

PHÂN TÍCH DỮ LIỆU HỌC TẬP - MÔ HÌNH LÝ THUYẾT VÀ ĐỀ XUẤT ỨNG DỤNG TRONG HỌC TẬP VÀ GIẢNG DẠY

● PHẠM THỊ QUỲNH LỆ

TÓM TẮT:

Trong bối cảnh khoa học công nghệ dành cho giáo dục đang phát triển mạnh mẽ đặt ra yêu cầu giáo dục trực tuyến elearning về việc thiết kế bài giảng thu hút, tăng cường tương tác giữa người dạy và người học, đặc biệt là phân tích dữ liệu học tập để nâng cao kết quả học tập và chất lượng giảng dạy. Bài viết nhằm giới thiệu phương pháp nghiên cứu dữ liệu học tập (learning analytics - LA), mô hình nghiên cứu lý thuyết và đề xuất ứng dụng cho các đối tượng liên quan trong giảng dạy.

Từ khóa: giáo dục trực tuyến, phân tích dữ liệu học tập, phương pháp giảng dạy.

1. Đặt vấn đề

Định hướng đổi mới trong giáo dục như lấy người học làm trung tâm, cá nhân hóa trải nghiệm học tập,... buộc các cơ sở giáo dục đại học phải đổi mới phương pháp giảng dạy và ứng dụng LA. Mặc dù đây là lĩnh vực nghiên cứu khá mới mẻ tại Việt Nam, nhưng LA hứa hẹn sẽ mang lại các thông tin hữu ích trong việc theo dõi quá trình học tập, phát hiện vấn đề để người học, người dạy và tổ chức giáo dục có sự điều chỉnh kịp thời.

2. Phương pháp nghiên cứu

Tác giả đã thu thập và sử dụng những tài liệu, số liệu liên đến hoạt động LA của các tác giả trong và ngoài nước. Tác giả sử dụng phương pháp luận: phân tích, tổng hợp, đối chiếu so sánh, thống kê, đồng thời kế thừa các nghiên cứu có liên quan đến bài viết.

3. Kết quả nghiên cứu

3.1. Khái quát về *Learning analytics*

Phân tích dữ liệu học tập - Learning Analytics (LA) là lĩnh vực nghiên cứu thuộc phân tích dữ liệu được áp dụng trong giáo dục để dự đoán thành công của người học hoặc thông báo cho người dạy về thời điểm và cách can thiệp với người học để giảm nguy cơ thất bại, nhằm mang lại lợi ích cho người học (Martin & Sherin, 2013).

Theo Siemens and Long (2011), LA là phép đo lường, thu thập, phân tích và báo cáo dữ liệu về người học và bối cảnh học tập, nhằm mục đích tìm hiểu và tối ưu hóa việc học cũng như môi trường học tập.

Theo Seat Software, LA bao gồm đo lường, thu thập, phân tích và báo cáo dữ liệu về sự tiến bộ của người học và bối cảnh học tập. LA sử dụng

dữ liệu lớn (Big Data) sẵn có xung quanh hoạt động của người học và dấu chân kỹ thuật số (digital footprints) do hoạt động của sinh viên để lại trong môi trường học tập để cải thiện việc học và giảng dạy.

Vậy có thể hiểu LA là tiến trình nghiên cứu dữ liệu từ hoạt động tương tác giữa người dạy và người học, giữa người học với nhau, quá trình tự học, các hoạt động khác trong môi trường học tập.... cung cấp thông tin cho các quyết định liên quan đến việc học tập, giảng dạy, quản lý nhằm cải thiện kết quả học tập và nâng cao chất lượng giáo dục.

3.2. Tầm quan trọng của phân tích dữ liệu học tập trong giáo dục trực tuyến

Đối với người học: Cá nhân hóa trải nghiệm elearning thông qua việc người học được lựa chọn chương trình học phù hợp với trình độ khả năng tiếp thu của bản thân; Cho phép người học kiểm soát quá trình học tập dựa trên các phản hồi của người dạy và dữ liệu về tiến độ học tập.

Đối với người dạy: Hỗ trợ đánh giá hoạt động người học bằng các công cụ đánh giá người học mà hệ thống LA cung cấp như chấm điểm, phân tích tỷ lệ lựa chọn câu trả lời,... Nâng cao chất lượng dạy học: việc đánh giá được từng người học với các chỉ số cụ thể sẽ giúp người dạy có thể điều chỉnh, cải tiến phương pháp giảng dạy, đánh giá phù hợp. Bên cạnh đó, các ứng dụng giảng dạy trực tuyến tích hợp LA còn cung cấp nhiều công cụ thiết kế bài giảng đa dạng, trực quan, thu hút người học tương tác, góp phần nâng cao chất lượng dạy học.

