HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG VIỆN KINH TẾ BƯU ĐIỆN



BÀI THI TIỀU LUẬN Phương pháp luận nghiên cứu khoa học

Nhóm môn học: 21

Giảng viên: Lê Vũ Điệp

Sinh viên: Trần Vân Anh

Mã sinh viên: B20DCCN075

Ngày sinh: 11/01/2002

Mã lớp: D20CNPM02

Số điện thoại: 0382386867

MỤC LỤC

Câu 1:	3
Câu 2:	
Câu 3:	
Câu 4:	10
Tài liêu tham khảo	13

<u>Câu 1:</u> Những thành phần chính mà một nghiên cứu khoa học cần có? Ứng dụng kiến thức này để phân tích các thành phần trong một công trình nghiên cứu khoa học trong ngành học của bạn?

1.1. Những thành phần chính mà một nghiên cứu khoa học cần có

- Phần Mở đầu
 - Tính cấp thiết của đề tài
 - Mục tiêu nghiên cứu
 - Câu hỏi và giả thuyết nghiên cứu
 - Đối tượng và phạm vi nghiên cứu
 - Phương pháp nghiên cứu
 - Lịch sử nghiên cứu
 - Tính mới của đề tài
 - Ý nghĩa lý luận và thực tiễn
- Phần nội dung
 - Chương 1: Cơ sở lý luận nghiên cứu của đề tài
 - Chương 2: Cơ sở thực tiễn nghiên cứu của đề tài
 - Chương 3: Đánh giá và kết luận
- Phần Kết luận: Tài liệu tham khảo và Phụ lục

1.2. Phân tích thành phần của 1 công trình nghiên cứu khoa học ngành CNTT

Bài báo khoa học: Hệ gợi ý mua sắm dựa theo phiên làm việc với mô hình mạng học sâu đồ thị.^[1]

- Phần Mở đầu
 - Tính cấp thiết của đề tài:

Các hệ thống gợi ý ngày nay càng được chú trọng. Hệ thống gợi ý hiệu quả sẽ là vấn đề sống còn đối với nhà cung cấp dịch vụ hoặc bán hàng. Khi một khách hàng vào một trang thương mại điện tử thì có hai xu hướng: hoặc họ đã định hướng được sản phẩm mà họ sẽ mua, hoặc là họ được định hướng được sản phẩm mà họ nên mua. Đối với kịch bản thứ hai, người dùng sẽ gặp khó khăn hơn nhiều vì họ sẽ phải chọn sản phẩm phù hợp nhất với nhu cầu của họ. Vấn đề đặt ra là làm sao họ có thể làm được điều đó trong vô số sản phẩm giống nhau mà họ đang tìm kiếm.

- Mục tiêu nghiên cứu: Nghiên cứu và đề xuất mô hình dự báo hành vi lựa chọn sản phẩm của khách hàng mô hình mạng học sâu.
- Câu hỏi và giả thuyết nghiên cứu

- Câu hỏi: "Với số lượng sản phẩm đã lựa chọn trong phiên tương tác hiện tại thì khả năng người dùng sẽ chọn sản phẩm nào tiếp theo"
- Giả thuyết: Khách hàng sẽ lựa chọn một trong danh sách gợi ý K sản phẩm có xác xuất cao nhất mà mô hình gợi ý đưa ra

• Đối tượng và phạm vi nghiên cứu

- Đối tượng nghiên cứu: Hành vi nhấp chuột (lựa chọn sản phẩm) của khách hàng trong một phiên mua hàng
- Phạm vi thời gian: trong 6 tháng, từ tháng 04/2014 đến tháng 09/2014
- Phạm vi không gian: tập hợp nhiều phiên làm việc của một trang web thương mại điện tử hoạt động trong lĩnh vực bán lẻ tại châu Âu
- Phạm vi nội dung: tập trung vào thông tin về chuỗi nhấp chuột và một danh sách sản phẩm mà khách hàng lựa chọn trong mỗi phiên làm việc

• Phương pháp nghiên cứu:

