

THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút):

<https://youtu.be/hle9s7iW900>

- Link slides (dạng .pdf đặt trên Github):

<https://github.com/giaonnq18uit/CS2205.CH181/blob/08c0f6d2f160650e35715cf57e2b73cf21b29d9b/IntegrateLIMEwithDKT.pdf>

<ul style="list-style-type: none">• Họ và Tên: Nguyễn Ngọc Quỳnh Giao• MSSV: 230201040 	<ul style="list-style-type: none">• Lớp: CS2205.CH181• Tự đánh giá (điểm tổng kết môn): 8/10• Số buổi vắng: 0• Số câu hỏi QT cá nhân: 0• Link Github: https://github.com/giaonnq18uit/CS2205.CH181
--	---

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

TÍCH HỢP LIME VỚI DEEP KNOWLEDGE TRACING ĐỂ CẢI THIẾN TÍNH MINH BẠCH TRONG DỰ ĐOÁN HỌC TẬP

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

IMPROVING TRANSPARENCY IN LEARNING PREDICTIONS: INTEGRATING LIME WITH DEEP KNOWLEDGE TRACING

TÓM TẮT *(Tối đa 400 từ)*

Trong lĩnh vực giáo dục trực tuyến và học tập thông minh, việc theo dõi và dự đoán trạng thái kiến thức của học viên là vô cùng quan trọng và cũng không kém phần thách thức. Deep Knowledge Tracing (DKT) [1] đã được phát triển như một phương pháp mạnh mẽ để mô hình hóa quá trình học tập của học viên và dự đoán hiệu suất học tập dựa trên sự thể hiện và thành tích đạt được trước đó của học viên. Tuy nhiên, một hạn chế lớn của DKT và các mô hình học sâu khác là tính minh bạch khá thấp, khiến giáo viên và học viên khó hiểu được lý do đằng sau các dự đoán mà mô hình đưa ra.

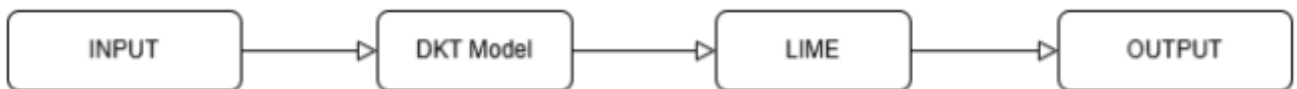
Để giải quyết vấn đề này, tôi đề xuất tích hợp phương pháp Local Interpretable Model-agnostic Explanations (LIME) [2] với DKT nhằm cải thiện tính minh bạch trong các dự đoán học tập. Bằng cách sử dụng LIME, ta có thể giải thích cụ thể những đặc trưng nào đã ảnh hưởng đến quyết định của DKT trong việc dự đoán trạng thái kiến thức của học viên. Điều này không chỉ giúp giáo viên và học viên hiểu rõ hơn về quá trình học tập mà còn cung cấp thông tin chi tiết để cải thiện phương pháp giảng dạy và học tập cá nhân hóa.

GIỚI THIỆU *(Tối đa 1 trang A4)*

Trong những năm gần đây, giáo dục trực tuyến hay học trực tuyến (e-learning) ngày càng trở nên phổ biến và trở thành một phần quan trọng của nền giáo dục hiện đại.

Không chỉ hỗ trợ việc học tập và giảng dạy từ xa mà công nghệ này còn cung cấp các công cụ mạnh mẽ để theo dõi và đánh giá quá trình học tập của học viên. Một trong những phương pháp tiên tiến nhất để mô hình hóa quá trình học tập và dự đoán hiệu suất của học viên là Deep Knowledge Tracing (DKT). DKT sử dụng các mạng neural sâu để nắm bắt các mẫu học tập phức tạp và dự đoán trạng thái kiến thức của học viên dựa trên dữ liệu lịch sử. Mặc dù DKT đã chứng minh được hiệu quả cao trong việc dự đoán, song một trong những hạn chế chính của nó là tính minh bạch thấp. Các mô hình học sâu, bao gồm DKT thường được coi là “hộp đen” vì khó hiểu được cách chúng đưa ra quyết định.

Tính minh bạch trong dự đoán học tập là một yếu tố quan trọng đối với cả giáo viên và học viên. Giáo viên cần hiểu rõ lý do đằng sau các dự đoán để điều chỉnh phương pháp giảng dạy phù hợp, trong khi học viên cần nhận thức được những yếu tố ảnh hưởng đến kết quả học tập của mình để có thể cải thiện tốt hơn. Để giải quyết vấn đề này, nghiên cứu này đề xuất tích hợp phương pháp Local Interpretable Model-agnostic Explanations (LIME) với DKT. LIME là một kỹ thuật giải thích cục bộ, không phụ thuộc vào mô hình, giúp tạo ra các giải thích dễ hiểu về các dự đoán của các mô hình phức tạp.



Hình 1. Mô hình Quy trình tích hợp LIME với DKT

Trong đó:

- INPUT: dữ liệu học tập của học viên
 - Lịch sử bài kiểm tra: dữ liệu về các bài kiểm tra mà học viên đã thực hiện.
 - Điểm số: Kết quả điểm số của các bài kiểm tra.
 - Thời gian làm bài: thời gian mà học viên dành cho mỗi bài kiểm tra.
 - Các đặc trưng khác liên quan như số lần xem lại bài học, mức độ khó của bài kiểm tra, v.v.
- OUTPUT:

- Giải thích các dự đoán của mô hình DKT: cung cấp các giải thích dễ hiểu và trực quan về các dự đoán, giúp giáo viên và học viên hiểu rõ hơn về quá trình dự đoán.
- Phản hồi và tinh chỉnh: thu thập phản hồi từ giáo viên và học viên về tính hữu ích của các giải thích, sau đó điều chỉnh mô hình DKT và phương pháp giảng dạy dựa trên phản hồi nhận được.

