TẠP CHÍ KHOA HỌC TRƯỜNG ĐẠI HỌC SƯ PHẠM HÀ NỘI
Khoa học Giáo dục 2026, Tập 71, Số 7, tr. 121-130
ISSN 2354-1075; https://hnuejs.edu.vn/es
DOI: 10.18173/2354-1075.2026-0108

MÔ HÌNH HÓA TRẠNG THÁI TRI THỨC CỦA HỌC SINH BẰNG BAYESIAN KNOWLEDGE TRACING: MỘT THỰC NGHIỆM VỚI CHỦ ĐỀ PHƯƠNG TRÌNH MẶT CẦU

Huỳnh Phú Sĩ[*1] và Lê Minh Cường[2]
[1]*Trường THCS và THPT Mỹ Thuận, Vĩnh Long, Việt Nam*
[2]*Trường Đại học Đồng Tháp, Đồng Tháp, Việt Nam*
[*]Tác giả liên hệ: Huỳnh Phú Sĩ, e-mail: huynhphusi@gmail.com
Ngày nhận bài: 18/6/2026. Ngày sửa bài: 9/7/2026. Ngày nhận đăng: 20/7/2026.

Tóm tắt. Cá nhân hóa việc học đang dần trở thành yêu cầu thiết yếu trong giáo dục hiện đại, kéo theo nhu cầu mô hình hóa trạng thái tri thức của người học một cách chính xác. Trong nghiên cứu này, chúng tôi lần lượt thử nghiệm mô hình Bayesian Knowledge Tracing (BKT) truyền thống cùng hai biến thể của nó là KT-IDEM (tích hợp độ khó câu hỏi) và Forgetting BKT (có tính đến yếu tố quên kiến thức), hướng đến việc theo dõi năng lực học sinh trong chủ đề "Phương trình mặt cầu" (Toán 12). Cụ thể, một web app đã được xây dựng để ghi nhận dữ liệu tương tác thực tế của 143 học sinh, trong đó toàn bộ 283 câu hỏi được gán nhãn chi tiết theo 11 kĩ năng thành phần. Kết quả kiểm định chéo 5 phần cho thấy KT-IDEM cho kết quả dự báo tốt hơn so với cả BKT chuẩn lẫn Forgetting BKT. Ngoài ra, việc trực quan hóa trạng thái tri thức từ chính các tham số đã được huấn luyện không chỉ giúp phát hiện ra những tình huống đoán mò hay sơ suất, mà còn đem lại một "bản đồ năng lực" cá nhân hóa đến từng kĩ năng, hỗ trợ đắc lực cho cả giáo viên lẫn học sinh. Những kết quả ban đầu này cho thấy KT-IDEM có tiềm năng hỗ trợ chẩn đoán trạng thái tri thức của học sinh, mở ra triển vọng xây dựng các hệ thống gợi ý bài tập thích ứng theo thời gian thực, góp phần thúc đẩy dạy học cá nhân hóa trong bối cảnh giáo dục Việt Nam.
Từ khóa: Bayesian Knowledge Tracing, KT-IDEM, truy vết tri thức, phương trình mặt cầu, dạy học cá nhân hóa.

MODELING STUDENT KNOWLEDGE STATES USING BAYESIAN KNOWLEDGE TRACING: AN EXPERIMENT WITH THE TOPIC OF SPHERE EQUATIONS

Huynh Phu Si[*1] and Le Minh Cuong[2]
[1]*My Thuan Middle and High School, Vinh Long, Vietnam*
[2]*Dong Thap University, Dong Thap, Vietnam*
[*]Corresponding author: Huynh Phu Si, e-mail: huynhphusi@gmail.com
Received June 18, 2026. Revised July 9, 2026. Accepted July 20, 2026.

Abstract. Personalized learning is increasingly becoming an essential requirement in modern education, driving the need for accurately modeling learners' knowledge states. In this study, we systematically evaluated the traditional Bayesian Knowledge Tracing (BKT) model alongside two of its variants KT-IDEM (which incorporates item difficulty) and Forgetting BKT (which accounts for knowledge decay) to track student performance on the topic of "Sphere Equations" (Grade 12 Mathematics). Specifically, a web application was developed to collect real interaction data from 143 students, with all 283 questions labeled in detail according to 11 constituent skills. Results from 5-fold cross-validation demonstrated that KT-IDEM yielded better predictive performance compared to both standard BKT and Forgetting BKT.

Furthermore, visualizing knowledge states using the trained model parameters not only helped identify instances of guessing and slipping but also provided a personalized "competency map" at the individual skill level, offering valuable support for both teachers and students. These preliminary findings suggest that KT-IDEM has the potential to support the diagnosis of students' knowledge states, opening up promising avenues for building real-time adaptive exercise recommendation systems that contribute to advancing personalized learning in the Vietnamese educational context.

Keywords: Bayesian Knowledge Tracing, KT-IDEM, knowledge tracing, sphere equations, personalized learning.

1. Mở đầu

Trong bối cảnh giáo dục hiện đại, nhu cầu cá nhân hóa (CNH) quá trình học tập đang trở thành một xu hướng tất yếu nhằm đáp ứng năng lực và nhịp độ tiếp thu đa dạng của từng học sinh (HS) (Alkhatlan & Kalita, 2018). Sự phát triển mạnh mẽ của trí tuệ nhân tạo (AI) đã thúc đẩy sự ra đời của các hệ thống dạy học thông minh, cho phép mô phỏng lại kiến thức và kinh nghiệm của người học để cung cấp các hỗ trợ tùy chỉnh (Corbett & Anderson, 1995). Để đạt được điều này, một thành phần cốt lõi không thể thiếu là kĩ thuật mô hình hóa người học (learner modeling), tức là quá trình ước lượng mức độ hiểu biết hoặc năng lực của người học đối với một chủ đề hay kĩ năng cụ thể (Desmarais & Baker, 2012). Bằng cách lập mô hình chính xác các đặc điểm ẩn của HS, các hệ thống học tập thích ứng có thể tự động điều chỉnh độ khó của tài liệu hoặc đưa ra các bài tập phù hợp, từ đó tối ưu hóa lộ trình và hiệu quả giáo dục (Yudelson & nnk, 2013).