- Phương pháp toán học
- Phương pháp nghiên cứu định lượng

• Lịch sử nghiên cứu

- Những năm 2000, JB Schafer và cộng sự đề xuất thuật toán sử dụng các thông tin trong quá khứ và sở thích cá nhân (personalization) của khách hàng.
- Badrul M. Sarwar và cộng sự (2000) đề xuất thuật toán phổ biến của học máy
 là SVD để đưa ra "top N list" gợi ý.
- o Năm 2004, Zan Huang và cộng sự đề xuất đồ thị có hướng gồm hai lớp.
- Năm 2006, Yehuda Koren, Robert Bell và Chris Volinsky trình bày mô hình phân tích ma trận thành nhân tử (matrix factorization).
- Balázs Hidasi và cộng sự (2015) đưa ra mô hình mạng hồi quy RNN trong việc xây dựng hệ gợi ý tập trung vào những phiên làm việc ngắn và hiện tại.
- yong Kiam Tan và cộng sự (2016) đã đề xuất cải tiến mô hình RNN với thuật toán xử lý dữ liêu làm việc theo chuỗi (phiên) phù hợp hơn cho mô hình RNN.
- 2016, HengTze Cheng và cộng sự, Khang Nguyen và cộng sự áp dụng mô hình mô hình mạng học sâu và rộng (Wide & Deep Learning).
- Shu Wu và cộng sự (2019) sử dụng mô hình mạng học sâu và đồ thị cùng khá nhiều kỹ thuật xử lý đồ thị và các biến thể khác nhau của GNN để phân tích bài toán gợi ý dựa vào phiên làm việc.
- Tính mới của đề tài: Xây dựng đồ thị từ bộ dữ liệu lựa chọn sản phẩm của người mua hàng và sử dụng mạng học sâu đồ thị GNN để học dữ liệu đồ thị để đưa ra các mô hình gợi ý dựa theo bài toán top-K

- Ý nghĩa lý luận và thực tiễn
 - Ý nghĩa lý luận
 - Xây dựng mô hình dự đoán hành vi người dùng: Cung cấp quy luật về mối quan hệ trong hành vi mua sắm trực tuyến và quyết định mua hàng của người tiêu dùng
 - Mô hình mạng học sâu: Giúp mô hình dự báo trở nên mạnh mẽ và linh hoạt hơn
 - Đồ thị và mối quan hệ trong dữ liệu: mang lại thông tin quan trọng về cách người tiêu dùng liên kết và chọn lựa sản phẩm trong các phiên làm việc.
 - Ý nghĩa thực tiễn
 - Tối ưu hóa trải nghiệm người dùng: Hệ gợi ý Top-K giúp người dùng dễ dàng lựa chọn sản phẩm phù hợp
 - Tăng cường hiệu suất kinh doanh và sự hài lòng của khách hàng.
 - Nghiên cứu trong lĩnh vực Recommendation Systems (RecSys): cung cấp cơ sở lý luận và thực nghiệm cho các nghiên cứu tương lai trong lĩnh vực này.
- Chương 1: Cơ sở lý luận nghiên cứu của đề tài
 - Các lý thuyết liên quan:
 - o Mô hình Session-based Recommendation (SR)
 - Mô hình mạng học sâu đồ thị (Graph Neural Network)
 - Các khái niệm liên quan đến đề tài nghiên cứu:
 - Đồ thị (graph)
 - o Phiên làm việc (session)
 - o Mô hình mạng (neural network)
 - Hệ gợi ý (recommendation system)
- Chương 2: Cơ sở thực tiễn nghiên cứu của đề tài
 - Nguồn dữ liệu(Dataset):
 - Nghiên cứu sử dụng bộ dữ liệu cung cấp bởi YOOCHOOSE GmbH
 - Vì lý do quyền riêng tư, toàn bộ thông tin về người sử dụng đã được ẩn đi khỏi bộ dữ liệu.
 - Bộ dữ liệu ghi lại tập hợp nhiều phiên làm việc của một trang web thương mại điện tử hoạt động trong lĩnh vực bán lẻ tại châu Âu, bao gồm:
 - Dữ liệu nhấp chuột
 - Dữ liệu nhấp chuột đánh giá

