

ĐỀ CƯƠNG KHOÁ LUẬN TỐT NGHIỆP

**MỘT PHƯƠNG PHÁP CẢI TIẾN MÔ
HÌNH NHÂN TỐ ẨN CHO BÀI TOÁN
TƯ VẤN NHÓM**

*A method to improve latent matrix factorization model in group
recommender system*

1. THÔNG TIN CHUNG

Người hướng dẫn:

– Th.S Tiết Gia Hồng (Khoa Công nghệ Thông tin)

Sinh viên thực hiện:

1. Nguyễn Lê Duy (MSSV: 20127482)

2. Lê Phan Duy Tùng (MSSV: 20127661)

Loại đề tài: Nghiên cứu

Thời gian thực hiện: Từ 1/2024 đến 7/2024

2. NỘI DUNG THỰC HIỆN

2.1 Giới thiệu về đề tài

Hiện nay, với sự phát triển vượt bậc của công nghệ, con người dễ dàng tiếp cận được với một lượng lớn thông tin đến từ nhiều nguồn khác nhau dẫn đến vấn đề quá tải thông tin. Người dùng ngày càng khắt khe hơn và đòi hỏi cao hơn khi lựa chọn sản phẩm. Điều này thúc đẩy các doanh nghiệp phải ngày càng chú trọng trong việc phát triển và nâng cấp hệ thống của mình giúp người dùng dễ dàng tìm được sản phẩm phù hợp, giúp nâng cao trải nghiệm người dùng khi sử dụng sản phẩm.

Nhiều doanh nghiệp lớn hiện nay không chỉ dừng lại ở việc phát triển và tích hợp các mô hình gợi ý cho từng cá nhân mà còn hướng tới xây dựng các hệ thống tư vấn dành cho nhóm người dùng, vì xu hướng mua hàng theo nhóm ngày càng tăng cao. Tuy nhiên, việc tư vấn nhóm đối diện với một số thách thức đặc biệt, bao gồm:

- **Sự đa dạng của nhóm:** Mỗi người dùng có thể có sở thích, yêu cầu và hành vi mua hàng khác nhau. Việc tìm ra các sản phẩm phù hợp cho một nhóm với nhiều thành viên và đa dạng sở thích như vậy đòi hỏi các phương pháp và mô hình phức tạp.
- **Sự tương tác phức tạp:** Sự tương tác giữa các thành viên trong nhóm, cũng như giữa nhóm và sản phẩm có thể phức tạp và khó dự đoán. Cần có các phương pháp hiệu quả để mô hình hóa và dự đoán các mẫu tương tác này.

Nhóm thực hiện đề tài này nhằm mục tiêu giải quyết các thách thức trên bằng cách đề xuất một phương pháp cải tiến hệ thống tư vấn nhóm trở nên chính xác và đáng tin cậy hơn khi đối mặt với vấn đề dữ liệu thừa thớt, giúp các doanh nghiệp cung cấp các gợi ý sản phẩm phù hợp nhất cho nhóm người dùng. Điều này có thể giúp tăng cường trải nghiệm người dùng, nâng cao lợi nhuận và củng cố vị thế cạnh tranh của các doanh

nghiệp trên thị trường.

2.2 Mục tiêu đề tài

Trong thời đại phát triển của công nghệ, việc đáp ứng đầy đủ và chính xác nhu cầu của người dùng có thể mang lại lợi ích to lớn cho các doanh nghiệp. Một trong những vấn đề được quan tâm chính là cải tiến các mô hình tư vấn sản phẩm cho không chỉ một người dùng mà còn có thể tư vấn cho một nhóm người dùng.

Hiện nay, các mô hình tư vấn sử dụng phương pháp nhân tố ẩn (Latent Factor) và các giải pháp cải tiến dựa trên đó đã mang lại kết quả tích cực [8]. Nhóm đã đề xuất một phương pháp mới để cải thiện hệ thống tư vấn nhóm bằng cách tích hợp thông tin đánh giá (review) của nhóm với ma trận rating để khắc phục tính thừa thớt của ma trận rating khi dự đoán trong mô hình nhân tố ẩn. Điều này giúp tăng độ chính xác của mô hình nhân tố ẩn khi được áp dụng cho các hệ tư vấn trên nhóm.