Đối với tổ chức giáo dục: Các phản hồi của

người học giúp cải tiến, xây dựng chương trình học phù hợp. Tăng tỷ lệ giữ chân người học khi được quan tâm, theo sát tình hình học tập. Tối ưu mặt chi phí, giúp cho việc sử dụng nguồn lực dạy học tốt hơn và nâng cao danh tiếng trong và ngoài tổ chức.

3.3. Phân loại các nhóm phân tích dữ liệu học tập

- Phân tích mô tả: tạo ra dữ liệu tổng quan (dashboard) về đặc điểm người học.
- Phân tích dự đoán: dự đoán xu hướng về sự thay đổi của người học trong tương lai.
- Phân tích đề xuất: đưa ra các khuyến nghị cho các hoạt động dạy và học tiếp theo.

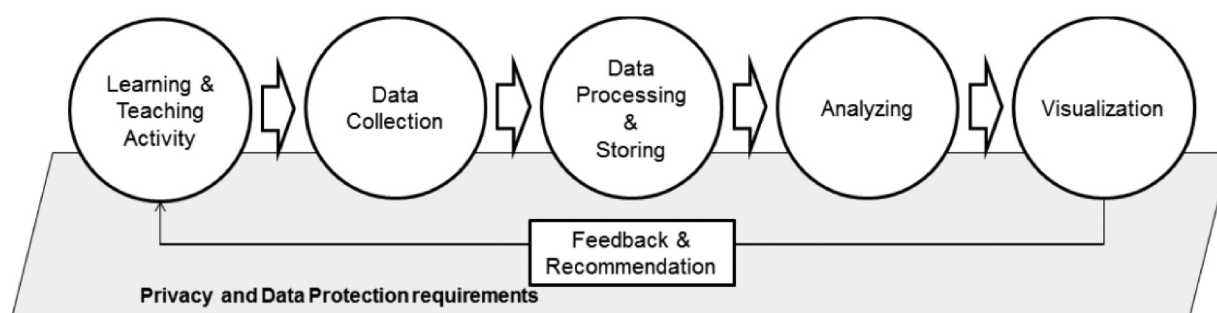
3.4. Quy trình thực hiện phân tích dữ liệu học tập LA

Quy trình thực hiện LA là một vòng lặp dựa trên việc thu thập và phân tích dữ liệu về hoạt động dạy và học trong bối cảnh học tập, giáo dục và đào tạo (the learning, education, and training-LET), đưa ra phản hồi, góp phần cải thiện được chất lượng giảng dạy, với điều kiện đảm bảo quyền riêng tư và bảo vệ dữ liệu của các bên liên quan. (Hình 1)

Các mục dữ liệu đầu vào có thể được lấy từ nhiều hoạt động dạy và học và kết quả đầu ra có thể cung cấp phản hồi và khuyến nghị để nâng cao kết quả dạy và học.

Learning and teaching activity process - Tiến trình hoạt động dạy và học: Nguồn dữ liệu để thu thập thông qua hoạt động dạy và học gồm: đọc bài giảng, làm bài tập, thực hiện dự án,... Môi trường học tập không đồng nhất và người học khác nhau về nhân khẩu học, trình độ, mức độ sử dụng công cụ

Hình 1: Quy thực hiện của Learning Analytics



Nguồn: ISO/IEC TR 20748-1

học tập nên các dữ liệu này cần được mô hình hóa (data modeling), tạo ra luồng dữ liệu có liên quan, được chuẩn hóa để phân tích. Mô hình hóa dữ liệu thực hiện bởi các câu hỏi sư phạm (pedagogical questions), phác thảo khía cạnh nào của việc học cần hỗ trợ. Có thể hiểu, đây là bước xây dựng các tiêu chí đánh giá người học, như: kết quả học tập, tiến độ và thái độ học tập, phát triển các kỹ năng,... hay trong mỗi bài học là mức độ tương tác, hiểu bài, chất lượng bài làm... Từ đó, xác định dữ liệu nào được lựa chọn để thu thập trong vô vàn dữ liệu mà hệ thống đã ghi nhận. Khi các nguồn dữ liệu cần thiết được xác định, cần được sự cho phép trước khi đi vào khai thác dữ liệu để đảm bảo tính riêng tư và bảo vệ tính tuyệt mật của chủ sở hữu. (Hình 2)

