to discuss recent improvements in KT models on mining information existing in exercises. This paper will briefly discuss DKT, a popular approach to knowledge tracing, and its limitation on mining information from exercises. And then, this review will introduce four exercise-enhanced knowledge tracing models including EERNN, EKT, HGKT and Concept-Aware Deep Knowledge Tracing. EERNN and EKT can track students' knowledge acquisition by using vectors to represent knowledge concepts. HGKT goes a step further by examining the interconnectedness among different exercises based on EERNN and EKT. Concept-Aware DKT is an improvement of DKVMN, the other approach of KT, by considering the influence of knowledge concepts and corresponding exercises in more detail. Finally, the applications of exercise-enhanced KT models will be covered.

Keywords: knowledge tracing, deep learning, exercise.

1. Introduction

Knowledge tracing constitutes a methodological approach aimed at modeling a learner's cognitive proficiency based on their interactions and engagement with learning materials. By leveraging this technique, it becomes feasible to predict a student's prospective academic performance during the course of their learning journey [1, 2]. The area of education can benefit through the utilization of this technique, especially for, intelligent tutoring systems (ITSs). ITSs can provide more individualized service and feedback for learners by determining the level of knowledge held by the pupils. For instance, in circumstances where the ITS discerns a student's struggle in comprehending novel concepts, it can promptly intervene by providing supplementary assistance or tutoring interventions, facilitating the consolidation of newly acquired knowledge.

Corbett et al. first applied Knowledge Component (KC) to evaluate the current knowledge of a learner [2]. A KC is a description of structure to finish a set of tasks and can connect with other KCs, usually including concepts, skills, principles and schemata [2, 3]. The prerequisites for applying and the student's response to a task decide the taxonomy of KCs. KCs with varied reactions under variable situations are the most challenging.; otherwise, the simplest KCs only take account of constant responses and conditions [3].

Traditionally, Bayesian Knowledge Tracing (BKT) is one approach for KT [4]. Probabilistic graphical models are utilized in BKT to predict the probability that the likelihood of students mastering a skill [4]. BKT is one of the applications of Bayes's theorem that calculate the conditional probability of an event A occurring if an event B occurs i.e., P(A|B) [3, 4]. Except for the standard BKT, there are many extensions. For example, the individualized BKT model takes account of students' individuality to weigh each student; the dynamic BKT model is considered as a remedy for the case of multiple skills.

As the increasing of number of educational large-scale datasets, Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNNs) are applied in Deep Knowledge Tracing (DKT), which has increasingly gained popularity as an approach for knowledge tracing. [1, 4, 5]. Compared with BKT, DKT is easily affected by overfitting on small-scale datasets; however, it has more advantages in dealing with large-scale datasets since DKT is good at modeling more intricate connections between the student learning interaction records and the correctness log-odds, especially in predicting student's sequential progress through the material [1].

Although DKT is powerful in solving the knowledge tracing problem, it still has some limitations. On the forecast of KCs' performance over time, DKT cannot remain constant. [6]. Besides, it fails to model the relationships between KCs [4]. The vast majority of DKT strategies overlook the abundance of information inherent within the exercises themselves and only treat exercise history as sequence of knowledge [7]; meanwhile, these approaches also ignore the prior relations between exercises [8].

Exercise is one of the crucial elements in the knowledge tracking process., constituting an integral facet of learner engagement within AI-driven educational frameworks. It is not only used to practice learners' skills or knowledge, but also a way to obtain learners' current performance helping Artificial Intelligence (AI) educational applications analyze learners' future performance. Due to the importance of exercise and its latent rich information, it is necessary to explore how to mine the information existing in exercise and how to use that information in knowledge tracing, Therefore, this paper aims to review recent development of exercise-enhanced knowledge tracing.

In the method section, this paper will discuss several KT models that focus on mining data from students' exercise history. And then, this paper will talk about the applications of exercise-enhanced KT models in the section of the discussion.