Để theo dõi và mô hình hóa trạng thái tri thức của HS một cách liên tục theo thời gian, truy vết tri thức (knowledge tracing - KT) được áp dụng rộng rãi như một kĩ thuật thiết yếu để dự đoán hiệu suất tương lai dựa trên chuỗi tương tác quá khứ của người học (Corbett & Anderson, 1995; Dai & nnk, 2021). Trong đó, Bayesian Knowledge Tracing (BKT) là một trong những mô hình nền tảng và thành công nhất (Corbett & Anderson, 1995). Về mặt toán học, BKT hoạt động dựa trên cấu trúc của một mô hình Markov ẩn có hướng, trong đó trạng thái kiến thức của HS là một biến ẩn (đã nắm vững hoặc chưa nắm vững), và câu trả lời quan sát được của HS (đúng hoặc sai) đóng vai trò là biến hiển thị (Shimada & Okada, 2025). Mô hình này liên tục cập nhật xác suất nắm vững kiến thức thông qua định lí Bayes dựa trên 4 tham số: xác suất có sẵn kiến thức ban đầu $P(L_0)$, xác suất học được kiến thức sau một cơ hội luyện tập $P(T)$, xác suất đoán mò đúng dù chưa biết $P(G)$, và xác suất sơ suất trả lời sai dù đã nắm vững kiến thức $P(S)$ (Bulut & nnk, 2023; Corbett & Anderson, 1995). Điểm mạnh lớn nhất của BKT chính là tính minh bạch và khả năng diễn giải trực quan, giúp các nhà giáo dục dễ dàng nắm bắt được động lực học tập của HS thay vì đối mặt với các mô hình "hộp đen" phức tạp (Khajah & nnk, 2016).

Trên thế giới, BKT đã thu hút được sự quan tâm rộng rãi và được mở rộng thành nhiều biến thể khác nhau nhằm giải quyết các giới hạn của mô hình tiêu chuẩn (Abdelrahman & nnk, 2022). Các biến thể này bao gồm việc CNH tham số theo từng HS như mô hình Individualized BKT (Yudelson & nnk, 2013), tích hợp độ khó của từng câu hỏi cụ thể như KT-IDEM (Pardos & Heffernan, 2011), hoặc bổ sung yếu tố quên kiến thức theo thời gian như PC-BKT (Nedungadi & Remya, 2015). Trong khi đó, tại Việt Nam, các nghiên cứu ứng dụng kĩ thuật KT đang bắt đầu được quan tâm, điển hình như các nỗ lực khám phá những mô hình học sâu phức tạp (như SAINT+) trên các bộ dữ liệu quy mô lớn của quốc tế (Phạm Hoàng An, 2025). Mặc dù các mô hình học sâu mang lại độ chính xác cao, nhưng cơ chế đưa ra quyết định thiếu minh bạch của chúng lại gây ra thách thức lớn đối với các nhà giáo dục trong việc diễn giải và ứng dụng thực tiễn (Phạm Hoàng An, 2025). Hơn nữa, vẫn còn thiếu vắng các nghiên cứu thực nghiệm triển khai và so sánh trực tiếp BKT trên các bộ dữ liệu nội địa được thu thập từ hệ thống học tập của HS phổ thông, đặc biệt là với các chuyên đề Toán học cụ thể.

Nhằm lấp đầy khoảng trống nghiên cứu này, bài báo tập trung thiết lập môi trường và thu thập dữ liệu thực nghiệm đối với chủ đề "Phương trình mặt cầu". Cụ thể, nghiên cứu tập trung

giải quyết hai câu hỏi chính: (1) Biến thể BKT nào (Standard BKT, KT-IDEM hay Forgetting BKT) cho khả năng dự báo trạng thái tri thức chính xác nhất đối với dữ liệu thực nghiệm? (2) Mô hình BKT sau khi huấn luyện cung cấp những chẩn đoán định lượng khác biệt như thế nào về năng lực HS so với phương pháp tính điểm truyền thống? Bằng cách trả lời hai câu hỏi trên, chúng tôi không chỉ đánh giá hiệu năng của từng biến thể BKT, mà còn bước đầu khám phá tiềm năng ứng dụng của chúng vào việc xây dựng các hệ thống dạy học CNH.

2. Phương pháp nghiên cứu

2.1. Xây dựng sơ đồ tri thức

Trong các hệ thống dạy học thông minh, việc mô hình hóa tri thức chuyên môn đóng vai trò nền tảng, cho phép biểu diễn cấu trúc kiến thức mà người học cần lĩnh hội thành các đơn vị nhỏ hơn, có thể đo lường và theo dõi được (Koedinger & nnk, 2012). Các đơn vị này, thường được gọi là các thành tố tri thức (knowledge components – KC), đại diện cho một kĩ năng, một khái niệm hoặc một quy tắc cụ thể mà HS cần nắm vững. Việc phân rã chính xác một chủ đề học tập thành các KC không chỉ giúp mô hình hóa quá trình học tập một cách chi tiết mà còn cung cấp cơ sở để giáo viên (GV) đưa ra những can thiệp sư phạm phù hợp.

Đối với chủ đề "Phương trình mặt cầu", chúng tôi đã tiến hành phân tích chuyên môn dựa trên yêu cầu cần đạt của Bộ Giáo dục và Đào tạo, kết hợp với kinh nghiệm giảng dạy thực tế của GV bộ môn. Qua quá trình phân tích, nội dung bài học được phân rã thành 11 KC có mối liên hệ chặt chẽ với nhau. Danh sách chi tiết các KC được trình bày trong Bảng 1.

