



Mô hình tư vấn học tập trong đào tạo trực tuyến dựa trên cộng đồng người học đa tiêu chí

BÙI XUÂN HUY^{a,*}, NGUYỄN AN TẾ^a, TRẦN THỊ SONG MINH^b

^a Trường Đại học Kinh tế TP. Hồ Chí Minh

^b Trường Đại học Kinh tế Quốc dân

THÔNG TIN	TÓM TẮT
Ngày nhận: 02/02/2021 Ngày nhận lại: 28/05/2021 Duyệt đăng: 03/06/2021 Mã phân loại JEL: C61; C63; C67 Từ khóa: Đào tạo trực tuyến; Hệ thống tư vấn. Keywords: E-learning; Recommended System.	Chuyển đổi từ hình thức đào tạo truyền thống sang môi trường trực tuyến là một xu hướng diễn ra mạnh mẽ vì những tính chất ưu việt mà nó đem lại cho cả người dạy và người học – đó là tính chia sẻ, tái sử dụng, tương tác và thích nghi... Dưới tác động của sự phát triển công nghệ cũng như ảnh hưởng của đại dịch Covid-19, từ đầu năm 2020, xu hướng trực tuyến ngày càng diễn ra mạnh mẽ hơn trước. Bài báo này sẽ trình bày về các nghiên cứu và ứng dụng của hệ thống tư vấn tự động trong môi trường đào tạo trực tuyến cũng như các hướng nghiên cứu khả thi. Abstract Shifting from traditional forms of training to online environments is a strong trend because of the preeminent properties brought to both teachers and learners. That are sharing, reusing, interactive and adaptation... under the influence of technology development as well as the influence of the Covid-19 pandemic in early 2020, this trend is even stronger than before. This paper will present research and applications of automated advisory systems in the online training environment as well as possible research directions.

* Tác giả liên hệ.

Email: huybx@ueh.edu.vn (Bùi Xuân Huy), tena@ueh.edu.vn (Nguyễn An Tế), minhst@neu.edu.vn (Trần Thị Song Minh).

Trích dẫn bài viết: Bùi Xuân Huy, Nguyễn An Tế, & Trần Thị Song Minh. (2020). Mô hình tư vấn học tập trong đào tạo trực tuyến dựa trên cộng đồng người học đa tiêu chí. *Tạp chí Nghiên cứu Kinh tế và Kinh doanh Châu Á*, 31(11), 21–35.

1. Giới thiệu

Trong cuộc cách mạng 4.0, cùng với sự phát triển mạnh mẽ của công nghệ, đào tạo trực tuyến (E-Learning) không chỉ là hoạt động chia sẻ tri thức diễn ra trên môi trường internet mà còn có thể đa dạng hóa về hình thức và trở nên linh hoạt hơn khi vượt ra khỏi giới hạn của thời gian và không gian, tăng cường tính tương tác và chủ động của cả người dạy và người học. Đào tạo trực tuyến hiện đang trở thành một làn sóng mạnh mẽ ở các cấp học khác nhau. Theo báo cáo khảo sát thường niên năm 2018 của tạp chí Training¹, thị trường học tập điện tử từ năm 2017 đến năm 2024 đã, đang và sẽ có tốc độ tăng trưởng tổng hợp hàng năm là 5%. Theo kết quả khảo sát năm 2018 của công ty đầu tư mạo hiểm chuyên về dịch vụ giáo dục Navitas Ventures², hơn 50% số người được hỏi (bao gồm lãnh đạo các trường đại học và các công ty khởi nghiệp trong lĩnh vực giáo dục ở Australia, Mỹ, Canada và Anh) tin rằng đến năm 2025 mô hình đào tạo đại học truyền thống sẽ bị phá vỡ. Tuy vậy, tác động của đại dịch Covid-19 đầu năm 2020 đã làm cho thời hạn đó bị rút ngắn lại và làn sóng triển khai các loại hình đào tạo trực tuyến ngày càng diễn ra mạnh mẽ hơn. Bên cạnh việc tương tác linh hoạt giữa người dạy và người học, khả năng truy xuất tri thức chủ động, thực hiện theo dõi và đánh giá tiến trình học tập theo một lộ trình cố định đã được quy định trước thì trên hệ thống E-Learning, người học cần có sự hỗ trợ kịp thời để tự xây dựng lộ trình học tập phù hợp với đặc điểm cá nhân, thực hiện các bài kiểm tra phù hợp với năng lực, điều chỉnh việc học dựa trên những dự báo được đề xuất từ việc phân tích kết quả kiểm tra thi cử và các hoạt động học tập khác...

Không chỉ trong lĩnh vực học tập, con người luôn luôn phải ra quyết định về nhiều vấn đề khác nhau trong cuộc sống như: Mua bán, giải trí, kinh doanh... Quá trình này phải dựa vào nguồn dữ liệu được thu thập một cách chủ động hoặc bị động và xử lý thành thông tin có ích. Ngày nay, sự phát triển của công nghệ đem đến cho con người khả năng tiếp cận với một lượng lớn dữ liệu mỗi ngày, mỗi giờ, thậm chí có thể tính đến từng giây thông qua sách vở, báo chí, kênh truyền hình và các mạng xã hội, và thậm chí các hành vi của con người cũng là một dạng dữ liệu... Các nghiên cứu cho thấy con người đọc tài liệu khoảng 10MB/ngày, nghe khoảng 400MB/ngày, và bình quân mỗi giây nhìn thấy 1MB thông tin; trong năm 2015, lượng thông tin mỗi người tiêu thụ tăng lên 74GB/ngày (UCSC, 2014). Điều này dẫn đến sự quá tải trong việc tiếp nhận và xử lý lượng dữ liệu đó để biến thành thông tin có ích hỗ trợ cho quá trình ra quyết định của người dùng. Từ thực tế đó, các hệ thống tư vấn thông tin ra đời.

Sau phần giới thiệu, bài viết được bố cục như sau: Phần 2 trình bày về các nghiên cứu và ứng dụng của hệ thống tư vấn tự động trong môi trường đào tạo trực tuyến cũng như một số vấn đề nghiên cứu khả thi; phần 3 đề xuất một mô hình tư vấn tự động trong đào tạo trực tuyến dựa trên cộng đồng người học đa tiêu chí; và cuối cùng là phần 4, kết luận.