2. Method

2.1. DKT

DKT illustrates the pupil's mater level by leveraging recurrent neural networks as learning skills or knowledge [2, 4, 9]. In DKT, the term "deep" pertains to the recurrent structure of the network and its capacity to retain information over time; it uses vectors of artificial neurons to represent the hidden knowledge states and enables the learning of the hidden variable representation of student knowledge directly from data [9]. When a student finishes a homework or assignment, DKT anticipates forthcoming student performance by leveraging historical temporal data, employing LSTM units for determining the optimal preservation of prior information and its amalgamation with contemporary input [10]. The data representation is one of the differences between DKT and BKT. BKT [3, 4] assumes that skill has two states (known or unknown) and uses four parameters for model definition: K_0 is a likelihood that a student already knew a skill before learning; K_i is the probability a student acquires a skill during learning; Guess is the probability a student guesses correctly despite not knowing the skill; Slip is the likelihood that a pupil will make a mistake despite having learned the skill. However, in DKT, information about learner's performance on a particular task that relates to the skills they are currently practicing is exploited as input of RNNs; a single 1 in vector (others are all zeros) is used to indicate which knowledge is involved and whether the student gives right response [9]. Empirical findings substantiate the overall superior efficacy of DKT compared to BKT in educational performance evaluations.

DOI: 10.54254/2755-2721/37/20230477

However, on utilizing the information existing in the exercise, numerous approaches (e.g., deep learning and BKT) merely exploit the performance history of students on each knowledge to predict students' knowledge state so that two exercises (e_1 and e_2) are said to be identical if two exercises have the same knowledge concept labeled on both of them, they are said to be identical [7]. However, these two exercises labeled with the identical knowledge concept may also exist differences: one exercise may be harder to be answered than another one. That means, if one student solves e_1 and the other solves e_2 , these approaches are unable to differentiate the knowledge acquisition of those two students. Hence, these approaches do not mine information of exercise well, existing information loss problem.

2.2. EERNN and EKT

In the pursuit of more efficacious utilization of exercise-derived information, Liu et al. [7] proposed a framework called Exercise-Enhanced Recurrent Neural Network (EERNN) to explicitly trace learners' knowledge acquisition. For each student, EERNN utilizes a single integrated hidden vector for condensing and tracking knowledge states across all concepts by *Exercise Embedding* and *Student Embedding*. *Exercise Embedding* is a modification of LSTM, using the exercise text content e_i as input. Here, e_i is a list with M words, denoted as $e_i=\{w_1,w_2,\ldots,w_M\}$. EERNN can use a vector to learn each exercise e_i 's semantic representation x_i , based on its associated exercise contents, via *Exercise Embedding*, automatically differentiating the differences between exercises based on the characteristics of exercises, such as difficulty. After getting the semantic representation x_i, *Student Embedding* takes the student's exercising process as input, denoted as $s=\{(x_1,r_1),(x_2,r_2),\ldots,(x_T,r_T)\}$. x_t is the semantic representation corresponding to each exercise. r_t is a binary variable representing the rightness of the corresponding exercise (right response is 1; wrong response is 0). *Student Embedding* will model the exercising process to learn the hidden representations of students at various exercise steps by considering students' performance on exercise records.

Although EERNN can forecast how well students will perform on upcoming exercises., it merely uses one integrated hidden vector to represent all knowledge concepts to summarize the knowledge states of the student, rather than clearly representing the degree to which the student has mastered a specific knowledge concept. To address the issue, Liu et al. [7] did modifications to EERNN and proposed an upgrade framework called Exercise-aware Knowledge Tracing (EKT). Instead of using one integrated vector, EKT incorporates the information of multiple vectors into a matrix. Each vector indicates a knowledge concept or knowledge component. After the process of Exercise Embedding, a new component, Knowledge Embedding, is introduced in EKT. Knowledge Embedding takes the sequence of exercise content, corresponding knowledge concept, and the rightness as input to calculate the impact weight that quantitatively describes how each exercise and its corresponding knowledge concept improve student states. EKT can monitor knowledge states of students under the circumstance of numerous knowledge concepts by including the impact weight into Student Embedding. Moreover, both EERNN and EKT have Markov property variants (EERNNM and EKTM) and attention mechanism variants (EERNNA and EKTA).

