

TRÍCH YẾU LUẬN ÁN

I. Tóm tắt mở đầu

- Tên luận án: Nghiên cứu mô hình hóa học sâu dữ liệu chính tắc và phụ trợ nhằm nâng cao tính vững chắc và thích nghi của hệ thống khuyến nghị.

II. Nội dung bản trích yếu

1. Mục tiêu và đối tượng nghiên cứu của luận án

Nghiên cứu của luận án hướng đến ba mục tiêu chính sau đây:

- **Thứ nhất**, phát triển khung hệ thống khuyến nghị thống nhất dựa trên Graph Neural Networks (GNNs) để giải quyết các thách thức về cold-start, tính thưa thớt dữ liệu và khả năng mở rộng. Khung này kết hợp kỹ thuật fusion thông tin phụ dựa trên attention, soft clustering và masked contrastive learning để tạo ra embedding người dùng-sản phẩm công bằng, bền vững và độc lập với ID.
- **Thứ hai**, đề xuất framework học liên tục (continual learning) dựa trên task masking cho khuyến nghị đa miền nhằm đảm bảo hiệu suất công bằng, không thiên vị giữa nhiều miền và phân khúc người dùng, đồng thời ngăn chặn hiện tượng catastrophic forgetting và cho phép thích ứng đa miền bền vững.
- **Thứ ba**, xây dựng khung khuyến nghị lai tích hợp deep learning dựa trên đồ thị với các mô hình ngôn ngữ lớn (LLMs) trong kiến trúc Retrieval-Augmented Generation (RAG), nhằm nâng cao khuyến nghị phim bằng cách kết hợp dữ liệu tương tác lịch sử người dùng-sản phẩm với phản hồi hội thoại thời gian thực để cải thiện độ chính xác, cá nhân hóa và nhận biết ngữ cảnh.

2. Các phương pháp nghiên cứu đã sử dụng

Luận án sử dụng phương pháp nghiên cứu kết hợp bao gồm: phương pháp nghiên cứu định tính và phương pháp nghiên cứu định lượng. Luận án tiến hành việc phân tích định tính các khái niệm và mô hình từ hệ thống tài liệu liên quan về hệ thống khuyến nghị, Graph Neural Networks, contrastive learning và continual learning. Từ đó, đề xuất các kỹ thuật và mô hình khuyến nghị mới dựa trên đồ thị.

Đồng thời, luận án tiến hành các nghiên cứu định lượng thông qua việc triển khai các hệ thống thực nghiệm tương ứng trên các bộ dữ liệu chuẩn như MovieLens-1M, TV360, Book-Crossing và Douban. Các kỹ thuật cụ thể được sử dụng bao gồm: Graph Neural Networks để mô hình hóa tương tác người dùng-sản phẩm; Self-supervised contrastive learning với masked graph learning; Continual learning với task masking để bảo vệ kiến thức miền cụ thể; Kỹ thuật fusion dựa trên attention cho thông tin phụ; và Retrieval-Augmented Generation kết hợp với LLMs.

3. Các kết quả chính và kết luận

Tham gia vào dòng nghiên cứu trên thế giới về hệ thống khuyến nghị thông minh và graph-based learning, luận án đã hoàn thiện mục tiêu nghiên cứu với ba đóng góp chính sau đây:

- **Đầu tiên**, luận án đề xuất nghiên cứu về framework khuyến nghị dựa trên GNNs thống nhất giải quyết cold-start, data sparsity và khả năng mở rộng quy mô lớn. Theo định hướng này, luận án phát triển ba mô hình: EfficientRec sử dụng soft clustering và interaction embedding profiles để loại bỏ phụ thuộc user ID; Masked Simple Graph Contrastive Learning (MS-GCL) áp dụng cơ chế masking có thể học được để nâng cao chất lượng biểu diễn; và GIFT4Rec tích hợp thông tin phụ đa nguồn thông qua attention mechanism để giải quyết vấn đề cold-start.
- **Tiếp đó**, luận án đã nghiên cứu và phát triển framework continual learning cho khuyến nghị đa miền. Bằng việc tổng hợp các nghiên cứu về task masking và domain adaptation, luận án đã đề xuất mô hình CL4Rec với hai thành phần chính: Domain Masking để xác định và bảo vệ tham số quan trọng cho từng miền, và Domain Specialization để ngăn chặn catastrophic forgetting. Framework này có tính model-agnostic và có thể áp dụng trên các mô hình tiềm ẩn khác nhau như matrix factorization, graph neural networks và auto-encoders.
- **Cuối cùng**, luận án nghiên cứu về hệ thống khuyến nghị hội thoại và phát triển kiến trúc lai GNN-RAG-LLM. Theo định hướng nghiên cứu này, luận án tích hợp sức mạnh cấu trúc của GNNs trong việc mô hình hóa tương tác người dùng-sản phẩm với khả năng hiểu ngữ cảnh của LLMs. Hệ thống bao gồm hai thành phần chính: conversational generator để hiểu intent người dùng và recommendation generator để tạo ra khuyến nghị được cá nhân hóa dựa trên cả dữ liệu lịch sử và phản hồi hội thoại thời gian thực.

Với các mô hình đề xuất mới, luận án đều tiến hành các thực nghiệm comprehensive và đưa ra so sánh kết quả với các mô hình state-of-the-art. Bằng các kết quả thực nghiệm trên nhiều bộ dữ liệu chuẩn, luận án đã chỉ ra hiệu quả vượt trội của các phương pháp đề xuất với cải thiện đáng kể về độ chính xác (84% trên HitRatio@20), hiệu suất tính toán (42-908 lần nhanh hơn) và khả năng mở rộng trong môi trường thực tế.

Danh mục công trình khoa học của Tác giả liên quan tới luận án:

1. “An incremental ensemble learning system for Vietnamese e-commerce product classification”, (KSE 2021), The 13th IEEE International Conference on Knowledge and Systems Engineering
2. “EfficientRec: An unlimited user scale recommendation system based on clustering and user’s interaction embedding profile”, (ACIIDS – 2022), 14th Asian Conference on Intelligent Information and Database Systems
3. “GIFT4Rec: An effective side Information Fusion Technique apply to Graph neural network for cold-start recommendation”, (ACIIDS – 2023), 15th Asian Conference on Intelligent Information and Database Systems, Pages 334 – 345
4. “The Masked Simple Graph Contrastive Learning for Recommendation” In 16th International Conference on Knowledge and System Engineering (KSE) (pp. 156-160). IEEE., 2024

5. “Continual Learning based on Task Masking for Multi-Domain Recommendation” In Asian Conference on Intelligent Information and Database Systems, 2024
6. “Summarizing Vietnamese Books Using a Multi-stage Hybrid Pipeline” In 17th International Conference on Computational Collective Intelligence, 2025
7. “Improving Retrieval-Augmented Generation for Scalable Movie Chatbots via Graph Based Recommendation Models” In – Submitted (In-peer-reviewing) – IEEE Access

Hà Nội, ngày tháng năm 2026