

NGHIÊN CỨU XÂY DỰNG HỆ THỐNG THÔNG TIN QUẢN LÝ THÀNH TÍCH THỂ THAO CỦA TRƯỜNG ĐẠI HỌC THỂ DỤC THỂ THAO BẮC NINH DỰA TRÊN KHAI PHÁ DỮ LIỆU

Ngô Trang Hưng*

1. Đặt vấn đề

Cách mạng công nghiệp 4.0 dựa trên nền tảng công nghệ số nhằm xây dựng thế giới siêu kết nối và tích hợp các công nghệ thông minh để tối ưu hóa quy trình, phương thức sản xuất. Những công nghệ tiên tiến nhất hiện nay là Trí tuệ nhân tạo (AI); Dữ liệu lớn (Big Data); Internet kết nối vạn vật (IoT), Công nghệ thực tế ảo (VR), Công nghệ sinh học... Cách mạng công nghiệp 4.0 đã tạo cơ hội và đặt ra yêu cầu tất yếu phải xây dựng nền giáo dục 4.0. Theo đó, giáo dục trở thành một hệ sinh thái mà mọi người có thể cùng học tập mọi lúc, mọi nơi với các thiết bị được kết nối. Nhà trường trở thành một hệ sinh thái tạo sản phẩm sáng tạo mang tính cá thể, với kiến thức và năng lực đổi mới, sáng tạo riêng của cá nhân. Trong giáo dục 4.0, lãnh đạo nhà trường, giảng viên có thể thu thập dữ liệu, phân tích và đánh giá chính xác về người học; theo dõi quá trình học tập tại nhà, kiểm tra mức độ hoàn thành bài tập và thông báo kết quả học tập tới sinh viên. Với công nghệ AI có thể thay thế giảng viên ở một số khâu như điểm danh, chấm bài, soạn bài và hỗ trợ dạy ngoại ngữ. Hiện nay, các tập đoàn công nghệ lớn đã cho ra đời nhiều thiết bị và phần mềm thông minh phục vụ giáo dục. Chẳng hạn, Bộ công cụ giáo dục Google (G-Suite for Education), Microsoft Office365, Chroma Key tạo trường quay thu nhỏ để người học tự sản xuất sản phẩm truyền thông, Công nghệ đám mây (Cloud) cung cấp những phần mềm học tập trên Internet có sử dụng tài khoản...

Giáo dục 4.0 sẽ mang lại sự hỗ trợ tối ưu cho công tác xây dựng thông tin trong trường đại học nói chung, đồng thời tác động đến công tác phát triển thể chất cho sinh viên nói riêng. Trường Đại học TDTT Bắc Ninh là cơ sở đào tạo cán bộ TDTT hàng đầu và đào năng khiếu TDTT, đào tạo vận động viên trẻ quốc gia. Công nghệ khai thác dữ liệu kết hợp với giảng dạy, huấn luyện ở các môn thể thao của Nhà trường sẽ là cơ sở để xây

dựng một hệ thống quản lý thành tích thể thao dựa trên khai thác dữ liệu và nâng cao chất lượng giảng dạy, huấn luyện các môn thực hành. Thông qua thực tế công tác quản lý TDTT và khai thác dữ liệu của nhà trường có thể xây dựng hệ thống thông tin quản lý thể thao trường đại học, trong đó việc phân tích các giá trị liên quan và sử dụng phương pháp giữ vai trò tiên quyết để xây dựng hệ thống thông tin quản lý thể thao dựa trên khai thác dữ liệu. Hệ thống thông tin này dựa trên ý tưởng khai thác dữ liệu để có thể cung cấp thông tin cho quá trình ra quyết định liên quan đến huấn luyện các môn thể thao. Cấu trúc hệ thống thông tin xây dựng dự kiến gồm ba lớp có khả năng thích ứng tốt và khả năng mở rộng chức năng, tiết kiệm khối lượng công việc bảo trì cho sau khi ứng dụng.