¹ Tạp chí Training là một tạp chí phát triển nghề nghiệp, ủng hộ việc đào tạo và phát triển lực lượng lao động như một công cụ kinh doanh. Tạp chí được viết cho các chuyên gia đào tạo, nhân sự và quản lý kinh doanh trong tất cả các ngành công nghiệp. Website: trainingmag.com

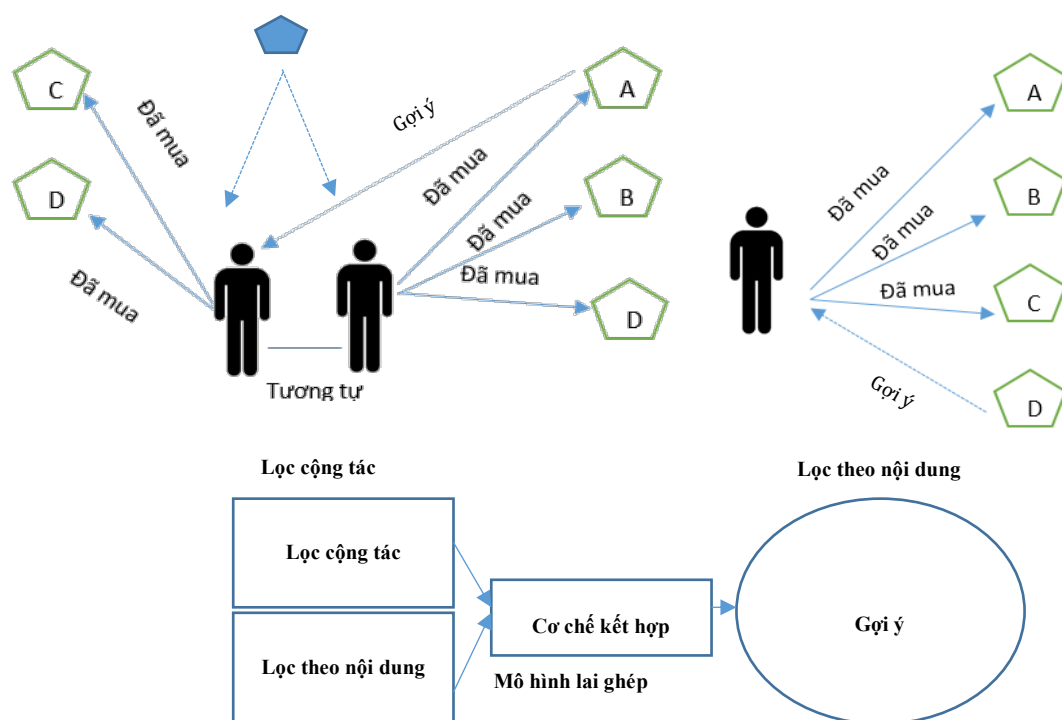
² Navitas Ventures là chi nhánh liên doanh của Navitas, cung cấp vốn đầu tư và mở rộng cũng như hỗ trợ cho các công ty khởi nghiệp trong lĩnh vực giáo dục. Website: <https://www.crunchbase.com/organization/navitas-ventures>

2. Tổng quan về hệ thống tư vấn

Trong phần này, bài viết sẽ trình bày cơ sở lý thuyết về hệ thống tư vấn và một số ứng dụng của hệ thống tư vấn trong các lĩnh vực, bao gồm cả lĩnh vực đào tạo trực tuyến, từ đó gợi mở một số vấn đề nghiên cứu mới cụ thể.

2.1. Cơ sở lý thuyết về hệ thống tư vấn

Các hệ thống tư vấn được định nghĩa là các công cụ phần mềm và kỹ thuật, cung cấp những gợi ý hỗ trợ quá trình ra quyết định đa dạng của người dùng (Ricci và cộng sự, 2011).



Hình 1. Mô hình hệ thống tư vấn

Ghi chú: A, B, C, D lần lượt là các sản phẩm, dịch vụ hoặc thông tin mà người dùng đã mua trong quá khứ, cũng như được hệ thống gợi ý.

Hệ thống tư vấn có hai mô hình cơ bản là: *Lọc cộng tác* (Collaborative Filtering), *lọc theo nội dung* (Content Based Filtering). Bên cạnh đó, mô hình lai ghép (Hybrid Filtering) là kết hợp từ hai mô hình trên để nâng cao hiệu quả tư vấn.

- *Lọc theo nội dung:* Mô hình này đưa ra gợi ý cho người dùng những sản phẩm, dịch vụ hoặc thông tin (Items) tương đồng với những Item mà họ ưa thích, đã mua hoặc cho đánh giá cao trong quá khứ (Montanner và cộng sự, 2003). Mô hình lọc theo nội dung dựa trên việc xây dựng hồ sơ đặc trưng (Profile) của người dùng, mô tả sở thích, thói quen của họ về một lĩnh vực nào đó như: Phim ảnh, sách báo, nhạc... sau đó, so khớp thông tin biểu diễn của Item với hồ sơ đặc trưng để dự đoán mức độ ưa thích của người dùng đối với Item đó và đưa ra gợi ý phù hợp. (Trong bài báo này, các cụm từ

User Profile, hồ sơ người dùng và hồ sơ đặc trưng người dùng sẽ được gọi chung là hồ sơ với ý nghĩa tương đương).

- *Lọc cộng tác*: Mô hình này định nghĩa cộng đồng của một người dùng bao gồm những người có cùng sở thích hoặc mối quan tâm với người dùng đó (Montanner và cộng sự, 2003) và những gợi ý được đưa ra dựa trên các ý kiến phản hồi (Feedback) hay đánh giá (Rating) từ cộng đồng này. Cộng đồng của người dùng được xây dựng thông qua một ma trận đánh giá (Rating Matrix) thể hiện mức độ ưa thích của mỗi người trong cộng đồng với một danh sách các Item.

Một số mô hình khác có thể kể đến là tư vấn dựa theo thông tin nhân khẩu học và tư vấn dựa trên tri thức. Cụ thể:

- *Tư vấn dựa theo thông tin nhân khẩu học* là mô hình dựa trên giả định những người dùng có cùng thông tin nhân khẩu học như: Độ tuổi, giới tính, trình độ học vấn... sẽ đánh giá tương tự nhau cho một Item. Mô hình tư vấn này sẽ phân nhóm người dùng theo thông tin nhân khẩu học (Deshpande & Karypis, 2004), xem xét mức độ ưa thích của mỗi nhóm đối với các Item và đưa ra tư vấn cho người dùng mới (Yang và cộng sự, 2014).

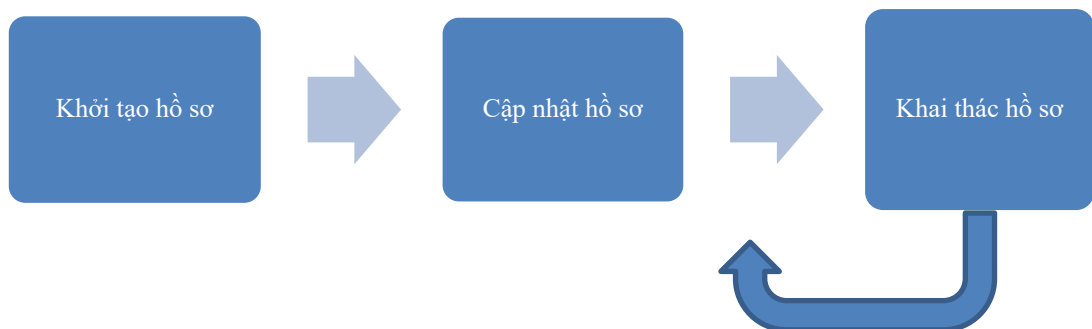
- *Tư vấn dựa trên tri thức* là mô hình tư vấn cho người dùng dựa trên dữ liệu hoạt động trong quá khứ của họ và dự báo mức độ yêu thích đối với một Item bằng cách suy diễn dựa trên những mối tương quan (Shani & Gunawardana, 2011).

Đặc điểm chung của các hệ thống tư vấn là dựa trên hồ sơ người dùng. Hình 2 mô tả hoạt động của một hệ thống tư vấn trải qua các giai đoạn như sau: (1) Khởi tạo hồ sơ người dùng, (2) cập nhật hồ sơ người dùng, (3) khai thác hồ sơ người dùng. Cụ thể:

- *Khởi tạo hồ sơ người dùng*: Cung cấp thông tin ban đầu của người dùng như: Kiến thức, sở thích, các chủ đề cần quan tâm... làm cơ sở để đánh giá, suy diễn và đưa ra gợi ý. Thông tin trong giai đoạn này thường kém chính xác.

