

ỨNG DỤNG DỮ LIỆU LỚN (BIG DATA) TRONG ĐÁNH GIÁ KẾT QUẢ HỌC TẬP CỦA SINH VIÊN

Trần Chí Hiếu
Trường Đại học Mở Thành phố Hồ Chí Minh
Huỳnh Thị Ngọc Diễm
Trường THCS Đồng Đen, Thành phố Hồ Chí Minh
Cao Hồng Vân
Trường Cao đẳng Nghề Thành phố Hồ Chí Minh

Tóm tắt: Trong thời đại số, Dữ liệu lớn (Big Data) đang cách mạng hóa nhiều lĩnh vực, trong đó có giáo dục đại học. Truyền thống đánh giá kết quả học tập sinh viên chủ yếu dựa vào điểm số và bài kiểm tra, nhưng với Big Data, quá trình này trở nên toàn diện, khách quan và cá nhân hóa hơn. Bài viết này phân tích ứng dụng Big Data trong đánh giá kết quả học tập, từ việc thu thập dữ liệu hành vi học tập, phân tích tương tác trên nền tảng số đến dự đoán hiệu suất học tập bằng trí tuệ nhân tạo (AI). Đồng thời, bài viết cũng làm rõ thách thức và đề xuất giải pháp triển khai Big Data trong giáo dục, hướng đến một mô hình đánh giá linh hoạt và hiệu quả hơn.

Từ khóa: Big Data, giáo dục, đánh giá học tập, trí tuệ nhân tạo, học tập cá nhân hóa

APPLICATION OF BIG DATA IN ASSESSING STUDENT LEARNING OUTCOMES

Tran Chi Hieu
Ho Chi Minh City Open University
Huynh Thi Ngoc Diem
Dong Den Secondary School, Ho Chi Minh City
Cao Hong Van
Ho Chi Minh City Vocational College

Abstract: In the digital age, Big Data is revolutionizing various fields, including higher education. Traditional student assessment methods rely mainly on grades and exams, but with Big Data, the evaluation process becomes more comprehensive, objective, and personalized. This paper explores the applications of Big Data in student performance assessment, from analyzing learning behaviors and digital interactions to predicting academic outcomes using artificial intelligence (AI). Additionally, it highlights the challenges and proposed solutions for implementing Big Data in education, aiming for a more adaptive and effective assessment model.

Keywords: Big Data, education, student assessment, artificial intelligence, personalized learning

Nhận bài: 20/02/2025

Phản biện: 20/03/2025

Duyệt đăng: 25/03/2025

I. ĐẶT VẤN ĐỀ

Theo nghiên cứu của Trần Thanh Điện và cộng sự (2023), việc ứng dụng kỹ thuật xử lý dữ liệu lớn giúp tăng tốc tìm kiếm tài nguyên học tập theo nội dung, cải thiện hiệu quả học tập của sinh viên. Ngoài ra, việc khai thác Big Data trong giáo dục và đào tạo du lịch cũng góp phần nâng cao năng lực quản trị và cải thiện chất lượng đào tạo (Vũ, 2020).

Tuy nhiên, việc triển khai Big Data trong giáo dục đối mặt với nhiều thách thức, bao gồm cơ sở hạ tầng công nghệ, bảo mật dữ liệu và kỹ năng số của giảng viên. Những thách thức này đòi hỏi các cơ sở giáo dục cần có chiến lược và giải pháp phù hợp để tận dụng tối đa lợi ích mà Big Data mang lại trong việc đánh giá và nâng cao chất lượng học tập của sinh viên.

