Recent Advances on Deep Learning based Knowledge Tracing

Zitao Liu
Guangdong Institute of Smart Education, Jinan University
Guangzhou, China
liuzitao@jnu.edu.cn

Jiahao Chen
TAL Education Group
Beijing, China
chenjiahao@tal.com

Weiqi Luo
Guangdong Institute of Smart Education, Jinan University
Guangzhou, China
lwq@jnu.edu.cn

ABSTRACT

Knowledge tracing (KT) is the task of using students' historical learning interaction data to model their knowledge mastery over time so as to make predictions on their future interaction performance. Recently, remarkable progress has been made of using various deep learning techniques to solve the KT problem. However, the success behind deep learning based knowledge tracing (DLKT) approaches is still left somewhat unknown and proper measurement and analysis of these DLKT approaches remain a challenge.

In this talk, we will comprehensively review recent developments of applying state-of-the-art deep learning approaches in KT problems, with a focus on those real-world educational data. Beyond introducing the recent advances of various DLKT models, we will discuss how to guarantee valid comparisons across DLKT methods via thorough evaluations on several publicly available datasets. More specifically, we will talk about (1) KT related psychometric theories; (2) the general DLKT modeling framework that covers recently developed DLKT approaches from different categories; (3) the general DLKT benchmark that allows existing approaches comparable on public KT datasets; (4) the broad application of algorithmic assessment and personalized feedback. Participants will learn about recent trends and emerging challenges in this topic, representative tools and learning resources to obtain ready-to-use models, and how related models and techniques benefit real-world KT applications.

CCS CONCEPTS

• **Information systems** → **Data mining**; *Web mining*; • **Applied computing** → **Computer-assisted instruction**; **E-learning**.

KEYWORDS

Knowledge tracing; cognitive diagnosis; deep learning; psychometric theory; AI in education

ACM Reference Format:
Zitao Liu, Jiahao Chen, and Weiqi Luo. 2023. Recent Advances on Deep Learning based Knowledge Tracing. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining (WSDM '23), February 27-March 3, 2023, Singapore, Singapore.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3539597.3575790

WSDM '23, February 27-March 3, 2023, Singapore, Singapore

ACM ISBN 978-1-4503-9407-9/23/02.
https://doi.org/10.1145/3539597.3575790

1 MOTIVATION

Knowledge tracing (KT) is the task of *using students' historical learning interaction data to model their knowledge mastery over time so as to make predictions on their future interaction performance*, shown in Figure 1. Such predictive capabilities can potentially help students learn better and faster when paired with high-quality learning materials and instructions, which is crucial for building next-generation smart and personalized education.

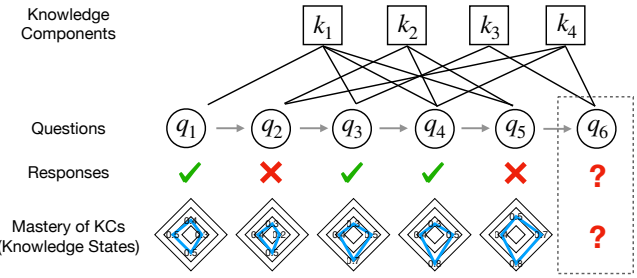

Figure 1: A graphical illustration of knowledge tracing.

The KT related research has been studied since 1990s where Corbett and Anderson, to the best of our knowledge, were the first to estimate students' current knowledge with regard to each individual knowledge component (KC) [5]. A KC is a description of a mental structure or process that a learner uses, alone or in combination with other KCs, to accomplish steps in a task or a problem[1]. Since then, many attempts have been made to solve the KT problem, such as probabilistic graphical models [7] and factor analysis based models [2, 8, 17]. Recently, due to the rapid advances of deep neural networks, deep learning based knowledge tracing (DLKT) models have become the de facto KT framework for modeling students' mastery of KCs [3, 4, 6, 10, 12, 13, 15, 16, 18].