- Dữ liêu mua sắm
- Chương 3: Đánh giá và kết luận
 - So sánh với các nghiên cứu trước đây
 - Nghiên cứu này sử dụng cả tập dữ liệu huấn luyện và kiểm thử từ bộ dữ liệu gốc. Các nghiên cứu trước đây không sử dụng bộ dữ liệu kiểm thử riêng biệt của bộ dữ liệu gốc, mà trích ra từ tập dữ liệu huấn luyện
 - Nghiên cứu này đề xuất và xây dựng mô hình có tính mở rộng cao khi hoạt động với đồ thị. Một số nghiên cứu liên quan trình bày không thể chạy được mô hình với bộ dữ liệu đầy đủ, do đó họ phải thực nghiệm với bộ dữ liệu nhỏ hơn.
 - Lợi ích: Thực nghiệm cho thấy hướng đề xuất của nghiên cứu này cho kết quả tốt hơn hẳn các nghiên cứu trước đây, đặc biệt là khả năng xử lý đồ thị với số lượng đỉnh rất lớn và có khả năng mở rộng hơn nữa, đây chính là điểm hạn chế của các nghiên cứu trước đây.
 - Hướng phát triển: Hướng mở rộng của nghiên cứu này là tiếp tục sử dụng một số kỹ thuật biến đổi đồ thị (ví dụ phép nhúng đồ thị) sao cho mô hình GNN hoạt động hiệu quả hơn, hoặc phương án thiết kế tối ưu lớp sâu của mô hình GNN để nó có thể hoạt động hiệu quả hơn với đồ thị sâu K.
- Phần Kết luận: Tài liệu tham khảo và phụ lục

<u>Câu 2:</u> Trình bày cấu trúc của 1 bài báo khoa học. Lấy ví dụ 1 bài báo khoa học trong ngành học của bạn và phân tích để thấy rõ các yêu cầu về cấu trúc đó (lưu ý đính kèm bài báo cùng câu trả lời).

2.1. Cấu trúc của một bài báo khoa học

- Tiêu đề (Title):
 - Phản ánh nội dung đề cập trong bài báo.
 - Dưới tiêu đề bài báo thường là tên tác giả, tập thể tác giả, email, cơ quan công tác, ngày nhận bài báo và ngày chấp nhận đăng bài báo.
- Tóm tắt (Abstract):
 - Thể hiện vấn đề/mục tiêu nghiên cứu, phương pháp nghiên cứu, thời gian, số liệu dùng cho nghiên cứu, kết quả tác giả mới tìm ra, và kết luận.
 - Dưới tóm tắt là từ khóa (Key words) gồm 3 5 từ quan trọng có tần suất lặp lại nhiều.
- Giới thiệu/Đặt vấn đề (Introduction):
 - Thường nói về lý do, tầm quan trọng của vấn đề nghiên cứu

- Phải nêu rõ câu hỏi nghiên cứu
- Lược sử về nghiên cứu trước đây (Literature review)
 - Có thể gộp với phần Introduction hoặc tách riêng
 - Phải nêu những nghiên cứu quan trọng trước đây trên thế giới đã làm liên quan đến vấn đề nghiên cứu
- Phương pháp nghiên cứu (Methodology): Đề cập cách thực hiện nghiên cứu
- Kết quả và bàn luận (Results and Discussion): Trình bày kết quả của nghiên cứu và so sánh với các nghiên cứu trước đây.
- Kết luận (Conclusions): Tóm tắt lại các điểm chính của bài báo, kết quả quan trọng và hướng phát triển tương lai.
- Tài liệu tham khảo (References): Liệt kê các nguồn tài liệu mà tác giả đã tham khảo
- Phụ lục(Appendices) (có thể có hoặc không): Liệt kê các thông tin chi tiết, dữ liệu bổ sung,... mà tác giả không muốn đặt trực tiếp trong văn bản chính

2.2. Phân tích cấu trúc của 1 bài báo khoa học ngành CNTT

– Tiêu đề (Title):

Hệ gợi ý mua sắm dựa theo phiên làm việc với mô hình mạng học sâu đồ thị

Nguyễn Tuấn Khang¹, Nguyễn Tú Anh², Mai Thúy Nga², Nguyễn Hải An³, Nguyễn Việt Anh¹