2.3. HGKT

Tong et al. focused on the hierarchical relationship between exercises and proposed Hierarchical Graph Knowledge Tracing (HGKT) architecture for the prediction of students' knowledge levels [8]. To illustrate the associations between exercises, this framework makes use of two hierarchical graphs (*Direct support graph* and *indirect support graph*). *Direct support graph* is used to link exercises whose solutions have strong relations. For example, assume Exercise e_1 is strongly supported by Exercise e_2, meaning there is a high probability to solving e_2 if students do e_1 correct. Problem schema is linked by *indirect support graph*. The concept of problem schema is introduced by Tong et al, which refer to the group of exercise with similar solutions [8]. The framework assumes that each exercise is associated with a concept of knowledge and a problem schema. It consists of two subsystems. The first system utilizes *direct support graph* to cluster exercises through Bert [8], generating problem schema and

DOI: 10.54254/2755-2721/37/20230477

corresponding hierarchical graphs automatically. The second system is fed with these hierarchical graphs generated from the first system and employed RNN to trace students' states.

2.4. DKVMN and its improvement

Dynamic Key-Value Memory Network (DKVMN) is another traditional design for DKT. For saving knowledge concepts and altering student mastery levels for knowledge concepts, it relies on *key* (static matrix) and *value* (dynamic matrix) respectively [11]. The model predicts students' knowledge state by reading value which is calculated by the correlation weight between *key* and exercises. Ai et al. improved DKVMN, considered the mapping relationship between exercise and concept, and built a knowledge tracing model called Concept-Aware DKT [12]. Specifically, when calculating the correlation weight between *key* and exercises, Concept-Aware DKT model only takes account into exercise and its corresponding knowledge concepts instead of all knowledge concepts stored in *key*; irrelevant knowledge concepts will be weighted as zeros. Besides, in the process of updating value, the student's time cost in doing exercises is also considered as a factor that affects the student's problem-solving skills, which is also neglected by DKVMN.

3. Application and discussion

3.1. Exercise recommendation system

Exercise recommendation system is a system designed to provide recommend relevant exercises for learners according to their knowledge states and interaction. In online learning platforms, exercising is one of the most important steps for students; however, the selection of exercises also bears weight, as exercises that are excessively challenging or overly simple might not provide optimal assistance to students [12, 13]. If a learning platform enables to provide exercises or learning resources that are the most suitable for students' current knowledge states, students will make the most progress. Most learning recommendation systems only consider short-term rewards or broadly recommend learning resources to students, but those approaches are not helpful for learners due to the lack of personalization and long-term consideration. Exercise-enhanced KT models can provide more specific targets for learning recommendation systems. The exercise recommendation systems can apply those exercise-enhanced KT models discussed above to mine information existing in students' exercising history maximumly and employ that information to recommend an optimal set of exercises for students by considering exercises' difficulty level, students' current knowledge states and long-term reward. For instance, Ai et al. [12] proposed an exercise recommendation system that introduces Concept-Aware DKT model to map the relationship between knowledge concepts and exercise and recommends a series of targeted exercises for students' weak knowledge concepts on math. Thus, exercise-enhanced KT models are beneficial for the improvement of personalization and long-term engagement of exercise recommendation systems.

