[](https://www.researchgate.net/?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_1&_esc=publicationCoverPdf)

Xem thảo luận, thống kê và hồ sơ tác giả cho công trình này tại: https://www.researchgate.net/publication/333072035

[Hệ thống Gợi ý cho Nền tảng Trò chơi Video Trực tuyến: Trường hợp của STEAM](https://www.researchgate.net/publication/333072035_Recommender_Systems_for_Online_Video_Game_Platforms_the_Case_of_STEAM?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_3&_esc=publicationCoverPdf)

**Bài báo Hội nghị · Tháng 5 năm 2019**

DOI: 10.1145/3308560.3316457

TRÍCH DẪN

66

ĐỌC

8,418

**3 tác giả:**

[](https://www.researchgate.net/profile/German-Cheuque?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_4&_esc=publicationCoverPdf)[Germán Cheuque](https://www.researchgate.net/profile/German-Cheuque?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_5&_esc=publicationCoverPdf)

[Đại học Công giáo Pontifical Chile](https://www.researchgate.net/institution/Pontifical-Catholic-University-of-Chile?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_6&_esc=publicationCoverPdf)

**3 CÔNG TRÌNH 98 TRÍCH DẪN**

[Jose Antonio Guzman Gomez](https://www.researchgate.net/profile/Jose-Antonio-Guzman-Gomez-2?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_5&_esc=publicationCoverPdf) [Đại học Công giáo Pontifical Chile](https://www.researchgate.net/institution/Pontifical-Catholic-University-of-Chile?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_6&_esc=publicationCoverPdf)

[](https://www.researchgate.net/profile/Jose-Antonio-Guzman-Gomez-2?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_4&_esc=publicationCoverPdf)**1 CÔNG TRÌNH 65 TRÍCH DẪN**

[XEM HỒ SƠ](https://www.researchgate.net/profile/German-Cheuque?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_7&_esc=publicationCoverPdf)

[XEM HỒ SƠ](https://www.researchgate.net/profile/Jose-Antonio-Guzman-Gomez-2?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_7&_esc=publicationCoverPdf)

[](https://www.researchgate.net/profile/Denis-Parra?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_4&_esc=publicationCoverPdf)[Denis Parra](https://www.researchgate.net/profile/Denis-Parra?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_5&_esc=publicationCoverPdf)

[Đại học Công giáo Pontifical Chile](https://www.researchgate.net/institution/Pontifical-Catholic-University-of-Chile?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_6&_esc=publicationCoverPdf)

**151 CÔNG TRÌNH 2,635 TRÍCH DẪN**

[XEM HỒ SƠ](https://www.researchgate.net/profile/Denis-Parra?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_7&_esc=publicationCoverPdf)

Tất cả nội dung sau trang này đã được tải lên bởi [Denis Parra](https://www.researchgate.net/profile/Denis-Parra?enrichId=rgreq-904dc760d7ea2dab617619862fe3d724-XXX&enrichSource=Y292ZXJQYWdlOzMzMzA3MjAzNTtBUzo4MDAxNDMwMDM2Mzk4MTBAMTU2Nzc4MDM2NjY2OQ%3D%3D&el=1_x_10&_esc=publicationCoverPdf) vào ngày 06 tháng 9 năm 2019.

Người dùng đã yêu cầu nâng cao tệp đã tải xuống.

**Hệ thống Gợi ý cho Nền tảng Trò chơi Video Trực tuyến: Trường hợp của STEAM**

# Germán Cheuque

IMFD &

Đại học Công giáo Pontifical Santiago, Chile gacheuque@uc.cl

# José Guzmán

Đại học Công giáo Pontifical Santiago, Chile jaguzman6@uc.cl

# Denis Parra

IMFD &

Đại học Công giáo Pontifical Santiago, Chile dparra@ing.puc.cl

## TÓM TẮT

Thế giới trò chơi video đã thay đổi đáng kể trong những năm gần đây. Sự đa dạng hóa của nó đã làm tăng đáng kể số lượng người dùng tham gia vào các cộng đồng trực tuyến trong lĩnh vực giải trí này, và do đó, số lượng và loại trò chơi có sẵn. Bối cảnh quá tải thông tin này là nền tảng cho sự phát triển của các hệ thống gợi ý có thể tận dụng thông tin mà các nền tảng trò chơi video thu thập, do đó theo xu hướng các trò chơi mới ra mắt mỗi năm. Trong công trình này, chúng tôi thử nghiệm tiềm năng của các mô hình gợi ý tiên tiến dựa trên Máy phân tích (FM), mạng nơ-ron sâu (DeepNN) và một mô hình được phát triển từ sự kết hợp của cả hai (DeepFM), được chọn vì khả năng nhận nhiều đầu vào cũng như các loại biến đầu vào khác nhau. Chúng tôi đánh giá kết quả của mình bằng cách đo độ chính xác xếp hạng của gợi ý và sự đa dạng/độc đáo của danh sách gợi ý. Tất cả các thuật toán đều đạt được kết quả tốt hơn so với một cơ sở dựa trên phản hồi ngầm (mô hình Bình phương tối thiểu thay thế). Thuật toán hoạt động tốt nhất là DeepNN, các tương tác bậc cao quan trọng hơn các tương tác bậc thấp cho nhiệm vụ gợi ý này. Chúng tôi cũng phân tích tác động của cảm xúc được trích xuất trực tiếp từ các đánh giá trò chơi và nhận thấy rằng nó không quan trọng cho việc gợi ý như người ta có thể mong đợi. Chúng tôi là những người đầu tiên nghiên cứu các hệ thống gợi ý nêu trên trong bối cảnh các nền tảng trò chơi video trực tuyến, báo cáo các kết quả mới có thể được sử dụng làm cơ sở trong các công trình tương lai.

## TỪ KHÓA

Hệ thống Gợi ý; Máy phân tích; Mạng nơ-ron sâu; Máy phân tích sâu; Độc đáo; Đa dạng

**Định dạng Tham khảo ACM:**

Germán Cheuque, José Guzmán, và Denis Parra. 2019. Hệ thống Gợi ý cho Nền tảng Trò chơi Video Trực tuyến: Trường hợp của STEAM. Trong Tài liệu Hội nghị của Hội nghị World Wide Web 2019 (WWW ’19 Companion), 13–17 tháng 5, 2019, San Francisco, CA, Hoa Kỳ. ACM, New York, NY, Hoa Kỳ, 8 trang. https://doi.org/10.1145/3308560.3316457

## GIỚI THIỆU

Theo Thị trường Trò chơi Di động Châu Âu, vào năm 2016, hơn 2.5 tỷ người đã dành một phần thời gian của họ để chơi trò chơi video.

Bài báo này được xuất bản theo giấy phép Creative Commons Attribution 4.0 Quốc tế (CC-BY 4.0). Các tác giả giữ quyền phân phối công trình trên các trang web cá nhân và công ty của họ với sự ghi nhận thích hợp.

*WWW ’19 Companion, 13–17 tháng 5, 2019, San Francisco, CA, Hoa Kỳ*

© 2019 IW3C2 (Ủy ban Hội nghị World Wide Web Quốc tế), xuất bản theo Giấy phép Creative Commons CC-BY 4.0.

ISBN ACM 978-1-4503-6675-5/19/05.

<https://doi.org/10.1145/3308560.3316457>

số lượng lớn người hâm mộ đã khiến ngành công nghiệp trò chơi video trở thành một trong những ngành có giá trị nhất trên thế giới. Chỉ riêng trong năm 2017, ngành này đã tăng trưởng 10.7% về doanh thu, đạt hơn 116 tỷ đô la. Một trong những lý do lớn nhất đứng sau thành công này là sự đa dạng hóa cao mà ngành này đã trải qua trong những năm gần đây. Chúng ta hiện có thể tìm thấy nhiều nền tảng dành riêng cho trò chơi, một số lượng khổng lồ các thể loại và danh mục trò chơi, thậm chí cả những nền tảng cung cấp sự tương tác xã hội giữa các người chơi. Sự thích ứng này đã tạo ra một số lượng lớn các mặt hàng với các thuộc tính khác nhau có khả năng thu hút người dùng đa dạng. Một ví dụ về sự thích ứng là STEAM.1Nền tảng, một công ty chuyên về phân phối kỹ thuật số các trò chơi video, cho phép người dùng mua, xem, chia sẻ ý kiến, chơi trực tuyến, chơi ngoại tuyến, cạnh tranh và hợp tác trên nền tảng và các trò chơi. Hơn 10 triệu người kết nối với máy chủ STEAM mỗi giờ để chơi trò chơi video. Tuy nhiên, việc có một sự đa dạng sản phẩm và số lượng người dùng lớn như vậy khiến cho việc chọn một trò chơi mới cụ thể mà người dùng thích trở nên khó khăn. Ngoài ra, theo các ghi chép của STEAM vào năm 2014, khoảng 37% trò chơi đã mua chưa bao giờ được người dùng chơi. Bối cảnh này tạo ra nhu cầu về các hệ thống gợi ý, là những hệ thống có khả năng đưa ra các gợi ý cá nhân hóa phù hợp, thông báo cho người dùng về các trò chơi chưa biết cũng như các bản phát hành mới.