Although DLKT approaches have constituted new paradigms of the KT problem [6, 14, 16, 19] and achieved promising results, recent studies [9, 10, 16] seem to resemble each other with very limited nuances from the methodological perspective. Most existing work only provides coarse evaluation and both the contributing factors leading to the success of DLKT and how the DLKT models perform in the real-world educational contexts still remain somewhat unknown. Furthermore, evaluations of existing DLKT work are not standardized and reported AUC results of the same approach on the same dataset vary surprisingly from 0.709 to 0.86 [1, 11]. Therefore, there is a substantial need for a systematical and comprehensive review about recent advances on the deep learning based knowledge tracing approaches. This will benefit both researchers and

[1] A KC is a generalization of everyday terms like concept, principle, fact, or skill.

practitioners to (1) be capable of differentiating advantages and disadvantages of the DLKT algorithms in real-world educational contexts; and (2) be able to evaluate their proposed approaches against a wide range of state-of-the-art methods on both publicly available and private datasets.

2 SOCIAL IMPACT

Recent years have witnessed growing efforts from AI research community devoted to advancing our education. It is a vibrant multi-disciplinary field of increasing importance and with extraordinary potential. AI is becoming a valuable tool for teachers in many ways and teachers are able to get insights from results of AI driven applications. Although it is still in the early stage, promising results have been achieved in solving various critical problems in education. The potential KT based AI in education applications is able to free up time for teachers, ensure students receive the timely assistance, and allow students to learn independently. It very well may help us to rethink and restructure how we approach compulsory education in the 21st century in a way that is beneficial to all parties.

3 INTENDED AUDIENCE

All the researchers and practitioners are welcome. The audiences are assumed to have basic knowledge in data mining and machine learning. Specially, those who have devoted to advanced methodologies in natural language processing and machine learning, etc., will be encouraged to participate to learn the new challenges when applying real-world educational scenarios. This tutorial will shed light on the opportunities of building successful DLKT models with the recent advanced algorithms of learning and encourages AI researchers from different domains to understand and work together to tackle some of the most exciting and challenging KT and AIED problems. This tutorial aims to be composed of a good balance between the introductory and advanced material (50% for beginners and 50% for intermediate and advanced).

4 BIOGRAPHY

Zitao Liu is the Dean of Guangdong Institute of Smart Education at Jinan University. His research is in the area of Artificial Intelligence (AI) with its application in education. He has published over 70 papers in highly ranked conference proceedings, such as NeurIPS, AAAI, WWW, AIED, etc. and his applied research has resulted in more than 30 technology transfer and patents. Zitao serves as the executive committee of the International AI in Education Society and Area Chair and Senior PC for many top conferences. He won the 1st/2nd places at over 10 international academic challenges and competitions including NeurIPS 2020 education challenge and EMNLP 2020 ClariQ challenge. Before joining Jinan University, Zitao was the Head of Engineering, ThinkAcademy at TAL Education Group (NYSE:TAL) and was a senior research scientist at Pinterest. Zitao received his Ph.D degree in Computer Science from University of Pittsburgh.

5 COMPANY PORTRAIT

TAL Education Group is a smart learning solutions provider in China. The acronym "TAL" stands for "Tomorrow Advancing Life", which reflects our vision to promote top learning opportunities for students through both high-quality teaching and content, as well as leading edge application of technology in the education experience. TAL Education Group offers comprehensive learning services to students from all ages through diversified class formats. TAL's learning services mainly cover enrichment learning programs and some academic subjects in and out of China. Its ADSs trade on the New York Stock Exchange under the symbol "TAL".

REFERENCES

[1] Ghodai Abdelrahman, Qing Wang, and Bernardo Pereira Nunes. 2022. Knowledge Tracing: A Survey. *ArXiv preprint* abs/2201.06953 (2022). https://arxiv.org/abs/2201.06953

[2] Hao Cen, Kenneth Koedinger, and Brian Junker. 2006. Learning factors analysis–a general method for cognitive model evaluation and improvement. In *International Conference on Intelligent Tutoring Systems*. Springer, 164–175.