3.2. Adaptive learning

According to a learner's reaction, adaptive learning is an educational process that continuously modifies the degree or style of instruction [13]. In traditional education, it is challenging for teachers to provide each student with individual assignments or homework based on the student's performance in classes, but introducing KT models in the learning process may be able to address this problem. Dissimilar to exercise recommendation systems that merely recommend helpful exercises for students and passively rely on autonomous learners, adaptive learning is more motivated. It can design customized learning plans for each student, dynamically adjust plans as progress, and remind students to implement exercises on time. Compared with other KT models, exercise-enhanced KT models have advantages in knowledge acquisition tracking that enable quantitively measuring the extent of students' master level to each knowledge concept. Measuring understanding of each knowledge concept is highly valuable because it can inspire students to actively participate in the learning plan enhancing its effectiveness [7]. Furthermore, exercise-enhanced KT models can be applied to evaluate the quality of learning resources.

Based on students' performance and feedback from exercises over some time, the system of adaptive learning can evaluate whether current learning materials are helpful for learners. If a learner is still weak in a certain knowledge concept, the system of adaptive learning can actively change to other learning materials to guarantee the learner's effectiveness.

4. Conclusion

This review introduced several proposed KT models that enhance the ability of mining information from students' exercise history records. First of all, DKT and its limitation were briefly discussed. And then, this review introduced several exercise-enhanced KT models. Compared with the original DKT, EERNN and EKT can track students' knowledge acquisition by using vectors to represent knowledge concepts. HGKT goes a step further by examining the interconnectedness among different exercises. Concept-Aware DKT is an improvement of DKVMN by considering the influence of knowledge concepts and corresponding exercises in more detail. In practice, applying those KT models is capable to provide more individualized services for each learner based on their current knowledge states. Eventually, two classic applications of KT models, exercise recommendation system and adaptive learning, were mentioned. Using exercise-enhanced KT models in those scenarios can facilitate making a dynamic study plan, recommending more targeted exercises, and boosting the learning effectiveness of students.

References

[1] Gervet T Koedinger K Schneider J and Mitchell T 2020 When is Deep Learning the Best Approach to Knowledge Tracing? (Journal of Educational Data Mining 12(3) 31-54)
[2] Liu Z Chen J and Luo W 2023 Recent Advances on Deep Learning based Knowledge Tracing (Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining 1295-1296)
[3] Pelánek R 2017 Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques (User Modeling and User-Adapted Interaction 27 313-350)
[4] Abdelrahman G Wang Q and Nunes B 2023 Knowledge Tracing: A Survey (ACM Computing Surveys 55(11) 1-37)
[5] Song X Li J Tang Y Zhao T Chen Y and Guan Z 2021 JKT: A joint graph convolutional network based Deep Knowledge Tracing (Information Sciences 580 510-523)
[6] Yeung C K and Yeung D Y 2018 Addressing two problems in deep knowledge tracing via prediction-consistent regularization (Proceedings of the fifth annual ACM conference on learning at scale 1-10)
[7] Liu Q Huang Z Yin Y Chen E Xiong H Su Y and Hu G 2019 Ekt: Exercise-aware knowledge tracing for student performance prediction (IEEE Transactions on Knowledge and Data Engineering 33(1) 100-115)
[8] Tong H Wang Z Zhou Y Tong S Han W and Liu Q 2020 HGKT: Introducing hierarchical exercise graph for knowledge tracing (arXiv preprint arXiv:2006.16915)
[9] Xiong X Zhao S Van Inwegen E G and Beck J E 2016 Going deeper with deep knowledge tracing (International Educational Data Mining Society)
[10] Zhang L Xiong X Zhao S Botelho A and Heffernan N T 2017 Incorporating rich features into deep knowledge tracing. (Proceedings of the fourth (2017) ACM conference on learning@ scale 169-172)
[11] Zhang J Shi X King I and Yeung D Y 2017 Dynamic key-value memory networks for knowledge tracing. (Proceedings of the 26th international conference on World Wide Web 765-774)
[12] Ai F Chen Y Guo Y Zhao Y Wang Z Fu G and Wang G 2019 Concept-aware deep knowledge tracing and exercise recommendation in an online learning system (International Educational Data Mining Society)
[13] Liu Q Shen S Huang Z Chen E and Zheng Y 2021 A survey of knowledge tracing (arXiv preprint arXiv:2105.15106)