Chơi trò chơi điện tử là một hoạt động lặp đi lặp lại, theo nghĩa là nó tương tự hơn với việc nghe nhạc so với việc xem phim. Một trò chơi ưa thích được chơi nhiều lần, nhưng người dùng cũng muốn khám phá các trò chơi mới. Điều này đại diện cho một thách thức kép cho ngành công nghiệp: nhu cầu về các trò chơi điện tử khuyến khích người dùng quay lại, cũng như giúp người dùng tìm ra các trò chơi mới lạ mà sẽ được tiêu thụ nhiều như những trò đã được yêu thích. Trực giác của chúng tôi là số lượng lớn các tính năng có sẵn trên nền tảng STEAM cho phép chúng tôi khám phá các tương tác giữa các tính năng của các mục, cũng như giữa người dùng và sở thích của họ. Những mối quan hệ này khiến chúng ta có lý do để lập luận rằng có thể đối mặt với thách thức thứ hai. Việc phát triển một thuật toán giải quyết các thách thức đã nêu có thể mang lại lợi ích lớn cho ngành công nghiệp, cộng đồng người dùng và thậm chí cho các nhà phát triển trò chơi, bằng cách dự đoán những gì người dùng mong muốn nhất cũng như quảng bá các sản phẩm mới.

Để đạt được những mục tiêu này, chúng tôi thử nghiệm ba thuật toán tiên tiến dựa trên các mô hình gợi ý khác nhau có thể được sử dụng với các loại dữ liệu đầu vào khác nhau:

* Hệ thống Gợi ý (RecSys) dựa trên Lọc Hợp tác (CF) sử dụng thuật toán ALS (Bình phương tối thiểu luân phiên) để đưa ra gợi ý dựa trên phản hồi ngầm [4]. Chúng tôi sử dụng nó như một cơ sở tham chiếu.

1<https://store.steampowered.com/>

* Hệ thống gợi ý dựa trên Máy phân tích (Factorization Machines) [10] tận dụng một lượng lớn dữ liệu, ngữ cảnh và các đặc tính từ người dùng, mặt hàng và giao dịch mua.
* Hệ thống gợi ý dựa trên FM và các kỹ thuật học sâu (DeepNN và DeepFM [3]), nơi chúng tôi giới thiệu phân tích cảm xúc để giải mã thông tin chứa đựng trong các đánh giá.

Chúng tôi cho rằng việc cung cấp cho các mô hình của chúng tôi chỉ dựa trên lịch sử sở thích của người dùng (phản hồi ngầm [4, 7]) là đủ để tạo ra một mô hình gợi ý cơ bản tốt. Hơn nữa, chúng tôi nghĩ rằng việc thêm thông tin nội dung và ngữ cảnh có thể được sử dụng hiệu quả bởi các mô hình FM và DeepNN để cải thiện hiệu suất xếp hạng và tính mới, do các hiệu ứng tương tác. Tuy nhiên, sự kết hợp của cả hai trong một mô hình chung (DeepFM) có thể tạo ra sự cải thiện hoặc có thể chỉ là sự dư thừa so với DeepNN.

Các đóng góp của chúng tôi như sau: (i) Chúng tôi là những người đầu tiên nghiên cứu các hệ thống gợi ý tiên tiến đã đề cập trong bối cảnh trò chơi video, báo cáo các kết quả mới để làm cơ sở cho các công trình trong tương lai, (ii) bên cạnh độ chính xác của xếp hạng, chúng tôi báo cáo và phân tích các kết quả về tính mới mẻ và sự đa dạng của các gợi ý, điều này rất quan trọng trong ngành công nghiệp trò chơi video trực tuyến, và (iii) chúng tôi là những người đầu tiên phân tích tác động của việc sử dụng phân tích cảm xúc từ các đánh giá văn bản đối với chất lượng, tính mới mẻ và sự đa dạng của các gợi ý trong trò chơi video trực tuyến.

Công trình này được tổ chức như sau: Phần 2 trình bày thảo luận về tình trạng hiện tại, các nghiên cứu đã truyền cảm hứng cho công việc của chúng tôi. Trong phần 3, chúng tôi trình bày Tài liệu (tập dữ liệu) và phương pháp (mô hình gợi ý) được sử dụng trong công trình này. Phần 4 cho thấy phương pháp thí nghiệm và các chỉ số đánh giá để đo lường hiệu suất của các hệ thống của chúng tôi, trong khi ở phần 5, chúng tôi trình bày phân tích độ nhạy của các tham số. Phần 6 trình bày các kết quả quan trọng nhất của chúng tôi và cuối cùng, phần 7 thảo luận về kết luận và ý tưởng cho công việc trong tương lai.

## CÔNG TRÌNH LIÊN QUAN

Nhiều nghiên cứu đã đề cập đến vấn đề gợi ý, tuy nhiên chỉ có một vài nghiên cứu xuất hiện trong ngành công nghiệp trò chơi điện tử. Các nền tảng mua sắm trò chơi điện tử hoặc trong các cộng đồng trò chơi điện tử là những nơi thú vị để phát triển và thử nghiệm tiềm năng của các hệ thống gợi ý hiện có. Mặc dù vậy, vẫn có một số nghiên cứu đáng chú ý mà ứng dụng của chúng có thể liên quan đến bối cảnh đã được giới thiệu.

Bertens et al. 2018 [1] ví dụ, cố gắng đo lường hiệu suất của hai hệ thống gợi ý dựa trên ý tưởng dự đoán mặt hàng tiếp theo có khả năng mua nhất. Bài báo nghiên cứu hành vi của các người chơi thẻ Nhật Bản trong trò chơi Age of Ishtaria, xem xét thời gian chơi, tiến độ hàng ngày và lịch sử mua sắm của họ. Họ thử nghiệm hai mô hình: mô hình đầu tiên được gọi là Cây cực kỳ ngẫu nhiên (ERT), là một phiên bản ngẫu nhiên của thuật toán cây quyết định. Phương pháp này có ưu điểm là hiệu quả tính toán cao nhờ khả năng song song hóa tốt, và nó cũng có tính chất ngăn chặn hiện tượng quá khớp. Mô hình thứ hai tương ứng với một thuật toán dựa trên mạng nơ-ron sâu, với đặc điểm.

người có hiệu suất tốt hơn cả về thời gian huấn luyện và khả năng mở rộng.

Quadrana et al. 2018 [9] đề xuất một hệ thống gợi ý dựa trên sự chú ý đến chuỗi sự kiện như một trong những cách tiếp cận tốt nhất để giải quyết nhiệm vụ gợi ý trong bối cảnh đã được giới thiệu. Họ khảo sát nhiều mô hình đã phát triển, xem xét các nhiệm vụ và mục tiêu khác nhau để đạt được. Trong cuộc thảo luận của họ, họ nhấn mạnh tiềm năng của mạng nơ-ron cho nhiệm vụ gợi ý. Mặc dù cuộc thảo luận của họ có liên quan đến bối cảnh ngành công nghiệp trò chơi điện tử, họ nhấn mạnh rằng các ứng dụng của họ cần được thử nghiệm trong bối cảnh chú ý theo phiên. Một cách tiếp cận đáng chú ý của loại ứng dụng này có thể được tìm thấy trong Wan et al. 2018 [13] nơi các tác giả nghiên cứu khái niệm chuỗi hành vi đơn điệu, được hiểu là một chuỗi sự kiện phản ánh sở thích rõ ràng hơn mà người dùng có về một mục. Chuỗi này được xây dựng từ thông tin ngầm và phong phú đến thông tin rõ ràng và hạn chế hơn. Sự đổi mới của công bố nằm ở khả năng giới thiệu một mô hình mới để đối mặt với nhiệm vụ gợi ý và có thể chứng minh tiềm năng của nó trên nhiều tập dữ liệu khác nhau, bao gồm một tập dữ liệu về trò chơi điện tử từ STEAM. Sự cải thiện về hiệu suất của họ dao động từ 1% đến 28% so với phương pháp Phổ biến Nhất, tùy thuộc vào tập dữ liệu.

Do bối cảnh của các mạng lưới và tương tác tồn tại trong các nền tảng trò chơi điện tử, hệ thống gợi ý dựa trên đồ thị có thể rất hữu ích như một phương pháp mô hình hóa. Trong chủ đề này, Shams et al. (2016)

[11] là một trong những nghiên cứu gần đây nhất. Trong ấn phẩm này, họ giới thiệu phương pháp GRank để mô hình hóa chính xác các ưu tiên của người dùng và phân biệt tốt hơn các kết nối liên quan giữa các nút.

Trong công trình này, chúng tôi đề xuất phát triển một hệ thống gợi ý cho trò chơi dựa trên các kỹ thuật tiên tiến, bằng cách so sánh kết quả của lọc cộng tác dựa trên phản hồi ngầm, máy phân tích yếu tố và mạng nơ-ron sâu. Khác với các công trình trước đây, hệ thống của chúng tôi tận dụng các tương tác giữa người dùng và mục tiêu bằng cách kết hợp thông tin ngầm từ người dùng và các đặc điểm từ mục tiêu, xem xét thông tin ngầm như thời gian chơi và thông tin rõ ràng như ý kiến trong các bài đánh giá. Ngoài ra, chúng tôi đánh giá không chỉ dựa trên độ chính xác dự đoán hoặc xếp hạng, mà còn sử dụng các chỉ số được đề xuất bởi Vargas và Castells để đo lường tính mới mẻ và sự đa dạng, vì những chỉ số này rất quan trọng cho các phát hành liên tục và đa dạng trong ngành công nghiệp trò chơi.

