



TRẮC NGHIỆM THÍCH ỨNG TRÊN MÁY TÍNH: GIẢI PHÁP MỚI ĐÁNH GIÁ NĂNG LỰC THÍ SINH

ThS Nguyễn Thị Hiền

Viện Văn hóa Nghệ thuật Quốc gia Việt Nam, Bộ Văn hóa, Thể thao và Du lịch

“

“Phương pháp kiểm tra - đánh giá năng lực người học bằng bài thi trắc nghiệm thích ứng trên máy tính” là công trình nghiên cứu của PGS.TS Lê Thái Hưng và nhóm nghiên cứu gồm các giảng viên thuộc Khoa Quản trị Chất lượng, Trường Đại học Giáo dục, Đại học Quốc gia Hà Nội. Nhóm nghiên cứu đã dành nhiều năm để phát triển và hoàn thiện phương pháp kiểm tra đánh giá năng lực người học bằng bài thi trắc nghiệm thích ứng trên máy tính (CAT), góp phần thúc đẩy chuyển đổi số trong lĩnh vực giáo dục. Sản phẩm của nhóm nghiên cứu (Hệ thống đánh giá năng lực người học bằng bài thi trắc nghiệm thích ứng trên máy tính của Trường Đại học Giáo dục - UEd-CAT) đã được Cục Sở hữu Trí tuệ cấp Bằng độc quyền sáng chế.

”



Hệ thống đánh giá năng lực người học bằng bài thi trắc nghiệm thích ứng trên máy tính của Trường Đại học Giáo dục đã được Cục Sở hữu Trí tuệ cấp Bằng độc quyền sáng chế.

Bằng sáng chế độc quyền đầu tiên ở Việt Nam thuộc lĩnh vực khoa học giáo dục

Các hình thức kiểm tra truyền thống với các câu hỏi cố định trên giấy có hạn chế rõ rệt: tất cả thí sinh thường làm một bài kiểm tra với các câu hỏi giống nhau, có thể chỉ thay đổi thứ tự câu hỏi để tránh sự gian lận. Điều này khiến đề kiểm tra phải dài và nhiều câu hỏi mới có thể đánh giá đúng và phân loại được năng lực của tất cả thí sinh, từ thấp đến cao.

Trắc nghiệm thích ứng trên máy tính (Computerized Adaptive Testing - CAT) được biết đến với nhiều ưu thế nổi bật: đánh giá chính xác, đầy đủ và toàn diện về năng lực người học; cho phép phân tích các chỉ số về năng lực của thí sinh ngay sau khi họ trả lời câu hỏi và liên tục cập nhật



thông tin về năng lực của học sinh trong quá trình làm bài. Các bài CAT thường ngắn hơn một nửa so với bài kiểm tra truyền thống. Chỉ cần máy tính kết nối internet, CAT có thể đánh giá trên diện rộng với số lượng lớn học sinh tham gia.

Ra đời từ năm 1960, CAT đã phát triển mạnh mẽ trên thế giới. Tuy nhiên, tại Việt Nam, các nghiên cứu về trắc nghiệm thích ứng trên máy tính vẫn còn rất mới mẻ, chủ yếu tập trung vào cơ sở lý luận và khung lý thuyết, chưa xây dựng được cơ sở dữ liệu để sử dụng CAT trong việc đánh giá năng lực người học.

Với sự phát triển của trí tuệ nhân tạo (AI) và học máy (ML), CAT ngày càng cải thiện độ chính xác, tăng tính bảo mật và được sử dụng rộng rãi trong giáo dục.

Nhóm nghiên cứu gồm các giảng viên thuộc Khoa Quản trị Chất lượng, Trường Đại học Giáo dục, Đại học Quốc gia Hà Nội đã dành nhiều năm để phát triển và hoàn thiện phương pháp kiểm tra đánh giá năng lực người học bằng bài thi trắc nghiệm thích ứng trên máy tính. Quá trình nghiên cứu bao gồm: nghiên cứu, thử nghiệm trên nhiều đối tượng học sinh ở các cấp độ khác nhau, phân tích dữ liệu và điều chỉnh thuật toán để đảm bảo độ chính xác cao nhất. Những nỗ lực của nhóm tác giả đã được Cục Sở hữu Trí tuệ cấp Bằng độc quyền sáng chế số 40296 ngày 30/05/2024. Sáng chế này đề cập đến phương pháp đánh giá thích ứng năng lực người học trên máy tính. Để đo lường chính xác năng lực người học, nhóm sử dụng các thuật toán ước lượng trong toán học thống kê và các thuật toán học máy tăng cường. Sáng chế còn bao gồm quy trình xây dựng ngân hàng câu hỏi và triển khai đánh giá năng lực người học. Cho đến thời điểm này, đây là Bằng sáng chế độc quyền đầu tiên ở Việt Nam được cấp cho một sản phẩm thuộc lĩnh vực phương pháp giảng dạy và đánh giá trong giáo dục.

Nguyên lý hoạt động của CAT dựa trên lý thuyết khảo thí hiện đại (Item Response Theory - IRT). Xác suất trả lời đúng một câu hỏi phụ thuộc vào rất nhiều yếu tố: năng lực thực sự của thí sinh (kinh nghiệm, sự thông minh), các tham số liên quan tới câu hỏi như độ khó (ký hiệu là b), độ phân biệt của câu hỏi (ký hiệu là a), xác suất trả lời đúng ngẫu nhiên một câu hỏi (ký hiệu là c). Khi một bài kiểm tra thích ứng bắt đầu, câu hỏi đầu tiên sẽ hiển thị trên màn hình cho thí sinh thường sẽ là câu hỏi ở mức trung bình (ứng với năng lực thí sinh bằng 0 trên thang logit). Câu hỏi khởi tạo cũng có thể là một câu hỏi khó hoặc dễ nếu ta có thêm thông tin về năng lực thí sinh.

Nếu thí sinh trả lời đúng câu hỏi này tức là năng lực của thí sinh ở mức cao hơn mức năng lực của câu hỏi đề xuất, câu hỏi tiếp theo thí sinh nhận được sẽ là một câu hỏi khó hơn. Ngược lại thí sinh sẽ nhận được câu hỏi dễ hơn. Quá trình này cứ lặp đi lặp lại cho tới khi một tiêu chuẩn dừng (độ dài của câu hỏi, độ chính xác của năng lực ước lượng...) được thỏa mãn. Khi đó, năng lực thực sự của thí sinh sẽ được ước lượng dựa trên các câu hỏi mà thí sinh đã trả lời. Một bài CAT được chia thành 4 bước: bước khởi tạo, bước kiểm tra, bước dừng và bước kết thúc.

Ưu điểm nổi trội và ứng dụng của hệ thống

Dựa trên nguyên tắc hoạt động của CAT, hệ thống UEd-CAT đã kế thừa, phát triển và tối ưu hóa các phương pháp đánh giá năng lực thích ứng của người học. Điểm khác biệt và ưu điểm nổi bật của UEd-CAT bao gồm:

Là hệ thống trắc nghiệm thích ứng dựa trên năng lực của người học: Hệ thống UEd-CAT không giống với một số hệ thống trắc nghiệm thích ứng khác, vốn sử dụng các gói câu hỏi tổng hợp sẵn theo nội dung và trình độ mà người học lựa chọn. Thay vào đó, UEd-CAT phản hồi trực tiếp theo thời gian thực dựa trên thông tin mà thí sinh tương tác qua mỗi câu trả lời, từ đó tạo ra vô vàn bộ câu hỏi khác nhau, phù hợp nhất để đánh giá năng lực.

Là hệ thống trắc nghiệm thích ứng có tính tùy biến cao: Thuật toán ước lượng năng lực của UEd-CAT có thể điều chỉnh để phù hợp với mục đích kiểm tra, đánh giá. Các tham số về "tốc độ học" (learning rate) có thể được điều chỉnh nhằm kéo dài hoặc rút ngắn tương đối bài kiểm tra, phù hợp với các mục đích kiểm tra khác nhau như đánh giá thường xuyên hay tổng kết. Đây là điểm khác biệt của UEd-CAT so với các hệ thống đánh giá thích ứng khác trên thế giới, khi các hệ thống này thường giữ cấu trúc bài kiểm tra cố định về nội dung và số lượng câu hỏi giữa các thí sinh.

