**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

**---------------------------------------**



**NGUYỄN BẢO TÙNG**

**ÁP DỤNG KỸ THUẬT HỌC SÂU CHO TƯ VẤN TRÊN SÀN THƯƠNG MẠI ĐIỆN TỬ POSTMART**

**Chuyên ngành: Hệ thống thông tin**

**Mã số: 8.48.01.04**

**TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ**

**Hà Nội – Năm 2024**

Đề án tốt nghiệp được hoàn thành tại:

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

Người hướng dẫn khoa học: PGS, TS Trần Đình Quế

Phản biện 1: …………………………………………………

Phản biện 2: …………………………………………………

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: ....... giờ ....... ngày ....... tháng ....... năm 2025

Có thể tìm hiểu đề án tốt nghiệp tại:

 - Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

**I. MỞ ĐẦU**

Ngày nay, thương mại điện tử (TMĐT) đã và đang trở thành lĩnh vực có ảnh hưởng cực kỳ quan trọng đến tăng trưởng kinh tế của các quốc gia. Sự phát triển của TMĐT không chỉ giúp các hoạt động kinh doanh thuận lợi mà còn cung cấp nhiều giá trị mới và đáp ứng những nhu cầu mới của các doanh nghiệp và người tiêu dùng. Chính vì vậy, mọi quốc gia trên thế giới đều quan tâm đến việc phát triển TMĐT. Có thể kể đến một số vai trò của TMĐT bao gồm: Thay đổi tính chất của nền kinh tế mỗi quốc gia và nền kinh tế toàn cầu, làm cho tính tri thức trong nền kinh tế ngày càng tăng lên, mở ra cơ hội phát huy ưu thế của các nước phát triển sau, để họ có thể đuổi kịp, thậm chí vượt các nước đi trước, rút ngắn khoảng cách về trình độ tri thức giữa các nước phát triển với các nước đang phát triển.

Trong những năm gần đây, thị trường TMĐT Việt Nam ngày càng được mở rộng và hiện đã trở thành phương thức kinh doanh phổ biến được doanh nghiệp, người dân biết đến. Sự đa dạng về mô hình hoạt động, về đối tượng tham gia, về quy trình hoạt động và chuỗi cung ứng hàng hóa, dịch vụ với sự hỗ trợ của hạ tầng Internet và ứng dụng công nghệ hiện đại đã đưa TMĐT trở thành trụ cột quan trọng trong tiến trình phát triển kinh tế số của quốc gia.

Mặc dù gặp những ảnh hưởng tiêu cực trong năm 2020 do đại dịch Covid-19, TMĐT Việt Nam vẫn có những bước tăng tốc mạnh mẽ, trở thành một trong những thị trường TMĐT tăng trưởng nhanh nhất trong khu vực Đông Nam Á. Theo Sách trắng Thương mại điện tử Việt Nam, năm 2020, tốc độ tăng trưởng của TMĐT đạt mức 18%, quy mô đạt 11,8 tỷ USD và là nước duy nhất ở Đông Nam Á có tăng trưởng TMĐT 2 con số. Theo tính toán của các tập đoàn lớn thế giới như Google, Temasek và Bain&Company, nhiều khả năng quy mô của nền kinh tế số Việt Nam sẽ vượt ngưỡng 52 tỷ USD và giữ vị trí thứ 3 trong khu vực ASEAN vào năm 2025.

Tuy nhiên, nguồn nhân lực của TMĐT nhìn chung vẫn còn thiếu và yếu, cũng như hạ tầng thương mại kỹ thuật cho TMĐT chưa thuận lợi, khiến TMĐT chưa tạo được sự tin tưởng cũng như chưa phát triển mạnh mẽ. Do vậy, việc xây dựng và liên tục học hỏi, cải thiện và nâng cấp các chức năng, hệ thống cho sàn Thương mại điện tử là một giải pháp vô cùng quan trọng. Đặc biệt là các tính năng phục vụ cho khách hàng, đối tượng chủ chốt tạo ra nguồn lợi nhuận cho doanh nghiệp.

Với mục đích đưa những tiến bộ công nghệ vào phục vụ cho sàn Thương mại điện tử cũng như việc kinh doanh cho doanh nghiệp, học viên xin chọn đề tài nghiên cứu “*Áp dụng kỹ thuật học sâu cho tư vấn trên Sàn thương mại điện tử Postmart*”.