- Viện Công Nghệ Thông Tin, Học Viện Khoa Học Công Nghệ, Hà Nội
- ² Trường Đại học Thăng Long, Hà Nội
- ³ Phòng Khoa Học Công Nghệ, Tổng Công Ty Thăm Dò Khai Thác Dầu Khí, Hà Nội

Tác giả liên hệ: Nguyễn Tuấn Khang, khang_nt@yahoo.com

Ngày nhận bài: 30/08/2022, ngày sửa chữa: 13/11/2022, ngày duyệt đăng: 25/11/2022

Định danh DOI: 10.32913/mic-ict-research-vn.v2022.n2.1135

Tóm tắt (Abstract):

Tóm tắt: Phân tích phiên làm việc của khách hàng để dự báo khả năng họ sẽ lựa chọn sản phẩm nào tiếp theo là một bài toán dự báo khá phổ biến trong ngành thương mại điện tử. Việc dự báo này giúp cho doanh nghiệp đưa ra các ý tưởng bán hàng phù hợp trong quá trình người dùng tương tác với hệ thống bán hàng của mình. Bài báo này đề xuất hướng sử dụng mạng học sâu đồ thị để xây dựng mô hình gợi ý dựa vào phiên làm việc của khách hàng. Kết quả thực nghiệm cho thấy việc sử dụng đồ thị rất phù hợp trong việc biểu diễn dữ liệu lựa chọn sản phẩm thông qua hành vi nhấp chuột trong phiên làm việc của khách hàng và mô hình gợi ý sử dụng GNN cho kết quả dự báo với 2 chỉ số đánh giá mô hình Recall@20 và MRR@20 tốt hơn so với các mô hình trước đây.

Từ khóa: dữ liệu nhấp chuột, hành vi mua sắm, hệ thống gợi ý, phiên làm việc, mạng học sâu đồ thị (Graph Neural Network).

Giới thiệu/Đặt vấn đề (Introduction):

I. TỔNG QUAN

1. Tổng quan bài toán

Khi một khách hàng vào một trang thương mại diện tử thì có hai xu hướng: hoặc họ đã định hướng được sản phẩm mà họ sẽ mua, hoặc là họ được định hướng được sản phẩm mà họ nên mua. Đối với kịch bản thứ hai, người dùng sẽ

2. Đặt vấn đề

Đối tượng nghiên cứu của bài báo này là hành vi nhấp chuột (lựa chọn sản phẩm) của khách hàng trong một phiên mua hàng. Mục tiêu của bài báo này là nghiên cứu và đề xuất mô hình dự báo hành vi lựa chọn sản phẩm của khách hàng mô hình mạng học sâu, cụ thể hơn là trả lời câu hỏi

- Lược sử về nghiên cứu trước đây (Literature review)

II. NGHIÊN CỨU LIÊN QUAN

Bài toàn gợi ý trong lĩnh vực thương mại điện tử không phải là vấn đề mới, ngay từ những năm 2000 JB Schafer và cộng sự [1] đã nêu ra vấn đề này để tìm cách nâng

- Phương pháp nghiên cứu (Methodology):

III. PHƯƠNG PHÁP LUẬN

1. Ý tưởng xây dựng đồ thị

IV. NGUỒN DỮ LIÊU NGHIÊN CỨU

1. Nguồn Dữ Liệu (Dataset)

- Kết quả và bàn luận (Results and Discussion):

- V. KẾT QUẢ THỰC NGHIÊM VÀ ĐÁNH GIÁ
- 1. Tham Số Đánh Giá Mô Hình

3. So sánh với các nghiên cứu liên quan

Để so sánh hướng tiếp cận và kết quả đạt được của nghiên cứu này, tác giả lựa chọn một số bài báo tương tự cùng giải

- Kết luận (Conclusions):

VI. KẾT LUẬN

Nghiên cứu này có hai đóng góp chính gồm: (1) sử dụng đồ thị để mô tả chuỗi nhấp chuột trong quá trình lựa chọn sản phẩm trong phiên làm việc hiện tại, bao gồm ba loại đồ thị trực tiếp G, gián tiếp H và sâu K và (2) sử dụng