## VẬT LIỆU VÀ PHƯƠNG PHÁP

* 1. **Tài liệu: Tập dữ liệu**

Để thực hiện việc triển khai các hệ thống gợi ý đã đề cập (được thảo luận chi tiết trong phần tiếp theo) và đánh giá khả năng dự đoán của chúng, chúng tôi làm việc trên ba tập dữ liệu, được lấy từ bộ sưu tập được chia sẻ bởi J. McAuley.2Bộ dữ liệu đầu tiên bao gồm lịch sử mua hàng của người dùng Australia trên nền tảng STEAM, một cơ sở dữ liệu được phân loại theo người dùng, chỉ ra cho mỗi người danh sách các mặt hàng đã mua kèm theo một bộ siêu dữ liệu nhỏ như thời gian chơi. Bộ dữ liệu thứ hai chứa ý kiến của các người dùng khác nhau trên nền tảng về các mặt hàng có sẵn, ý kiến này phục vụ như một đánh giá cho.

của việc được xây dựng như một mạng lưới hồi tiếp do tính chất thiết yếu

của dữ liệu. Cả hai thuật toán đều cho kết quả rất tương tự, với ERT là

2<http://cseweb.ucsd.edu/~jmcauley/datasets.html#steam_data>

toàn bộ cộng đồng. Tập dữ liệu thứ ba và cuối cùng cung cấp cho chúng ta thông tin chi tiết về các đặc điểm khác nhau của trò chơi điện tử, chẳng hạn như thể loại mà chúng thuộc về hoặc khả năng tương thích với nền tảng (như đã thảo luận trong [5]). Dữ liệu được thu thập từ tháng 10 năm 2010 đến tháng 1 năm 2018.

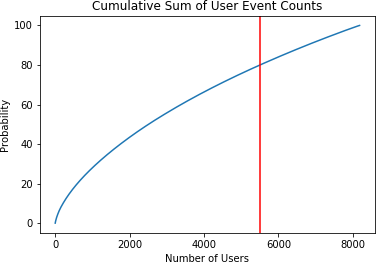
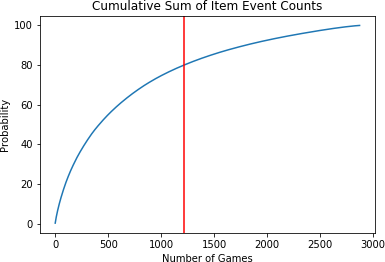
Các tập dữ liệu khác nhau được kết hợp với mục đích tạo ra các bộ dữ liệu liên quan đến từng cặp người dùng và mặt hàng. Bảng 1 trình bày tóm tắt các thống kê mô tả cuối cùng. Mỗi bộ dữ liệu bao gồm 32 đặc trưng, trong đó có nhiều thông tin ngầm phong phú, như thời gian chơi, và thông tin rõ ràng hiếm hoi như đặc trưng gợi ý. Như chúng ta thấy, không có thước đo đánh giá chung nào có thể được sử dụng để đo lường mức độ mà người dùng thích hoặc ưu tiên một mặt hàng. Vì lý do này, phân tích của chúng tôi và mô hình mà chúng tôi xem xét dựa trên các thước đo phản hồi ngầm hoặc được tạo ra từ sự kết hợp của thông tin mà chúng tôi cho là hữu ích trong việc đưa ra gợi ý. Sự phong phú của thông tin mang đến cho chúng tôi cơ hội độc đáo để thử nghiệm các loại dữ liệu khác nhau và sự tương tác rõ ràng cũng như ngầm của chúng, cũng như tính phi tuyến tính của chúng.

Chú trọng đặc biệt được dành cho thời gian chơi, xem xét tầm quan trọng của biến này được báo cáo trong các nghiên cứu trước đây về hệ thống gợi ý [7, 14].

|  |  |
| --- | --- |
| Thống kê mô tả |  |
| Số lượng đăng ký | 2,149,858 |
| Số lượng người dùng khác nhau | 8,183 |
| Số lượng các loại mặt hàng khác nhau | 2,872 |
| Số lượng đánh giá | 9,823 |
| Mật độ tập dữ liệu | 9,14% |
| Số lượng mua sắm trung bình trên mỗi người dùng | 262.72 |
| Số lượng mua trung bình trên mỗi mặt hàng | 748.55 |
| Số giờ trung bình chơi game mỗi người dùng | 161,317.9 |
| Giờ chơi trung bình mỗi mục | 476,727.9 |

### Bảng 2: Tóm tắt thống kê của cơ sở dữ liệu cuối cùng sau khi cắt.

Với việc lọc dữ liệu, chúng tôi đã thu được một tập dữ liệu mới có mật độ gần 10%, giảm kích thước ban đầu xuống còn một nửa, nhưng vẫn đủ lớn để thực hiện các thí nghiệm. Chúng tôi cũng phân tích phân phối tiêu thụ trong tập dữ liệu, nhằm loại bỏ các vấn đề tiềm ẩn do thiên lệch mất cân bằng: quá ít người dùng chiếm quá nhiều giao dịch tổng thể. Hình 1 cho thấy thiên lệch người dùng là nhỏ: khoảng 67% trong số họ giải thích cho 80% số bộ dữ liệu. Tuy nhiên, Hình 2 cho thấy một số lượng nhỏ các mặt hàng (40%) chịu trách nhiệm cho hầu hết các bản ghi (80%). Tình huống này khuyến khích chúng tôi phân tích thêm các hệ thống của mình để đảm bảo rằng chúng đề xuất các loại mặt hàng khác nhau. Để giải quyết mối quan tâm này, chúng tôi sẽ sử dụng các chỉ số mới mẻ và đa dạng. Các chi tiết thêm được thảo luận trong phần 4.

**Hình 1: Tổng hợp tích lũy đóng góp của người dùng vào số lượng hồ sơ. Sáu mươi bảy phần trăm người dùng giải thích 80% dữ liệu.**

|  |  |  |
| --- | --- | --- |
| Tính năng | Loại | Mô tả |
| *mã người dùng* | Chuỗi | Mã định danh duy nhất của người dùng |
| *mã\_hàng\_hóa* | Chuỗi | Mã định danh duy nhất của mục |
| *đếm* | Số nguyên | Số lượng trò chơi được người dùng mua |
| *thời gian chơi* | Số nguyên | Thời gian mà trò chơi đã diễn ra.  thời gian chơi của người dùng, tính bằng giờ |
| *Số lượng bản ghi* | Số nguyên | Số lần một mục đã được.  được khuyến nghị |
| *Metacritic* | Số nguyên | Thẩm định giá trò chơi, đã được đồng ý.  nhà phê bình chuyên ngành. |
| *Thể loại là*  |G| = 13 | [Mảng Bool] | Đúng nếu mục thuộc về một trong những  13 thể loại trò chơi khác nhau. |
| *Danh mục*  |C| = 8 | [Mảng Bool] | Đúng nếu trò chơi thuộc một trong các  8 danh mục trò chơi, ví dụ: nhiều người chơi |
| *Nền tảng*  |P| = 3 | [Mảng Bool] | Đúng nếu trò chơi được hỗ trợ bởi  một trong 3 hệ điều hành được liệt kê. |
| *khuyến nghị* | [Mảng Bool] | Đúng nếu người dùng khuyến nghị  mục này. |
| *đánh giá* | Chuỗi | Văn bản tự do báo cáo ý kiến của người dùng. |

Hình 2: Tổng hợp tích lũy đóng góp của các mục vào số lượng hồ sơ. Bốn mươi phần trăm các mục giải thích 80% dữ liệu.

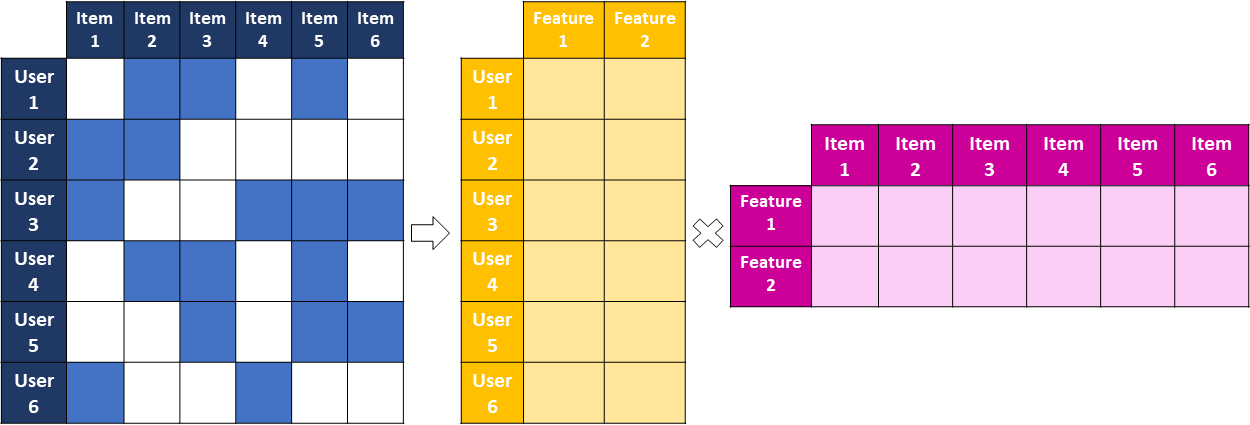
**Bảng 1: Tóm tắt các đặc điểm.**

Khoảng 5, 153, 209 bộ dữ liệu được chứa trong tập dữ liệu của chúng tôi với tổng cộng 70, 912 người dùng khác nhau và 10, 978 mục khác nhau (trò chơi video) để lựa chọn. Những con số này cho chúng ta biết về sự thưa thớt lớn của

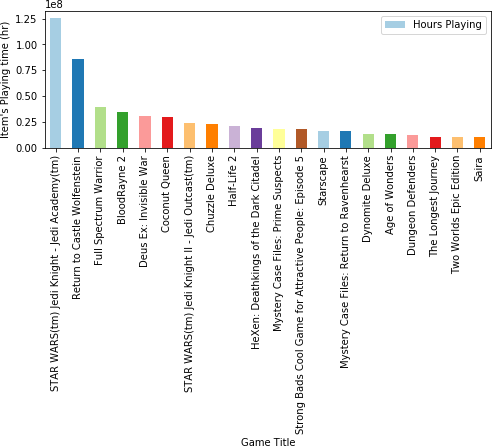
0,66% hồ sơ đã thấy. Sự thưa thớt lớn không lý tưởng cho việc đào tạo một mô hình tương tác giữa người dùng và các mục. Vì lý do này, chúng tôi lọc tập dữ liệu; chúng tôi tin rằng mật độ khoảng 10% có thể hữu ích cho các mục tiêu của chúng tôi. Tập dữ liệu cuối cùng của chúng tôi chỉ xem xét các mục được mua ít nhất 200 lần và người dùng có ít nhất 100 mục đã mua. Một tóm tắt về cơ sở dữ liệu cuối cùng của chúng tôi được trình bày trong Bảng 2.