Khai phá dữ liệu (Data mining) là quá trình tính toán để tìm ra các mẫu trong các bộ dữ liệu lớn liên quan đến các phương pháp tại giao điểm của máy học, thống kê và các hệ thống cơ sở dữ liệu. Đây là một lĩnh vực liên ngành của khoa học máy tính. Mục tiêu tổng thể của quá trình khai thác dữ liệu là trích xuất thông tin từ một bộ dữ liệu và chuyển nó thành một cấu trúc dễ hiểu để sử dụng tiếp. Ngoài bước phân tích thô, nó còn liên quan tới cơ sở dữ liệu và các khía cạnh quản lý dữ liệu, xử lý dữ liệu trước, suy xét mô hình và suy luận thống kê, các thước đo thứ vị, các cân nhắc phức tạp, xuất kết quả về các cấu trúc được phát hiện, hiện hình hóa và cập nhật trực tuyến. Khai thác dữ liệu là bước phân tích của quá trình “khám phá kiến thức trong cơ sở dữ liệu”. Data mining cùng với Data science (khoa học dữ liệu) chính là hai lĩnh vực công nghệ đang được sử dụng phổ biến nhất hiện nay. Với khả năng tập hợp, sắp xếp số lượng dữ liệu lớn từ đó đưa ra những phân tích chính xác nhất. Quá trình khai phá dữ liệu của Data mining được diễn ra với công nghệ tính toán tiên tiến không chỉ giới hạn trong việc trích xuất dữ liệu mà còn được sử dụng để chuyển đổi, làm sạch, tích hợp

*PGS.TS, Trường Đại học TDTT Bắc Ninh

dữ liệu và phân tích mẫu. Như vậy, khai phá dữ liệu chính là việc sử dụng các thuật toán và công nghệ liên quan để đưa ra các dự báo dựa trên dữ liệu hiện có, nhằm đưa ra quyết định chủ động và dựa trên kiến thức. Công nghệ khai thác dữ liệu hiện đang là lĩnh vực có ảnh hưởng quốc tế và là trọng tâm nghiên cứu hàng đầu về xử lý thông tin và ra quyết định, các ngành nghề khác nhau đều ứng dụng để phân tích thông tin dựa trên các hoạt động hỗ trợ ra quyết định. Mức độ phổ biến của nó ngày càng sâu rộng và đóng vai trò quan trọng trong chuyển đổi số. Để nâng cao năng lực cạnh tranh cốt lõi của Trường Đại học TĐTT Bắc Ninh cần phân bổ tối ưu được các nguồn lực, cũng như sử dụng tài nguyên. Ngoài ra, Trường cần phân tích công nghệ khai thác dữ liệu và thực trạng quản lý thông tin nhằm rút ra các bài học kinh nghiệm phục vụ tự chủ đại học và kinh doanh thể thao hiện nay.

Mục tiêu cuối cùng của khai phá dữ liệu huấn luyện thể thao là tìm thấy một lượng lớn dữ liệu từ kiến thức có giá trị, có ý nghĩa. Dưới góc độ lý thuyết cơ bản, khai phá dữ liệu bao gồm các tính năng sau: Phân tích dữ liệu tự động và dự báo xu hướng; Khai thác dữ liệu có khả năng thực hiện phân tích tự động dựa trên cơ sở dữ liệu lớn; Tìm kiếm thông tin dự đoán. Với việc phân tích mối quan hệ, tương quan: Có mối quan hệ giữa hai tập dữ liệu hoặc tính thường xuyên hoặc các nhóm dữ liệu. Phân tích tương quan là phân tích các mối quan hệ này và có thể cao hơn là ứng dụng các mô hình hồi quy qua các phần mềm R, SPSS. Phân cụm: Dữ liệu (Data) trong cơ sở dữ liệu (Database) là một đơn vị lưu trữ dữ liệu duy nhất, song dữ liệu không thể mang ý nghĩa duy nhất, mà các cụm dữ liệu này được phân tán, quy nạp, hợp nhất để tạo thành một tập hợp con có ý nghĩa. Để tạo một hệ cơ sở dữ liệu phân tán trong huấn luyện thể thao, các tệp (File) không chỉ quan hệ với nhau về mặt logic mà cần có một cấu trúc giao diện chung giữa chúng để có thể truy cập lẫn nhau. Thực tế có rất nhiều ứng dụng yêu cầu các hệ quản trị dữ liệu thao tác trên dữ liệu bán cấu trúc hoặc không cấu trúc, như các File Web trên mạng Internet.