MỤC TIÊU

- Phát triển hệ thống tích hợp LIME với DKT để cung cấp giải thích cho các dự đoán của DKT.
- Đánh giá hiệu quả của hệ thống này trong việc cải thiện tính minh bạch và hiểu biết về các dự đoán học tập.
- Kiểm tra tác động của các giải thích này đối với quá trình giảng dạy và học tập.

NỘI DUNG VÀ PHƯƠNG PHÁP

1. Nội dung

- Khảo sát và lựa chọn dữ liệu học tập: thu thập bộ dữ liệu học tập từ các hệ thống học trực tuyến như lịch sử bài kiểm tra, điểm số, thời gian làm bài, và các đặc trưng khác liên quan.
- Huấn luyện mô hình Deep Knowledge Tracing (DKT):
 - Chuẩn bị dữ liệu: tiền xử lý dữ liệu để đưa vào mô hình DKT.
 - Xây dựng mô hình DKT: sử dụng thư viện học máy như TensorFlow hoặc PyTorch để xây dựng và huấn luyện mô hình với dữ liệu đã chuẩn bị.
- Tích hợp LIME vào mô hình DKT:
 - Phát triển phương pháp sử dụng LIME để tạo giải thích cục bộ cho các dự đoán của DKT.
 - Triển khai LIME: sử dụng LIME để giải thích các dự đoán từ mô hình DKT, hiển thị các đặc trưng ảnh hưởng đến các dự đoán.
- Xác định các tiêu chí đánh giá tính minh bạch và hiểu biết của các dự

đoán.

- Thu thập và phân tích phản hồi:
 - Thu thập phản hồi từ giáo viên và học viên về tính minh bạch và dễ hiểu của các giải thích do LIME cung cấp.
 - Sử dụng các phương pháp khảo sát, phỏng vấn và bảng câu hỏi để thu thập dữ liệu phản hồi.
- Đo lường tác động của giải thích đến giảng dạy và học tập:
 - Theo dõi sự thay đổi trong phương pháp giảng dạy của giáo viên dựa trên các giải thích của LIME và đánh giá sự cải thiện trong việc hỗ trợ học viên.
 - Theo dõi tiến bộ học tập của học viên khi được cung cấp giải thích về các dự đoán và đánh giá sự thay đổi trong động lực và hiệu suất học tập của học viên.

2. Phương pháp

- Thu thập và xử lý dữ liệu: sử dụng các nguồn dữ liệu mở để thu thập dữ liệu học tập cần thiết, như bộ dataset 2009-2010 ASSISTment Data¹. Nếu có thể hợp tác với các tổ chức giáo dục để thu thập dữ liệu thì càng tốt.
- Phát triển mô hình DKT: xây dựng mô hình dựa trên các nghiên cứu hiện có, điều chỉnh các tham số mô hình để đạt hiệu suất tốt nhất.
- Triển khai LIME:¹
 - Tích hợp LIME vào mô hình DKT để tạo các giải thích cho từng dự đoán.
 - Hiển thị giải thích: phát triển giao diện để hiển thị các giải thích của LIME cho giáo viên và học viên.
- Xây dựng bảng khảo sát và phỏng vấn: phát triển các câu hỏi khảo sát và phỏng vấn để thu thập phản hồi từ người dùng về tính minh bạch và hữu ích của LIME.
- Phân tích dữ liệu:

¹ <https://sites.google.com/site/assistmentsdata/home/2009-2010-assistment-data>

- Sử dụng các phương pháp phân tích định tính và định lượng để đánh giá hiệu quả của hệ thống.
- So sánh với các mô hình DKT không có giải thích để đánh giá sự cải thiện về tính minh bạch.
- Thu thập dữ liệu trước và sau can thiệp, sử dụng các phương pháp thống kê để phân tích sự thay đổi và tác động của giải thích.

KẾT QUẢ MONG ĐỢI

(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)

- Hệ thống DKT được tích hợp với LIME, có khả năng đưa ra các dự đoán về trạng thái kiến thức của học viên và cung cấp các giải thích cục bộ cho những dự đoán này.
- Đạt được sự cải thiện đáng kể trong hiểu biết của người dùng về các dự đoán học tập.
- Thu thập phản hồi liên tục từ giáo viên và học viên để tinh chỉnh và cải tiến hệ thống. Từ đó, đề xuất các hướng phát triển và tối ưu hệ thống dựa trên phản hồi và kết quả thực nghiệm, nhằm nâng cao hiệu quả giảng dạy và học tập.

TÀI LIỆU THAM KHẢO *(Định dạng DBLP)*

- [1] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J. Guibas, Jascha Sohl-Dickstein: Deep Knowledge Tracing. NIPS 2015: 505-513
- [2] Marco Túlio Ribeiro, Sameer Singh, Carlos Guestrin: "Why Should I Trust You?": Explaining the Predictions of Any Classifier. HLT-NAACL Demos 2016: 97-101
- [3] Ghodai Abdelrahman, Qing Wang, Bernardo Pereira Nunes: Knowledge Tracing: A Survey. ACM Comput. Surv. 55(11): 224:1-224:37 (2023)