Personalised Programming Education with Knowledge Tracing

Martha Shaka
Centre for Research Training in Artificial Intelligence
University College Cork
Cork, Ireland
121127347@umail.ucc.ie

Diego Carraro, Kenneth N. Brown
Insight Centre for Data Analytics
University College Cork
Cork, Ireland
{diego.carraro,ken.brown}@insight-centre.org

ABSTRACT

In traditional programming education, addressing diverse student needs and providing effective and scalable learning experiences is challenging. Conventional methods struggle to adapt to varying learning styles and offer personalised feedback. AI-based Programming Tools (AIPTs) have shown promise in automating feedback, simplifying programming concepts, and guiding students. Their widespread adoption is hindered by limitations related to accuracy, explanation, and personalisation. Conversely, AIPTs tailored for expert programmers, such as ChatGPT and Copilot, have gained popularity for their productivity-enhancing capabilities, but they still fall short in terms of personalisation, neglecting individual students' unique knowledge and skills. Our research aims to leverage AI to create AIPTs that offer personalised feedback through adaptive learning, accommodating diverse student backgrounds and proficiency levels. In particular, we explore using Knowledge Tracing (KT) to anticipate specific syntax errors in programming assignments, addressing the challenges novices face in acquiring syntactical knowledge. The findings suggest the KT's potential to transform programming education by enabling timely interventions for students dealing with specific errors or misconceptions, automating personalised feedback, and informing tailored instructional strategies.

KEYWORDS

Knowledge Tracing, Syntax Errors, Programming Assignments, Personalisation, Automated Feedback

ACM Reference Format:
Martha Shaka and Diego Carraro, Kenneth N. Brown. 2023. Personalised Programming Education with Knowledge Tracing. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 PROGRAMMING KNOWLEDGE TRACING

AI can create tools that anticipate and adapt to future student behaviours, providing intelligent feedback in response to these behaviours, especially in the context of practical programming assignments. Novice programmers often grapple with self-assessment and debugging, driving the need for personalised guidance, especially in the growing landscape of online courses. Knowledge Tracing (KT) is a data mining technique that can track and predict students' evolving comprehension during the learning process. It captures and analyses their performance over time, offering valuable insights into their knowledge state and enabling the timely and tailored interventions crucial for effective learning.

KT models, which leverage historical student performance data to predict evolving knowledge states, are a cornerstone in personalised education. While traditional models like Bayesian Knowledge Tracing (BKT) have played a pivotal role, Deep Knowledge Tracing (DKT), empowered by recurrent neural networks, excels, particularly with large datasets. DKT [1] autonomously learns and predicts a student's evolving knowledge state with each interaction with learning materials, such as exercises, thereby enhancing its understanding of their skill mastery. However, the prevalent reliance on binary correctness to assess student performance poses limitations, as it fails to capture the nuances of the learning process. Moreover, these models often overlook domain-specific features like code representations and question embeddings, which are instrumental in understanding a student's approach to problem-solving.

2 PROPOSED ERROR TRACING FRAMEWORK

Our study extends DKT models by integrating additional input features, such as code representations, reference solutions, and question embeddings. Our primary focus is on the prediction of syntax errors in programming assignments, a multi-label classification task that encompasses highly imbalanced and co-occurring error classes. Drawing upon a dataset from an introductory Java programming course, we explore diverse training approaches for breaking down the multi-label task, enhancing the DKT models' efficacy in predicting specific errors.

Our proposed solution involves a two-phase prediction strategy, first gauging whether a student is likely to commit an error or not and subsequently identifying the specific errors they have made, building upon the results from the initial phase. We evaluate prediction performance across various dimensions: overall performance, specific error classes, and individual questions. These assessments yield vital insights into how historical data informs predictions and cater to the sequential nature of students' problem-solving journeys. In the future, for example, we plan to explore the integration of Large Language Models, and we envision building a comprehensive system that offers personalised interventions in the realm of programming education.

REFERENCES

[1] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas Guibas, Jascha Sohl-Dickstein, Stanford University, and Khan Academy. 2015. Deep Knowledge Tracing. *Advances in Neural Information Processing Systems* 28 (2015).

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn