

NGHIÊN CỨU VAI TRÒ CỦA BIG DATA TRONG VIỆC HIỂU VÀ TỐI ƯU HÓA HÀNH VI HỌC TẬP CỦA NGƯỜI HỌC

Nguyễn Thành Khoa
Trường Cao đẳng nghề Cần Thơ
Võ Văn Liệt
Trường Cao đẳng Cộng đồng Hậu Giang

Tóm tắt: Trong bối cảnh giáo dục số hóa, hệ thống Quản lý Học tập (LMS) tạo ra một lượng lớn “dữ liệu lớn” từ nhật ký tương tác của người học. Việc khai thác các log hành vi học tập này thông qua các kỹ thuật Learning Analytics (phân tích học tập) hiện đại, như phân cụm, phân loại, khai phá mẫu tuần tự và học sâu, giúp hiểu rõ hơn về thói quen học tập và cá thể hóa quá trình học. Bài báo này đề xuất một khung phân tích dữ liệu hành vi học tập từ log LMS bằng công nghệ Big Data. Trước hết, chúng điểm lại cơ sở lý thuyết về dữ liệu lớn trong giáo dục và mô hình Learning Analytics (từ mô tả, chẩn đoán đến dự đoán và định hướng can thiệp). Tiếp đến, chúng giới thiệu các kỹ thuật phân tích dữ liệu học tập đặc thù: phân cụm giúp nhóm người học theo kiểu hành vi, phân loại để dự đoán kết quả học tập (như phát hiện học sinh gặp khó), và khai phá mẫu tuần tự để nắm bắt chuỗi hành vi diễn ra theo thời gian. Học máy nâng cao (Machine Learning) và đặc biệt là học sâu (Deep Learning) đã được áp dụng thành công để dự đoán chính xác kết quả học tập, với các mô hình như mạng neural nhiều lớp hay CNN-LSTM đạt độ chính xác cao. Trong phần phương pháp, chúng trình bày một quy trình (pipeline) tổng thể xử lý log LMS – từ thu thập, tiền xử lý, đến trích xuất đặc trưng và phân tích (Hình 1) – nhằm minh họa cách ứng dụng Big Data và trí tuệ nhân tạo trong Learning Analytics. Kết quả phân tích tổng hợp cho thấy Big Data đóng vai trò nền tảng trong việc tối ưu hóa trải nghiệm học tập cá nhân, cho phép nhà trường can thiệp kịp thời (phát hiện sớm học sinh yếu) và xây dựng môi trường học tập thích ứng. Cuối cùng, bài báo đề xuất định hướng tích hợp AI và hệ thống học tập thích ứng (Adaptive Learning) dựa trên phân tích dữ liệu lớn để nâng cao kết quả giáo dục.

Từ khóa: Dữ liệu lớn, Hành vi học tập, Phân tích học tập, LMS, Phân cụm, Phân loại, Khai phá mẫu tuần tự, Học sâu.

RESEARCH ON THE ROLE OF BIG DATA IN UNDERSTANDING AND OPTIMIZING LEARNERS' LEARNING BEHAVIORS

Abstract: In the era of digital education, Learning Management Systems (LMS) generate massive big data from learners' interaction logs. Exploiting these logs through modern Learning Analytics techniques – such as clustering, classification, sequential pattern mining, and deep learning – enhances understanding of learning behaviors and enables personalization of the learning process. This paper proposes a framework for analyzing learner behavior from LMS logs using Big Data technologies. First, we review the theoretical foundations of Big Data in education and the Learning Analytics pipeline (from descriptive to prescriptive analytics). Then, we present specific analytical techniques for learning behavior data: clustering to group learners by behavior, classification to predict outcomes and identify struggling students, and sequential pattern mining to uncover temporal sequences of actions. Advanced Machine Learning, especially deep learning models (e.g., DNN, CNN-LSTM), have proven very effective, achieving high accuracy in performance prediction. In the methodology section, we describe an end-to-end pipeline for processing LMS logs – from data collection, preprocessing, feature extraction, to analytics (Figure 1) – illustrating how Big Data and AI are applied in Learning Analytics. The synthesis shows that Big Data is fundamental for optimizing personalized learning experiences, enabling timely interventions (early detection of at-risk students) and adaptive learning environments. Finally, we propose a Big Data-driven behavioral analysis framework and discuss future directions integrating AI and adaptive learning to improve educational outcomes.