Dữ liệu phản hồi ngầm. Chúng tôi sẽ làm việc với phản hồi ngầm để học sở thích của người dùng, cụ thể là, sử dụng thời gian chơi như một thước đo proxy của chúng tôi. Sau khi phân tích sơ bộ về toàn bộ phân phối thời gian chơi (xem Hình 3), chúng tôi đề xuất một giới hạn 5 giờ là đủ để phân biệt giữa sở thích hay không. Đây thực sự là một giả định mạnh mẽ, vì lý tưởng sẽ là biết thời gian chơi thông thường của mỗi trò chơi như một thước đo chính xác hơn. Nhưng vì lý do đơn giản, chúng tôi để ý tưởng này cho công việc trong tương lai. Hình 3 và 4 cho thấy phân phối thời gian chơi cho những người dùng hoạt động nhất và các trò chơi được mua nhiều nhất.

Hình 3 cho thấy sự khác biệt nhỏ về thời gian chơi giữa những người dùng hoạt động nhất chơi trò chơi video trên STEAM. Trong khi đó, các trò chơi được mua nhiều nhất



### Hình 3: Thời gian chơi của những người dùng hoạt động nhất.



**Hình 4: Thời gian chơi của các trò chơi được mua nhiều nhất.**

cho thấy sự biến đổi lớn hơn về thời gian chơi, nơi chúng tôi tìm thấy các trò chơi được chơi rộng rãi giữa các người dùng và cũng có những trò chơi không được chơi chút nào. Khi chúng tôi xem thời gian chơi trung bình trong bảng 2, chúng tôi có thể suy ra một đặc điểm khác về phân phối của người dùng và trò chơi. Khoảng 161,317.9 giờ trung bình được chơi mỗi người dùng, trong khi tổng cộng 476,727.85 giờ là thời gian trung bình mà một mục được chơi trong lịch sử ghi chép. Một lần nữa, chúng tôi thấy rằng người dùng có phân phối đồng nhất hơn so với các mục, mà một phần nhỏ các mục đại diện cho hầu hết các tương tác trong tập dữ liệu. Mặc dù những con số này có thể gợi ý cho chúng tôi rằng có ít tiềm năng để tạo ra cả danh sách khuyến nghị chính xác và đa dạng, khi chúng tôi áp dụng một ngưỡng thời gian chơi 5 giờ lên các hồ sơ của chúng tôi, chúng tôi thấy rằng 46% cơ sở dữ liệu vượt quá ngưỡng này và do đó, tương ứng với các mục liên quan cho nhiệm vụ khuyến nghị.

## Phương pháp: Mô hình Hệ thống Khuyến nghị

* + 1. ***Phương pháp Bình phương Nhỏ Xen kẽ (ALS)..***

### Hình 5: Biểu diễn mô hình phân tích ma trận.

Hình 5 cho thấy ý tưởng dưới phương pháp phân tích, tức là, tìm một biểu diễn mới cho người dùng (x *u* ) và các mục (y *i* ) tương tác, để một sở thích cụ thể của một người dùng về một mục được cho bởi tích vô hướng của biểu diễn yếu tố tiềm ẩn của mỗi người. Một đổi mới mà thuật toán này giới thiệu phân tích thước đo phản hồi ngầm thông qua việc chèn hai thước đo mới đại diện cho sở thích và độ tin cậy mà tồn tại đằng sau việc tiếp nhận một mục. Sở thích cho chúng ta biết liệu một mục có được tiêu thụ hay không.

*p ui* = ( )

1ifr *ui* > 0 1

0ifr *ui* = 0

Nơi r *ui* là một thước đo nào đó của phản hồi ngầm. Thước đo thứ hai đại diện cho độ tin cậy của sở thích đó,

c *ui* = 1 + α ∗ r *ui*

Trong đó α là một hệ số tỷ lệ tuyến tính, đặt nhiều trọng số hơn cho các mục liên quan so với những mục chưa từng được chơi. Giá trị α = 40 thường được sử dụng như một kết quả từ bài báo gốc [4]. Dưới những định nghĩa này, việc tìm kiếm yếu tố tiềm ẩn cho người dùng và các mục được thực hiện thông qua việc tối ưu hóa hàm mất mát sau đây,

min *y* ∗ I: *u* ,*i c ui* (p *ui* − x y *i* ) + λ( I: *u* ||x *u* || + I: *i* ||y *i* || ),

*T 2*22

*u*

Tối ưu hóa thông qua bình phương tối thiểu, điều này đã đặt tên cho thuật toán, là rõ ràng. Theo cách này, các cập nhật từng phần của các yếu tố tiềm ẩn được đưa ra bởi,

x*u* = (V*TV + VT*(C*u* − I )V + λI )−1*VTCu p(u)*

y *i* = (U*TU + UT*(C*i* − I )U + λI )−1*UTCi p(i)*

Một số tham số liên quan như số lượng yếu tố tiềm ẩn được chọn thông qua phân tích trong nhiều lần lặp. Trong bài viết này, chúng tôi đã sử dụng triển khai ALS được tìm thấy trong pyreclab.3thư viện gợi ý

#### *Máy phân tích yếu tố (FM).*

Mô hình Máy Phân Tích (FMs) là một loại mô hình cho hệ thống gợi ý, giúp các mô hình yếu tố tiềm ẩn trở nên dễ sử dụng như các mô hình hồi quy hoặc SVM. FMs có khả năng xử lý các loại đầu vào khác nhau, từ biến liên tục đến biến rời rạc, và quan trọng hơn, là các mô hình yếu tố tiềm ẩn. FMs có thể mô hình hóa các tương tác của các bậc khác nhau (bậc-n) giữa các biến này. Trong trường hợp của chúng tôi, n = 2 được xác định vì nó đã được sử dụng rộng rãi theo cách này, mang lại kết quả tốt trong khi vẫn duy trì thời gian huấn luyện hợp lý. Với cấu hình này, đầu ra là một dự đoán được tạo ra từ sự tương tác tuyến tính của các đầu vào (bậc-1) cộng với sự tương tác của các yếu tố tiềm ẩn, được xác định bởi các mối quan hệ giữa các cặp đầu vào (bậc-2), như đã quan sát trong phần 3.2.2.

Dưới sự đa dạng của các mô hình phân tích ma trận, mô hình ALS

nổi bật vì có khả năng làm việc dựa trên phản hồi ngầm.3<https://github.com/gasevi/pyreclab>

y(x) := w0cộngI: *n*

*i=0*

w*i xi* cộngI: *n*

*n*

*j = i + 1*

I:

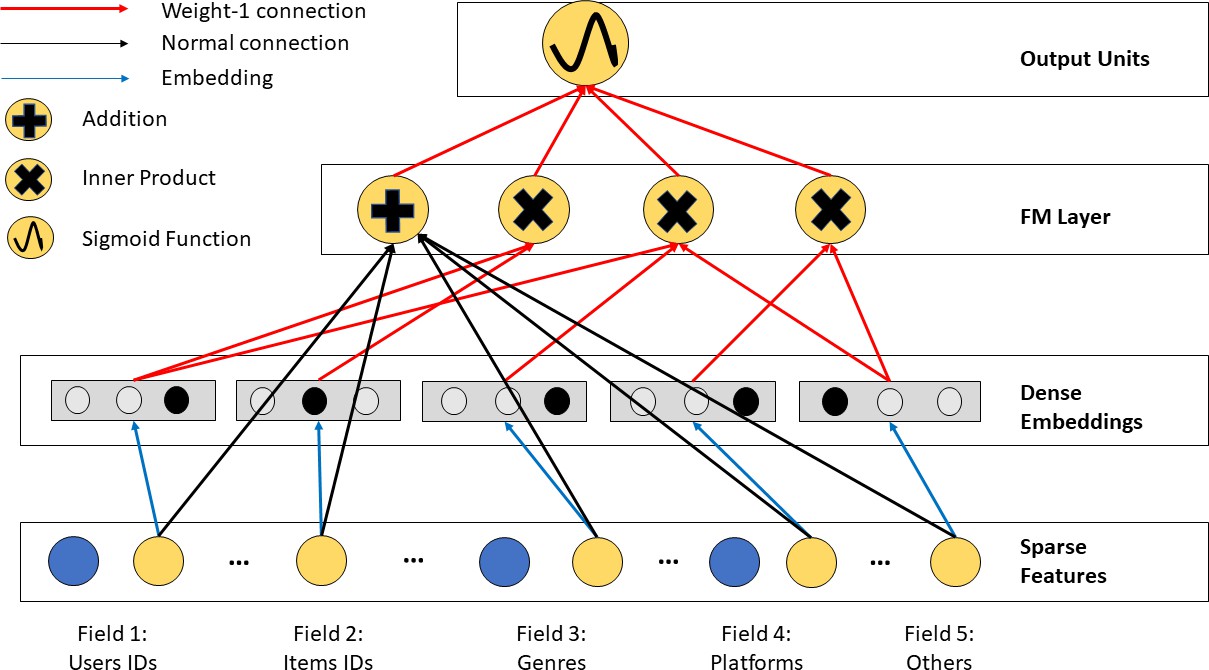
< v*i* , v*j*> x*i xj*

Chúng tôi quyết định phân loại các mục theo loại của chúng (người dùng, mặt hàng, danh mục, nền tảng, v.v.) và sử dụng một embedding cho các mục.