Hệ thống trắc nghiệm thích ứng UEd-CAT thực sự mang lại một bước đột phá trong việc đánh giá năng lực người học, đảm bảo tính chính xác và công bằng cao hơn trong quá trình kiểm tra.

Trong bối cảnh giáo dục đại chúng và phổ cập, việc các hệ thống giáo dục đủ năng lực về công nghệ và chuyên môn để tổ chức các kỳ thi diện rộng, kiểm tra được nhiều người cùng lúc là nhu cầu cấp thiết. Điều này thay thế cho các kỳ thi truyền thống vốn tốn kém, gây áp lực và có mức độ tin cậy, chính xác thấp. Hướng nghiên



Học sinh làm bài kiểm tra trên hệ thống đánh giá năng lực người học bằng bài thi trắc nghiệm thích ứng trên máy tính của Trường Đại học Giáo dục.

cứu ứng dụng này sẽ góp phần hiện thực hóa xu thế đổi mới dạy học và kiểm tra đánh giá người học theo hướng phát triển năng lực, cá nhân hóa trong môi trường học tập kết hợp và đổi mới thi trung học phổ thông ở Việt Nam.

Bộ công cụ đánh giá của hệ thống UEd-CAT đã được các chuyên gia trong lĩnh vực đo lường và đánh giá trong giáo dục nhận định: Hệ thống bài kiểm tra thích ứng đã đưa ra được bằng chứng thực nghiệm về việc áp dụng lý thuyết hồi đáp và phương pháp cân bằng trong xây dựng và chuẩn hoá ngân hàng câu hỏi trong lĩnh vực đo lường và đánh giá trong giáo dục, công trình này cũng được đánh giá là có thể chuyển giao cho các cơ sở đào tạo tiếng Anh hoặc được nhân rộng, tham khảo cho các nội dung đào tạo khác ngoài tiếng Anh.

Với những thành tựu đã đạt được, trong tương lai, hệ thống UEd-CAT có thể tiếp tục triển khai theo hai hướng:

Đối với nhà trường: UEd-CAT có thể sử dụng hệ thống hiện có để phát triển ngân hàng câu hỏi thích nghi cho nhiều lĩnh vực khác nhau, phục vụ các nhiệm vụ kiểm tra, đánh giá như thi tuyển sinh, đánh giá quá trình và đánh giá tổng kết. Hệ thống có thể phát triển hướng tới kết hợp với quá trình học tập thích nghi: xây dựng các bài kiểm tra theo mô đun (gói câu hỏi theo nhóm kỹ

năng hoặc năng lực, theo tiến độ học tập...) và phát triển hệ thống học tập đi kèm. Điều này giúp theo dõi kết quả học tập của học sinh và cá nhân hóa lộ trình đào tạo cho từng học sinh.

Đối với học sinh: UEd-CAT cung cấp cho học sinh khả năng theo dõi quá trình học tập của mình và đóng góp vào việc xây dựng chiến lược học tập tiến bộ. Hệ thống phần mềm trắc nghiệm thích ứng trên máy tính cũng là một môi trường học tập chủ động cho học sinh.

*

* *

Có thể khẳng định, UEd-CAT là sản phẩm nghiên cứu khoa học tiêu biểu; là minh chứng khẳng định tính đúng đắn trong chiến lược điều hướng sản phẩm nghiên cứu mang tính ứng dụng cao, có tính chất liên ngành và xuyên ngành; phù hợp với sứ mệnh và tầm nhìn của một đại học định hướng nghiên cứu; đáp ứng nhu cầu của giáo dục hiện đại và tiếp cận tiêu chuẩn quốc tế. Đây cũng là một trong những sản phẩm tiêu biểu, góp phần thúc đẩy chuyển đổi số nói chung và chuyển đổi số trong lĩnh vực giáo dục nói riêng.