Keywords: Big Data, learner behavior, learning analytics, LMS, clustering, classification, sequential pattern mining, deep learning.

Nhận bài: 27.12.2025

Phản biện: 19.01.2026

Duyệt đăng: 24.01.2026

I. ĐẶT VẤN ĐỀ

Hệ thống quản lý học tập trực tuyến (LMS) như Moodle, Canvas, EdX ghi nhận lượng dữ liệu khổng lồ (nhật ký log) phản ánh mọi tương tác của người học: từ việc xem bài giảng, truy cập tài nguyên, gửi bài tập đến tham gia diễn đàn. Những dữ liệu này không chỉ có kích thước lớn (độ “Volume” hàng TB–PB) mà còn đa dạng và phát sinh liên tục (độ “Velocity” cao, đa dạng về cấu trúc). Chúng vượt xa khả năng xử lý của các công cụ dữ liệu truyền thống. Theo Định nghĩa 5V, Big Data

trong giáo dục có các đặc tính về độ lớn, tính đa dạng, tốc độ cập nhật và giá trị tiềm năng.

Việc phân tích hành vi học tập từ các log này là rất cần thiết để cải thiện kết quả giảng dạy và học tập. Các nghiên cứu chỉ ra rằng mô hình học tập của sinh viên (chẳng hạn kiểu lướt tài liệu, thời gian học, số lượt làm bài kiểm tra) thường liên quan chặt chẽ đến hiệu quả học tập. Hơn nữa, trong các khóa học trực tuyến, khả năng dự đoán và can thiệp sớm đối với học sinh gặp khó mang

ý nghĩa then chốt, giúp tăng tỷ lệ hoàn thành và giảm bỏ học. Nhờ vậy, Big Data và Learning Analytics được xem là “cần câu” công nghệ để thúc đẩy giáo dục hướng dữ liệu: từ mô tả thực trạng (Descriptive) đến chẩn đoán vấn đề (Diagnostic), dự báo kết quả (Predictive) và cuối cùng là đề xuất giải pháp can thiệp (Prescriptive).

Tuy nhiên, thách thức là cần áp dụng các kỹ thuật phân tích dữ liệu tiên tiến để biến các log hỗn tạp thành tri thức hữu ích. Mục tiêu của bài viết này là nghiên cứu vai trò của Big Data trong việc hiểu và tối ưu hóa hành vi học tập của người học thông qua phân tích log LMS. Cụ thể, chúng tôi đề xuất một khung phân tích tổng thể bao gồm các bước từ thu thập, xử lý đến ứng dụng mô hình dữ liệu hành vi. Đồng thời, bài viết đi qua các nghiên cứu nổi bật liên quan đến phân cụm, phân loại, khai phá mẫu tuần tự và học sâu trong bối cảnh dữ liệu học tập. Trên cơ sở lý thuyết và tổng quan tài liệu quốc tế gần đây (2018–2025), chúng tôi làm rõ ý nghĩa khoa học và thực tiễn của mô hình đề xuất, cũng như hướng ứng dụng AI trong giáo dục thích ứng tương lai.

II. NỘI DUNG NGHIÊN CỨU

2.1. Mục tiêu nghiên cứu

Mục tiêu chung: Khảo sát và tổng hợp vai trò của công nghệ Big Data trong việc phân tích và tối ưu hóa hành vi học tập của người học trong môi trường LMS.

Mục tiêu cụ thể: Xác định các đặc tính, nguồn dữ liệu hành vi học tập (log LMS, clickstream, điểm số, v.v.) và thách thức khi xử lý chúng. Tổng hợp cơ sở lý thuyết liên quan đến Big Data trong giáo dục và các mô hình Learning Analytics (phân tích học tập). Trình bày và đánh giá các kỹ thuật phân tích dữ liệu hành vi học tập: phân cụm (clustering), phân loại (classification), khai phá mẫu tuần tự (sequential pattern mining), học sâu (deep learning), ... Đề xuất khung (pipeline) xử lý log LMS để phân tích hành vi học tập, bao gồm quy trình xử lý dữ liệu từ thu thập đến tạo mô hình. Định hướng ứng dụng AI (machine learning, adaptive learning) từ phân tích Big Data để cải thiện hiệu quả học tập.