- *Cập nhật hồ sơ người dùng*: Trong quá trình khai thác, hệ thống sẽ cập nhật lại thông tin của người dùng thông qua hoạt động của họ, kết quả tư vấn cũng như phản hồi của người dùng về hệ thống.

- *Khai thác hồ sơ người dùng*: Giai đoạn này sử dụng các cơ chế lọc cộng tác hoặc lọc theo nội dung... để đưa ra các gợi ý phù hợp với mô hình người dùng.



Hình 2. Hoạt động của hệ thống tư vấn

Việc quyết định hai người dùng có cùng thuộc một cộng đồng hay không phụ thuộc vào độ tương đồng (độ đo) giữa các Rating của họ (ví dụ, có sở thích tương tự nhau về một đối tượng nào đó, hoặc cùng mức điểm đối với môn học, tần suất truy cập hệ thống...). Các thuật toán phổ biến để tính độ tương đồng này có thể kể đến như: Pearson – Correlation, Cosine, Euclidean... (Koren, 2008). Cách tiếp cận để xây dựng cộng đồng người dùng như trên gọi là cách tiếp cận dựa trên láng giềng (Neighborhood-Based). Ngoài ra, cũng theo Breese và cộng sự (1998), còn có cách tiếp cận dựa trên mô hình (Model-Based Approach) (Breese và cộng sự, 1998). Mô hình này sử dụng các thuật toán gom cụm (Clustering) để chia tập các người dùng thành k cộng đồng dựa theo mức độ tương đồng. Khi có người sử dụng mới, người đó sẽ được hệ thống phân vào cộng đồng phù hợp dựa trên việc xem xét hồ sơ của họ và các đặc trưng của từng cộng đồng. Từ đó, hệ thống sẽ đưa ra gợi ý phù hợp. Những thuật toán phân cụm tiêu biểu có thể kể đến như: K-Mean, Fuzzy C-mean, Expectation-Maximization. Cách tiếp cận thứ ba là dựa trên mô hình đặc trưng ngầm (Latent Factor Model) với thuật toán thừa số hóa ma trận (*Matrix Factorization*) (Koren & Bell, 2011; Nguyễn Thúy Ngọc & Nguyễn An Tê, 2011). Tuy nhiên, việc xem xét người dùng chỉ trong một cộng đồng vẫn còn hạn chế vì trong thực tế, một người có thể thuộc về nhiều cộng đồng khác nhau, như vậy, đây cũng là một nguồn tri thức có khả năng ảnh hưởng đến kết quả tư vấn (Chen và cộng sự, 2013).

2.2. Một số ứng dụng của hệ thống tư vấn và các hướng nghiên cứu khả thi

Con người luôn phải ra quyết định trong cuộc sống hàng ngày ở nhiều lĩnh vực khác nhau như: Kinh doanh, mua sắm, đào tạo, giải trí... Tuy nhiên, với sự bùng nổ thông tin thì việc ra quyết định một cách hiệu quả là vấn đề nan giải nếu không có sự hỗ trợ của các hệ thống tư vấn tự động. Dưới đây là ứng dụng của hệ thống tư vấn trong một số lĩnh vực.

Trước hết là mô hình lọc cộng tác sử dụng vị trí như một yếu tố ảnh hưởng. Ở mô hình này, người dùng được hệ thống phân loại theo vị trí hiện tại của họ và các dịch vụ website mà họ sử dụng. Khi có người dùng mới, hệ thống sẽ tư vấn các dịch vụ website phù hợp dựa theo vị trí và hồ sơ của những người dùng trước cùng ở vị trí đó (Chen và cộng sự, 2013), hay tư vấn dịch vụ website được cá nhân hóa theo vị trí (Liu và cộng sự, 2016).

Trong một nghiên cứu khác ở lĩnh vực giải trí, thời gian được xem xét như một yếu tố ảnh hưởng đến sở thích của người dùng, trong đó, Shen và cộng sự (2016) đề xuất hệ thống gợi ý có tên là TOTAR (Temporal Overlapping Community Detection Using Time-Weighted Association Rules). Hệ thống được thử nghiệm trên dữ liệu của MovieLen và NetFlix. Kết quả thử nghiệm cho thấy TOTAR vượt trội so với các thuật toán trước đây về độ chính xác và tính đa dạng.

Trong lĩnh vực du lịch, một nghiên cứu đề xuất hệ thống tư vấn về các điểm du lịch thu hút dựa trên tri thức thu thập được từ các website truyền thông xã hội (Social Media), trong đó, Shen và cộng sự (2016) sử dụng kỹ thuật phân lớp để tìm kiếm các điểm du lịch có mức độ hấp dẫn tương tự nhau dựa trên phản hồi của khách hàng, sau đó, kết hợp với thông tin du lịch để tư vấn cho khách hàng mới. Nguyen và Ricci (2018) đề xuất một hệ thống tư vấn du lịch dựa trên việc ghi nhận và phân tích các hành vi, xu hướng sở thích của người dùng trong các nhóm chat. Ở một nghiên cứu khác, Polatidis và Georgiadis (2016) đề xuất mô hình tư vấn đa cấp độ cải thiện chất lượng của gợi ý và loại bỏ dữ liệu nhiễu bằng cách sử dụng các chỉ số tương quan Pearson Correlation và hệ số tương quan Pearson có trọng số.

Trong lĩnh vực kinh doanh, Xu và cộng sự (2016) đề xuất một mô hình tư vấn dịch vụ thông tin tùy chỉnh hai cấp độ sử dụng kỹ thuật AHP để tổng hợp thứ hạng cho các mục được gợi ý. Mô hình này tư vấn thông tin dự án nghiên cứu và phát triển (R&D) được cá nhân hóa cho khách hàng. Một số hệ thống tư vấn khác, như: Hệ thống tư vấn người dùng tìm kiếm nhà ở (Badriyah và cộng sự, 2018) đề xuất một hệ thống tư vấn dựa trên nội dung, hệ thống này phân nhóm nhà ở theo các tính chất như: Vị trí, giá tiền, diện tích... và xây dựng hồ sơ người dùng dựa trên các hoạt động truy cập vào những mẫu quảng cáo của họ – hồ sơ này được xử lý bằng thuật toán Apriori để tìm ra những mẫu nhà phù hợp; hay Panda và cộng sự (2020) đề xuất một hệ thống lọc cộng tác chuẩn hóa để khắc phục nhược điểm không đồng nhất của các ma trận đánh giá, sự không đồng nhất này là do một số mục được người dùng đánh giá nghiêm ngặt, nhưng một số khác thì không – kết quả thực nghiệm được kiểm chứng bằng kiểm định t-test với độ tin cậy hơn 95%.

- *Ứng dụng của hệ thống tư vấn trong môi trường E-Learning*

Trong phần tiếp theo, nghiên cứu này tập trung vào hệ thống tư vấn tự động trong môi trường E-Learning. Do đó, từ “người học” sẽ được sử dụng thay thế cho từ “người dùng” với ý nghĩa tương đương.