[3] Jiahao Chen, Zitao Liu, Shuyan Huang, Qiongqiong Liu, and Weiqi Luo. 2023. Improving Interpretability of Deep Sequential Knowledge Tracing Models with Question-centric Cognitive Representations. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

[4] Youngduck Choi, Youngnam Lee, Junghyun Cho, Jineon Baek, Byungsoo Kim, Yeongmin Cha, Dongmin Shin, Chan Bae, and Jaewe Heo. 2020. Towards an appropriate query, key, and value computation for knowledge tracing. In *Proceedings of the Seventh ACM Conference on Learning@Scale*. 341–344.

[5] Albert T Corbett and John R Anderson. 1994. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-adapted Interaction* 4, 4 (1994), 253–278.

[6] Aritra Ghosh, Neil Heffernan, and Andrew S Lan. 2020. Context-Aware Attentive Knowledge Tracing. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*.

[7] Tanja Käser, Severin Klingler, Alexander G Schwing, and Markus Gross. 2017. Dynamic Bayesian networks for student modeling. *IEEE Transactions on Learning Technologies* 10, 4 (2017), 450–462.

[8] Elise Lavoué, Baptiste Monterrat, Michel Desmarais, and Sébastien George. 2018. Adaptive gamification for learning environments. *IEEE Transactions on Learning Technologies* 12, 1 (2018), 16–28.

[9] Jinseok Lee and Dit-Yan Yeung. 2019. Knowledge query network for knowledge tracing: How knowledge interacts with skills. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. 491–500.

[10] Qi Liu, Zhenya Huang, Yu Yin, Enhong Chen, Hui Xiong, Yu Su, and Guoping Hu. 2019. EKT: Exercise-aware knowledge tracing for student performance prediction. *IEEE Transactions on Knowledge and Data Engineering* 33, 1 (2019), 100–115.

[11] Zitao Liu, Qiongqiong Liu, Jiahao Chen, Shuyan Huang, Jiliang Tang, and Weiqi Luo. 2022. pyKT: A Python Library to Benchmark Deep Learning based Knowledge Tracing Models. In *Thirty-sixth Conference on Neural Information Processing Systems*.

[12] Ting Long, Yunfei Liu, Jian Shen, Weinan Zhang, and Yong Yu. 2021. Tracing Knowledge State with Individual Cognition and Acquisition Estimation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 173–182.

[13] Koki Nagatani, Qian Zhang, Masahiro Sato, Yan-Ying Chen, Francine Chen, and Tomoko Ohkuma. 2019. Augmenting knowledge tracing by considering forgetting behavior. In *The World Wide Web Conference*. 3101–3107.

[14] Hiromi Nakagawa, Yusuke Iwasawa, and Yutaka Matsuo. 2019. Graph-based knowledge tracing: modeling student proficiency using graph neural network. In *2019 IEEE/WIC/ACM International Conference on Web Intelligence*. IEEE, 156–163.

[15] Shalini Pandey and Jaideep Srivastava. 2020. RKT: relation-aware self-attention for knowledge tracing. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 1205–1214.

[16] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J Guibas, and Jascha Sohl-Dickstein. 2015. Deep knowledge tracing. *Advances in neural information processing systems* 28 (2015).

[17] Nguyen Thai-Nghe, Lucas Drumond, Tomáš Horváth, Artus Krohn-Grimberghe, Alexandros Nanopoulos, and Lars Schmidt-Thieme. 2012. Factorization techniques for predicting student performance. In *Educational Recommender Systems and Technologies: Practices and Challenges*. IGI Global, 129–153.

[18] Zhiwei Wang, Xiaoqin Feng, Jiliang Tang, Gale Yan Huang, and Zitao Liu. 2019. Deep knowledge tracing with side information. In *International Conference on Artificial Intelligence in Education*. Springer, 303–308.

[19] Jiani Zhang, Xingjian Shi, Irwin King, and Dit Yan Yeung. 2017. Dynamic Key-Value Memory Networks for Knowledge Tracing. In *Proceedings of the 26th International Conference on World Wide Web*. 765.