- Tài liệu tham khảo (References):

TÀI LIÊU THAM KHẢO

- J. B. Schafer, J. Konstan, and J. Riedl, "Recommender systems in e-commerce," in *Proceedings of the 1st ACM* conference on Electronic commerce, 1999, pp. 158–166.
- [2] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

<u>Câu 3</u>: Hãy tìm hiểu 1 đồ án/khóa luận tốt nghiệp trong ngành học của mình và tóm tắt lại các nội dung sau: Tên đồ án; Tác giả, Nơi công bố, Năm công bố; Mục tiêu nghiên cứu; Cơ sở lý thuyết; Phương pháp thu thập và xử lý thông tin; Kết quả đạt được; Hạn chế của công trình.

- Tên đồ án: Hệ thống tư vấn món ăn dựa trên thói quen và đánh giá của khách hàng^[2]
- Tác giả: PGS.TS Trần Đình Quế, sinh viên Hoàng Tăng Khải
- Nơi công bố: Thư viện số Học viện CNBCVT
- Năm công bố: 2022
- Mục tiêu nghiên cứu:
 - Nghiên cứu và đề xuất món ăn dựa trên thói quen và đánh giá của khách hàng
 - Tích hợp hệ thống tư vấn vào một ứng dụng mobile giúp người dùng xem những công thức nấu ăn được hệ thống tư vấn
- Cơ sở lý thuyết:
 - Các lý thuyết liên quan:
 - Hệ thống tư vấn Recommendation Systems (RecSys)
 - o Phương pháp phân tích suy biến SVD
 - Các khái niệm trong đề tài nghiên cứu:
 - o Món ăn (meal)
 - Thói quen (habit)
 - Đánh giá (review)
 - Khách hàng (customer)
- Phương pháp thu thập và xử lý thông tin:
 - Phương pháp toán học
 - Phương pháp nghiên cứu định lượng
- Kết quả đạt được
 - Tim hiểu được tổng quan hệ thống tư vấn, hiểu rõ và triển khai được kỹ thuật tư vấn dưa trên thuật toán SVD
 - Cài đặt thành công bộ dữ liệu mẫu để đánh giá thuật toán SVD và xây dựng hệ thống tư vấn thử nghiệm
- Han chế
 - Quá trình thử nghiệm mới chỉ thực hiện trên một bộ dữ liệu

 Chưa có so sánh thuật toán SVD với các thuật toán khác để đưa ra nhận xét khách quan hơn

Câu 4: Hãy xây dựng đề cương nghiên cứu cho 1 đề tài nghiên cứu khoa học trong ngành học của bạn mà bạn yêu thích (theo mẫu)

Đề tài: Hệ thống Dự đoán và Phân loại Tình trạng Tâm lý dựa trên Nội dung Viết trên Mạng Xã Hội

- Phần Mở đầu
 - Tính cấp thiết của đề tài:

Trong môi trường mạng xã hội, người dùng thường chia sẻ những suy nghĩ, cảm xúc và trạng thái tâm lý của họ qua các bài viết và bình luận. Điều này tạo ra một nguồn lượng dữ liệu lớn, đa dạng và có tiềm năng lớn để hiểu biết về tâm lý con người. Việc phát triển một hệ thống dự đoán và phân loại tình trạng tâm lý từ nội dung viết trên mạng xã hội sẽ mang lại lợi ích lớn cho việc nghiên cứu và ứng dụng trong lĩnh vực tâm lý học, y tế tâm thần, và quản lý tâm trạng cá nhân. Vấn đề đặt ra là làm thế nào có thể dự đoán tình trạng tâm lý của người dùng thông qua nội dung viết trên mạng xã hội.