Một hệ thống tư vấn học tập có các hình thức vô cùng phong phú, từ hệ thống tư vấn tài nguyên cho người học, hệ thống hướng dẫn người học giải quyết các vấn đề của môn học đến hệ thống tư vấn khóa học phù hợp (Verbert và cộng sự, 2012). Các thuật toán được dùng trong hệ thống này thường sử dụng thông tin về người học và thông tin về tài nguyên để đưa ra gợi ý thích hợp. Mức độ tri thức về người học cũng như: tài nguyên học tập (như: Kiến thức, thành tích học tập trong quá khứ, hoạt động trực tuyến, phong cách học tập...) giúp cá nhân hóa quá trình tư vấn. Trong khảo sát của mình, nhóm tác giả cũng chỉ ra các thách thức trong việc xây dựng một hệ thống tư vấn học tập bao gồm:

- *Thu thập và biểu diễn thông tin ngữ cảnh của người học*: Các thông tin ngữ cảnh của người học (như: Thông tin cá nhân, trình độ học vấn, sở thích, vị trí, thiết bị sử dụng, quan hệ xã hội, độ nhiễu của thông tin) nếu chỉ dựa vào kết quả khảo sát ban đầu thì thường không chính xác, chưa kể đến việc nó có thể thường xuyên thay đổi mà sự thay đổi này có thể chính người học cũng không nhận biết đầy đủ.

- *Đánh giá hiệu quả của hệ thống tư vấn*: Khó khăn của việc xác định mức độ ảnh hưởng của các yếu tố trong ngữ cảnh đến kết quả tư vấn và mức độ phù hợp của kết quả tư vấn đó đối với người học.

- *Vấn đề xử lý dữ liệu lớn*: Trong bối cảnh phát triển công nghệ thông tin và truyền thông, việc thu thập, xử lý và lưu trữ tập dữ liệu của người học trong các ngữ cảnh cũng là một thách thức cần giải quyết.

- *Vấn đề tương tác*: Không chỉ là xây dựng một giao diện ứng dụng dễ hiểu mà còn cho phép người học hiểu rõ và thay đổi các thông số đầu vào để trải nghiệm các gợi ý khác nhau hơn là tiếp nhận các gợi ý một cách bị động, từ đó tăng tính tin cậy của hệ thống tư vấn.

- *Tính riêng tư*: Đây cũng là một thách thức đối với một hệ thống hoạt động dựa trên việc thu thập thông tin của người học.

Manouselis và cộng sự (2012) đã đưa ra khung phân loại các hệ thống tư vấn học tập dựa trên các yếu tố sau:

Các tác vụ được hỗ trợ, bao gồm: Tìm kiếm/tư vấn nội dung hoặc một chuỗi các nội dung dựa trên những nội dung hiện tại của người học; tư vấn người học có đặc điểm tương tự dành cho hoạt động đào tạo từ xa; dự đoán kết quả học tập hoặc tư vấn hoạt động học tập.

Cách tiếp cận, bao gồm: Xây dựng mô hình người học, mô hình tri thức, cá nhân hóa. Trong đó, mô hình người học được xây dựng dựa trên các phương pháp như: Không gian Vector, Ontology, nhân khẩu học... Phương pháp khởi tạo mô hình người học cũng được nghiên cứu như: Nhập tay từ người học, tự động thu thập trong quá trình người học tương tác với hệ thống hoặc dùng các thuật toán khai phá dữ liệu (phân cụm, phân lớp)... Mô hình tri thức cũng được xây dựng dựa trên những phương pháp tương tự, ngoài ra, có những phương pháp đặc thù như: Phân loại, chỉ mục, đồ thị, luật suy diễn... và cũng dùng các thuật toán khai phá dữ liệu, phân tích tuần tự để khởi tạo mô hình. Cách tiếp cận cá nhân hóa chính là các cơ chế thích nghi để đưa ra kết quả tư vấn cho người học như: Lọc cộng tác, lọc dựa trên nội dung, dựa trên luật suy diễn, phương pháp đồ thị...

- *Vận hành*: Bao gồm kiến trúc tập trung hay phân tán; các gợi ý được đưa ra từ máy chủ hay tích hợp ngay trong thiết bị của người học; người học tiếp nhận gợi ý một cách chủ động hoặc bị động...

Một số nghiên cứu về xây dựng mô hình người học có thể kể đến bao gồm: Brusilovsky và Millán (2007) giới thiệu hai cách tiếp cận để xây dựng mô hình người học trong hệ thống tư vấn: Cách tiếp cận lớp phủ và cách tiếp cận dựa trên thông tin không chắc chắn; Pan và cộng sự (2021) đề xuất thuật toán cải thiện độ chính xác của hệ thống tư vấn dựa trên thẻ TRS (Tag Aware Recommender System) bằng cách tái cấu trúc hồ sơ người học – việc này được thực hiện bằng cách nhận biết các thẻ (Tag) người học gắn lên tài nguyên và từ đó suy diễn ra các thẻ có liên quan.

Một số nghiên cứu áp dụng lọc cộng tác có thể kể đến bao gồm: Vesin và cộng sự (2013) giới thiệu ứng dụng Protus, đánh giá và tư vấn các khối kiến thức phù hợp cho sinh viên học môn lập trình – hệ thống này áp dụng phương pháp lọc công tác và xây dựng mô hình người học để phân loại sinh viên thành các nhóm tương đồng dựa trên thông tin khách quan, bao gồm: Dữ liệu do sinh viên trực tiếp cung cấp (chẳng hạn như: Dữ liệu cá nhân, kiến thức trước đây, sở thích...), thành tích của sinh viên (bao gồm: Dữ liệu về mức độ hiểu biết lĩnh vực đang học, quan niệm, sự tiến bộ và hiệu suất tổng thể của mỗi người), lịch sử học tập (bao gồm: Thông tin về các bài học và bài kiểm tra mà sinh viên đã học, tương tác của họ với hệ thống, các bài đánh giá mà họ đã trải qua...); John và cộng sự (2018) đã đề xuất hệ thống tư vấn cộng tác dựa trên việc kết hợp thông tin ngữ cảnh (như: Trình độ kiến thức và mục tiêu học tập...) vào hồ sơ người học và sử dụng thuật toán phân tích chuỗi tuần tự để khai thác nhật ký website nhằm khám phá các mẫu truy cập tuần tự của người học để tăng độ chính xác của kết quả tư vấn; Albatayneh và cộng sự (2018) giới thiệu một kiến trúc tư vấn cộng tác có thể đề xuất các bài đăng thú vị cho người học trong diễn đàn thảo luận trực tuyến bằng kỹ thuật lọc nội dung dựa trên ngữ nghĩa và xếp hạng tiêu cực của người học về các bài đăng. Brik và Touhria (2020) giả thuyết rằng thông tin ngữ cảnh là một yếu tố ảnh hưởng đến kết quả tư vấn và từ đó đề xuất mô hình lọc cộng tác có thể tự động thu thập và phân tích dữ liệu về hành vi của người học để xác định những thông tin ngữ cảnh phù hợp hoặc không phù hợp, bên cạnh đó, nghiên cứu này cũng đề xuất phương pháp xây dựng hồ sơ người học bằng phương pháp Ontology; Yassine và cộng sự (2021) đề xuất hệ thống lọc cộng tác kết hợp thuật toán K-means – thuật toán thực hiện hiệu quả trên bộ dữ liệu thử nghiệm GroupLens.