Chúng tôi quyết định chọn mô hình này vì nó là một loại hệ thống mà

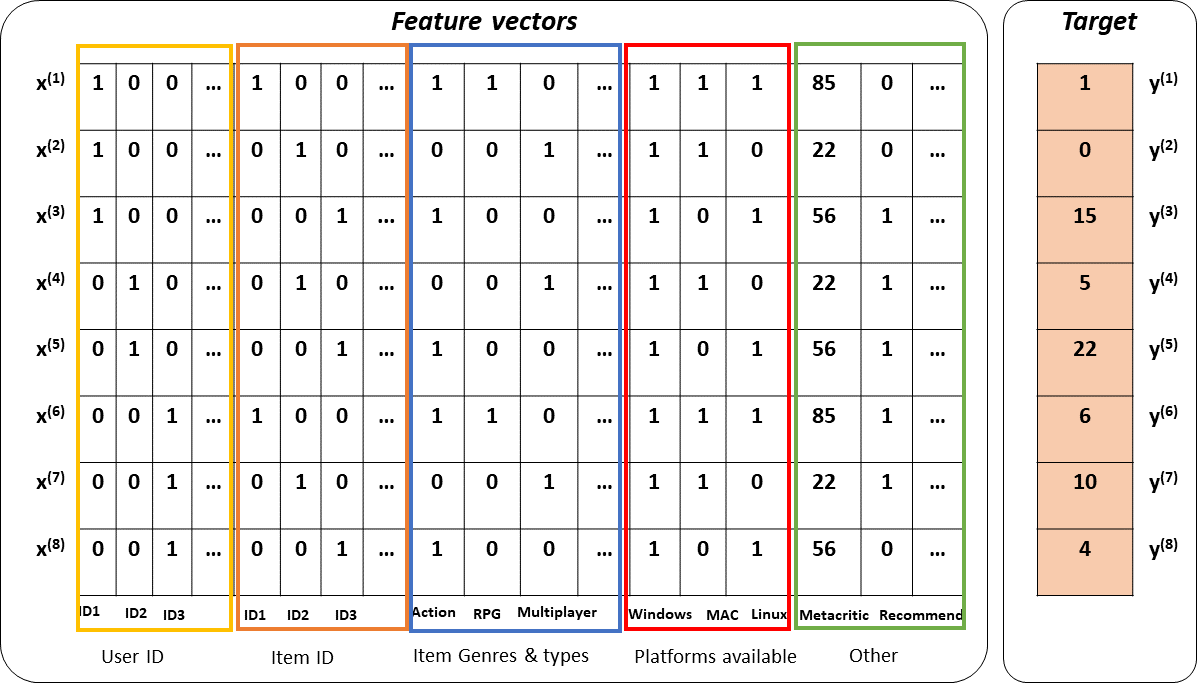
*i=1*

có khả năng mở rộng nhanh chóng, đã thể hiện hiệu suất xuất sắc khi làm việc với các tập dữ liệu thưa thớt, và hỗ trợ dữ liệu đầu vào số của bất kỳ loại nào. Đây đều là những tính năng góp phần vào quá trình gợi ý trong một công ty như trò chơi điện tử, nơi sự đa dạng sản phẩm và tăng trưởng là những yếu tố quan trọng trong quy trình.



### Hình 6: Kiến trúc Máy phân tích yếu tố. Hình ảnh được điều chỉnh từ [3].

Hình 6 (như trong [3]) cho thấy một ví dụ về kiến trúc của mô hình máy phân tích và Hình 7 (như trong [10]) là một ví dụ về các đầu vào và đầu ra mà nó có thể nhận.

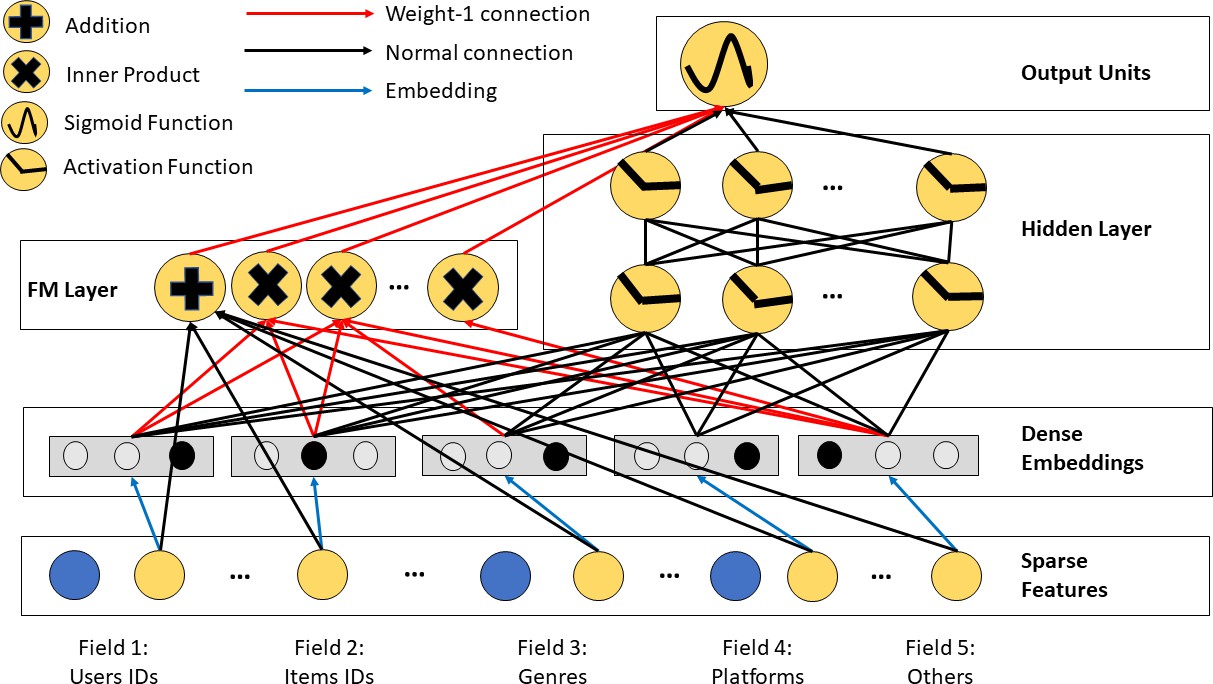


### Hình 7: Ví dụ đầu vào của các đặc trưng nhận được bởi FM. Hình ảnh được điều chỉnh từ [10].

#### *DeepFM.*

DeepFM là một mô hình mới hơn nhằm tận dụng tính linh hoạt của Máy phân tích (Factorization Machines) để mô hình hóa các tương tác bậc thấp giữa các biến đầu vào, với khả năng mô hình hóa các tương tác sâu hơn của Mạng nơ-ron sâu (Deep Neural Networks - DNN). Để thực hiện điều này, một DNN được triển khai song song với FM và kết quả của cả hai được kết hợp tại một nút cuối cùng bằng cách sử dụng hàm sigmoid, làm cho đầu ra trở thành một dự đoán kết quả từ sự tương tác của nhiều bậc đầu vào, đại diện cho xác suất người dùng có thuộc về một lớp cụ thể nào đó được định nghĩa bằng các nhãn huấn luyện hay không. Trong trường hợp của chúng tôi, lớp này.

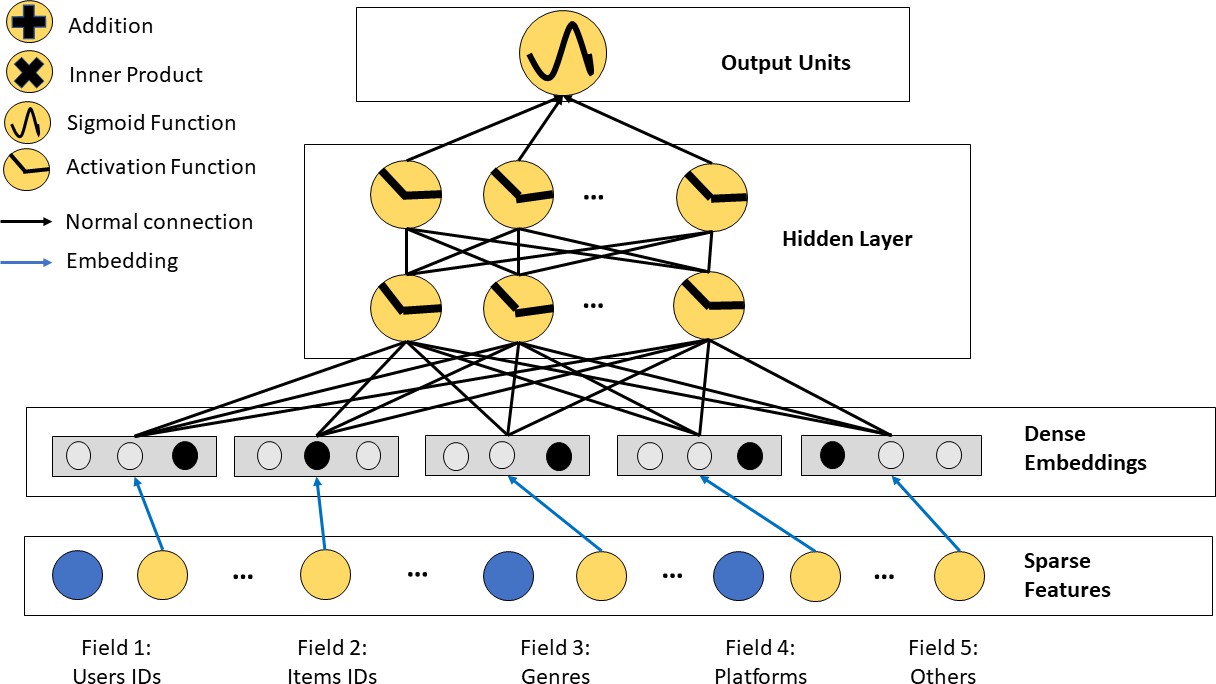
điều này đảm bảo rằng mỗi nhóm được coi trọng như nhau, để hiệu ứng của các biến "thưa" không can thiệp một cách đáng kể vào hiệu ứng của các biến khác. Việc nhúng này bao gồm việc tạo ra nhiều ma trận cho phép chuyển đổi N.*i* các biến của mỗi lớp đầu vào vào một số K cố định được định nghĩa là tham số của mô hình. Ý tưởng này được lấy từ cùng một tài liệu [3] từ nơi mà thông tin mô hình được thu thập, cung cấp cho chúng tôi một phần lớn của việc triển khai nó.4, mà chủ yếu được thực hiện bằng cách sử dụng thư viện tensorflow. Hình 8 (như trong [3]) cho thấy một biểu diễn của tất cả các thành phần của mô hình và cách chúng được kết nối.