- Mục tiêu nghiên cứu: Nghiên cứu và phát triển hệ thống có khả năng dự đoán và phân loại tình trạng tâm lý của người dùng dựa trên nội dung viết trên mạng xã hôi.
- Câu hỏi và giả thuyết nghiên cứu
 - Câu hỏi nghiên cứu: "Làm thế nào nội dung viết trên mạng xã hội có thể được sử dụng để dự đoán tình trạng tâm lý của người dùng?"
 - Giả thuyết: Hệ thống phân loại sẽ có khả năng dự đoán chính xác tình trạng tâm lý dựa trên mô hình máy học và xử lý ngôn ngữ tự nhiên.
- Đối tượng và phạm vi nghiên cứu
 - Đối tượng nghiên cứu: Bài đăng trên mạng xã hội có nội dung văn bản liên quan đến tâm trang
 - Phạm vi thời gian: trong 6 tháng, từ tháng 12/2023 đến tháng 5/2024
 - Phạm vi không gian: tập hợp các bài viết trên mạng xã hội Facebook
- Phương pháp nghiên cứu:
 - Phương pháp thu thập và xử lý dữ liệu: sử dụng API, web scraping
 - Phương pháp toán học
 - o Phương pháp nghiên cứu định lượng

- Lịch sử nghiên cứu
 - 2005, Gilad Mishne và cộng sự đề xuất phương pháp PMI-IR sử dụng các bài đăng của người viết để phân loại blog theo tâm trạng
- Tính mới của đề tài: Áp dụng các phương pháp máy học tiên tiến để dự đoán và phân loại tâm trạng dựa trên nội dung viết trên mạng xã hội.
- Ý nghĩa lý luận và thực tiễn
 - Ý nghĩa lý luận
 - Nâng cao hiểu biết về cách nội dung trên mạng xã hội phản ánh tâm trạng của người dùng.
 - Ý nghĩa thực tiễn
 - Phát triển ứng dụng thực tế hỗ trợ quản lý tâm lý và tư vấn tâm sự dựa trên nội dung mạng xã hội.
- Chương 1: Cơ sở lý luận nghiên cứu của đề tài
 - Các lý thuyết liên quan:
 - Các lý thuyết liên quan về xử lý ngôn ngữ tự nhiên và phân loại tâm trạng.
 - o Mô hình học máy dựa trên NLP đã được áp dụng trong lĩnh vực này.
 - Các khái niệm trong đề tài nghiên cứu:
 - o Tâm lý (metal state)
 - Bài viết (status)
 - o Mạng xã hội (social media)
- Chương 2: Cơ sở thực tiễn nghiên cứu của đề tài
 - Nguồn dữ liệu(Dataset):
 - API công cộng cung cấp bởi mạng xã hội để thu thập dữ liệu theo thời gian thực
 - Vì lý do quyền riêng tư, toàn bộ thông tin về người sử dụng đã được ẩn đi khỏi bộ dữ liệu.
 - Bộ dữ liệu bao gồm:
 - Dữ liệu về bài đăng và bình luận
 - Dữ liệu về tương tác
 - Dữ liệu văn bản và hình ảnh
- Chương 3: Đánh giá và kết luận
 - Hạn chế

- Tâm trạng tinh thần là một khái niệm phức tạp và có thể được hiểu khác nhau bởi mỗi người. Các thuật toán phân loại tâm trạng có thể gặp khó khăn khi đối mặt với sự phức tạp này.
- Hướng phát triển
 - o Mở rộng trên các nền tảng khác nhau
 - o Tối ưu hóa đa ngôn ngữ và văn hóa
- Phần Kết luận: Tài liệu tham khảo và phụ lục

Tài liệu tham khảo

[1] Nguyễn Tuấn Khang và nnk (2022). Hệ gợi ý mua sắm dựa theo phiên làm việc với mô hình mạng học sâu đồ thị. *Tạp chí Thông tin và Truyền thông: Chuyên san Các công trình nghiên cứu, phát triển và ứng dụng CNTT và truyền thông / Bộ Thông tin và Truyền thông, Tập 2022, Số 2, Tháng 12*, Trang 73-82. DOI: https://doi.org/10.32913/mic-ict-research-vn.v2022.n2.1135

[2] Hoàng Tăng Khải (2022). *Hệ thống tư vấn món ăn dựa trên thói quen và đánh giá của khách hàng* (Đồ án tốt nghiệp, Học viện Công nghệ Bưu chính Viễn thông). http://dlib.ptit.edu.vn/handle/HVCNBCVT/2505