2.2. Đối tượng nghiên cứu

Log LMS và Clickstream: Các sự kiện người dùng trên hệ thống LMS như lượt truy cập nội dung (page views), lượt xem video, lượt làm bài kiểm tra (quiz attempts), tham gia diễn đàn, v.v. “Mỗi click hay sự kiện đều được ghi lại kèm dấu thời gian, tạo nên chuỗi log chi tiết phản ánh quá trình học tập của từng cá nhân”.

Thông tin người học: Dữ liệu cấu trúc như

hồ sơ học sinh (thông tin nhân khẩu, lớp chuyên ngành), kết quả đánh giá (quiz, assignment, điểm cuối khóa) và thông tin mạng xã hội giáo dục. Khi kết hợp với lao động, các dữ liệu này tạo thành một bộ dữ liệu đa chiều về tiến trình học tập.

Dữ liệu phụ trợ khác: Ví dụ cảm biến (nếu có), phản hồi của giáo viên, tính năng hệ thống (ví dụ thời gian hoạt động), v.v. Trong phạm vi này, chủ yếu tập trung vào dữ liệu số thu được từ LMS mà không xét đến dữ liệu ngoại vi như hoạt động offline.

2.3. Ý nghĩa khoa học và thực tiễn

Về khoa học, việc nghiên cứu Big Data trong giáo dục và phân tích hành vi học tập góp phần phát triển các lý thuyết về học tập điện tử và phân tích học tập. Phân tích dữ liệu quy mô lớn cho phép phát hiện các quy luật hành vi và yếu tố ảnh hưởng đến hiệu quả học tập mà trước đây khó thu thập được. Ví dụ, việc áp dụng phân tích dự đoán giúp hiểu rõ mối liên hệ giữa các hành vi trong LMS với kết quả học tập. Nghiên cứu này còn đóng góp một khung lý thuyết tổng hợp về cách tiếp cận phân tích hành vi (từ mô tả đến định hướng can thiệp), và so sánh hiệu quả các thuật toán (clustering, classification, sequential mining, deep learning). Qua đó, kết quả này mở rộng hiểu biết về cách ứng dụng Big Data và AI trong bối cảnh giáo dục, đồng thời chỉ ra các hướng nghiên cứu mới (như tính công bằng trong phân tích, bảo vệ quyền riêng tư khi xử lý Big Data).

Về thực tiễn, kết quả phân tích học tập từ Big Data có tác động trực tiếp tới cải tiến giáo dục. Khi triển khai Learning Analytics, nhà trường và giáo viên có thể can thiệp kịp thời: ví dụ phát hiện sớm học sinh yếu hoặc có nguy cơ bỏ học dựa trên mô hình dự đoán, từ đó áp dụng biện pháp hỗ trợ (tư vấn, tăng tải nguyên học tập, điều chỉnh chương trình). Big Data còn tạo nền tảng cho việc cá nhân hoá trải nghiệm học tập – giới thiệu nội dung phù hợp với từng nhóm hành vi học tập nhờ mô hình phân cụm.

2.4. Cơ sở lý thuyết

Big Data trong giáo dục: Big Data trong giáo dục được định nghĩa bởi các đặc tính 5V: Volume (khối lượng), Velocity (tốc độ), Variety (đa dạng), Variability (tính biến động) và Value (giá trị). Trong LMS, dữ liệu được sinh ra liên tục từ hàng trăm đến hàng nghìn sự kiện mỗi ngày, với kích cỡ tổng thể có thể lên tới TB–PB. Bên cạnh đó, có nhiều loại dữ liệu (có cấu trúc, bán cấu trúc và phi cấu trúc), ví dụ text (nhật ký truy cập), video (ghi hình bài giảng), hình ảnh, âm thanh, kết quả đánh giá, v.v.