Learning and teaching activity process - Tiến trình thu thập dữ liệu: Giai đoạn này bao gồm tập hợp và đo lường thông tin về các biến quan tâm trong hoạt động dạy và học. Hoạt động dạy và học cùng nguồn dữ liệu liên quan như thiết bị học tập, phần mềm, ứng dụng học tập và tài liệu, công cụ học tập, câu hỏi và bài giao, diễn đàn thảo luận, tin nhắn, mạng xã hội, bài tập về nhà, số tín chỉ đã hoàn thành, thành tích, nhật ký hệ thống (system

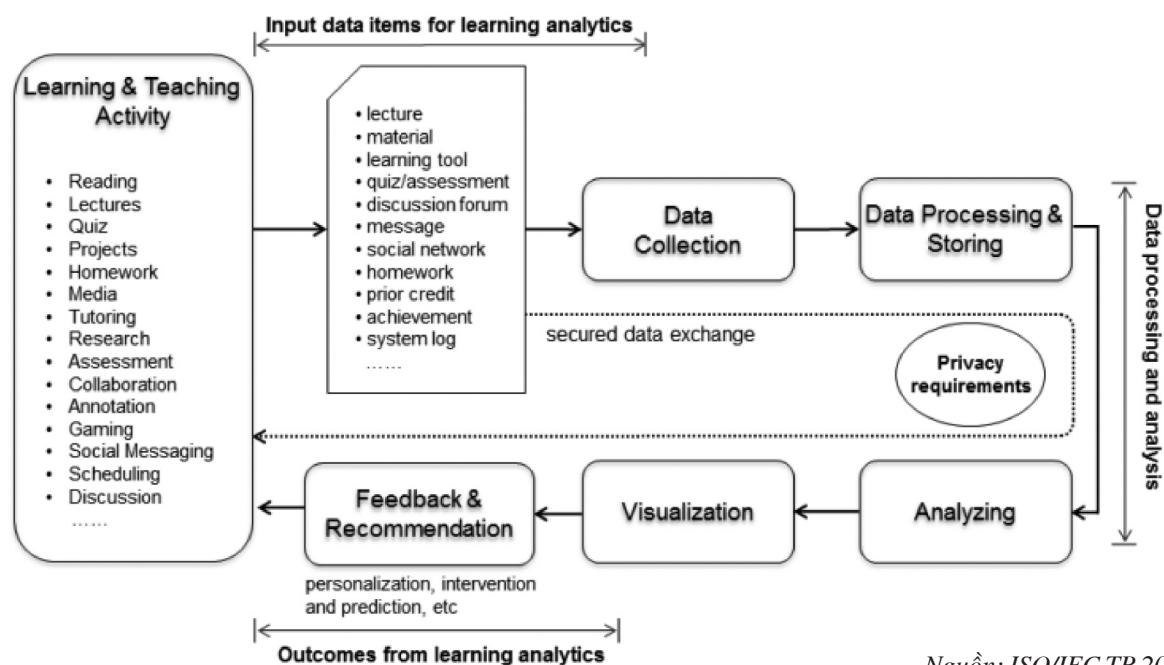
logs),... cần được thu thập hoặc chuyển đổi thành các dữ liệu API¹ như Experience APITM (xAPI) và IMS Caliper AnalyticsTM.

Dữ liệu từ hoạt động LET chứa thông tin người học phải đảm bảo quyền riêng tư, tức là thông báo đến người học (hoặc cha mẹ, người giám hộ hợp pháp) về việc thu thập dữ liệu, cam kết chỉ được sử dụng cho mục đích LA và được bảo vệ bằng các phương tiện cần thiết như mã hóa, khử nhận dạng -de-identification, phân hủy theo thời gian chrono-degradability, giả danh - pseudomization và ẩn danh. Quy trình thu thập có thể phải kiểm tra sự phù hợp hoặc chuyển đổi dữ liệu trước khi lưu trữ trong kho dữ liệu tạm thời, như kho lưu trữ sự kiện trong IMS Caliper² hoặc lưu trữ hồ sơ học tập trong xAPI.