Một số nghiên cứu áp dụng lọc theo nội dung có thể kể đến bao gồm: Ghauth và Abdulla (2011) xây dựng mô hình tư vấn theo nội dung – mô hình này tư vấn tài liệu học tập cho sinh viên dựa trên

dữ liệu đánh giá của những sinh viên có kết quả học tốt, các tài liệu được phân thành các nhóm tương tự nhau và điểm đánh giá của từng tài liệu trong nhóm là điểm trung bình của nhóm đó; Jiangbo và cộng sự (2018) đề xuất hệ thống tư vấn theo nội dung hiện thực bằng thuật toán mạng Neuron phức hợp để tư vấn tài nguyên học tập cho người học (Jiangbo và cộng sự, 2018).

Một số nghiên cứu và ứng dụng có liên quan khác, bao gồm: Shishehchi và cộng sự (2012) đề xuất một nghiên cứu xây dựng hệ thống tư vấn dựa trên tri thức bằng tiếp cận Ontology để cải thiện chất lượng tư vấn trên hệ thống E-Learning; Obeid và cộng sự (2018) đề xuất hệ thống tư vấn dựa trên Ontology và kỹ thuật máy học cho phép đánh giá điểm mạnh và điểm yếu, sở thích và năng lực nghề nghiệp của học sinh để giới thiệu chuyên ngành và trường đại học phù hợp cho từng người; trợ lý ảo Jill Watson tích hợp hệ thống đào tạo trực tuyến MOOCs để cải thiện việc học và tăng cường sự tương tác giữa giáo viên và sinh viên hệ thống có khả năng tự động trả lời các câu hỏi của sinh viên thường xuyên và liên tục (Ashok & Lalith, 2020); Niranjana và cộng sự (2012) đề xuất xây dựng hệ thống trả lời tự động (Chatbot) bằng cách sử dụng mô hình mạng Bayes, trong đó, hệ thống sẽ trích xuất các từ khóa nằm trong câu hỏi, sau đó phân tích cú pháp bằng phương pháp xác suất Bayes để lựa chọn câu trả lời phù hợp nhất với câu hỏi.

Tại Việt Nam, trong những năm gần đây cũng có rất nhiều nghiên cứu về lĩnh vực E-Learning, hầu hết đều đi vào việc áp dụng triển khai cho các môn học cụ thể. Tuy nhiên, bên cạnh đó, có hai luận án mà mục tiêu và nội dung nghiên cứu có tính tương đồng rất cao với ý tưởng nghiên cứu của đề tài này. Lê Đức Long (2014) đưa ra cách thức biểu diễn mô hình nội dung tri thức (Knowledge Graph) để tổ chức cấu trúc một chương trình đào tạo hay một học phần từ những đơn vị tri thức riêng biệt, sau đó khai thác nội dung tri thức đó trong các ngữ cảnh học tập khác nhau. Lê và cộng sự (2009) đưa ra một cách tiếp cận khác để xây dựng mô hình người học dựa trên việc đánh giá tính hữu ích của các đặc trưng của họ. Tính hữu ích này phụ thuộc vào cộng đồng mà người học đó thuộc về và phải có tính phân loại. Ví dụ: Trong một cộng đồng đa giới tính, thì đặc trưng về giới là đặc trưng hữu ích, nhưng trong một cộng đồng chỉ toàn nam hoặc nữ thì đặc trưng này là vô ích. Trong một nghiên cứu khác, Lê Đức Long và cộng sự (2009) cũng đề xuất một cấu trúc của mô hình người học, bao gồm: Các thông tin về nhân khẩu, kinh nghiệm học vấn, hoạt động tự học, nhu cầu và động cơ học tập và đưa ra cách khởi tạo và cập nhật thông tin trong hồ sơ của mỗi cá nhân bằng các luật suy diễn để tạo ra một hồ sơ phản ánh đúng hơn các đặc trưng của người học.

Trong các phần trên, bài báo đã trình bày tổng quan về hệ thống tư vấn tự động và ứng dụng trong đào tạo trực tuyến trong và ngoài nước và phân tích, đánh giá các công nghệ mới được triển khai trong đào tạo trực tuyến (E-Learning), từ đó gợi mở một số vấn đề nghiên cứu mới cụ thể như:

- *Thứ nhất*, việc xây dựng một hệ thống tư vấn theo mô hình lọc cộng tác đặt người học trong một cộng đồng cụ thể để đưa ra gợi ý, nhờ đó có khả năng phát hiện những thông tin phù hợp mà chính người học cũng không nhận biết. Tuy nhiên, hầu hết các nghiên cứu hiện nay đều chỉ xem xét người học trong một cộng đồng. Trong thực tế, một người có thể thuộc về nhiều cộng đồng khác nhau, như vậy, đây cũng là một nguồn tri thức có khả năng ảnh hưởng đến kết quả tư vấn.

- *Thứ hai*, việc xem xét người học trong nhiều cộng đồng khác nhau có thể dẫn đến những tư vấn khác nhau, như vậy sẽ đặt ra vấn đề đánh giá sự phù hợp và mức độ chính xác của kết quả tư vấn.

- *Thứ ba*, tại Việt Nam, mặc dù chủ trương chung là đẩy mạnh việc giảng dạy trực tuyến nhưng gần như tất cả các trường đều đang áp dụng phối hợp cả hai hình thức học offline và online. Như vậy,

việc xác định các tiêu chí hữu ích để xây dựng hồ sơ người học và các cộng đồng người học thuộc về là một thách thức lớn.

- *Thứ tư*, hệ thống tư vấn tự động hướng đến có thể tư vấn về khối kiến thức, tài liệu học, xây dựng nội dung kiểm tra, dự báo kết quả học tập trong một môn học hoặc các môn cùng ngành/chuyên ngành... Như vậy, vấn đề thứ tư liên quan đến việc xây dựng cấu trúc khối kiến thức của môn học, bài kiểm tra, bài toán dự báo...

Trên đây là các khoảng trống nghiên cứu và là động lực để nhóm tác giả đề xuất mô hình tư vấn tự động trong đào tạo trực tuyến ở mục tiếp theo.

3. Đề xuất mô hình tư vấn tự động trong đào tạo trực tuyến dựa trên cộng đồng người học đa tiêu chí