### Hình 8: Ví dụ về Kiến trúc DeepFM. Hình ảnh được điều chỉnh từ [3].

#### *DeepNN.*

Thành phần mới của DeepFM là mạng nơ-ron sâu (còn được chúng tôi gọi là DeepNN) phân tích song song các tương tác giữa người dùng và các mục. Thành phần sâu là một mạng nơ-ron truyền thẳng và được sử dụng để học các tương tác bậc cao. Một tóm tắt về kiến trúc này được trình bày trong Hình 9.



### Hình 9: Ví dụ về thành phần DeepNN của Kiến trúc DeepFM. Hình ảnh được điều chỉnh từ [3].

đề cập đến việc người dùng có thích trò chơi hay không.

4<https://github.com/ChenglongChen/tensorflow-DeepFM>

Thành phần đầu tiên của nó là một lớp nhúng (embedding layer) giúp nén các trường đầu vào thành một vector thực có kích thước thấp và dày đặc. Lưu ý rằng các trường này có thể có độ dài khác nhau trong khi lớp nhúng luôn trả về một mã hóa có độ dài cố định là K đơn vị. Chúng tôi ký hiệu đầu ra của lớp nhúng là:

a(0) = [e1, e2, ..., e*m*],

nơi e*i* là trường i-th được nhúng từ tổng cộng m trường. Những

Các đầu vào được đưa vào các lớp ẩn, nơi các đặc trưng tương tác. Quá trình truyền tiến là:

a(l +1)= *σ*(*W*(l )*a*(l )cộng*b*(l )),

nơi l là độ sâu của lớp và σ là một hàm kích hoạt. Các biến a(*l*),W(*l*), b(*l*)là đầu ra, trọng số mô hình và độ chệch của lớp thứ l. Sau đó, một vector dày đặc cuối cùng được tạo ra và các kết quả này

được truyền qua một hàm sigmoid trả về đầu ra cuối cùng

Chúng tôi chỉ sử dụng ba bộ dữ liệu cho DeepFM. Bộ dữ liệu kiểm tra cũng được lấy ngẫu nhiên, tuy nhiên, do việc cắt giảm trước đó trên tập dữ liệu, chúng tôi có các mặt hàng được mua ít nhất 100 lần, vì vậy chúng tôi xác minh rằng ít nhất mười trong số những lần mua này vẫn nằm trong bộ dữ liệu kiểm tra để chúng tôi cố gắng dự đoán sở thích của người dùng về chúng và xác minh những gì mà các mô hình của chúng tôi gợi ý.

Để đo lường hiệu suất của các thuật toán, chúng tôi sử dụng một số chỉ số. Thời gian huấn luyện được coi là một thước đo liên quan để cho chúng tôi có ý tưởng về khả năng mở rộng của các mô hình. Ngoài ra, chúng tôi muốn đo lường Độ chính xác trung bình (Mean Average Precision - MAP) tại vị trí hàng đầu thứ 10 trong danh sách. Do đó, Độ chính xác trung bình được định nghĩa như sau,

I:

AP= *kP*@*k*∗ *mối quan hệ*(*k*)#các\_mục\_liên\_quan

và chỉ số độ chính xác trung bình (Mean Average Precision) được cho bởi,

*u=1*

cho nhiệm vụ gợi ý.

BẢN ĐỒ =I: *n*

AP (u)

#người\_dùng

## Phân tích cảm xúc

Một trong những đặc điểm thú vị nhất trong tập dữ liệu của chúng tôi tương ứng với các đánh giá. Đánh giá là một đoạn văn bản chứa đựng ý kiến của một người.

Tương tự, chúng ta muốn tính toán chỉ số nDCG tại vị trí thứ 10 của danh sách. Chỉ số này cố gắng chuẩn hóa lợi ích và giá trị của một danh sách gợi ý thông qua việc xem xét độ liên quan của các mục và vị trí xếp hạng của chúng trong danh sách (giới thiệu một mức giảm theo logarit). Định nghĩa của tham số này là tỷ lệ giữa Tổng lợi ích tích lũy đã giảm tại vị trí k, được cho bởi:

Người dùng có và đã cung cấp. Tuy nhiên, một đoạn văn bản không tự nó hữu ích cho.

DCG*k*= I: *k*

mối quan hệ (i)

nuôi dưỡng các mô hình của chúng tôi. Để giải quyết vấn đề này, chúng tôi đề xuất các kỹ thuật phân tích cảm xúc để chuyển đổi ý kiến này thành một thước đo số có thể cung cấp cho chúng tôi một đại diện về mức độ tích cực hoặc tiêu cực của ý tưởng được bày tỏ. Việc sử dụng nó có thể giúp tăng cường độ tin cậy thu được từ tính năng khuyến nghị rõ ràng, đồng thời cũng có khả năng suy diễn khi các giá trị đó không khớp. Chúng tôi thực hiện thí nghiệm với ba công cụ khác nhau.

Đầu tiên, chúng tôi sử dụng thuật toán mã nguồn mở Tweetment. Thuật toán này tồn tại dưới dạng thư viện Python và có khả năng phân loại cảm xúc nhị phân thành các nhãn tích cực hoặc tiêu cực, được biểu diễn bằng các giá trị 1 hoặc 0, tương ứng. Thuật toán này được huấn luyện trên một tweet.

cơ sở dữ liệu, đạt được điểm F1 là 69,02% trong phân loại nhị phân

*i* = 1*log2(1+i)*

và giá trị lý tưởng của nó tại vị trí k, còn được gọi là Ideal DCG. Do đó, thước đo nDCG được đưa ra bởi,

NDCG*k*= *DCGk*

*Tôi DCGk*

Cuối cùng, chúng tôi quan tâm đến việc đo lường mức độ đa dạng và mới mẻ của các gợi ý được đưa ra bởi hệ thống của chúng tôi. Như đã thảo luận, một số lượng lớn trò chơi mới được phát hành mỗi năm, làm cho các chỉ số này trở nên rất quan trọng. Để đo lường các yếu tố này, chúng tôi sử dụng các mối quan hệ dựa trên độ tương đồng giữa các mục trong danh sách gợi ý. Chi tiết được trình bày trong [12].

Độ mới (R|u) =Tôi:I: *n*,*j*thuộc về*u đĩa(n)p(tiến trình |in*, u)p(rel |j*n*, u)d(i, j)

Độ đa dạng (R|u) = 2

*k*<*ndisc(n)disc(k)p(rel |in*, u)d(i, k)

Một thuật toán khác được tham khảo là SentiWordNet, được phát triển như một ứng dụng khai thác ý kiến có khả năng phân loại văn bản trong một khoảng liên tục [-1, 1], cho thấy mức độ tích cực hoặc tiêu cực của ý tưởng được diễn đạt. Thuật toán này hoạt động dựa trên một từ điển đã được huấn luyện trước, tổng hợp các mối quan hệ n-gram, cung cấp một điểm số về tính tích cực, tiêu cực và khách quan, mà không bị giới hạn. Cuối cùng, chúng tôi đã thử một mô hình dựa trên mạng nơ-ron và nhúng word2vec. Thuật toán này cố gắng chú ý đến văn bản liên quan đến trò chơi cụ thể. Tuy nhiên, chúng tôi đã gặp phải hai nhược điểm trong quá trình phát triển: chúng tôi chỉ có thể thực hiện phân loại nhị phân và việc triển khai đạt được điểm F1 thấp hơn so với Tweetment. Cuối cùng, chúng tôi quyết định làm việc với SentiWordNet vì lợi thế của việc đo lường liên tục cho việc đánh giá cảm xúc. Điều này có thể phù hợp hơn cho các ngữ cảnh bình luận khác nhau (châm biếm, v.v.).

## PHƯƠNG PHÁP THÍ NGHIỆM

Các thí nghiệm được trình bày trong các phần sau bao gồm việc đào tạo và đánh giá các mô hình gợi ý. Để đảm bảo tính ổn định, chúng tôi tiến hành tạo các tập dữ liệu ngẫu nhiên (k-folding), áp dụng các kỹ thuật kiểm tra chéo để trung bình hóa các điểm số này. Để đào tạo thuật toán ALS, chúng tôi sử dụng năm tập.