**2. Tổng quan về vấn đề nghiên cứu:**

Hệ thống tư vấn (Recommendation system) dần trở thành một thành phần không thể thiếu của các sản phẩm điện tử có lượng người dùng lớn. Các sản phẩm cá nhân hóa điện tử ngày càng phổ biến với mục đích mang sản phẩm phù hợp tới người dùng hoặc giúp người dùng có các trải nghiệm tốt hơn trên nền tảng. Nếu quảng cáo sản phẩm tới đúng người dùng, khả năng các món hàng được mua nhiều hơn. Nếu tư vấn một video mà người dùng nhiều khả năng thích hoặc tư vấn kết bạn đến đúng đối tượng, họ sẽ ở lại trên nền tảng của bạn lâu hơn.

Thương mại điện tử là nơi hệ thống tư vấn được sử dụng nhiều nhất dưới dạng quảng cáo hướng đúng đối tượng. Quảng cáo điện tử ngoài việc giúp các doanh nghiệp bán được nhiều hàng còn giúp họ tiết kiệm được chi phí kho bãi. Họ sẽ không cần các cửa hàng ở vị trí thuận lợi để thu hút khách hàng hay phải trưng ra mọi mặt hàng ở vị trí đắc địa nhất trong cửa hàng. Mọi thứ có thể được cá nhân hóa sao cho mỗi người dùng nhìn thấy những sản phẩm khác nhau phù hợp với nhu cầu và sở thích của họ.

**3. Mục đích nghiên cứu:**

Nghiên cứu và xây dựng Hệ thống Tư vấn (Recommendation system) cho Sàn thương mại điện tử Postmart, cụ thể là tư vấn, gợi ý các sản phẩm thông qua việc mua sắm và tìm kiếm của khách hàng.

**4. Đối tượng và phạm vi nghiên cứu:**

Đề án tập trung vào nghiên cứu phương pháp xây dựng Hệ thống tư vấn (Recommendation system) dành cho sàn thương mại điện tử Postmart.

Trong đó, phạm vi nghiên cứu ở đây chỉ tập trung vào xây dựng hệ thống tư vấn gồm xử lí ngôn ngữ tự nhiên, phân loại ý định và xây dựng giao diện hỗ trợ xây dựng tìm kiếm sản phẩm, tư vấn sản phẩm giữa người dùng và hệ thống.

**5. Phương pháp nghiên cứu:**

Hai phương pháp hệ thống tư vấn:

* Hệ thống tư vấn dựa trên nội dung - Content based recommender systems: tức là hệ thống sẽ quan tâm đến nội dung, đặc điểm của mục tin hiện tại và sau đó tư vấn cho người dùng các mục tin tương tự.
* Hệ thống tư vấn dựa trên các user - lọc cộng tác - Collaborative filtering recommender systems: tức là hệ thống sẽ phân tích các user có cùng đánh giá, cùng mua mục tin hiện tại. Sau đó tìm ra danh sách các mục tin khác cũng được đánh gía bởi các user này, xếp hạng và tư vấn cho người dùng. Tư tưởng của phương pháp này chính là dựa trên sự tương đồng về sở thích giữa các người dùng để đưa ra các tư vấn.

Có một điều dễ nhận thấy thì phương pháp tư vấn dựa trên nội dung đòi hỏi chúng ta phải thu thập rất nhiều thông tin về các mục tin tương tự . Chính việc xác định xem một mục tin nào là tương tự với mục tin hiện tại đòi hỏi chúng ta phải thu thập và phần tích, xử lý toàn bộ các mục tin trong cơ sở dữ liệu. Tuy nhiên với phương pháp lọc công tác chúng ta không cần quá nhiều thông tin. Đơn giản chỉ là item\_id của item hiện tại, các user\_id và các feedback trên item đó mà thôi nên thực tế thì phương pháp lọc cộng tác được sử dụng phổ biến hơn để xây dựng các hệ thống tư vấn

**II. NỘI DUNG NGHIÊN CỨU**

**Chương 1: Tổng quan hệ thống tư vấn sản phẩm trên TMĐT**

**1.1 Khái niệm, vai trò và phân loại hệ thống tư vấn**

Hệ thống tư vấn sản phẩm (recommender system – RS) là tập hợp các thuật toán và mô hình nhằm dự đoán, gợi ý những sản phẩm mà người dùng tiềm năng sẽ quan tâm dựa trên dữ liệu lịch sử, hành vi tương tác, đặc điểm cá nhân hoặc bối cảnh truy cập. Trong TMĐT, RS không chỉ giúp tăng tỷ lệ chuyển đổi, giá trị giỏ hàng trung bình mà còn là vũ khí giữ chân, tạo dựng lòng trung thành của khách hàng.