Điều này làm tăng độ phức tạp cho quá trình xử lý và khai phá thông tin. Tuy nhiên, chính tính khối lượng và đa dạng đó cũng mang lại giá trị lớn: “Volume lớn giúp phát hiện xu hướng chung và cá biệt; Variety và Variability giúp khám phá mối liên hệ giữa các loại dữ liệu đa phương tiện; còn Velocity đòi hỏi hệ thống phân tích phải xử lý theo thời gian thực hoặc gần thực”.

Vì thế, Big Data Analytics (BDA) – tức các kỹ thuật và hệ thống chuyên biệt – trở thành nền tảng cho Learning Analytics. BDA bao gồm kho dữ liệu (Data Warehouse, Data Lake), kỹ thuật xử lý (Hadoop MapReduce, Spark), thuật toán khai phá dữ liệu (Data Mining, ML/DL) và trực quan hoá. Khi áp dụng cho giáo dục, Big Data Analytics hỗ trợ thu thập và xử lý hiệu quả lượng dữ liệu học tập khổng lồ, tạo điều kiện cho việc xây dựng mô hình thông minh và ra quyết định dựa trên dữ liệu.

Mô hình Learning Analytics: Learning Analytics (LA) là lĩnh vực sử dụng dữ liệu lớn để hiểu và cải thiện trải nghiệm học tập. Theo tiếp cận phân tích dữ liệu, LA thường được chia thành bốn loại chính:

- *Phân tích mô tả (Descriptive)*: Tóm tắt các dữ liệu đã ghi nhận, ví dụ báo cáo số lượng lượt truy cập, điểm trung bình, mức độ tham gia.

- *Phân tích chẩn đoán (Diagnostic)*: Xác định nguyên nhân hoặc mối quan hệ (ví dụ tại sao học sinh trễ hạn nộp bài, mối liên hệ giữa tần suất đăng nhập và điểm số).

- *Phân tích dự đoán (Predictive)*: Xây dựng mô hình máy học để dự đoán kết quả tương lai, ví dụ xác suất sinh viên trượt môn hoặc bỏ học. Một số nghiên cứu đã dùng ML/DL để dự đoán học sinh yếu sớm nhất ở giai đoạn đầu của khóa học.

- *Phân tích định hướng (Prescriptive)*: Đề xuất hành động can thiệp (ví dụ khuyến nghị tài nguyên học tập, thiết kế lộ trình học tùy chỉnh) dựa trên kết quả phân tích.

“Theo Belrzaeg et al., phân tích học tập được tổ chức theo một pipeline: đầu tiên mô tả dữ liệu, sau đó chẩn đoán vấn đề, tiếp đến dự đoán kết quả và cuối cùng là đề xuất giải pháp.” Đây là mô hình lý thuyết quan trọng trong Learning Analytics, nhấn mạnh vai trò của Big Data ở mỗi giai đoạn – từ việc thu thập và làm sạch dữ liệu đến ứng dụng mô hình trí tuệ nhân tạo.

Kỹ thuật phân tích dữ liệu hành vi học tập: Các kỹ thuật phổ biến trong Learning Analytics được xây dựng dựa trên Machine Learning và Data Mining để xử lý các tập dữ liệu lớn từ LMS:

- *Phân cụm (Clustering)*: Thuật toán phân nhóm người học có hành vi tương đồng dựa trên đặc trưng

log. Ví dụ, Franke và cộng sự (2024) đã áp dụng nhiều phương pháp phân cụm để xác định nhóm học sinh với kiểu tương tác LMS khác nhau. Kết quả cho thấy có thể xác định những cluster người học với độ tương tác cao/ thấp khác nhau, giúp xây dựng các lộ trình học tập cá nhân hoá.