Learning and teaching activity process - Tiến trình lưu trữ và xử lý dữ liệu: Là quá trình chuẩn bị và lưu trữ dữ liệu từ các nguồn dữ liệu đa dạng và không đồng nhất bằng cách sử dụng mô hình và biểu diễn dữ liệu được chuẩn hóa.

Dữ liệu học tập được xử lý bởi bộ chuyển dịch và lọc dữ liệu - Data Translator and Filter. Các kết quả đã xử lý được lưu trữ vào kho dữ liệu phân tích -

Hình 2: Quy trình thực hiện của Learning Analytics



Nguồn: ISO/IEC TR 20748-1

Analytics Data Store. Trình chuyển dịch dữ liệu hợp nhất - Unified Data Translator chuyển các dữ liệu không đồng nhất thành một biểu diễn thống nhất, ví dụ dữ liệu liên kết mở (LOD)³. Nguồn dữ liệu chính bao gồm dữ liệu về ngôn ngữ như diễn ngôn - discourse, ngôn ngữ viết, hội thoại,... được xử lý bằng xử lý ngôn ngữ tự nhiên - Natural Language Processing trước chuyển thành một biểu diễn thống nhất. Dữ liệu lưu trữ có thể được truy cập thông qua giao diện truy vấn dữ liệu được chuẩn hóa - Data Query và dữ liệu đã xử lý được lưu trữ vào kho dữ liệu phân tích thông qua di chuyển dữ liệu được chuẩn hóa - Data Migration.

Learning and teaching activity process - Tiến trình phân tích dữ liệu: Là quá trình điều tra một cách có hệ thống dữ liệu học tập bằng cách kiểm tra và mô hình hoá dữ liệu học tập với mục tiêu tạo ra tri thức mô tả và tri thức dự đoán.

Micro Data⁴ lưu trữ trong Analytics Data Store, miền dữ liệu chung như chương trình giảng dạy, tài nguyên,... lưu trữ trong Constant Information. Thông qua Analysis Interface, các thuật toán phân tích bên ngoài như phân tích dự đoán, phân tích thích ứng, khuyết tật học tập,... sử dụng ICT⁵. Xử lý phân tích bao gồm phân tích thống kê, phân tích chủ đề, phân tích xã hội,... Kết quả phân tích có thể đưa vào mẫu nghiên cứu, mô hình động và phân tích liên kết trước khi đề xuất nội dung và lộ trình học tập. Kết quả phân tích được tinh chỉnh bởi giao diện xử lý dữ liệu - Data Manipulation và lưu trữ vào Analytics Data Store cho các chu kỳ phân tích sâu hơn hoặc các bước xử lý sau đó.

Visualization process - Tiến trình trực quan hóa dữ liệu: Trực quan hóa là truyền đạt thông tin một cách rõ ràng và hiệu quả cho người dùng thông qua đồ họa thống kê, đồ thị, đồ họa trực quan, bảng, biểu đồ,... làm cho dữ liệu phức tạp dễ tiếp cận hơn, dễ hiểu và có thể sử dụng được.

Dữ liệu trong Analytics Data Store truy cập thông qua Data Query. Biểu diễn trực quan bao gồm thông tin tổng quan, hồ sơ điện tử - ePortfolio, sơ đồ xã hội, lộ trình học tập, tài nguyên. Thông tin tổng quan hiển thị các so sánh, tiến độ, đánh giá đề

xuất dựa trên chủ đề, biểu đồ mạng xã hội,... Dữ liệu hiển thị - Data Interface cung cấp dưới dạng mở cho hệ thống báo cáo và trang tổng quan bên ngoài để cung cấp thông tin phản hồi, chẳng hạn như cá nhân hóa, can thiệp hoặc dự đoán cho người dùng cá nhân.

Feedback process - Tiến trình phản hồi: Phản hồi đưa kết quả của chu trình phân tích học tập trở lại người học người thiết kế khóa học, giảng viên, nhà quản lý giáo dục để đề xuất các hành động khắc phục.

Số liệu thống kê là cơ sở để đưa ra phản hồi như kiểm soát tốc độ học tập, thay đổi thái độ hoặc phương pháp dạy và học. Điều này cho phép tạo ra môi trường học tập thích ứng cao, cung cấp các lộ trình học phi tuyến tính liên quan đến các tài nguyên kỹ thuật số nội dung hoặc tự đánh giá học tập được cá nhân hóa.