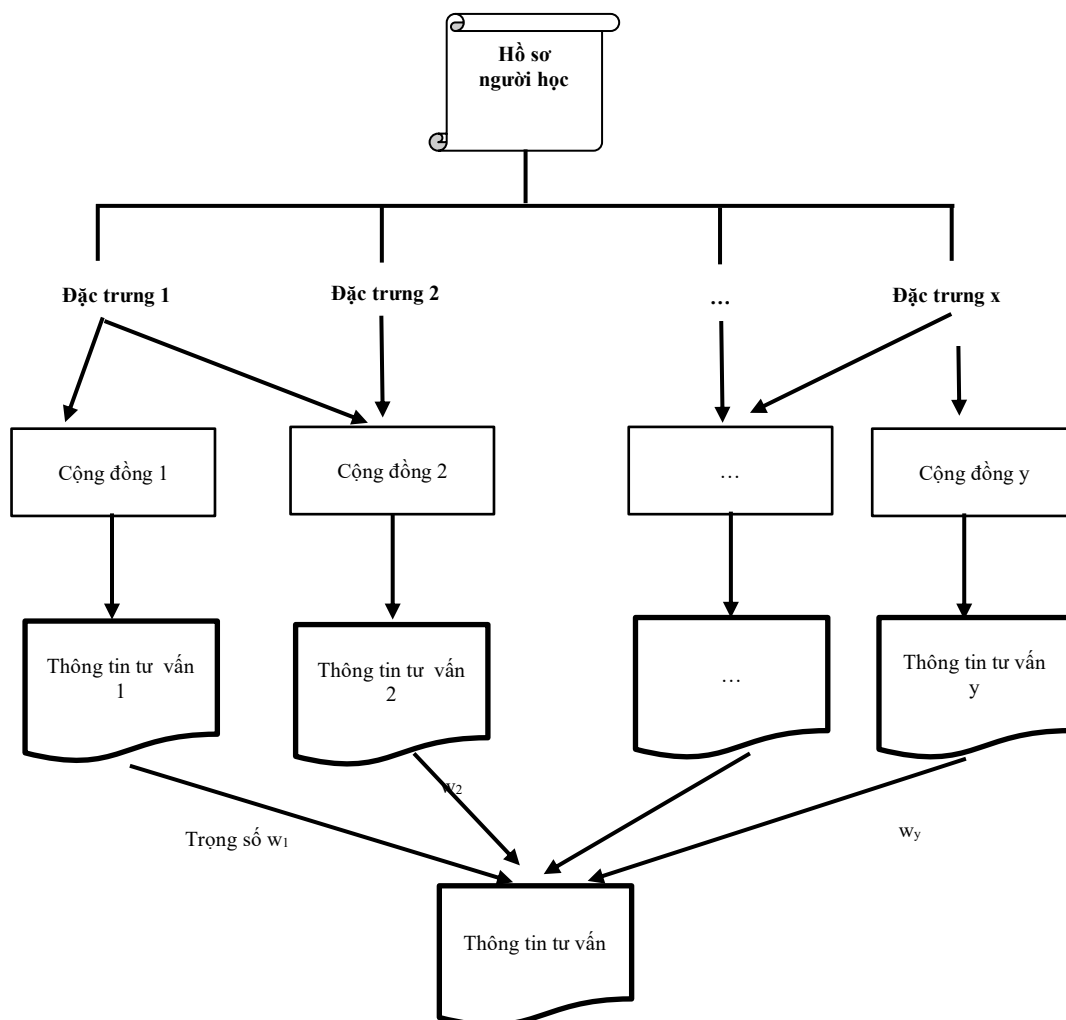
Những đặc trưng của người học được thể hiện qua cấu trúc của hồ sơ trong các hệ thống E-Learning phần lớn bao gồm những thông tin về nhân khẩu học và trình độ, nền tảng kiến thức như bằng cấp hay các tín chỉ đã đạt được trong các khóa học – đây là những thông tin được khởi tạo và cập nhật chủ yếu ở đầu hay cuối một giai đoạn học tập; còn lại, những đặc trưng khác dựa trên sự phân tích quá trình học tập của người học thì chưa được nghiên cứu, khai thác nhiều. Điều này khiến cho sự tiến hóa của hồ sơ kém linh hoạt. Ngoài ra, sự thích nghi trong các hệ thống E-Learning cũng chỉ dừng lại ở mức độ đáp ứng nhu cầu cho từng cá nhân riêng lẻ, trong khi vấn đề tương tác cộng đồng đóng một vai trò rất quan trọng trong E-Learning thì chưa được quan tâm đúng mức. Gần đây, các nghiên cứu về khai thác yếu tố cộng đồng trong E-Learning cũng chỉ tập trung chủ yếu vào việc ứng dụng phương pháp chia sẻ tài nguyên dựa trên sự cộng tác (Collaborative Filtering) (Breese và cộng sự, 1998; Burke, 2007; Madhour và Forte, 2006) mà vẫn chưa chú trọng đến sự tương tác trong cộng đồng, nghĩa là hệ thống hoàn toàn không có sự hỗ trợ nào đáng kể cho người học trong việc hình thành các nhóm học tập trong các môn học dựa trên những đặc trưng của họ. Vì vậy, mô hình tư vấn mà nhóm tác giả đề xuất hướng tới là việc giải quyết sự tương tác giữa người học với hệ thống cũng như sự tương tác giữa người học với nhau để cung cấp nhiều hình thức tư vấn cho các đối tượng trong một hệ thống E-Learning, đặc biệt là trong những hoạt động học tập của người học như: Tư vấn cách thức học tập và gợi ý nhóm học tập, cung cấp tài nguyên, tư vấn lựa chọn môn học, thiết kế bài kiểm tra, đánh giá kết quả học tập... bằng cách dựa trên những đặc trưng (nền tảng kiến thức, điều kiện và năng lực học tập...) của chính cá nhân họ cũng như dựa trên những bạn học có những đặc trưng tương tự.

3.1. Cộng đồng người học (nhóm học tập) đa tiêu chí

Một cộng đồng trong không gian cộng đồng đa tiêu chí được thành lập dựa trên nhóm người có đặc trưng/tiêu chí chung nào đó (tuổi tác, trình độ kiến thức, tần suất truy cập mạng xã hội...). Như vậy, một người có thể thuộc nhiều cộng đồng khác nhau nếu xét về các đặc trưng/tiêu chí khác nhau. Ví dụ: Một sinh viên có thể thuộc cộng đồng những sinh viên công nghệ thông tin, đồng thời cũng có thể thuộc về cộng đồng sinh viên học song ngành. Hiển nhiên, chúng ta cũng có thể hình dung ra những tiêu chí phức tạp hơn, kết hợp từ nhiều đặc trưng trong hồ sơ. Ví dụ: Một tiêu chí thành lập nhóm học tập với sự phối hợp của ba đặc trưng về nghề nghiệp, độ tuổi và nơi cư trú. Trong cách tiếp cận lọc cộng tác cổ điển được áp dụng trong các hệ thống tư vấn (Breese và cộng sự, 1998; Montanner và cộng sự, 2003), các cộng đồng chỉ được xem như loại đối tượng tồn tại nhất thời, là kết quả trung

gian của quá trình tính toán để tạo ra kết quả tư vấn sau cùng. Ngược lại, cộng đồng người học được xem xét ở đây sẽ được tổ chức như những đối tượng tồn tại một cách tường minh và bền vững: Từ khi bắt đầu được tạo lập và luôn cả quá trình tiến hóa về sau (thêm, bớt thành viên, cập nhật những đặc trưng). Một người học có thể thuộc về nhiều cộng đồng khác nhau nên hiển nhiên sẽ nhận được nhiều gợi ý khác nhau nên cần xem xét bộ hệ số để đánh giá mức độ phù hợp của các gợi ý trước khi đưa ra gợi ý cuối cùng cho người học.

Hình 3 minh họa mô hình cộng đồng người học đa tiêu chí với cơ chế đưa ra gợi ý. Trong đó, w_i là các hệ số đánh giá mức độ phù hợp mỗi gợi ý từ các cộng đồng khác nhau.