Khai phá dữ liệu là các quy tắc trích xuất dựa trên cơ sở dữ liệu hoặc kho dữ liệu được sử dụng để hướng dẫn các hành động hiện tại và tương lai. Khai thác dữ liệu nâng cao hiệu quả trình độ

quản lý thể thao trong công tác huấn luyện thể thao là một vấn đề quan trọng mà các giảng viên, huấn luyện viên, nhà quản lý của ngành TĐTT phải đối mặt. Mỗi một ngày trôi qua sẽ sinh ra một lượng dữ liệu rất lớn, đặc biệt là đối với công tác huấn luyện học sinh, sinh viên, vận động viên. Nếu dữ liệu đó không có phương tiện cập nhật, phân tích, dự báo thì những mô hình dự báo thành tích với các tình huống của ngày hôm nay đã có thể bị giảm xuống và như vậy ảnh hưởng lớn để việc đạt thành tích cao. Vì vậy, quá trình mô hình hóa cần lặp lại đối với công tác huấn luyện thể thao một cách thường xuyên liên tục là hết sức cấp thiết. Kỹ thuật khai phá dữ liệu có thể nhanh chóng thực hiện các thay đổi đối với dữ liệu phản ánh và cung cấp hỗ trợ quyết định. Khai phá dữ liệu là cần thiết để xác định các quy tắc tiềm năng, cũng như quản lý và duy trì các quy tắc. Quy tắc không nhất thiết phải áp dụng cho tất cả dữ liệu khám phá khi nó đạt đến ngưỡng tin cậy cần thiết.

2. Kết quả nghiên cứu

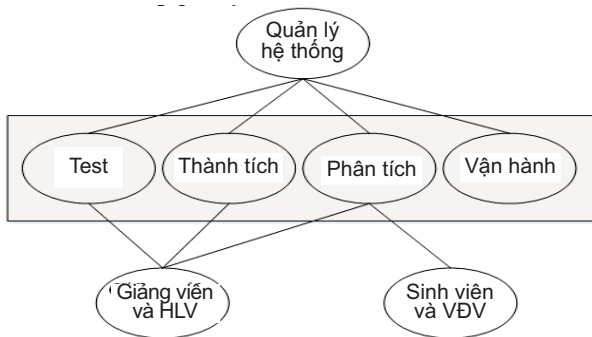
2.1. Phân tích chức năng Quản lý thể thao của trường đại học

Phát triển hệ thống thông tin quản lý thể thao trường đại học sẽ góp phần giúp các trường thể thao hoàn thiện các quy trình nghiệp vụ quản lý khoa học khác nhau, có lợi cho việc hoàn thành các chức năng và mục tiêu khác nhau của các trường đại học thể thao để giải quyết các vấn đề đào tạo sinh viên, vận động viên và kinh doanh thể thao. Hệ thống thông tin quản lý cùng một số phương pháp phân tích thông tin thông qua khai phá dữ liệu, sẽ là công cụ hỗ trợ hiệu quả quá trình ra quyết định và giữ vị trí vai trò thực sự để mang lại hiệu quả cao.