3.5. Đề xuất ứng dụng LA trong giảng dạy

Hiện tại, chưa có hệ thống LA nào áp dụng chung cho tất cả tổ chức giáo dục, bởi quá trình nghiên cứu, phát triển và ứng dụng LA cần xem xét các yếu tố đặc thù về kỹ thuật, môi trường học tập, hoàn cảnh sử dụng, cá nhân người dùng và mục tiêu học tập. Vì vậy, tác giả có một số đề xuất việc ứng dụng LA trong hoạt động dạy và học như sau:

3.5.1. Đối với người học

Xác định mục tiêu học tập rõ ràng: thông qua các tiêu chí định lượng để khai thác các chỉ số của LA cung cấp, từ đó điều chỉnh lộ trình học thích hợp.

Sử dụng các ứng dụng học tập tích hợp LA: có rất nhiều ứng dụng giúp người học tự chọn lộ trình, bài học dựa vào năng lực và thời gian biểu tự chọn. Ví dụ ứng dụng học tiếng Anh Duolingo cung cấp các trải nghiệm cá nhân hóa, như: cung cấp bài học theo cấp độ người học, báo cáo tiến độ, đưa ra lời cảnh báo, nhắc nhở khi chưa hoàn thành mục tiêu hoặc lơ là cũng như dành lời khen, phần thưởng khi người học có thành tích tốt.

3.5.2. Đối với người dạy

Sử dụng một số công cụ hỗ trợ trong giáo dục trực tuyến như: Google Classroom; Perusall-nền tảng cho phép giảng viên biết được mức độ tương

tác của sinh viên với các tài liệu học tập đã tải lên; GradeCraft - hệ thống quản lý học tập hỗ trợ các khóa học được thiết kế theo kiểu trò chơi. Người dạy có thể đưa ra nhiều hoạt động, bài tập khác nhau, tùy chọn thang điểm, còn người học có quyền lựa chọn nhiệm vụ nào để thực hiện và hệ thống sẽ cho biết được người học đã đạt ở giai đoạn nào trong lộ trình đạt được mục tiêu; Canvas Learning Management System - phần mềm ứng dụng dành cho quản trị, thống kê, theo dõi, báo cáo và phân phối các khóa học, chương trình học tập, đào tạo.

Xác định các tiêu chí đánh giá: trước khi quyết định lựa chọn ứng dụng hỗ trợ học tập tích hợp LA nào, cần xây dựng các tiêu chí này nên cụ thể, được xác định tùy theo đặc điểm bài tập, môn học, lớp học và mong muốn của người dạy. Đồng thời, kết hợp cùng với chuyên gia/nhà nghiên cứu thiết kế LA để thiết kế các tiêu chí nghiên cứu LA phù hợp và bám sát với đặc thù giảng dạy.

3.5.3. Đối với các tổ chức giáo dục

Tự thiết kế ứng dụng LA: với tổ chức giáo dục có nguồn lực lớn, có thể tự thiết kế các công cụ LA dựa trên sự phối hợp giữa người học, người dạy, nhà nghiên cứu, nhà phát triển ứng dụng để đưa ra các tiêu chí dữ liệu cần thu thập phù hợp và phản ánh đúng đặc thù của môn học, người học, môi trường học tập.

Kết hợp với các tổ chức về công nghệ: Đối với các tổ chức giáo dục chưa thể thiết kế riêng các công cụ LA, có thể mua các sản phẩm LA đã có hoặc hợp tác với các công ty về đầu tư công nghệ và chuyển đổi số trong giáo dục để có thể tạo ra các công cụ LA phù hợp với đặc thù môi trường học tập của tổ chức đó.

4. Kết luận

Phân tích dữ liệu học tập LA tuy là lĩnh vực nghiên cứu chưa phổ biến rộng rãi tại Việt Nam, nhưng không thể phủ nhận vai trò to lớn của nó trong việc phản ánh hoạt động học tập, nâng cao chất lượng giảng dạy và xây dựng mối quan hệ tốt hơn giữa sinh viên và trường học. Vì vậy, các tổ chức giáo dục nên phát triển và tích hợp hoạt động này trong hệ thống elearning hiện có. Tuy nhiên, cần lưu ý rằng, không nên kết luận hoàn toàn dựa trên số liệu phân tích. Bởi kết quả LA dựa trên các hoạt động của người học lưu lại trên hệ thống trực tuyến, nhưng trong một số trường hợp, người học vì lý do khách quan không thể tương tác trực tuyến, nhưng vẫn có các hoạt động học bên ngoài và hoàn thành bài được giao. Vì vậy, cần đầu tư kỹ lưỡng khi xây dựng các tiêu chí nghiên cứu trong LA và sử dụng nhiều công cụ LA khác nhau để có cái nhìn toàn cảnh khi đánh giá người học ■

TÀI LIỆU TRÍCH DẪN:

¹API là viết tắt của Giao diện lập trình ứng dụng (Application Programming Interface), là một trung gian phần mềm cho phép hai ứng dụng giao tiếp với nhau (Theo Hybrid-technologies.vn).