Bảng 1. Danh sách 11 thành tố tri thức của chủ đề "Phương trình mặt cầu"

KC	Tên kĩ năng	Kĩ năng điều kiện
KC1	Tìm tâm và bán kính của phương trình mặt cầu dạng chính tắc	
KC2	Tìm tâm và bán kính của phương trình mặt cầu dạng tổng quát	
KC3	Nhận biết phương trình tổng quát của mặt cầu	KC2
KC4	Viết phương trình mặt cầu khi biết tâm và bán kính	
KC5	Viết phương trình mặt cầu khi biết tâm và một điểm đi qua	KC4
KC6	Viết phương trình mặt cầu khi biết đường kính	KC4, KC5
KC7	Viết phương trình mặt cầu khi biết tâm và mặt phẳng tiếp xúc	KC4
KC8	Viết phương trình mặt cầu đi qua 4 điểm	
KC9	Tìm vị trí tương đối của một điểm và mặt cầu	
KC10	Tìm giao điểm của đường thẳng và mặt cầu	
KC11	Bài toán thực tiễn về phương trình mặt cầu	KC4, KC5, KC7, KC8, KC9, KC10

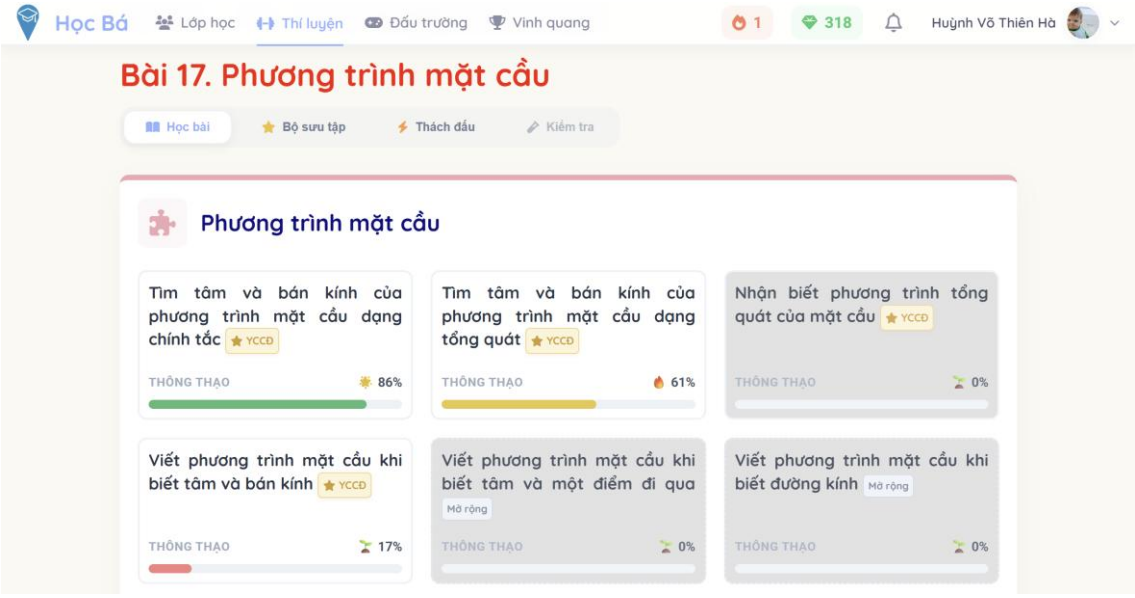
Dựa trên 11 KC đã xác định, chúng tôi đã xây dựng một ngân hàng câu hỏi gồm 283 câu hỏi (trung bình 26 câu hỏi/KC), được thiết kế theo cấu trúc định dạng mới của Chương trình giáo dục phổ thông 2018. Mỗi câu hỏi được gán nhãn một cách tường minh vào đúng một KC mà nó hướng tới đánh giá. Các câu hỏi trong một KC có độ khó tương đồng, được tổ Toán – Tin của trường THCS&THPT Mỹ Thuận thẩm định, bao gồm ba định dạng trắc nghiệm là trắc nghiệm nhiều lựa chọn, trắc nghiệm đúng – sai, trắc nghiệm trả lời ngắn. Trong đó, câu trắc nghiệm đúng – sai được thiết kế thành các nhóm câu hỏi, mỗi ý đúng – sai trong một nhóm được coi là một câu hỏi độc lập. Mỗi ý có thể ánh xạ đến một KC tùy thuộc vào yêu cầu của ý đó.

2.2. Thiết kế hệ thống thực nghiệm

Để triển khai thực nghiệm sư phạm và thu thập dữ liệu tương tác của HS, chúng tôi đã xây dựng một hệ thống học tập trực tuyến chuyên biệt. Hệ thống được xây dựng với mục tiêu kép:

vừa là môi trường học tập CNH cho HS, vừa là công cụ quản trị dữ liệu tích hợp thuật toán BKT. Giao diện web app được thiết kế hướng tới sự tập trung, minh bạch hóa tiến trình học tập và ứng dụng các yếu tố trò chơi hóa để duy trì động lực học tập cho HS:

Khi truy cập vào một bài học, HS không chỉ thấy danh sách các kĩ năng cần học mà còn nhìn thấy trực quan độ thông thạo của bản thân đối với từng KC (Hình 1). Con số này chính là xác suất $P(L_n)$ được mô hình BKT tính toán và trả về sau mỗi lượt tương tác, giúp HS nhận thức rõ "lỗ hổng" kiến thức của mình để có chiến lược ôn tập phù hợp.



Hình 1. Giao diện lựa chọn kĩ năng luyện tập của HS

HS chỉ có thể tham gia luyện tập một kĩ năng nếu đó là kĩ năng cơ bản hoặc tất cả kĩ năng điều kiện đều đã thông thạo trên 70%. Cơ chế này đóng vai trò như một "gia sư ảo", ngăn chặn tình trạng quá tải nhận thức do nhảy cóc kiến thức, đồng thời định hình một lộ trình học tập được CNH theo nhịp độ làm chủ tri thức của từng HS.

Trước khi bước vào luyện tập, HS có thể tự ôn tập kiến thức nền tảng thông qua các hoạt động đa dạng do web app cung cấp như: đọc văn bản lí thuyết, xem video bài giảng, làm bài tập điền khuyết hoặc ghép cặp tương ứng. Mức độ hoàn thành các hoạt động này được hệ thống quy đổi thành điểm kinh nghiệm, tạo cơ chế khen thưởng tức thời.

Sau bước ôn tập kiến thức, hệ thống chỉ hiển thị duy nhất một câu hỏi (hoặc một ý đúng/sai của một nhóm dữ kiện) trên màn hình. Thiết kế này bắt buộc HS phải đưa ra quyết định cho từng vấn đề cụ thể, đồng thời giúp hệ thống ghi nhận chính xác thời gian và kết quả đúng/sai cho từng KC riêng biệt. Sau khi trả lời, hệ thống lập tức cập nhật độ thông thạo và cung cấp phản hồi để HS tự điều chỉnh nhận thức trước khi chuyển sang câu tiếp theo.