Đề tài có thể mang lại nhiều ảnh hưởng tích cực đối với cả vấn đề cụ thể về tư vấn nhóm và lĩnh vực nghiên cứu hệ thống tư vấn trong tổng thể, như sau:

Đối với vấn đề cụ thể, một hệ thống tư vấn nhóm hiệu quả có thể giúp tăng cường trải nghiệm mua sắm của người dùng, tăng doanh số bán hàng cho các doanh nghiệp và củng cố vị thế cạnh tranh trên thị trường.

Trong lĩnh vực nghiên cứu, phương pháp cải tiến có thể cung cấp một bước tiến quan trọng trong việc tăng cường hiệu suất của các hệ thống tư vấn thông qua việc áp dụng các kỹ thuật tiên tiến và phức tạp hơn trong mạng neural và học sâu. Điều này có thể mở ra cánh cửa cho nhiều nghiên cứu và ứng dụng các kiến thức về trí tuệ nhân tạo trong hệ thống tư vấn.

2.3 Phạm vi của đề tài

Đề tài tập trung vào phương pháp cải tiến mô hình tư vấn nhóm bằng phương pháp tích hợp rating dự đoán từ mô hình ngôn ngữ lớn đã được huấn luyện để dự đoán rating sản phẩm từ text review vào trong mô hình nhân tố ẩn dành cho nhóm để dự đoán, đề xuất sản phẩm cho nhóm người dùng

2.4 Cách tiếp cận dự kiến

Có hai chiến lược gợi ý dành cho nhóm phổ biến đó là tổng hợp gợi ý và tổng hợp profile [7]:

- Tổng hợp gợi ý: Tổng hợp gợi ý cho từng thành viên trong nhóm sau đó tổng hợp lại danh sách này để tạo thành một danh sách dành cho nhóm
- Tổng hợp profile nhóm người dùng: Tổng hợp sở thích của các thành viên trong nhóm để tạo thành profile nhóm duy nhất còn được gọi là người dùng ảo. Các gợi ý cho người dùng ảo sẽ chính là gợi ý dành cho nhóm

Trong hai cách tiếp cận trên thì cách tiếp cận thứ hai có những lợi thế so với cách tiếp cận trước đó như sau [7]:

- Chiến lược tổng hợp profile nhóm người dùng hữu ích hơn trong những tình huống mà các thành viên trong nhóm có cơ hội để phân tích, điều chỉnh các ưu tiên của nhóm.
- Chiến lược tổng hợp profile nhóm người dùng giảm thiểu các mối lo ngại về quyền riêng tư của người dùng có thể được giảm bớt do không cần phải ghi lại và duy trì hồ sơ người dùng cá nhân.

Để cải tiến độ hiệu quả trong cách tiếp cận gợi ý dựa trên phương pháp tổng hợp profile người dùng, đã có một số phương pháp như Latent Group Model [5] đề xuất áp

dùng latent group factor trong việc biểu diễn nhóm người dùng thành một profile người dùng ảo (virtual-user) trong không gian nhân tố ẩn thay vì ma trận user-item rating trong tư vấn nhóm. Theo sau đó là phương pháp Observed-filled rating [8] và Observed-filled rating kết hợp với trọng số của người dùng [6] đã tiếp tục cải tiến mô hình nhóm nhân tố ẩn trong giai đoạn tổng hợp profile nhóm và khắc phục điểm yếu của mô hình khi làm việc với bộ dữ liệu thưa có kích thước lớn.

Tuy nhiên các nghiên cứu [5][6][8] trên vẫn dựa trên điểm rating từ 1 đến 5 cho sản phẩm để tổng hợp gợi ý dành cho nhóm. Trong khi dữ liệu rating của sản phẩm thường rất thưa và có thể khiến kết quả phân tích kém chính xác trong nhiều trường hợp. Để giải quyết vấn đề này, ngày càng có nhiều mô hình phát triển để kết hợp thêm các loại feedback khác của người dùng thông qua các dạng văn bản, hình ảnh, hay nhãn dán để gia tăng độ tin cậy cho hệ thống [9][10]. Trong đó, phản hồi dạng văn bản hay còn gọi là bình luận (review) có độ tin cậy cao nên được nhiều mô hình tư vấn áp dụng để tạo đề xuất.

Hướng tiếp cận mà nhóm đang thực hiện sẽ áp dụng mô hình ngôn ngữ lớn BERT để dự đoán điểm đánh giá của người dùng dựa trên text review [11] và Observed-filled rating kết hợp với trọng số của người dùng [6] để xây dựng được hai profile người dùng ảo đại diện cho nhóm người dùng thay vì chỉ là một như mô hình trước [5][6][8]. Sau đó kết hợp hai profile người dùng ảo để đưa ra gợi ý cho nhóm người dùng.