Hình 3. Mô hình cộng đồng người học đa tiêu chí

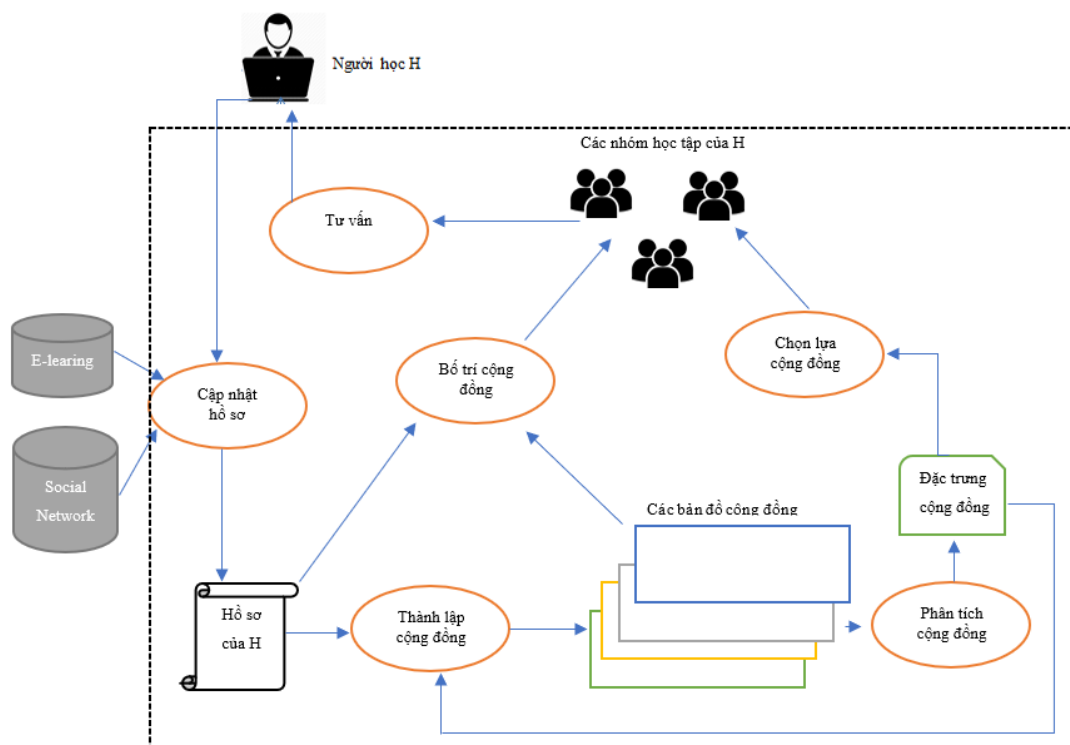
Nguồn: Nguyễn Thúy Ngọc và Nguyễn An Tế (2015).

3.2. Mô hình tư vấn tự động dựa trên cộng đồng người học đa tiêu chí

Từ những nội dung đã trình bày ở trên, nhóm tác giả đề xuất mô hình tư vấn tự động dựa trên cộng đồng người học đa tiêu chí bao gồm các Module tư vấn và các Module xử lý hoạt động dựa trên việc xây dựng và khai thác hồ sơ cũng như cộng đồng của người học (Hình 4).

- *Bản đồ cộng đồng*: Toàn bộ những người học có thể được tổ chức thành những cộng đồng theo nhiều tiêu chí khác nhau được rút trích từ những đặc trưng trong hồ sơ, ví dụ, toàn thể người học có thể được phân chia thành tập hợp các cộng đồng theo nghề nghiệp, theo nơi cư trú... Nói khác đi, ứng với mỗi tiêu chí cụ thể, hệ thống có thể tạo ra một phân hoạch gồm nhiều cộng đồng, và bản đồ cộng đồng là hình thức thể hiện trực quan cho một phân hoạch. Thông thường, mỗi tiêu chí khác nhau sẽ tạo nên một phân hoạch khác nhau trên tập hợp những người học của hệ thống.

- *Đặc trưng cộng đồng*: Về nguyên tắc, bản thân tiêu chí thành lập cũng có thể được dùng làm đặc trưng của cộng đồng. Ví dụ: Cộng đồng A gồm những giáo viên trung học, cộng đồng B gồm những người cư trú tại Cà Mau... Tuy nhiên, theo thời gian, việc thêm, bớt thành viên sẽ dẫn đến vấn đề là tiêu chí để thành lập ban đầu không còn phản ánh đúng hoàn toàn đặc trưng của cộng đồng. Vì bản thân mỗi thành viên đã có sự cập nhật liên tục hồ sơ của mình. Lúc này, hệ thống cần phải xây dựng đặc trưng thật sự cho cộng đồng. Đây còn gọi là sự tiến hóa của cộng đồng. Việc xây dựng đặc trưng cộng đồng là một trong những cách có thể giúp bố trí người học vào một nhóm học tập phù hợp. Thay cho việc so khớp hồ sơ của người học với hồ sơ từng thành viên của nhóm, hệ thống sẽ so khớp với đặc trưng chung của nhóm đó. Ngoài ra, chính người học cũng có thể tự đánh giá (một cách chủ quan) mức độ phù hợp giữa mình với nhóm thông qua việc xem xét đặc trưng của nhóm đó, để từ đó có thể chủ động quyết định có nên tham gia vào nhóm hay không.



Hình 4. Mô hình tư vấn tự động dựa trên cộng đồng người học đa tiêu chí

- *Module tư vấn*: Đây là Module quan trọng nhất bao gồm các chức năng: Tư vấn đăng ký môn học, tư vấn cách thức học tập, tư vấn tài nguyên học tập, tư vấn nhóm học tập... Những chức năng tư vấn của Module này đều dựa vào những khái niệm đã đề cập ở trên.

- *Các Module xử lý*: Bao gồm cập nhật hồ sơ, thành lập cộng đồng, phân tích cộng đồng, chọn lựa cộng đồng và bố trí cộng đồng. Trong đó, Module phân tích cộng đồng đảm nhận nhiệm vụ phân tích và xác định đặc trưng trong quá trình tiến hóa của cộng đồng; Module bố trí cộng đồng tự động xác định cộng đồng cho người học; Module chọn lựa cộng đồng cho phép người học tự chọn cộng đồng mà mình thấy phù hợp.

4. Kết luận

Bài báo này đã trình bày tổng quan về hệ thống tư vấn tự động và ứng dụng trong đào tạo trực tuyến trong và ngoài nước. Trong đó, nhóm tác giả đã phân tích, đánh giá các công nghệ mới được triển khai trong đào tạo trực tuyến (E-Learning), từ đó gợi mở một số vấn đề nghiên cứu mới; bên cạnh đó, nhóm tác giả cũng đưa ra ý tưởng về mô hình tư vấn tự động trong E-Learning dựa trên cộng đồng người học đa tiêu chí. Mô hình đặt ra nhiều hướng phát triển cho hệ thống như:

- Xây dựng, cập nhật, khai thác hồ sơ của người học.
- Tích hợp mô hình vào E-Learning.
- Khai thác thông tin từ các hoạt động của người học nếu họ tham gia vào mạng xã hội.
- Hiện thực hóa các Module tư vấn.
- Đánh giá chất lượng của tư vấn vì có thể xảy ra trường hợp người học nhận được những tư vấn trái ngược từ các cộng đồng khác nhau.
- Phát triển các Module hỗ trợ giảng dạy và quản lý chuyên môn■

Tài liệu tham khảo

- Albatayneh, N. A., Ghauth, K. I., & Chua, F.-F. (2018). Utilizing learners' negative ratings in semantic content-based recommender system for e-learning forum. *Educational Technology & Society*, 21(1), 112–125.
- Ashok, K. G., & Lalith, P. (2020). *Jill Watson: A virtual teaching assistant for online education*. Retrieved from <http://hdl.handle.net/1853/59104>
- Bradiyah, T., Azvy, S., Yuwono, W., & Syarif, I. (2018). *Recommendation system for property search using content based filtering method*. International Conference on Information and Communications Technology (ICOIACT) (pp.25–29). Yogyakarta: IEEE.
- Breese, J. S. Heckerman, D., & Kadie, C. (1998). *Empirical Analysis of Predictive Algorithms of Collaborative Filtering*. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (pp. 43–52).
- Brik, M., & Touahria, M. (2020). Contextual information retrieval within recommender system: Case study “E-learning system”. *TEM Journal*, 9(3), 1150–1162.