II. NỘI DUNG NGHIÊN CỨU

2.1. Tổng quan về Big Data trong giáo dục

Big Data, hay Dữ liệu lớn, là thuật ngữ dùng để chỉ các tập dữ liệu có dung lượng khổng lồ, tốc độ tạo ra nhanh và có độ đa dạng cao, vượt quá khả năng xử lý của các hệ thống quản lý dữ liệu truyền thống. Đặc trưng của Big Data được mô tả qua mô hình 5V, bao gồm Volume (khối lượng dữ liệu lớn), Velocity (tốc độ tạo và xử lý nhanh), Variety (đa dạng về nguồn dữ liệu), Veracity (độ tin cậy của dữ liệu) và Value (giá trị khai thác) (Gandomi & Haider, 2015). Trong bối cảnh giáo dục hiện đại, sự gia tăng nhanh chóng của các nền tảng học tập trực tuyến, hệ thống quản lý học tập (LMS – Learning Management System), và các công nghệ hỗ trợ giảng dạy đã tạo ra một lượng dữ liệu khổng lồ, đặt ra nhu cầu sử dụng Big Data để tối ưu hóa quá trình dạy và học.

Ứng dụng của Big Data trong giáo dục chủ yếu xoay quanh hai lĩnh vực quan trọng: Phân tích học

tập (Learning Analytics) và khai phá dữ liệu giáo dục (Educational Data Mining – EDM). Phân tích học tập tập trung vào việc thu thập, xử lý và diễn giải dữ liệu về hành vi học tập của sinh viên để cải thiện quá trình giảng dạy, trong khi khai phá dữ liệu giáo dục sử dụng các phương pháp tính toán để phát hiện các mô hình và xu hướng trong dữ liệu học tập nhằm hỗ trợ ra quyết định (Siemens & Long, 2011). Những công nghệ này cho phép các trường đại học theo dõi tiến trình học tập của sinh viên theo thời gian thực, xác định các yếu tố ảnh hưởng đến kết quả học tập và điều chỉnh nội dung giảng dạy một cách linh hoạt.

Tại Việt Nam, việc ứng dụng Big Data trong giáo dục vẫn đang trong giai đoạn thử nghiệm, nhưng đã có những bước tiến đáng kể. Nhiều trường đại học đã bắt đầu tích hợp LMS và sử dụng các công cụ phân tích dữ liệu để đánh giá kết quả học tập của sinh viên (Nguyễn & Trần, 2021). Bộ Giáo dục và Đào tạo cũng đã đưa ra các chính sách khuyến khích chuyển đổi số trong giáo dục đại học, trong đó nhấn mạnh việc tận dụng dữ liệu lớn để nâng cao chất lượng giảng dạy và quản lý đào tạo.

Nhìn chung, Big Data đã và đang tạo ra những thay đổi quan trọng trong giáo dục, từ việc thu thập dữ liệu, phân tích xu hướng đến hỗ trợ ra quyết định trong giảng dạy. Việc nghiên cứu tổng quan về Big Data không chỉ giúp hiểu rõ hơn về bản chất của công nghệ này mà còn là nền tảng quan trọng để đánh giá những ứng dụng tiềm năng và định hướng phát triển trong tương lai.

2.2. Ứng dụng Big Data trong đánh giá kết quả học tập của sinh viên

Trong bối cảnh giáo dục hiện đại, việc áp dụng Big Data đã mở ra những cơ hội mới trong việc đánh giá và nâng cao chất lượng học tập của sinh viên. Thông qua việc thu thập và phân tích dữ liệu từ nhiều nguồn khác nhau, các cơ sở giáo dục có thể hiểu rõ hơn về hành vi học tập, từ đó đưa ra các chiến lược giảng dạy hiệu quả hơn.

Một trong những ứng dụng quan trọng của Big Data trong giáo dục là khả năng phân tích hành vi học tập của sinh viên. Bằng cách theo dõi các hoạt động như thời gian truy cập tài liệu, tần suất tham gia thảo luận trực tuyến và kết quả các bài kiểm tra, các hệ thống phân tích có thể xác định được

mô hình học tập của từng cá nhân.

Ngoài ra, Big Data còn hỗ trợ việc phát hiện sớm các vấn đề trong học tập. Thông qua việc phân tích dữ liệu, các hệ thống có thể nhận diện những sinh viên có nguy cơ gặp khó khăn, từ đó đưa ra cảnh báo và đề xuất các biện pháp hỗ trợ kịp thời.