Trước khi có kết quả huấn luyện từ dữ liệu thực tế, hệ thống vận hành dựa trên một bộ tham số BKT khởi tạo mặc định. Cụ thể, đối với mỗi KC, xác suất thành thạo ban đầu được thiết lập là $P(L_0) = 0,2$ và xác suất học được kiến thức sau mỗi lượt luyện tập là $P(T) = 0,1$. Đối với các tham số liên quan đến câu hỏi, nhằm phản ánh sự khác biệt về khả năng đoán mò và sơ suất giữa các định dạng trắc nghiệm, chúng tôi thiết lập phân hóa như sau:

- Câu hỏi trắc nghiệm nhiều lựa chọn: $P(G) = 0,25$, $P(S) = 0,1$.
- Câu hỏi trắc nghiệm đúng – sai: $P(G) = 0,5$, $P(S) = 0,1$.
- Câu hỏi trắc nghiệm trả lời ngắn: $P(G) = 0,05$, $P(S) = 0,15$.

Bộ tham số mặc định này được sử dụng để tính toán và hiển thị trực quan độ thông thạo ban đầu cho HS, đồng thời đóng vai trò là mốc tham chiếu để đánh giá mức độ cải thiện của mô hình sau khi được huấn luyện trên dữ liệu thực tế.

2.3. Dữ liệu nghiên cứu

Dữ liệu được thu thập từ kết quả làm bài của 143 HS lớp 12, trường THCS&THPT Mỹ Thuận, từ ngày 9/4/2026 đến 16/4/2026. Các em làm bài trước khi học chủ đề này trên lớp, kiến thức nền được sử dụng là phương trình đường tròn và phương pháp tọa độ trong không gian. Bộ dữ liệu thô ban đầu bao gồm 11.813 lượt làm bài (trung bình 83 lượt/HS, 1074 lượt/KC), trong đó mỗi bản ghi bao gồm các thông tin: mã HS, mã kĩ năng, mã câu hỏi, kết quả trả lời, thời điểm làm bài và thời gian phản hồi.

Để đảm bảo chất lượng đầu vào cho mô hình BKT, chúng tôi đã tiến hành một bước tiền xử lí nhằm loại bỏ các câu hỏi có dữ liệu tương tác cực đoan, không cung cấp đủ thông tin cho quá trình suy luận. Cụ thể, dựa trên phân tích thống kê mô tả, các câu hỏi thỏa mãn một trong ba điều kiện sau sẽ bị loại bỏ: (i) có ít hơn 5 lượt làm bài, (ii) có tỉ lệ trả lời đúng tuyệt đối 100%, hoặc (iii) có tỉ lệ trả lời đúng là 0%. Kết quả thống kê cho thấy có 17 câu hỏi thuộc diện cực đoan, dẫn đến việc loại bỏ 320 lượt làm bài liên quan. Sau quá trình làm sạch, bộ dữ liệu cuối cùng được sử dụng cho các phân tích tiếp theo bao gồm 11.493 lượt làm bài. Toàn bộ quá trình tiền xử lí và huấn luyện được thực hiện bằng ngôn ngữ Python với sự hỗ trợ của thư viện pandas và pyBKT (chi tiết được chia sẻ trong tệp pyBKT_test.ipynb tại https://github.com/huynhphusi/BKT-Test).

2.4. Quy trình huấn luyện và độ đo đánh giá

Sau khi bộ dữ liệu đã được làm sạch, chúng tôi tiến hành huấn luyện và đánh giá các biến thể của mô hình BKT bằng thư viện pyBKT (Bulut & nnk, 2023) trên nền tảng ngôn ngữ Python. Thư viện pyBKT cung cấp một bản cài đặt hoàn chỉnh của thuật toán Expectation–Maximization để ước lượng các tham số của mô hình từ dữ liệu tương tác của HS.

2.4.1. Mô hình toán học của BKT

Trong BKT, trạng thái tri thức của HS đối với một KC được biểu diễn dưới dạng xác suất HS đã nắm vững KC đó, liên tục được cập nhật sau mỗi câu trả lời. Gọi $P(L_{n-1})$ là xác suất HS đã nắm vững KC trước khi thực hiện câu hỏi thứ n. Sau khi HS trả lời câu hỏi thứ n, hệ thống dựa trên kết quả (obs_n) và áp dụng định lí Bayes để suy luận ngược về trạng thái tri thức.

(1) Pha 1 – Cập nhật sau khi quan sát kết quả trả lời:

Nếu HS trả lời đúng ($obs_n = 1$), xác suất nắm vững kĩ năng được tính theo công thức:

$$P(L_n \mid obs_n = 1) = \frac{P(L_{n-1}) \cdot \big(1 - P(S)\big)}{P(L_{n-1}) \cdot \big(1 - P(S)\big) + \big(1 - P(L_{n-1})\big) \cdot P(G)} \quad (1)$$

Nếu HS trả lời sai ($obs_n = 0$), xác suất nắm vững kĩ năng được cập nhật theo công thức:

$$P(L_n \mid obs_n = 0) = \frac{P(L_{n-1}) \cdot P(S)}{P(L_{n-1}) \cdot P(S) + \big(1 - P(L_{n-1})\big) \cdot \big(1 - P(G)\big)} \quad (2)$$

(2) Pha 2 – Dự báo trạng thái tri thức trước khi thực hiện câu hỏi tiếp theo: Trạng thái tri thức trước khi bắt đầu câu hỏi thứ $n+1$ sẽ được ngoại suy như sau:

$$P(L_{n+1}) = P(L_n \mid obs_n) + \big(1 - P(L_n \mid obs_n)\big) \cdot P(T) \quad (3)$$

Đối với mô hình BKT có yếu tố quên, thư viện pyBKT bổ sung thêm tham số $P(F)$ – xác suất HS quên kiến thức giữa hai lần thực hành liên tiếp. Đối với mô hình KT-IDEM, các tham số $P(S)$ và $P(G)$ được mô hình hóa riêng biệt cho từng câu hỏi, thay vì sử dụng chung cho toàn bộ KC như trong BKT tiêu chuẩn.