Các định nghĩa giới thiệu nhiều yếu tố chiết khấu, được sử dụng như các yếu tố logarithmic. Thuật ngữ d(i,k) đề cập đến khoảng cách giữa các cặp mục, được tính toán dưới dạng khoảng cách cosine. Cuối cùng, các thuật ngữ p(rel |i*n*, u) đề cập đến xác suất mà người dùng u ưa thích mục i.*n*,

một yếu tố có thể được hiểu như một độ liên quan. Thước đo này là

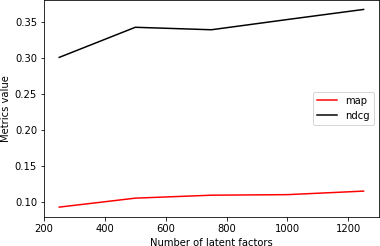
được thu thập trực tiếp từ đầu ra của hệ thống gợi ý.

## Phân tích độ nhạy tham số

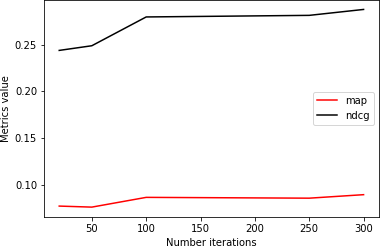
Để đạt được kết quả tốt nhất, chúng tôi tiến hành phân tích các tham số có thể điều chỉnh khác nhau. Thuật toán ALS có ba tham số mà chúng tôi có thể thay đổi: Số lượng yếu tố tiềm ẩn được sử dụng để đại diện cho phân tích, số lượng vòng lặp huấn luyện còn được gọi là số lần lặp; cuối cùng là yếu tố λ, được sử dụng để điều chỉnh trong các phương trình 3 và 4. Phân tích độ nhạy được trình bày trong các hình tiếp theo.

Chúng tôi cuối cùng đã làm việc với sự kết hợp của các tham số tốt nhất được tìm thấy. Chúng tôi đặt hệ số điều chỉnh là 0.01, số lượng 500 yếu tố tiềm ẩn ưu tiên thời gian huấn luyện so với những cải tiến nhỏ hơn 800. Ngoài ra, 300 vòng lặp đã được sử dụng cho quá trình huấn luyện.

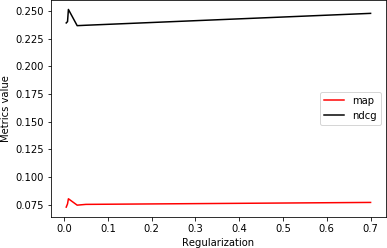
Các mô hình học sâu có hai tham số có thể điều chỉnh. Tham số đầu tiên ảnh hưởng trực tiếp đến kiến trúc mạng, đề cập đến số lượng nơ-ron trong mỗi lớp. Tham số thứ hai là kích thước lô, mà



### Hình 10: Huấn luyện với số lượng yếu tố tiềm ẩn khác nhau.



### Hình 11: Đào tạo với số lượng vòng lặp khác nhau.



### Hình 12: Huấn luyện với các tham số điều chỉnh khác nhau.

Có thể giải thích do số lượng đặc trưng trên mỗi tuple (32), vì vậy khi chúng ta sử dụng một số lượng nơ-ron nhỏ hơn, chúng ta đang ép buộc một nhúng có lợi từ các tương tác. Cuối cùng, chúng tôi kết hợp những cấu hình đó để tạo ra một mô hình có hiệu suất tốt nhất để sử dụng trong tập kiểm tra.

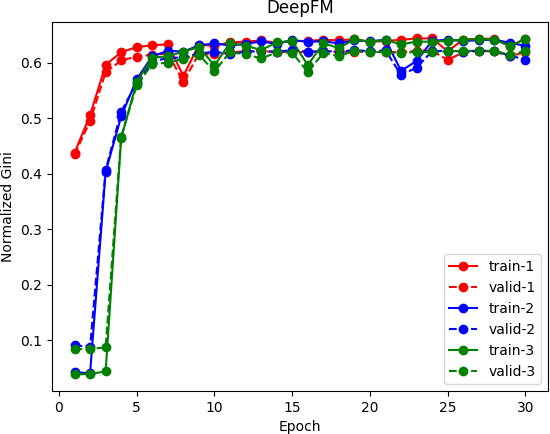
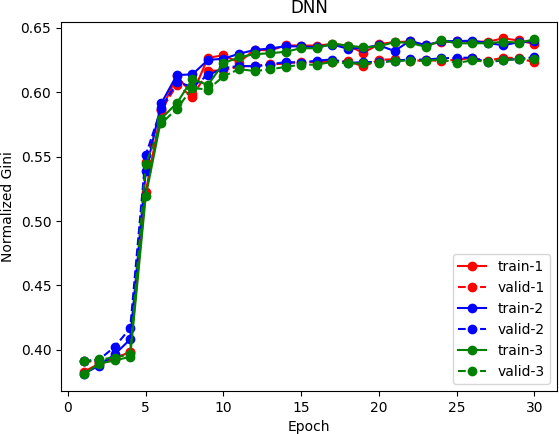
## KẾ QUẢ

Bảng 4 trình bày kết quả đánh giá tất cả các mô hình hệ thống gợi ý đã được thảo luận trước đó.

bao gồm số lượng bộ dữ liệu được phân tích cùng nhau. Yếu tố cuối cùng là rất quan trọng do trách nhiệm của nó đối với việc quá khớp và khả năng thay đổi tỷ lệ học. Quá trình đào tạo của chúng tôi tạo ra các đồ thị như những đồ thị được trình bày trong hình 13 và 14. Điểm số được trình bày được đo lường theo chuẩn Gini, một chỉ số tốt cho phân loại nhị phân tương tự như ROC-AUC nhưng với phổ độ phân giải lớn hơn.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Cấu hình-  tỷ lệ | Thời gian  (seg) | BẢN ĐỒ  @10 | NDCG  @10 | Đặc tính mới | Người lặn  thành phố |
| *ALS* | 1421,8 | 0,107 | 0,332 | - | - |
| *FM* | 1450,86 | 0,893 | 0,944 | 0,176 | 0,45 |
| *Mạng nơ-ron sâu* | 1889,58 | 0,897 | 0,947 | 0,186 | 0,49 |
| *DeepFM* | 2027,52 | 0,891 | 0,943 | 0,167 | 0,43 |
| *FM(WS)* | **1416,04** | 0,894 | 0,945 | 0,18 | 0,46 |
| *DeepNN(WS)* | 1866,35 | **0,897** | **0,948** | **0,197** | **0,54** |
| *DeepFM(WS)* | 1984,17 | 0,892 | 0,944 | 0,19 | 0,53 |

### Bảng 4: Kết quả kiểm tra cho từng mô hình được trình bày, bao gồm kết quả có và không (WS) phân tích cảm xúc được xem xét.

**Hình 13: Đào tạo Mạng Nơ-ron Sâu 14: Đường cong huấn luyện DeepFM.**

Trước hết, chúng ta thấy rằng mô hình cơ sở ALS là mô hình nhanh nhất để huấn luyện, nhưng lại có hiệu suất kém nhất trong tất cả các chỉ số đã thảo luận. Hiệu suất của FM, DNN và DeepFM gần gấp tám lần tốt hơn về mặt MAP@10, cho thấy rằng dữ liệu nội dung và ngữ cảnh rất quan trọng cho vấn đề này, và

Sử dụng các đường cong đó và các chỉ số đã được xác định, chúng tôi đã thực hiện một phân tích độ nhạy như được trình bày trong bảng 3.

### Bảng 3: Phân tích độ nhạy cho mô hình DeepFM. Chúng tôi thay đổi số lượng nơ-ron mỗi lớp (H*ixj*và kích thước lô (B*n*Chúng tôi sử dụng làm cơ sở một cấu hình mặc định với kích thước lô là 1024 và 32 nơ-ron mỗi lớp.

Chúng ta có thể thấy rằng khi kích thước lô (batch size) giảm từ 1024 xuống 512, chúng ta nhận được kết quả tốt hơn trên tất cả các chỉ số. Tuy nhiên, nếu tiếp tục giảm giá trị này xuống 256, kết quả sẽ hơi kém hơn ngoại trừ chỉ số Độ mới (Novelty) có tăng nhẹ. Điều này có thể được giải thích bởi vì khi giảm kích thước lô, chúng ta tạo ra hiện tượng quá khớp (overfitting) trên tập huấn luyện, và các mối quan hệ được học nhanh chóng với một số lượng bản ghi ít. Ngoài ra, khi giảm số lượng nơ-ron (neurons) mỗi lớp từ 32 xuống 8, chúng ta cũng nhận được kết quả tốt hơn trên tất cả các chỉ số. Hiệu ứng này có thể.

Các đồng phạm ngầm không đủ để mang lại kết quả tốt. Một kết quả quan trọng khác là trái với các giả định của chúng tôi, mô hình Deepnn đạt được kết quả tốt hơn trên tất cả các số liệu vượt trội so với FM và thậm chí DEEPFM bao gồm kiến ​​trúc này trên đó. Điều này có thể chỉ ra rằng các tương tác bậc cao được đóng góp từ mô hình DeepNN cung cấp nhiều thông tin hơn các tương tác thấp hơn từ FM, do đó các tương tác này đang tạo ra tiếng ồn cho kết quả cuối cùng của DeepFM. Mặc dù sự khác biệt về mặt MAP@10 và NDCG@10 là nhỏ, nhưng chúng luôn tốt hơn cho Deepnn. Hơn nữa, sự mới lạ và đa dạng cũng tốt hơn đối với Deepnn, cho thấy rằng ngay cả với kết quả xếp hạng tương tự, các khuyến nghị về khả năng đó là mới lạ và đa dạng hơn, đó là những khía cạnh rất quan trọng trong ngành công nghiệp trò chơi.