Tuy nhiên, việc áp dụng Big Data trong giáo dục cũng đặt ra những thách thức về bảo mật và quyền riêng tư. Việc thu thập và xử lý lượng lớn dữ liệu cá nhân đòi hỏi các biện pháp bảo vệ nghiêm ngặt để đảm bảo thông tin của sinh viên không bị lạm dụng. Do đó, các cơ sở giáo dục cần tuân thủ các quy định về bảo vệ dữ liệu và triển khai các công nghệ bảo mật tiên tiến.

Tổng kết lại, việc ứng dụng Big Data trong đánh giá kết quả học tập của sinh viên không chỉ giúp nâng cao chất lượng giáo dục mà còn tạo điều kiện cho việc cá nhân hóa trải nghiệm học tập. Tuy nhiên, để tận dụng tối đa lợi ích này, các cơ sở giáo dục cần chú trọng đến việc bảo mật dữ liệu và đào tạo nhân lực có kỹ năng phân tích dữ liệu.

2.3. Thách thức của Big Data trong đánh giá học tập

Thứ nhất, bảo mật và quyền riêng tư là một thách thức lớn khi áp dụng Big Data trong giáo dục. Việc thu thập và lưu trữ lượng lớn dữ liệu cá nhân của sinh viên đòi hỏi các biện pháp bảo vệ nghiêm ngặt để tránh rò rỉ thông tin và đảm bảo quyền riêng tư. Theo nghiên cứu của Abdalla (2022), việc sử dụng thông tin cá nhân không đúng mục đích, đặc biệt khi các thông tin đó được tổng hợp từ nhiều nguồn khác nhau, gây ra nhiều mối lo ngại về bảo mật và quyền riêng tư.

Thứ hai, chất lượng và độ tin cậy của dữ liệu là một vấn đề quan trọng. Dữ liệu không chính xác hoặc lỗi thời có thể dẫn đến kết quả phân tích sai lệch, ảnh hưởng tiêu cực đến quyết định giáo dục. Như được đề cập trong bài viết trên ICT Vietnam, sự thay đổi về chất lượng hoặc dữ liệu cũ và lỗi thời có thể tạo ra kết quả ngoài ý muốn và phá vỡ toàn bộ ứng dụng (ICT Vietnam, 2020).

Thứ ba, khả năng phân tích và xử lý dữ liệu còn hạn chế trong nhiều cơ sở giáo dục. Việc thiếu nguồn lực và chuyên môn để khai thác hiệu quả

Big Data có thể dẫn đến việc dữ liệu không được sử dụng tối ưu hoặc bị lãng phí. Việc tổ chức lưu trữ, xử lý và phân tích các tập dữ liệu lớn tiêu tốn rất nhiều thời gian do độ phức tạp tính toán lớn và đặt ra các yêu cầu rất cao về các nền tảng phần cứng, phần mềm với các kỹ thuật hỗ trợ tiên tiến (Học viện Nông nghiệp Việt Nam, 2021).

Thứ tư, tính tổng quát của dữ liệu có thể gây bất lợi cho khả năng ứng dụng của nó. Dữ liệu được rút ra từ một tập hợp con của dân số hoặc lực lượng lao động có thể không phù hợp với các sinh viên được nhắm mục tiêu. Hơn nữa, các giải pháp học tập trí tuệ nhân tạo đã được sử dụng cho sinh viên trong một trường đại học hàng đầu có thể không có kết quả và mức độ phù hợp như nhau đối với sinh viên trong một trường cao đẳng cộng đồng ở nơi khác (ICT Vietnam, 2020).