Nghiên cứu dựa trên khai phá dữ liệu của trường đại học thể thao rất hữu ích để cải thiện hiệu quả làm việc của giảng viên, huấn luyện viên và có độ chính xác cao, do đó nâng cao chất lượng giảng dạy – huấn luyện của nhà trường. Thông qua nền tảng thông tin chính xác và toàn diện hơn để đáp ứng nhu cầu của ba đối tượng: người quản trị, giảng viên – huấn luyện viên, sinh viên – vận động viên. Dưới góc độ vai trò người sử dụng, chức năng hệ thống thông tin chủ yếu có ba loại: Quản trị hệ thống; Giảng viên – Huấn luyện viên; Sinh viên – Vận động viên. Ngoài ra còn những người dùng khác song chủ yếu khác nhau ở vai trò

LÝ LUẬN VÀ THỰC TIỄN THỂ DỤC THỂ THAO

phân quyền điều hành hệ thống. Mỗi người dùng được cung cấp tài khoản và mật khẩu để đăng nhập vào hệ thống thông tin. Hình 1 là sơ đồ hóa hệ thống quản lý thể thao.



Hình 1. Hệ thống quản lý thể thao

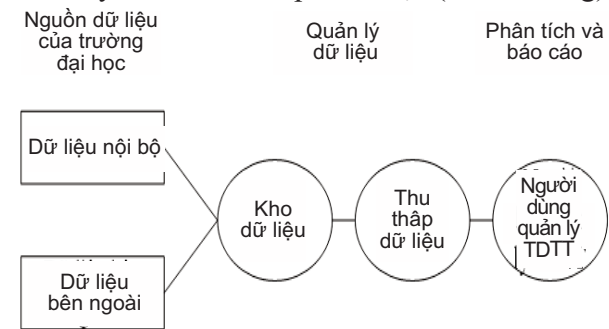
Hệ thống tuân theo nguyên tắc từ trên xuống dưới, đầu tiên cần xác định được các trường hợp có quyền cao nhất, điều này nhằm giới hạn phạm vi cho các phân hệ và người tham gia khác nhau của hệ thống. Sau đó, phân tích từng phạm vi cho người dùng ở mỗi phân hệ, quá trình phân tích này cứ như vậy được áp dụng cho đến hết. Nó phản ánh mối quan hệ giữa các bên tham gia và các chi tiết khác nhau của từng phân hệ chức năng của hệ thống. Chẳng hạn với phân hệ “Thành tích” khi áp dụng nguyên tắc này, thì đầu tiên phải phân tích tổng thể các trường hợp sử dụng, và sau đó tiếp tục đi từ trên xuống dưới để phân tích. Những người dùng khác nhau có các vai trò khác nhau trong phân quyền điều hành hệ thống. Khi người dùng đăng nhập, họ cần khai báo tài khoản và mật khẩu để xác minh quyền truy cập và các thao tác liên quan đến mỗi phân hệ của hệ thống.

2.2. Hệ thống khai thác dữ liệu và quy trình khai thác

Khai phá dữ liệu là để thích ứng với nhu cầu mới của xử lý thông tin và sự phát triển xã hội về tất cả các khía cạnh của nhu cầu cấp thiết. Từ đó phát triển một kỹ thuật phân tích thông tin mới để tìm ra thông tin hoặc tập hợp có ý nghĩa và quan trọng từ quá trình thăm dò dữ liệu lớn. Trong đó bao gồm một công cụ cơ sở dữ liệu quan hệ hiệu suất cao và một loạt các tích hợp dữ liệu trên cơ sở trí tuệ nhân tạo, công nghệ thống kê và cơ sở dữ liệu, theo hình thức nghiên cứu liên ngành.

Khai phá dữ liệu là việc sử dụng công nghệ khai thác dữ liệu, quá trình khai thác các kiến thức quý giá từ cơ sở dữ liệu, kho dữ liệu và các thông

tin khác trong một cơ sở dữ liệu lớn. Toàn bộ quy trình khai thác dữ liệu bao gồm chuẩn hóa dữ liệu, tích hợp dữ liệu, chuyển đổi dữ liệu, khai thác dữ liệu, đánh giá mẫu, biểu diễn tri thức... Hình 2 cho thấy tiến trình khai phá dữ liệu (Data mining).