- *Phân loại (Classification) và Dự đoán*: Sử dụng các mô hình học có giám sát (ví dụ Decision Tree, SVM, Neural Network) để dự đoán kết quả học tập (như điểm số, nguy cơ trượt) dựa trên dữ liệu log lịch sử. Riestra-González et al. (2021) đã tạo ra các mô hình phân loại để phát hiện sớm nhóm “cần hỗ trợ” (at-risk), nhóm “xuất sắc” và nhóm “bị điểm kém” chỉ dựa vào log tại các giai đoạn đầu của khóa học. Học sinh được gắn nhãn (trượt/đỗ) dựa vào điểm kiểm tra, và các thuật toán (DecTree, Naive Bayes, MLP...) được huấn luyện để dự đoán nhãn này. Kết quả cho thấy MLP (Multi-layer Perceptron) đạt độ chính xác cao nhất (tới 90% khi dự đoán sau 50% khóa học). Tương tự, Dobashi và cộng sự (2022) phân loại “mẫu học tập” của sinh viên qua log Moodle nhằm phát hiện sớm những sinh viên cần can thiệp.

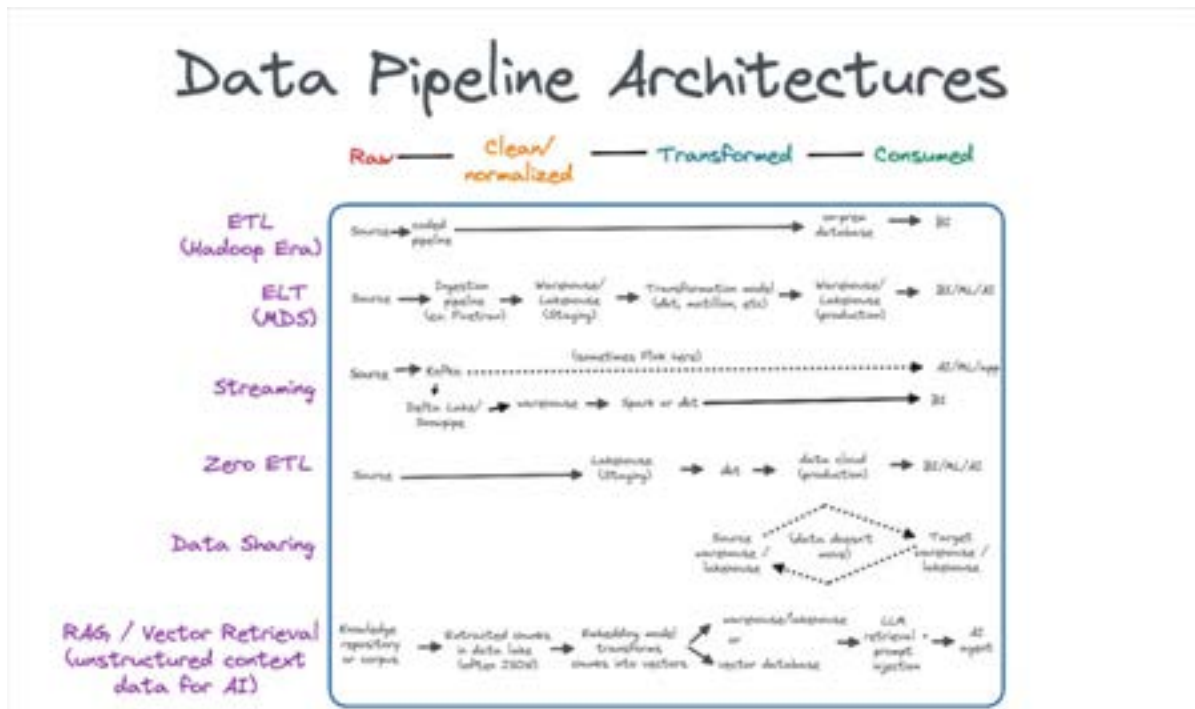
- *Khai phá mẫu tuần tự (Sequential Pattern Mining)*: Mục tiêu tìm ra các chuỗi hành vi lặp lại quan trọng trong dữ liệu thời gian. Đây là cách để nắm bắt thứ tự và cấu trúc tiềm ẩn trong các hành động học tập. Ví dụ, Deeva et al. (2022) đã chứng minh rằng các mẫu tuần tự hành vi của học sinh có thể dùng để dự đoán học sinh không đạt kết quả học tập ở giai đoạn đầu khóa học. Kỹ thuật này cho phép xây dựng mô hình “phân loại chuỗi” (sequence classification) để phát hiện học sinh yếu mà không cần dựa vào thông tin cá nhân nhạy cảm.