Thứ năm, chi phí triển khai và duy trì hệ thống Big Data có thể cao, đặc biệt đối với các cơ sở giáo dục có ngân sách hạn chế. Việc đầu tư vào hạ tầng công nghệ, phần mềm và đào tạo nhân lực đòi hỏi nguồn lực tài chính đáng kể. Như được đề cập trong bài viết trên PACE, việc thu thập, bảo quản và phân tích những bộ dữ liệu lớn để khai phá thông tin sâu rộng, giá trị đã trở nên khả thi về mặt kỹ thuật và kinh tế nhờ vào công nghệ Big Data; tuy nhiên, điều này cũng đặt ra thách thức về chi phí (PACE, 2020).

2.4. Giải pháp triển khai Big Data trong đánh giá học tập

Thứ nhất, cần áp dụng các tiêu chuẩn bảo mật dữ liệu chặt chẽ. Cụ thể, việc mã hóa dữ liệu, thiết lập quyền truy cập hạn chế và sử dụng các giao thức bảo mật như xác thực hai yếu tố (2FA) sẽ giúp ngăn chặn rò rỉ thông tin sinh viên. Đồng thời, các cơ sở giáo dục cần tuân thủ các quy định về bảo vệ dữ liệu, chẳng hạn như GDPR (General Data Protection Regulation) hoặc các tiêu chuẩn tương đương trong nước, đảm bảo rằng thông tin của sinh viên chỉ được sử dụng cho mục đích học tập và nghiên cứu.

Thứ hai, xây dựng hệ thống chuẩn hóa dữ liệu nhằm giảm thiểu sai sót trong quá trình thu thập, xử lý. Việc áp dụng các phương pháp lọc và làm sạch dữ liệu trước khi đưa vào phân tích sẽ giúp tăng độ chính xác của kết quả đánh giá. Bên cạnh

đó, việc cập nhật và kiểm định dữ liệu thường xuyên cũng là một biện pháp quan trọng để tránh tình trạng dữ liệu lỗi thời hoặc sai lệch ảnh hưởng đến quá trình phân tích và ra quyết định.

Thứ ba, đầu tư vào đào tạo nhân sự có kỹ năng về khoa học dữ liệu và công nghệ Big Data. Cụ thể, giảng viên và cán bộ quản lý giáo dục cần được trang bị kiến thức về phân tích dữ liệu, sử dụng công cụ như Python, R, hoặc Power BI, cũng như cách ứng dụng kết quả phân tích vào giảng dạy và quản lý học tập. Ngoài ra, việc hợp tác với các tổ chức nghiên cứu hoặc doanh nghiệp công nghệ sẽ giúp các trường đại học tiếp cận các giải pháp tiên tiến hơn mà không cần phải tự phát triển từ đầu.

Thứ tư, xây dựng các mô hình phân tích dữ liệu linh hoạt, có khả năng điều chỉnh theo từng nhóm sinh viên và bối cảnh giảng dạy khác nhau. Việc áp dụng học máy và AI để cá nhân hóa dữ liệu học tập có thể giúp các mô hình đánh giá trở nên phù hợp hơn với từng đối tượng sinh viên. Ngoài ra, dữ liệu cần được thu thập từ nhiều nguồn đa dạng để đảm bảo tính đại diện cao hơn, tránh thiên lệch trong đánh giá.

Thứ năm, để giảm chi phí triển khai và duy trì hệ thống Big Data, các trường có thể tận dụng các nền tảng điện toán đám mây (Cloud Computing) thay vì đầu tư vào hạ tầng phần cứng tốn kém. Các giải pháp lưu trữ và phân tích dữ liệu trên nền tảng đám mây không chỉ giúp giảm chi phí vận hành mà còn đảm bảo khả năng mở rộng hệ thống khi nhu cầu tăng cao. Ngoài ra, các trường có thể tìm kiếm các mô hình hợp tác công – tư, liên kết với các doanh nghiệp công nghệ lớn để triển khai các giải pháp Big Data với mức chi phí hợp lý hơn.