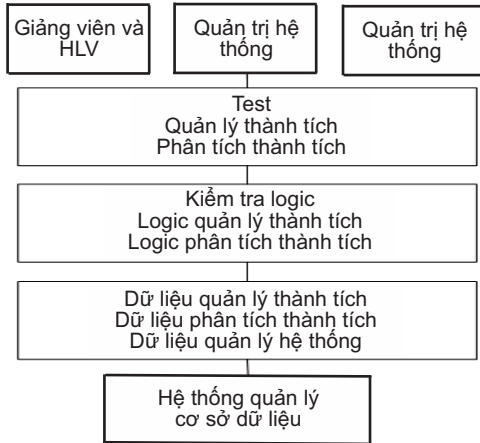
Hình 2. Tiến trình khai phá dữ liệu (Data mining)

Sau khi các trường đại học đặc tả dữ liệu quản lý thể thao, loại bỏ dữ liệu không chính xác và các quy trình quản lý dữ liệu khác, thông qua việc triển khai thiết kế kho dữ liệu, có thể xây dựng một kho dữ liệu với danh mục. Từ đó chuyển đổi thành một tập hợp các tập dữ liệu hoặc các trung tâm dữ liệu để khai thác thông qua các công cụ và phương pháp khai phá dữ liệu. Chẳng hạn như mạng nơ-ron nhân tạo (Neural Network), giải thuật di truyền, cây quyết định (DT - mô hình đưa ra quyết định dựa trên các câu hỏi), tìm kiếm lân cận gần nhất. Kết quả thu được, kiến thức hoặc chế độ khám phá và giải thích được gửi tới lãnh đạo trường đại học.

2.3. Xây dựng hệ thống thông tin quản lý thể thao của trường đại học

Việc xây dựng hệ thống thông tin quản lý TĐTT của trường đại học phải là một quy hoạch thống nhất, nhằm đáp ứng nhu cầu của các cấp quản lý khác nhau trong các trường đại học TĐTT. Cấu trúc phát triển của hệ thống được chia thành ba cấp, đó là lớp trình bày, lớp logic nghiệp vụ và lớp truy cập dữ liệu. Lớp trình bày là lớp ứng dụng hướng người dùng trực tiếp, lớp này xây dựng hệ thống thông tin quản lý kinh doanh, quản lý công việc hàng ngày, thu thập dữ liệu hoạt động giao dịch. Lớp logic nghiệp vụ ở giữa lớp trình bày và lớp truy cập dữ liệu, là phần cốt lõi của quá trình xử lý logic nghiệp vụ của toàn bộ hệ thống, chịu trách nhiệm tính toán và xử lý các yêu cầu logic trong khi nhận dữ liệu từ lớp truy cập dữ liệu của người sử dụng. Với mỗi lần truy cập dữ liệu thì các thủ tục được thực hiện

và các kết quả được trả lại cho người dùng. Hình 3 mô tả việc xây dựng hệ thống thông tin quản lý thể thao.



Hình 3. Mô tả xây dựng hệ thống thông tin quản lý thể thao

Hình trên cho thấy, hệ thống mục tiêu sẽ được thực hiện trên toàn bộ ba lớp ứng dụng như đã phân cấp, cấp cao nhất của trang điều khiển, đối tượng người dùng mà nó trở đến là giảng viên – huấn luyện viên, Sinh viên – Vận động viên và quản trị viên; lớp logic nghiệp vụ nằm ở giữa của lớp ứng dụng, công việc chính chịu trách nhiệm là phán đoán logic và thực thi, và các lớp dữ liệu đồng thời tương tác; tương tác lớp truy cập dữ liệu chủ yếu liên quan đến việc hoàn thành cơ sở dữ liệu, thực hiện nhiều yêu cầu dữ liệu. Chương trình được liên kết với cơ sở dữ liệu. Chẳng hạn với SQL Server dựa trên ASP.NET là một nền tảng ứng dụng web (web application framework), đó là hệ thống quản lý hiệu suất thể thao dựa trên nền tảng Visual Studio của Microsoft và các kỹ thuật khai phá dữ liệu khác để phân tích thành tích thể thao và đưa ra đánh giá. Chương trình được phát triển bởi hệ thống quản lý hiệu suất thể thao, có chức năng và thực tế, dễ vận hành với khả năng mở rộng, chức năng tốt và các mô-đun chức năng độc lập với nhau và phối hợp, phản ánh tư tưởng “gắn kết cao, ghép nối đơn giản” của công nghệ phần mềm.