- *Học sâu (Deep Learning)*: Các mạng neural sâu (DNN), CNN, RNN/LSTM hay kết hợp CNN-LSTM đã được sử dụng để trích xuất đặc trưng phức tạp và dự đoán chính xác. Trong môi trường học tập trực tuyến, học sâu giúp xử lý các tập dữ liệu khổng lồ, bao gồm cả tín hiệu thời gian hoặc đa phương tiện. Tổng quan hệ thống của Alnasyan et al. (2024) cho thấy DNN và mô hình CNN-LSTM phổ biến nhất và đạt độ chính xác >90% trong dự đoán kết quả học tập. Nghiên cứu cũng ghi nhận đặc trưng hành vi học tập (như mức độ tương tác, tham gia diễn đàn) thường là yếu tố dự báo quan trọng nhất. Học sâu kết hợp với khả năng giải thích (XAI) đang được quan tâm để làm cho các mô hình này có thể tin cậy và dễ hiểu hơn trong giáo dục.

2.5. Phương pháp nghiên cứu

Chúng tôi thực hiện tổng quan hệ thống các công trình đã công bố quốc tế (2018–2025) về Big Data và Learning Analytics trong giáo dục, tập trung vào các bài báo tạp chí có DOI. Việc tìm kiếm được thực hiện trên các cơ sở dữ liệu học thuật (Scopus, IEEE Xplore, ScienceDirect, Google Scholar) với từ khoá liên quan đến Big

Data, Learning Analytics, LMS log, clustering, classification, sequence mining, deep learning. Sau đó, tài liệu thu thập được được sàng lọc theo tiêu chí mức độ phù hợp và độ tin cậy. Phương pháp PRISMA (Preferred Reporting Items for Systematic Reviews) được tham khảo để đảm bảo quá trình chọn lọc minh bạch.



Hình 1. Mô hình pipeline xử lý log học tập từ LMS (mô hình tổng quát).

Pipeline gồm các giai đoạn: (i) Thu thập dữ liệu từ LMS (log clickstream, điểm số, hồ sơ học sinh); (ii) Tiền xử lý (lọc nhiễu, chuẩn hóa dữ liệu, phân tách session); (iii) Trích xuất đặc trưng (tính toán các chỉ số hành vi, rút trích trình tự hành động); (iv) Phân tích mô hình (phân cụm người học, phân loại kết quả, khai phá mẫu tuần tự, áp dụng mô hình học sâu); và (v) Xuất kết quả dưới dạng bảng điều khiển (dashboards) hoặc đề xuất can thiệp. Mỗi bước đều có thể ứng dụng các công cụ Big Data (ví dụ Spark) và thuật toán ML/DL tương ứng, nhằm tối ưu hoá hiệu năng và độ chính xác phân tích. Hình 1 minh họa cách thức một hệ thống Learning Analytics sử dụng Big Data pipeline để hiểu và tối ưu hoá hành vi học tập của người học.

Như vậy, bài báo kết hợp phân tích mô hình (model analysis) với tổng quan lý thuyết. Cụ thể, sau khi giới thiệu pipeline (Hình 1), các đoạn sau sẽ trình bày chi tiết về cách các kỹ thuật phân tích (clustering, classification, pattern mining, deep learning) được áp dụng trong từng giai đoạn và dẫn chứng qua các nghiên cứu điển hình. Chúng tôi

cũng thảo luận về các công cụ Big Data (hadoop, spark, Hive, TensorFlow, v.v.) phù hợp để triển khai pipeline này, và nhấn mạnh các chỉ tiêu hiệu quả (độ chính xác dự đoán, tốc độ xử lý).

III. KẾT LUẬN

Bài báo đã trình bày một nghiên cứu tổng quan về vai trò của Big Data trong hiểu và tối ưu hoá hành vi học tập dựa trên log hệ thống LMS. Kết quả khẳng định Big Data và Learning Analytics là chìa khóa cho việc chuyển đổi số trong giáo dục: nhờ khả năng xử lý và phân tích dữ liệu khổng lồ, ta có thể phát hiện kịp thời các xu hướng và vấn đề trong quá trình học tập. Việc ứng dụng kỹ thuật phân tích đa dạng (phân cụm, phân loại, khai phá mẫu tuần tự, học sâu) giúp khai thác hiệu quả các log hành vi, từ đó xây dựng các giải pháp giáo dục dữ liệu định hướng (data-driven). Bài báo cũng đã đề xuất một khung pipeline minh họa quy trình xử lý log LMS (Hình 1), làm cơ sở cho việc triển khai thực tế các hệ thống Learning Analytics, từ khâu chuẩn bị dữ liệu đến đầu ra là các công cụ giám sát và khuyến nghị.