Một kết quả thú vị khác là khi chúng ta không xem xét các giá trị phân tích cảm xúc được tạo ra từ các đánh giá, tất cả các mô hình đều cho thấy hiệu suất tốt hơn so với chính chúng khi xem xét tính năng này. Một nguyên nhân hợp lý cho điều này là tỷ lệ nhỏ của các đánh giá so với tổng số hồ sơ tương tác người dùng-trò chơi. Các chỉ số Đổi mới và Đa dạng cũng được hưởng lợi từ việc không bao gồm cảm xúc, đạt tới 0,54 về mặt đa dạng cho DeepNN mà không có us-.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Cấu hình-  tỷ lệ | Thời gian  giây (sec) | BẢN ĐỒ  @10 | NDCG  @10 | Đặc tính mới | Người lặn  thành phố |
| *B*1024*HL*32x 32 | 2027,52 | 0,8911 | 0,9437 | 0,1667 | 0,4245 |
| *B*512 | 2186,77 | 0,8940 | 0,9456 | 0,1845 | 0,4967 |
| *B*256 | 2692,96 | 0,8919 | 0,9440 | 0,1847 | 0,4996 |
| *HL*16x 16 | 1815,70 | 0,8939 | 0,9454 | 0,1815 | 0,4775 |
| *HL*8x 8 | **1708,57** | 0,8941 | 0,9455 | 0,1837 | 0,4869 |
| *B*512*HL*8x 8 | 2180,6 | **0,8945** | **0,9458** | **0,1857** | **0,5149** |

cảm xúc. Như chúng tôi đã đề cập, trong đề xuất trực tuyến

Vấn đề trò chơi điện tử, luôn tốt hơn khi đạt được những cải tiến về sự đa dạng và tính mới mẻ bên cạnh các dự đoán chính xác được đánh giá bằng MAP và NDCG.

## KẾ LUẬN

Trong công trình này, chúng tôi đã thử nghiệm với các tùy chọn khác nhau của hệ thống gợi ý trong bối cảnh gợi ý trò chơi video, sử dụng một tập dữ liệu từ công ty STEAM làm nền tảng thử nghiệm. Chúng tôi đã kết hợp ba cơ sở dữ liệu chứa thông tin về các giao dịch mua của người dùng, số giờ dành cho việc chơi từng mục, sự tương tác xã hội (đánh giá) và các đặc điểm của trò chơi video. Mô hình ALS được chọn làm mô hình cơ sở để so sánh, và các thử nghiệm đã được thực hiện với các mô hình Máy phân tích (FM), Mạng nơ-ron sâu (DeepNN) và DeepFM. Mô hình cuối cùng sử dụng một mạng nơ-ron sâu (DeepNN) hoạt động song song với một lớp được cấu thành từ FM, nhằm giới thiệu các tương tác bậc cao giữa các đầu vào, với mục tiêu cải thiện hiệu suất dự đoán của chúng tôi về các yếu tố mới mẻ, đa dạng và độ chính xác.

Tất cả các mô hình được nghiên cứu đều vượt trội hơn mô hình cơ sở ALS. DeepNN nổi bật hơn so với các mô hình khác. Mặc dù là một mô hình đơn giản hơn so với DeepFM, nhưng nó đã khai thác mối quan hệ giữa người dùng và sản phẩm tốt hơn, và mặc dù mất nhiều thời gian hơn so với FM và DeepFM để đạt được kết quả cạnh tranh, nó đạt được kết quả nhất quán trên các tập dữ liệu khác nhau. Ngoài ra, nó đạt được kết quả tốt hơn với các chỉ số đánh giá, điều này ngụ ý rằng các tương tác bậc cao mà mô hình DeepNN cung cấp cung cấp nhiều thông tin hơn so với các tương tác bậc thấp mà mô hình FM cung cấp. Như chúng tôi đã lập luận, sự tương tác phi tuyến từ các đặc trưng của sản phẩm được khai thác bởi các phương pháp dựa trên học sâu, nhưng thú vị là sự khác biệt này được nhận thấy rõ hơn trong việc cải thiện tính mới mẻ và đa dạng của các danh sách gợi ý.

Để thu thập thông tin bổ sung, chúng tôi đã sử dụng các đánh giá của người dùng với các phương pháp phân tích cảm xúc khác nhau, tìm ra một thước đo liên tục có thể trừu tượng hóa ý kiến của người dùng về trò chơi điện tử. Chúng tôi kết luận rằng tỷ lệ đánh giá có sẵn không đủ lớn để góp phần cải thiện kết quả, mà thay vào đó lại tạo ra hiệu ứng nhiễu. Các chỉ số giảm giá trị của chúng trong cả MAP@10 và NDCG@10, thậm chí trong các thước đo về Tính mới và Độ đa dạng.

Trong công việc tương lai, chúng tôi dự định thực hiện các bài kiểm tra phân tích tham số trên mô hình DNN đã cho kết quả tốt nhất để xem liệu chúng có thể được cải thiện hay không. Chúng tôi cũng có thể tiến hành một nghiên cứu người dùng để xác thực các kết quả ngoại tuyến trên một hệ thống thực, nơi mà các biến khác có thể có tác động mạnh đến kết quả cuối cùng. Cuối cùng, chúng tôi gần đây đã phát hiện rằng cơ sở dữ liệu STEAM đã được tác giả cập nhật. Trong phiên bản mới nhất, có một tập dữ liệu mới gồm 7.793.069 đánh giá, vì vậy chúng tôi mong đợi trong công việc tương lai sẽ tận dụng chúng cho một phân tích văn bản lớn hơn. Điều này sẽ giúp chúng tôi kiểm tra xem phân tích cảm xúc và các đặc điểm văn bản khác có thực sự là công cụ hữu ích cho việc gợi ý trò chơi video hay không.

## LỜI CẢM ƠN

Công trình này đã được tài trợ một phần bởi Viện Thiên niên kỷ về Nghiên cứu Cơ sở Dữ liệu (IMFD). Chúng tôi cũng đã nhận được sự hỗ trợ một phần từ NIC Chile và từ Đại học Công giáo Pontificia de Chile.

## TÀI LIỆU THAM KHẢO

1. Paul Bertens, Anna Guitart, Pei Pei Chen, và África Periáñez. 2018. Hệ thống gợi ý mục dựa trên học máy cho trò chơi điện tử. arXiv preprint arXiv:1806.04900 (2018).
2. Andrea Esuli và Fabrizio Sebastiani. 2007. SentiWordNet: một tài nguyên từ vựng có độ phủ cao cho khai thác ý kiến. Tạp chí Đánh giá 17 (2007), 1–26.
3. Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, và Xiuqiang He. 2017. DeepFM: một mạng nơ-ron dựa trên máy phân tích cho dự đoán CTR. Trong Kỷ yếu của Hội nghị Quốc tế lần thứ 26 về Trí tuệ Nhân tạo. AAAI Press, 1725–1731.
4. Yifan Hu, Yehuda Koren và Chris Volinsky. 2008. Lọc cộng tác cho các tập dữ liệu phản hồi ngầm. Trong Khai thác dữ liệu, 2008. Hội nghị Quốc tế lần thứ tám về Khai thác Dữ liệu IEEE. Ieee, 263–272.
5. Wang-Cheng Kang và Julian McAuley. 2018. Đề xuất tuần tự tự chú ý. Trong Hội nghị Quốc tế về Khai thác Dữ liệu IEEE 2018 (ICDM). IEEE, 197–206.
6. Saif M Mohammad, Svetlana Kiritchenko, và Xiaodan Zhu. 2013. NRC-Canada: Xây dựng công nghệ tiên tiến nhất trong phân tích cảm xúc của tweet. Trong Kỷ yếu của Hội nghị chung lần thứ hai về Từ vựng và Tính toán. 321–327.
7. Denis Parra và Xavier Amatriain. 2011. Đi bộ theo lời nói: phân tích mối quan hệ giữa phản hồi ngầm và rõ ràng để thu thập sở thích. Trong Kỷ yếu của hội nghị quốc tế lần thứ 19 về Mô hình người dùng, thích ứng và cá nhân hóa. Springer-Verlag, 255–268.
8. Denis Parra và Shaghayegh Sahebi. 2013. Hệ thống gợi ý: Nguồn kiến thức và các chỉ số đánh giá. Trong Kỹ thuật tiên tiến trong trí tuệ web-2. Springer, 149–175.
9. Massimo Quadrana, Paolo Cremonesi, và Dietmar Jannach. 2018. Hệ thống gợi ý nhận thức theo chuỗi. ACM Comput. Surv. 51, 4, Bài viết 66 (Tháng 7 năm 2018),

36 trang. https://doi.org/10.1145/3190616

1. Steffen Rendle. 2012. Máy phân tích với libfm. Tạp chí Giao dịch về Hệ thống Thông minh và Công nghệ (TIST) 3, 3 (2012), 57.
2. Bita Shams và Saman Haratizadeh. 2017. Xếp hạng hợp tác dựa trên đồ thị.

*Hệ thống Chuyên gia với Ứng dụng 67 (2017), 59–70.*

1. Saúl Vargas và Pablo Castells. 2011. Xếp hạng và độ liên quan trong các chỉ số mới mẻ và đa dạng cho hệ thống gợi ý. Trong Kỷ yếu của hội nghị ACM lần thứ năm về hệ thống gợi ý. ACM, 109–116.
2. Mengting Wan và Julian McAuley. 2018. Gợi ý mục trên chuỗi hành vi đơn điệu. Trong Kỷ yếu của Hội nghị ACM lần thứ 12 về Hệ thống Gợi ý. ACM, 86–94.
3. Xing Yi, Liangjie Hong, Erheng Zhong, Nanthan Nan Liu, và Suju Rajan. 2014. Vượt qua nhấp chuột: thời gian lưu lại cho cá nhân hóa. Trong Kỷ yếu của Hội nghị ACM lần thứ 8 về Hệ thống Gợi ý. ACM, 113–120.

[Xem thống kê xuất bản.](https://www.researchgate.net/publication/333072035)