2.4.2. Chiến lược huấn luyện và đánh giá

Để đảm bảo tính khách quan và hạn chế hiện tượng quá khớp, chúng tôi áp dụng phương pháp kiểm định chéo 5 phần trong toàn bộ quá trình đánh giá. Cụ thể, dữ liệu làm bài của HS được chia ngẫu nhiên thành 5 phần bằng nhau theo đơn vị HS, đảm bảo tất cả lượt làm bài của cùng một HS chỉ nằm trong một phần. Tại mỗi vòng lặp, 4 phần được sử dụng để huấn luyện mô hình, phần còn lại được dùng để kiểm tra độ chính xác của dự đoán. Quy trình chuẩn bị dữ liệu và huấn luyện được thực thi trong môi trường Jupyter Notebook.

Ba biến thể mô hình được huấn luyện và đánh giá một cách độc lập trên cùng một cách chia dữ liệu: (i) Standard BKT – mô hình nền tảng với các tham số đồng nhất cho mỗi kĩ năng và câu hỏi; (ii) KT-IDEM – mô hình phân biệt độ khó theo từng câu hỏi; (iii) Forgetting BKT – mô hình có bổ sung cơ chế quên kiến thức theo thời gian. Do quy mô mẫu thử nghiệm (N=143) có thể tạo ra phương sai cá nhân lớn, mô hình Individualized BKT không được đưa vào đối sánh nhằm đảm bảo sự ổn định của quá trình ước lượng tham số. Sau khi hoàn tất quá trình đánh giá chéo, mô hình nào cho kết quả dự đoán chính xác nhất sẽ được chọn để huấn luyện lại trên toàn bộ tập dữ liệu nhằm trích xuất bộ tham số tối ưu cuối cùng.

2.4.3. Các độ đo đánh giá

Hiệu năng dự đoán của các mô hình được lượng hóa thông qua hai độ đo phổ biến trong các bài toán phân lớp nhị phân và đánh giá mô hình dự báo:

- Diện tích dưới đường cong ROC (AUC – Area Under the ROC Curve): Chỉ số này phản ánh khả năng phân biệt giữa các trường hợp trả lời đúng và trả lời sai của mô hình trên toàn bộ ngưỡng phân loại. Giá trị AUC càng gần 1 cho thấy mô hình có năng lực phân biệt càng tốt.

- Sai số bình phương trung bình (RMSE – Root Mean Square Error): Đo lường mức độ chênh lệch trung bình giữa xác suất trả lời đúng do mô hình dự đoán và kết quả quan sát thực tế (0 hoặc 1). RMSE càng gần 0 đồng nghĩa với mức độ chính xác trong dự báo của mô hình càng cao.

2.5. Đánh giá trạng thái tri thức

2.5.1. Quy trình cập nhật trạng thái tri thức

Sau khi các mô hình BKT được huấn luyện bằng thư viện pyBKT và thu được bộ tham số tối ưu ($P(L_0)$, $P(T)$, $P(S)$, $P(G)$) cho từng KC và câu hỏi, chúng tôi thực hiện quy trình cập nhật lại toàn bộ trạng thái tri thức của HS thông qua chức năng "Cập nhật tham số" của web app.

Sau khi bộ tham số mới được cập nhật, chúng tôi sử dụng chức năng "Cập nhật BKT" trong giao diện quản trị để tính toán lại độ thông thạo của tất cả HS đối với 11 KC của bài học.

2.5.2. Phương pháp so sánh và đối chứng

Để khẳng định tính ưu việt và độ tin cậy của việc đánh giá trạng thái tri thức dựa trên dữ liệu thực chứng, chúng tôi tiến hành so sánh đối chứng giữa ba phương pháp tiếp cận:

- Phương pháp tính điểm truyền thống: Dựa trên tỷ lệ số câu trả lời đúng trên tổng số câu hỏi thực hiện (số câu đúng/số câu đã thực hiện). Đây là cách tiếp cận "tĩnh", xem mọi câu hỏi có giá trị như nhau và không tính đến thứ tự câu hỏi, xác suất may rủi hay sơ suất của HS.

- Mô hình BKT với tham số cố định: Sử dụng bộ tham số mặc định ban đầu ($P(L_0) = 0,2$, $P(T) = 0,1$ và các giá trị $P(G)$, $P(S)$ cố định theo loại câu hỏi). Phương pháp này đã bắt đầu tính đến tiến trình học tập nhưng chưa phản ánh được đặc thù riêng của từng KC, từng câu hỏi.

- Mô hình BKT với tham số đã huấn luyện: Sử dụng bộ tham số được tối ưu hóa từ 11.493 lượt tương tác thực tế. Phương pháp này được kỳ vọng sẽ mang lại kết quả đánh giá sát thực nhất với năng lực tiềm ẩn của HS bằng cách triệt tiêu các yếu tố nhiễu từ việc đoán mò hoặc sơ suất dựa trên dữ liệu thực chứng.

Việc so sánh giữa các phương pháp này giúp làm rõ liệu việc huấn luyện mô hình có thực sự mang lại cái nhìn sâu sắc và chính xác hơn về tiến trình nhận thức của HS hay không, từ đó đưa ra các khuyến nghị về triển vọng áp dụng BKT trong dạy học CNH tại trường phổ thông.

3. Kết quả và thảo luận

3.1. Hiệu năng của các mô hình BKT

Sau khi ba mô hình được huấn luyện và kiểm định chéo trên bộ dữ liệu đã làm sạch gồm 11.493 lượt làm bài. Kết quả được tổng hợp như Bảng 2.

Bảng 2. Hiệu năng của các mô hình BKT trên dữ liệu thực tế

Mô hình	AUC	RMSE
Standard BKT	0,6262	0,3459
KT-IDEM	0,6591	0,3420
Forgetting BKT	0,6405	0,3447

Kết quả ở Bảng 2 cho thấy KT-IDEM là mô hình tối ưu nhất trên dữ liệu thực tế với $AUC = 0,6591$ và sai số $RMSE = 0,3420$. Biến thể Forgetting BKT ($AUC = 0,6405$) tuy cao hơn mô hình truyền thống nhưng không vượt qua được KT-IDEM. Điều này phù hợp với đặc thù của việc luyện tập ngắn hạn trong một chủ đề hẹp như Phương trình mặt cầu, nơi HS thường làm bài liên tục và hiện tượng "quên" kiến thức chưa tác động quá mạnh mẽ.