2.5 Kết quả dự kiến của đề tài

Sau khi hoàn hành đề tài nhóm chúng em mong muốn đạt được những tiêu chí sau:

- Tích hợp thành công “text review” của người dùng bên cạnh điểm rating của sản phẩm vào trong mô hình tư vấn nhóm
- Tích hợp được trọng số giữa các thành viên trong nhóm người dùng để cải

thiện quá trình tổng hợp profile người dùng

- Thực nghiệm trên tập dữ liệu để đánh giá và so sánh hiệu suất của mô hình đề xuất so với mô hình cũ.

2.6 Kế hoạch thực hiện

Mốc thời gian	Nội dung công việc	Thành viên
1/2024 – 2/2024	Tìm hiểu mô hình nhân tố ẩn	Nguyễn Lê Duy
	Tìm hiểu về kiến trúc Transformer Encoder	Lê Phan Duy Tùng
2/2024 – 3/2024	Tìm hiểu về các LLM (Large Language Model) dựa trên kiến trúc Transformer Encoder	Lê Phan Duy Tùng
	Tìm hiểu phương tư vấn nhóm bằng mô hình nhân tố ẩn dựa trên các xếp hạng dự đoán của tất cả thành viên trong nhóm	Nguyễn Lê Duy
3/2023 – 4/2024	Tìm hiểu về phương pháp huấn luyện LLM trong tác vụ phân loại văn bản	Lê Phan Duy Tùng
	Tìm hiểu phương tư vấn nhóm bằng mô hình nhân tố ẩn dựa trên các xếp hạng dự đoán của tất cả thành viên trong nhóm có kết hợp trọng số	Nguyễn Lê Duy

4/2024 – 5/2024	Đề xuất mô hình tích hợp cải tiến.	Lê Phan Duy Tùng Nguyễn Lê Duy
5/2024 – 6/2024	Cài đặt mô hình và đánh giá.	Lê Phan Duy Tùng Nguyễn Lê Duy
6/2024 – 7/2024	Viết báo cáo phân tích đề tài	Lê Phan Duy Tùng Nguyễn Lê Duy


Tài liệu

- [1] BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, by Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova
- [2] Attention Is All You Need by Ashish Vaswani, Noam Shazeer, and Niki Parmar
- [3] Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT by Sudharsan Ravichandiran
- [4] Natural Language Processing with Transformers, Revised Edition by Lewis Tunstall, Leandro von Werra, Thomas Wolf
- [5] J. Shi, B. Wu and X. Lin, "A Latent Group Model for Group Recommendation," 2015 IEEE International Conference on Mobile Services, New York, NY, USA, 2015, pp. 233-238
- [6] Le Nguyen Hoai Nam, Towards comprehensive profile aggregation methods for group recommendation based on the latent factor model, Expert Systems with



Applications, Volume 185, 2021.

- [7] Alexander Felfernig, Ludovico Boratto, Martin Stettinger, Marko Tkalčič, Group Recommender Systems: An Introduction (Signals and Communication Technology) 2nd ed. 2024 Edition
- [8] Le Nguyen Hoai Nam, Ho Thi Hoang Vy, Le Hoang My, Le Thi Tuyet Mai, Hong Tiet Gia, and Ho Le Thi Kim Nhung. 2019. An approach to improving group recommendation systems based on latent factor matrices. In Proceedings of the 10th International Symposium on Information and Communication Technology (SoICT '19). Association for Computing Machinery, New York, NY, USA, 98–105
- [9] G. Adomavicius and A. Tuzhilin. “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions”. In: IEEE Transactions on Knowledge and Data Engineering 17.6 (2005), pp. 734–749.
- [10] Claudiu S. Firan, Wolfgang Nejdl, and Raluca Paiu. “The Benefit of Using TagBased Profiles”. In: 2007 Latin American Web Conference (LA-WEB 2007). 2007, pp. 32–41.
- [11] Yuanyuan Zhuang and Jaekyeong Kim A BERT-Based Multi-Criteria Recommender System for Hotel Promotion Management

XÁC NHẬN
CỦA NGƯỜI HƯỚNG DẪN
(Ký và ghi rõ Họ tên)


Tiet Gia Hong

TP. Hồ Chí Minh, ngày 10 tháng 9 năm 2024
NHÓM SINH VIÊN THỰC HIỆN
(Ký và ghi rõ họ tên)


Nguyen Lê Duy Anh

Tran Duy Hung