²IMS Caliper: Caliper Analytics - Ứng dụng của Tổ chức giáo dục toàn cầu IMS, cho phép các tổ chức thu thập dữ liệu học tập từ các tài nguyên kỹ thuật số để hiểu và hình dung rõ hơn về hoạt động học tập, đồng thời trình bày thông tin này cho sinh viên, người hướng dẫn và cố vấn theo những cách có ý nghĩa để giúp cung cấp thông tin.

³Hoạt động dạy và học cùng với các nguồn dữ liệu liên quan như thiết bị học tập, phần mềm, ứng dụng học tập và mạng xã hội mang đến nhiều dữ liệu khác nhau. Các nguồn bao gồm bài giảng, tài liệu học tập, công cụ học tập, câu hỏi và bài giao, diễn đàn thảo luận, tin nhắn, mạng xã hội, bài tập về nhà, số tín chỉ đã hoàn thành, thành tích, nhật ký hệ thống (system logs),... Những dữ liệu học tập này cần được thu thập hoặc chuyển đổi thành các dữ liệu API như Experience API™ (xAPI) và IMS Caliper Analytics™ (Caliper).

⁴Microdata là một kiểu định dạng lại ngôn ngữ html mà google thường hay cập nhật và hỗ trợ nhất hiện nay. Microdata hiểu một cách đơn giản là để thêm dữ liệu có cấu trúc vào website. Microdata nó định nghĩa thuộc tính, có thể đặt vào trong HTML để cho biết trang web đó viết về mục đích gì.

⁵ICT - Information & Communication Technologies: Công nghệ thông tin và truyền thông.

TÀI LIỆU THAM KHẢO:

1. Claudette Adamma Kika. (2018). *Supporting student experience management with learning analytics in the UK higher education sector*. PhD thesis, University of Bedfordshire, UK.
2. Lisa Janine Berry. (2017). Using learning analytics to predict academic success in online and face-to-face learning environments. [Online] Available at <https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=2317&context=td>
3. Martin, T., & Sherin, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences*, 22(4), 511-520.
4. Phil Long and George Siemens. (2011). Penetrating the Fog: Analytics in Learning and Education. *TD Tecnologie Didattiche*, 22(3), 132-137.
5. Seatssoftware. (2017). The Importance of Learning Analytics for Universities and Students. [Online] Available at https://www.seatssoftware.com/wp-content/uploads/2017/09/SEAtSImportanceofLearningAnalytics_N.pdf
6. Trịnh Minh Giang (2021). Phân tích dữ liệu học tập và cách áp dụng hiệu quả <https://trinhminhgiang.com/vi-VN/bai-viet/phan-tich-du-lieu-hoc-tap-va-cach-ap-dung-hieu-qua>
7. Yong-Sang Cho (2016). Issue Report: Prospect for Learning Analytics to achieve adaptive learning model. [Online] Available at <https://www.slideshare.net/zzosang/prospect-for-learning-analytics-to-achieve-adaptive-learning-model>

Ngày nhận bài: 5/4/2022

Ngày phản biện đánh giá và sửa chữa: 3/5/2022

Ngày chấp nhận đăng bài: 13/5/2022

Thông tin tác giả:

ThS. PHẠM THỊ QUỲNH LÊ

Khoa Quản trị kinh doanh

Trường Đại học Duy Tân, Đà Nẵng

LEARNING ANALYTICS - THE THEORETICAL RESEARCH MODEL AND APPLICATIONS IN LEARNING

● Master. **PHAM THI QUYNH LE**

Faculty of Business Administration, Duy Tan University

ABSTRACT:

The strong development of educational science and technology requires the design of e-learning become more attractive and more interactive to better connect teachers and learners. Especially, learning data should be analyzed to improve the learning outcome and teaching quality. This paper introduces the method of learning analytics, the theoretical research model of learning analytics, and proposes applications of this method for relevant stakeholders.

Keywords: online education, learning analytics, teaching method.