Mặc dù kết quả kiểm định không quá cao, nhưng với chỉ số $AUC \approx 0,66$, mô hình KT-IDEM vẫn cho thấy năng lực dự báo vượt xa ngưỡng ngẫu nhiên (0,5). Đây là một kết quả thực chứng rất có giá trị, có tiềm năng để sử dụng bộ tham số đã huấn luyện cho việc đánh giá lại trạng thái tri thức tiềm ẩn của HS. Sự nhất quán giữa AUC và RMSE củng cố niềm tin rằng việc mô hình hóa độ khó từng câu hỏi thực sự cải thiện chất lượng dự báo.

3.2. Trực quan hóa trạng thái tri thức của học sinh

3.2.1. Bộ tham số mô hình sau huấn luyện

Quá trình huấn luyện KT-IDEM cung cấp bốn nhóm tham số chính: $P(L_0)$, $P(T)$ cho mỗi KC và $P(S)$, $P(G)$ cho mỗi câu hỏi. Bảng 3 dưới đây trình bày kết quả đối với 11 KC, minh họa cho sự phân hóa rõ rệt về mức độ khó và tốc độ học của các KC.

Bảng 3. Tham số $P(L_0)$ và $P(T)$ của các KC do KT-IDEM ước lượng

KC	Tên kĩ năng	$P(L_0)$	$P(T)$
KC1	Tìm tâm và bán kính của phương trình mặt cầu dạng chính tắc	0,7736	0,0368
KC2	Tìm tâm và bán kính của phương trình mặt cầu dạng tổng quát	0,5495	0,0965
KC3	Nhận biết phương trình tổng quát của mặt cầu	0,5820	0,0809
KC4	Viết phương trình mặt cầu khi biết tâm và bán kính	0,7230	0,1712
KC5	Viết phương trình mặt cầu khi biết tâm và một điểm đi qua	0,6686	0,0618
KC6	Viết phương trình mặt cầu khi biết đường kính	0,7756	0,1504
KC7	Viết phương trình mặt cầu khi biết tâm và mặt phẳng tiếp xúc	0,6073	0,4267
KC8	Viết phương trình mặt cầu đi qua 4 điểm	0,8592	0,0016
KC9	Tìm vị trí tương đối của một điểm và mặt cầu	0,9617	0,0301
KC10	Tìm giao điểm của đường thẳng và mặt cầu	0,1442	0,0099
KC11	Bài toán thực tiễn về phương trình mặt cầu	0,0737	0,2712

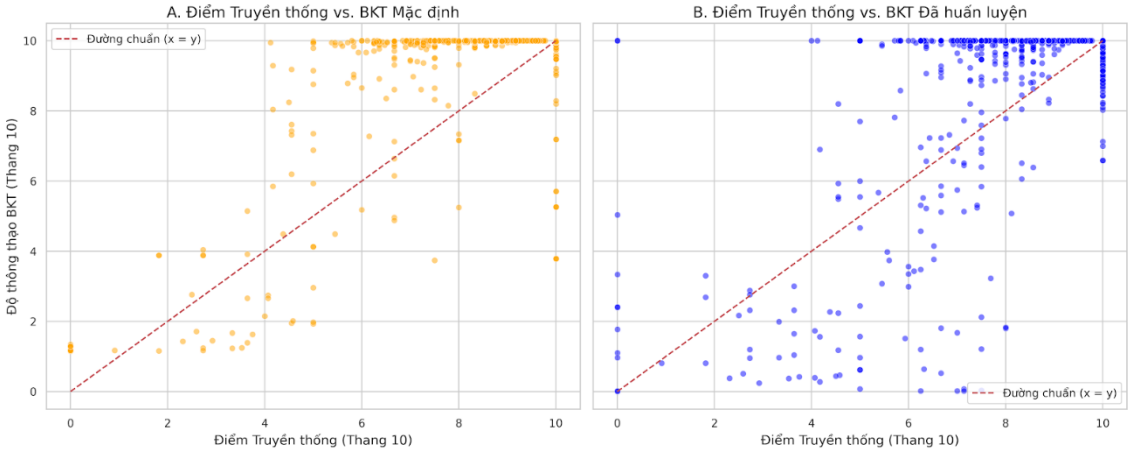
Số liệu từ Bảng 3 cho thấy sự tương hợp cao với thực tiễn dạy học: những kĩ năng nền tảng, thường gặp như KC1, KC8, KC9 có $P(L_0)$ khá cao, phản ánh việc HS đã có nền tảng tốt về phương trình đường tròn và phương pháp tọa độ. Ngược lại, các kĩ năng phức tạp, đòi hỏi vận dụng tổng hợp như KC10 hay KC11 có $P(L_0)$ rất thấp, cho thấy đa số HS khởi đầu ở trạng thái gần như chưa có kiến thức. Đặc biệt, $P(T)$ của KC7 (0,4267) cao vượt trội, ám chỉ rằng chỉ cần một vài lượt luyện tập, HS có thể nhanh chóng cải thiện năng lực viết phương trình mặt cầu khi biết tâm và mặt phẳng tiếp xúc. Ngược lại, $P(T)$ của KC8 và KC10 rất thấp, phản ánh đây là những kĩ năng "khó học", đòi hỏi nhiều cơ hội luyện tập hơn để đạt được sự thành thạo.

Đối với các tham số ở cấp độ câu hỏi, $P(G)$ trung bình trên toàn bộ ngân hàng câu hỏi là 0,5568, trong khi $P(S)$ trung bình là 0,1212. Các câu hỏi trắc nghiệm trả lời ngắn có $P(G)$ thấp nhất (trung bình 0,02), xác nhận hiệu quả của việc hạn chế đoán mò mà thiết kế này mang lại.