III. KẾT LUẬN

Việc ứng dụng Big Data trong đánh giá kết quả học tập của sinh viên mang lại những cải tiến vượt bậc, giúp quá trình đánh giá trở nên toàn diện, chính xác và cá nhân hóa hơn. Tuy nhiên, để tận dụng tối đa lợi ích mà Big Data mang lại, các cơ sở giáo dục cần giải quyết đồng bộ các thách thức về bảo mật, chất lượng dữ liệu, năng lực phân tích và chi phí triển khai. Bằng cách đầu tư vào cơ sở hạ tầng công nghệ, đào tạo nhân sự, xây dựng hệ thống dữ liệu chuẩn hóa và hợp tác với doanh nghiệp công nghệ, Big Data sẽ không chỉ hỗ trợ

cải thiện kết quả học tập mà còn góp phần đổi mới phương pháp giảng dạy, tối ưu hóa quản lý giáo dục. Khi được triển khai một cách bài bản, Big Data không chỉ là một công cụ hỗ trợ mà còn là nền tảng quan trọng giúp nâng cao chất lượng giáo dục trong thời đại số.

TÀI LIỆU THAM KHẢO

- Abdalla, M. (2022). Challenges of Big Data Analytics in Higher Education Institutions. *International Journal of Educational Technology in Higher Education*, 19(1), 1-15.
- Chango, W., Cerezo, R., & Romero, C. (2024). Multi-source and multimodal data fusion for predicting academic performance in blended learning university courses. *arXiv preprint arXiv:2403.05552*.
- Chua, W. H., Tay, E., & Toh, Y. (2020). The implementation of learning analytics in Singapore schools: Opportunities and challenges. *Journal of Learning Analytics*, 7(2), 105-123.
- Dede, C. (2018). Big Data in higher education: The Harvard experience. *Educational Technology & Society*, 21(1), 34-41.
- Ferguson, R. (2019). Horizon 2020 and the future of learning analytics in Europe. *Computers & Education*, 140, 103594.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Học viện Nông nghiệp Việt Nam. (2021). Dữ liệu lớn (Big Data) và một số thách thức đối với phân tích dữ liệu lớn. *Học viện Nông nghiệp Việt Nam*. Truy cập ngày 16 tháng 3 năm 2025, từ <https://vnua.edu.vn/>...
- ICT Vietnam. (2020). AI và Big Data trong giáo dục đại học: Triển vọng hay rủi ro? *ICT Vietnam*. Truy cập ngày 16 tháng 3 năm 2025, từ <https://ictvietnam.vn/>...
- Nguyễn, T. H. L., & Trần, T. N. M. (2021). Dữ liệu lớn (Big Data) trong giáo dục và giải pháp ứng dụng trong các học viện, trường đại học ở Việt Nam. *Tạp chí Giáo dục*, 508(2), 8-12.
- PACE. (2020). Big data là gì? Đặc điểm & ứng dụng Big data vào các ngành. PACE Institute of Management. Truy cập ngày 16 tháng 3 năm 2025, từ <https://www.pace.edu.vn/>...
- Schmucker, R., Wang, J., Hu, S., & Mitchell, T. M. (2021). Assessing the performance of online students -- *New data, new approaches, improved accuracy*. *arXiv preprint arXiv:2109.01753*.
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30-32.
- Trần, T. Đ., Nguyễn, N. T., Nguyễn, T. H., & Nguyễn, T. N. (2023). Tăng tốc tìm kiếm tài nguyên học tập theo nội dung bằng kỹ thuật xử lý dữ liệu lớn. *Tạp chí Khoa học và Công nghệ*, 15(2), 170-180.
- Vũ, H. N. (2020). Khai thác dữ liệu lớn trong giáo dục và đào tạo du lịch: Cơ hội và thách thức. *Hội thảo Quốc gia về Ứng dụng Công nghệ Thông tin trong Giáo dục*, 1(1), 45-53.
- Wang, C., Chen, J., Xie, Z., & Zou, J. (2024). Research on education big data for students academic performance analysis based on machine learning. *arXiv preprint arXiv:2407.16907*.