Mô hình phân chia thành ba tầng giúp cho các tầng ứng dụng được tách biệt và chỉ tập trung vào nhiệm vụ xử lý các điều khoản của riêng chúng, không chỉ giảm sự ghép nối giữa các mô-đun mà còn cải thiện tính liên kết. Tận dụng khung phát triển này giúp giảm khối lượng công việc lập trình một cách hiệu quả khi nhu cầu thay đổi. Dễ dàng điều chỉnh để đạt được những yêu cầu bằng một số logic nghiệp vụ mà không cần viết lại các

mã lệnh hoặc tiến hành công việc sửa đổi toàn diện. Vì vậy, chỉ cần tập trung cho việc nâng cao hiệu quả và phát triển chương trình, đồng thời kiểm soát tốt hơn chi phí phát triển, điều này đóng một vai trò rất quan trọng trong việc giảm chi phí.

3. Kết luận

Với sự phát triển của cách mạng công nghiệp 4.0, gắn với kết quả thành tích thể thao qua các năm của nhiều đối tượng (học sinh, sinh viên, vận động viên), quản lý thể thao là một bộ phận quan trọng của nền giáo dục quốc dân và ngành TDTT không thể bỏ qua. Cần thiết phải triển khai phân tích yêu cầu của hệ thống và nghiên cứu về công nghệ phần mềm liên quan đến quản lý thành tích thể thao, cũng như triển khai hệ thống quản lý thành tích thể thao dựa trên khai phá dữ liệu. Trong các trường đại học TDTT, do sự phát triển và mở rộng của các dịch vụ xã hội kéo theo nhu cầu mới của chức năng quản lý thể thao đã đòi hỏi sự đổi mới mạnh mẽ và cấp bách với công tác quản lý thể thao. Quản lý bằng công nghệ thông tin là xu thế tất yếu, tính tích hợp của hệ thống quản lý thể thao hiện có chưa cao. Do vậy, xây dựng hệ thống quản lý có ý nghĩa thiết thực đối với công tác giảng dạy, huấn luyện của trường Đại học TDTT Bắc Ninh và cơ sở đào tạo khác dựa trên hệ thống thông tin quản lý khai phá dữ liệu.

4. Tài liệu tham khảo

1. Agnes Bogairdi - Meszoly, Tihamer Levendovszky, Hassan Charaf. Performance Factors in ASP.NET Web Applications with Limited Queue Model .INES International Conference on Intelligent Engineering Systems. 2008, 253-257.
2. Breiman L Friedman, H R A, Stone C J. Classification Regression Trees.Wadsworth International Group,2010,134-139.
3. Fawaz A. Masoud, Dana H. Halabi, Deema H. Halabi. asp. net frameworks and JSP frameworks in model view controller implementation .IEEE, 2009, 3593-3598.
4. Mohammad Jafar Tarokh. Javad Soroor. Supply Chain Management Information Systems Critical Failure Factors .IEEE. 2010. 425-431.
5. <http://tuyengiao.vn/nghien-cuu/ly-luan/doi-moi-giao-duc-dai-hoc-trong-boi-can-h-cach-mang-cong-nghiep-4-0-123652>, Tạp chí của Ban Tuyên giáo Trung ương (2019), “Đổi mới giáo dục đại học trong bối cảnh Cách mạng Công nghiệp 4.0”.