3.2.2. So sánh ba phương pháp đánh giá năng lực

Để phân tích sâu hơn sự chuyển biến về chất của thuật toán BKT sau khi được tối ưu hóa, Hình 2 trình bày biểu đồ phân tán đối chiếu điểm số truyền thống với độ thông thạo do mô hình đánh giá. Trục ngang là điểm truyền thống, trục dọc là độ thông thạo BKT (theo thang điểm 10), đường đứt nét màu đỏ là đường chuẩn (điểm truyền thống = điểm BKT).

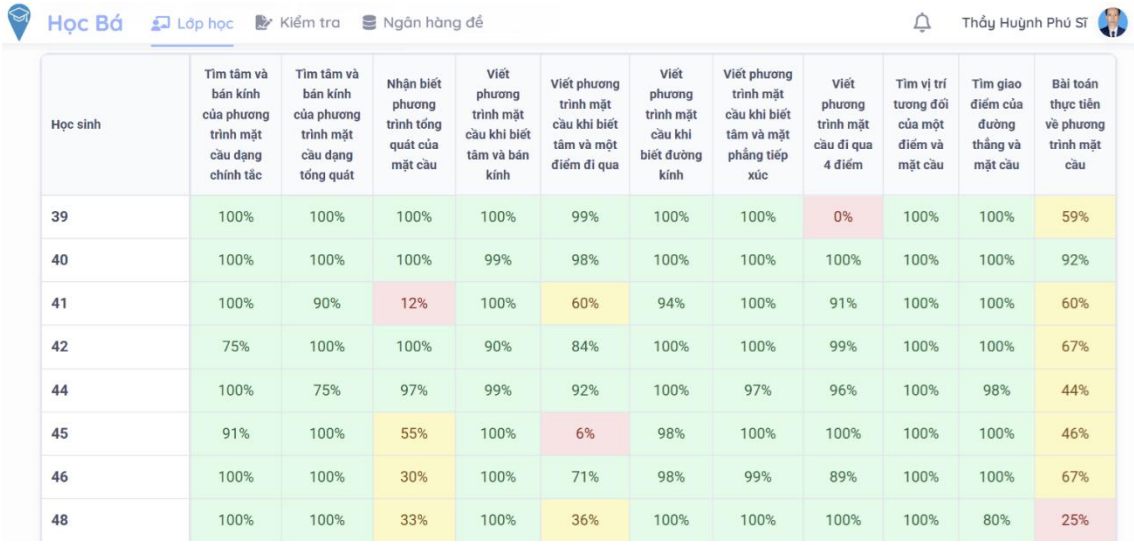
Các điểm nằm dưới đường nét đứt là HS có điểm truyền thống cao nhưng BKT đánh giá thấp, BKT đã phát hiện ra HS này có thể đang đoán mò hoặc làm đúng lúc đầu nhưng lại sai nhiều ở các câu sau. Các điểm nằm phía trên đường nét đứt là HS có điểm truyền thống thấp nhưng BKT lại đánh giá độ thông thạo rất cao. Chẳng hạn, HS làm sai 6 câu đầu nhưng đúng liên tiếp 4 câu cuối. BKT dự đoán HS đã học được và đã thành thạo, trong khi điểm truyền thống vẫn trừng phạt các lỗi sai trong quá khứ.



Hình 2. So sánh cách đánh giá giữa điểm truyền thống với BKT trước và sau khi huấn luyện. (A) So sánh với BKT Mặc định. (B) So sánh với BKT đã huấn luyện.

Quan sát hình 2A, có thể thấy các điểm phân bố theo những dải ngang rất cứng nhắc, đặc biệt là sự tập trung dày đặc ở mức thông thạo tối đa (10 điểm) dù điểm truyền thống của HS chỉ đạt mức 6-8 điểm. Điều này phản ánh nhược điểm của việc sử dụng bộ tham số mặc định. Ngược lại, hình 2B cho thấy một kết quả hoàn toàn khác. Sau khi tham số $P(G)$ và $P(S)$ được điều chỉnh dựa trên dữ liệu thực tế, các dải điểm ngang đã ít đi. Quần thể dữ liệu phân tán rộng hơn và bám sát đường chuẩn hơn, nhưng vẫn giữ được những độ lệch mang tính chẩn đoán. Điều này cho thấy việc áp dụng BKT không chỉ là cài đặt một thuật toán, mà là quá trình dùng dữ liệu hành vi để "dạy" hệ thống cách thấu hiểu sự phức tạp trong tư duy của HS.

Để minh họa cho kết quả đầu ra cuối cùng có ý nghĩa thực tiễn như thế nào, hình 3 dưới đây thể hiện giao diện "bản đồ năng lực" của một số HS lớp 12A2 (chúng tôi đã hiển thị mã HS thay cho tên HS) sau khi hệ thống cập nhật trạng thái tri thức bằng mô hình KT-IDEM đã huấn luyện.



Học sinh	Tìm tâm và bán kính của phương trình mặt cầu dạng chính tắc	Tìm tâm và bán kính của phương trình mặt cầu dạng tổng quát	Nhận biết phương trình tổng quát của mặt cầu	Viết phương trình mặt cầu khi biết tâm và bán kính	Viết phương trình mặt cầu khi biết tâm và một điểm đi qua	Viết phương trình mặt cầu khi biết đường kính	Viết phương trình mặt cầu khi biết tâm và mặt phẳng tiếp xúc	Viết phương trình mặt cầu đi qua 4 điểm	Tìm vị trí tương đối của một điểm và mặt cầu	Tìm giao điểm của đường thẳng và mặt cầu	Bài toán thực tiễn về phương trình mặt cầu
39	100%	100%	100%	100%	99%	100%	100%	0%	100%	100%	59%
40	100%	100%	100%	99%	98%	100%	100%	100%	100%	100%	92%
41	100%	90%	12%	100%	60%	94%	100%	91%	100%	100%	60%
42	75%	100%	100%	90%	84%	100%	100%	99%	100%	100%	67%
44	100%	75%	97%	99%	92%	100%	97%	96%	100%	98%	44%
45	91%	100%	55%	100%	6%	98%	100%	100%	100%	100%	46%
46	100%	100%	30%	100%	71%	98%	99%	89%	100%	100%	67%
48	100%	100%	33%	100%	36%	100%	100%	100%	100%	80%	25%