- Brusilovsky, P., & Millán, E. (2007). User Models for adaptive hypermedia and adaptive educational systems. In: P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The Adaptive Web. Lecture Notes in Computer Science* (vol. 4321). Berlin: Springer. doi: 10.1007/978-3-540-72079-9_1
- Burke, R. (2007). Hybrid web recommender systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl, *The Adaptive Web: Methods and Strategies of Web Personalization* (vol. 4321). Switzerland, AG.: Springer Nature.
- Chen, X., Zheng, Z., Liu, X., Hoang, Z., & Sun, H. (2013). Personalized QoS aware web service recommendation and visualization. *IEEE Transactions on Services Computing*, 6(1), 35–47.
- Deshpande, M., & Karypis, G. (2004). Item-based top-N recommendation algorithms. *ACM Transactions on Information Systems*, 22(1), 143–177.
- Ghauth, K. I., & Abdulla, N. A. (2011). The effect of incorporating good learners' ratings in E-learning content based recommender system. *Educational Technology and Society*, 14(2), 248–257.
- Jiangbo, S., Xiaoxuan, S., Hai, L., Baolin, Y., & Zhaoli, Z. (2018). A content-based recommendation algorithm for learning resources. *Multimedia Systems*, 24, 163–173.
- John, K. T., Zhendong, N., & Dorothy, K. (2018). A hybrid recommender system for E-learning based on context awareness and sequential pattern mining. *Soft Computing*, 22(8), 2449–2461.
- Koren, Y. (2008). *Factorization meets the neighborhood: A multifaceted collaborative filtering model*. International Conference on Knowledge Discovery and Data Mining, Nevada USA.
- Koren, Y., & Bell, R. (2011). Advances in Collaborative Filtering. *Recommender Systems Handbook* (pp. 145–186). Boston: Springer.
- Le, T. H, Le, D. L., Nguyen, A. T., & Nguyen, D. T. (2009). *An approach to building profile based on the utility of features in the adaptive systems*. Proceedings of the IADIS International Conference on WWW/Internet.
- Lê Đức Long. (2014). *Mô hình biểu diễn tri thức cho hệ tương tác tích cực*. Luận án Tiến sĩ, Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia TP.HCM.
- Lê Đức Long, Võ Thành C, Nguyễn An Tế, & Trần Văn Hạo. (2009). Xây dựng mô hình đặc trưng người học trong đào tạo trực tuyến thích nghi. *Tạp chí Khoa học Giáo dục Kỹ thuật*, 9, 30–41.
- Liu, J., Tang, M., & Zheng, Z. (2016). Location aware and personalized collaborative filtering for web service recommendation. *IEEE Trans Serv Computing*, 9, 686–699.
- Madhour, H., & Forte, M. W. (2006). Lausanne model: Towards an AHAM-Based reference model for adaptive learning object systems. *Advanced Technology for Learning*, 3(3), 159–169.
- Manouselis, N., Drachler, H., Verbert, K., & Duval, E. (2012). *Recommender Systems for Learning*. Berlin: Springer.
- Montanner, M., López, B., & Rosa, J. (2003). A taxonomy of recommender agents on the internet. *Artificial Intelligence Review*, 19, 285–330.

- Niranjan, M., Saipreethy, M. S., & Kumar, T. G. (2012). *An intelligent question answering conversational agent using Naïve Bayesian classifier*. IEEE International Conference on Technology Enhanced Education (ICTEE). Amritapuri: IEEE. doi: 10.1109/ICTEE.2012.6208614
- Nguyen, T. N., & Ricci, F. (2018). A chat-based group recommender system for tourism. *Information Technology and Tourism*, 18, 5–28.
- Nguyễn Thúy Ngọc, & Nguyễn An Tế. (2011). Phương pháp xây dựng hệ thống thông tin tư vấn học tập trong E-learning. *Tạp chí Khoa học Giáo dục Kỹ thuật*, 17, 28–41.
- Nguyễn Thúy Ngọc, & Nguyễn An Tế. (2015). Weighing the role of multi-criteria communities for recommender systems. *International Journal of Intelligent Engineering Informatics*, 3(4), 330–348.
- Obeid, C., Lahoud, I., Khoury, H., & Champin, P.-A. (2018). *Ontology-based recommender system in higher education*. WWW '18: Companion Proceedings of the The Web Conference 2018 (pp.1031–1034). doi: 10.1145/3184558.3191533
- Panda, S. K., Bhoi, S. K., & Singh, M. A. (2020). Collaborative filtering recommendation algorithm based on normalization approach. *Journal of Ambient Intelligence and Humanized Computing*, 11, 4643–4665. doi: 10.1007/s12652-020-01711-x
- Polatidis, N., & Georgiadis, C. K. (2016). A multi-level collaborative filtering method that improves recommendations. *Expert Systems with Applications*, 48, 100–110.
- Pan, Y., Huo, Y., Tang, J., Zeng, Y., & Chen, B. (2021). Exploiting relational tag expansion for dynamic user profile in a tag-aware ranking recommender system. *Information Sciences*, 545, 448–464.
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. (2011). *Recommender Systems Handbook*. Boston: Springer.
- Shani, G., & Gunawardana, A. (2011). Evaluating Recommendation System. *Recommender Systems Handbook*. Boston: Springer.
- Shen, J., Deng, C., & Gao, X. (2016). Attraction recommenation: Towards personalized tourism via collective intelligence. *Neuro Computing*, 173(3), 789–798.
- Shishehchi, S., Banihashem, S. Zin, N. A. M., & Noah, S. A. M. (2012). Ontological approach in knowledge based recommender system to develop the quality of elearning system. *Australian Journal of Basic Applied Sciences*, 6(2), 115–123.
- UCSC. (2014). *Announcing the 2014 UCSD Study*. San Dieago: CCES.
- Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsre, H., Bosnic, I., & Duval, E. (2012). Context-aware recommender systems for learning: A survey and future challenges. *IEEE Transactions on Learning Technologies*, 5(4), 318–335.
- Vesin, B., Klačnja-Milicević, A., Ivanović, M., & Budimac, Z. (2013). Applying Recommender Systems and Adaptive hypermedia for e-learning personalization. *Computing and Informatics*, 32(3), 629–659.

- Xu, W., Sun, J., Ma, J., & Du, W. (2016). A personalized information recommendation system for R&D project opportunity finding in big data contexts. *Journal of Network and Computer Application*, 59, 362–369.
- Yang, X., Guo, Y., Liu, Y., & Steck, H. (2014). A survey of collaborative filtering based social recommender systems. *Computer Communication*, 41, 1–10.
- Yassine, A., Mohamed, L., & Achhab, M. A. (2021). Intelligent recommender system based on unsupervised machine learning and demographic attributes. *Simulation Modelling Practice and Theory*, 107. 102198. doi: 10.1016/j.simpat.2020.102198