Bên cạnh đó, chúng đề xuất hướng phát triển

tương lai: tích hợp AI tiên tiến vào pipeline và phát triển các hệ thống học tập thích ứng (Adaptive Learning). Ví dụ, trí tuệ nhân tạo có thể tự động điều chỉnh nội dung học phù hợp với kiểu hành vi vừa được phân cụm, hoặc hệ thống sẽ đề xuất bài học bổ sung khi dự báo thấy học sinh yếu sắp gặp khó. Các công nghệ giải thích mô hình (Explainable AI) cũng nên được đưa vào để đảm bảo kết quả phân tích dễ hiểu với giảng viên và học sinh.

Tóm lại, dữ liệu lớn cho phép giáo dục trở nên cá nhân hoá và dữ liệu hóa hơn bao giờ

hết. Việc xây dựng khung phân tích hành vi như đã trình bày sẽ là tiền đề để các trường đại học và tổ chức giáo dục phát triển hệ thống hỗ trợ quyết định dựa trên dữ liệu học tập. Trong tương lai, xu hướng AI và học tập thích ứng sẽ được đẩy mạnh: hệ thống tự động điều chỉnh tiến trình học theo tình trạng học tập và tương tác của từng cá nhân. Những định hướng này hứa hẹn giúp tối ưu hoá kết quả học tập, nâng cao trải nghiệm người học và đáp ứng mục tiêu giáo dục bền vững.

TÀI LIỆU THAM KHẢO

- Adam, K., Abu Bakar, N. A., Ibrahim Fakhreldin, M. A., & Abdul Majid, M. (2018). Big Data and Learning Analytics: A Big Potential to Improve e-Learning. *Journal of Computational and Theoretical Nanoscience*, 24(10), 7838–7843. <https://doi.org/10.1166/asl.2018.13028>
- Riestra-González, M., del Puerto Paule-Ruíz, M., & Ortín, F. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Computers & Education*, 163, 104108. <https://doi.org/10.1016/j.compedu.2020.104108>
- Dobashi, K., Ho, C. P., Fulford, C. P., Lin, M.-F. G., & Higa, C. (2022). Learning pattern classification using Moodle logs and the visualization of browsing processes by time-series cross-section. *Computers & Education: Artificial Intelligence*, 3, 100105. <https://doi.org/10.1016/j.caeai.2022.100105>
- Deeva, G., De Smedt, J., Saint-Pierre, C., Galbraith, R. D., Vanthienen, J., & Verhaest, I. (2022). Predicting student performance using sequence classification with time-based windows. *Expert Systems with Applications*, 209, 118182. <https://doi.org/10.1016/j.eswa.2022.118182>
- Alnasyan, B., Basher, M., & Alassafi, M. (2024). The power of Deep Learning techniques for predicting student performance in Virtual Learning Environments: A systematic literature review. *Computers & Education: Artificial Intelligence*, 6, 100231. <https://doi.org/10.1016/j.caeai.2024.100231>
- Franke, S., Trommer, M., Hellbach, S., & Teich, T. (2024). Clustering of Learners' Data – A Case Study on Learning Analytics. *IFAC-PapersOnLine*, 58(26), 118–123. <https://doi.org/10.1016/j.ifacol.2024.10.281>
- Boujmiraz, S., Darhmaoui, H., & Drissi El Maliani, A. (2026). Predicting student performance: A comprehensive review of machine learning, deep learning, and explainable AI approaches. *Computers & Education: Artificial Intelligence*, 6, 100548. <https://doi.org/10.1016/j.caeai.2026.100548>
- Du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(11), e39630. <https://doi.org/10.1016/j.heliyon.2024.e39630>