Hình 3. Giao diện trực quan hóa trạng thái tri thức của một số HS đối với 11 KC

Kết quả thống kê này cung cấp một công cụ chẩn đoán tức thời cho cả GV và HS. Thông qua đó, GV có thể thấy rõ sự phân hóa năng lực của HS một cách có hệ thống. Xu hướng chung của toàn bộ mẫu cho thấy đa số HS đã nắm vững các kĩ năng nền tảng như KC1, KC2, KC4, KC8 và KC9 (màu xanh lá). Ngược lại, cột KC11 và KC3 lại nổi bật với sắc cam và đỏ, phản ánh rằng đây là những "vùng phát triển" mà nhiều HS của lớp 12A2 cần tập trung ôn luyện. Sự khác biệt này là minh chứng rõ ràng cho thấy mô hình không đánh đồng mọi kĩ năng hay mọi câu trả lời, mà ước lượng xác suất thành thạo dựa trên chính hành vi thực tế.

4. Kết luận

Nghiên cứu đã triển khai thành công mô hình BKT và hai biến thể để mô hình hóa trạng thái tri thức của HS với chủ đề "Phương trình mặt cầu" (Toán 12) trên dữ liệu tương tác của 143 HS. Kết quả thực nghiệm cho thấy KT-IDEM có hiệu năng nhỉnh hơn Standard BKT và Forgetting BKT. Việc trực quan hóa trạng thái tri thức từ tham số đã huấn luyện cho thấy ưu điểm so với điểm truyền thống và mô hình tham số mặc định: phát hiện đoán mò/sơ suất, ghi nhận tiến bộ thực chất, và cung cấp "bản đồ năng lực" CNH giúp HS tự xác định lỗ hổng và GV thiết kế can thiệp sư phạm phù hợp. Nghiên cứu cũng đã xác định bộ tham số BKT đặc thù bao gồm $P(L_0)$, $P(T)$ cho 11 KC; $P(G)$, $P(S)$ cho 283 câu hỏi của chủ đề "Phương trình mặt cầu".

Việc phát hiện một số câu hỏi có $P(G)$ và $P(S)$ cao bất thường cho thấy các ngân hàng câu hỏi cần được chuẩn hóa chặt chẽ nhằm giảm thiểu tác động của đoán mò và sơ suất. Bên cạnh đó, AUC và RMSE đạt được phản ánh những thách thức từ dữ liệu thực tế như nhiễu hành vi, kích thước mẫu chưa đủ lớn, sự mất cân bằng giữa tỉ lệ đúng/sai. Các nghiên cứu tiếp theo cần mở rộng quy mô dữ liệu và phát triển hệ thống gợi ý thích ứng dựa trên trạng thái tri thức của HS.

Ghi chú về tác giả: *ThS Huỳnh Phú Sĩ là giáo viên Trường THCS và THPT Mỹ Thuận, tỉnh Vĩnh Long, Việt Nam. PGS.TS Lê Minh Cường là giảng viên cao cấp Khoa Sư phạm Toán – Tin, Trường Sư phạm, Trường Đại học Đồng Tháp, Việt Nam. Tác giả 1: xây dựng ý tưởng, phương pháp nghiên cứu, thực nghiệm, viết bản thảo; Tác giả 2: tư vấn, phản biện và chỉnh sửa.*

Tuyên bố về xung đột lợi ích: *Các tác giả tuyên bố không có xung đột lợi ích.*

TÀI LIỆU THAM KHẢO

Abdelrahman, G., Wang, Q., & Nunes, B. (2022). Knowledge Tracing: A Survey. *ACM Computing Surveys*, *55*(11), 1–37. https://doi.org/10.1145/3569576

Alkhatlan, A., & Kalita, J. (2018). *Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments*. http://arxiv.org/abs/1812.09628

Bulut, O., Shin, J., Yildirim-Erbasli, S. N., Gorgun, G., & Pardos, Z. A. (2023). An Introduction to Bayesian Knowledge Tracing with pyBKT. *Psych*, *5*(3), 770–786. https://doi.org/10.3390/psych5030050

Corbett, A. T., & Anderson, J. R. (1995). Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. *User Modeling and User-Adapted Interaction*, *4*, 253–278.

Dai, M., Hung, J., Du, X., Tang, H., & Li, H. (2021). Knowledge Tracing: A Review of Available Techniques. In *Journal of Educational Technology Development and Exchange* (Vol. 14, Number 2).

Desmarais, M. C., & Baker, R. S. J. d. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, *22*(1–2), 9–38. https://doi.org/10.1007/s11257-011-9106-8

Khajah, M., Lindsey, R. V., & Mozer, M. C. (2016). *How deep is knowledge tracing?* http://arxiv.org/abs/1604.02416

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction Framework: Bridging the Science-Practice Chasm to Enhance Robust Student Learning. *Cognitive Science*, *36*(5), 757–798. https://doi.org/10.1111/j.1551-6709.2012.01245.x

Nedungadi, P., & Remya, M. S. (2015). Incorporating forgetting in the Personalized, Clustered, Bayesian Knowledge Tracing (PC-BKT) model. *2015 International Conference on Cognitive Computing and Information Processing(CCIP)*, 1–5. https://doi.org/10.1109/CCIP.2015.7100688

Pardos, Z. A., & Heffernan, N. T. (2011). *KT-IDEM: Introducing Item Difficulty to the Knowledge Tracing Model* (pp. 243–254). https://doi.org/10.1007/978-3-642-22362-4_21

Phạm Hoàng An. (2025). Mô hình truy vết tri thức người học kết hợp động lực thời gian: phân tích khả năng dự đoán từ SAINT+ và dữ liệu EdNet. *Tạp Chí Khoa Học Trường Đại Học Mở Hà Nội*, 549. https://doi.org/10.59266/houjs.2025.811

Shimada, D., & Okada, K. (2025). Reliability Coefficient for Bayesian Knowledge Tracing Models. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-025-09829-7.

Yudelson, M. V., Koedinger, K. R., & Gordon, G. J. (2013). *Individualized Bayesian Knowledge Tracing Models* (pp. 171–180). https://doi.org/10.1007/978-3-642-39112-5_18