Contribution Title

First Author[1[0000−1111−2222−3333]], Second Author[2,3[1111−2222−3333−4444]], and Third Author[3[2222−−3333−4444−5555]]

[1] Princeton University, Princeton NJ 08544, USA
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany
lncs@springer.com
http://www.springer.com/gp/computer-science/lncs
[3] ABC Institute, Rupert-Karls-University Heidelberg, Heidelberg, Germany
{abc,lncs}@uni-heidelberg.de

Abstract. xxxxx

Keywords: xxx · xxx · xxx

1 Introduction

Knowledge tracing (KT) is a fundamental task in educational data mining that aims to estimate learners' knowledge states and predict their future performance by modeling historical interaction data. With billions of daily learning interactions generated on global educational platforms, accurate knowledge tracing helps identify struggling learners and supports personalized learning path planning to improve learning outcomes.

With the rapid development of deep learning technology, the field of KT has made breakthrough progress. The introduction of sequence models, such as RNN [3], can effectively capture the dynamic evolution of learners' knowledge states, but faces challenges in dealing with long sequences, such as gradient vanishing and computational efficiency. In recent years, attention-based [8] models(such as SAINT [1], AKT [2], simpleKT [5]) have gradually become the mainstream solution in this field by establishing direct correlations within sequences. These models not only overcome the problem of long-range dependencies but also adaptively focus on key historical information.

Despite the success of attention-based KT models, several limitations remain. First, students' knowledge mastery is hierarchical, progressing from surface to deep understanding through cognitive processing. However, existing models suffer from over-smoothing [9], where the attention distribution tends to be uniform, making it difficult to distinguish between different levels of knowledge features and learning abilities. Second, current methods inadequately handle noisy data from real-world scenarios, such as random answering behaviors (such as carelessness, guessing, or fatigue). Third, while learning effectiveness relies on connecting new knowledge with existing knowledge, current approaches focus too heavily on question and KC IDs, neglecting the semantic information of the questions and the association between KCs. Although there are attempts to improve modeling using graph neural networks, the joint modeling of semantic information and graph structure remains insufficient.

Therefore, in this paper, we propose HCGKT (Hierarchical Contrastive Graph Knowledge Tracing) to address the aforementioned challenges. Specifically, we use the hierarchical graph filtering attention mechanism to address over-smoothing by gradually extracting multi-level feature representations of students, capturing the hierarchical knowledge relationships between questions and KCs. In addition, we combine contrastive learning with adversarial perturbations to handle noisy student interaction data. By injecting random perturbations during training, we simulate complex educational data scenarios and leverage contrastive learning to extract robust features. Finally, we employ Graph Convolutional Network (GCN) [4] to jointly model semantic information and structural relationships between questions and KCs. By introducing real-world KC semantics and relationship graphs, we effectively address the challenge of insufficient semantic and structural modeling, enhancing knowledge tracing capabilities. We conduct comprehensive experiments on three benchmark datasets (Algebra2005, Bridge2algebra2006 [7], XES3G5M [6]) and achieve superior performance in terms of AUC and Accuary metrics.

The main contributions of this paper are:

- We address the over-smoothing problem by capturing hierarchical knowledge relationships between questions and KCs through multi-level feature representations.
- We enhance model robustness against noisy student interaction data by combining adversarial perturbations with contrastive learning.
- We improve knowledge tracing performance by jointly modeling semantic information and structural relationships in question-KC graph.

2 Related Work

related

3 The HCGKT Framework

method

4 Experiments

exp

5 Conclusion

conclusion

References

1. Choi, Y., Lee, Y., Cho, J., Baek, J., Kim, B., Cha, Y., Shin, D., Bae, C., Heo, J.: Towards an appropriate query, key, and value computation for knowledge tracing. In: Proceedings of the Seventh ACM Conference on Learning@ Scale. pp. 341–344 (2020)
2. Ghosh, A., Heffernan, N., Lan, A.S.: Context-aware attentive knowledge tracing. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 2330–2339 (2020)
3. Graves, A., Graves, A.: Long short-term memory. Supervised sequence labelling with recurrent neural networks pp. 37–45 (2012)
4. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
5. Liu, Z., Liu, Q., Chen, J., Huang, S., Luo, W.: simplekt: a simple but tough-to-beat baseline for knowledge tracing. arXiv preprint arXiv:2302.06881 (2023)
6. Liu, Z., Liu, Q., Guo, T., Chen, J., Huang, S., Zhao, X., Tang, J., Luo, W., Weng, J.: Xes3g5m: A knowledge tracing benchmark dataset with auxiliary information. Advances in Neural Information Processing Systems **36** (2024)
7. Stamper, J., Koedinger, K., Pavlik Jr, P., Baker, R., Rossi, L.: A temporal analysis of student learning trajectories using kdd cup 2010 data. Journal of Educational Data Mining **2**(1), 107–128 (2010)
8. Vaswani, A.: Attention is all you need. Advances in Neural Information Processing Systems (2017)
9. Wu, X., Ajorlou, A., Wu, Z., Jadbabaie, A.: Demystifying oversmoothing in attention-based graph neural networks. Advances in Neural Information Processing Systems **36** (2024)