Phân loại hệ thống tư vấn:

* Lọc cộng tác (Collaborative Filtering – CF): Dựa vào lịch sử tương tác của nhiều người dùng để gợi ý.
* Lọc nội dung (Content-based Filtering): Gợi ý dựa vào đặc trưng sản phẩm và hồ sơ người dùng.
* Hệ thống lai (Hybrid): Kết hợp các phương pháp trên.
* Các hướng tiếp cận mới: Sử dụng mô hình học sâu, khai thác dữ liệu phi cấu trúc, log hành vi, dữ liệu ngữ nghĩa…

**1.2 Những thách thức trong xây dựng hệ thống tư vấn TMĐT**

* Quy mô dữ liệu lớn, tốc độ cập nhật liên tục.
* Dữ liệu thưa thớt (sparsity), cold-start (user/sản phẩm mới).
* Dữ liệu đa dạng: bảng, văn bản, hình ảnh, log hành vi, dữ liệu vị trí…
* Đa dạng đối tượng người dùng (vùng miền, độ tuổi, giới tính, hành vi, trình độ công nghệ…).
* Nhu cầu cá nhân hóa sâu sắc, đáp ứng thời gian thực.

**1.3 Xu hướng ứng dụng học sâu trong hệ thống tư vấn**

* Khả năng trích xuất đặc trưng tự động từ dữ liệu lớn, phi cấu trúc.
* Học biểu diễn phức tạp, phi tuyến, tích hợp nhiều loại thông tin (deep feature learning).
* Dễ dàng mở rộng với các loại dữ liệu mới (review text, ảnh, metadata).
* Tăng hiệu quả cá nhân hóa, phát hiện những sở thích tiềm ẩn, xu hướng mua hàng mới.

**Kết luận chương 1**:

Chương I đã cung cấp một nền tảng lý thuyết vững chắc về các hệ tư vấn, từ lịch sử phát triển, phân loại các phương pháp cổ điển như lọc cộng tác, lọc dựa trên nội dung, cho tới các xu hướng lai ghép (hybrid). Phần này cũng làm rõ những thách thức lớn trong lĩnh vực TMĐT, như dữ liệu lớn, tính thưa thớt, cold start và sự thay đổi hành vi người dùng. Qua đó, chương này đặt ra cơ sở khoa học cho việc ứng dụng các kỹ thuật hiện đại nhằm nâng cao chất lượng tư vấn cá nhân hóa trên sàn thương mại điện tử.

**Chương 2: Kỹ thuật học sâu cho tư vấn TMĐT**

**2.1. Tổng quan về một số mô hình mạng nơ-ron học sâu**

**2.1.1. Giới thiệu học sâu và vai trò trong TMĐT**

Học sâu (Deep Learning) là một nhánh của học máy (machine learning), sử dụng các mô hình mạng nơ-ron nhiều tầng để học biểu diễn dữ liệu phức tạp và trích xuất các đặc trưng tiềm ẩn tự động mà không cần can thiệp thủ công. Trong lĩnh vực TMĐT, học sâu được ứng dụng rộng rãi để giải quyết các bài toán như phân tích hành vi khách hàng, dự đoán nhu cầu mua sắm, phân loại sản phẩm, nhận diện hình ảnh, xử lý ngôn ngữ tự nhiên và đặc biệt là trong hệ thống tư vấn sản phẩm.

Những ưu điểm vượt trội của các mô hình học sâu trong TMĐT bao gồm:

* Khả năng trích xuất và học các đặc trưng phức tạp, phi tuyến từ dữ liệu lớn.
* Xử lý hiệu quả dữ liệu phi cấu trúc như văn bản đánh giá, mô tả sản phẩm, log hành vi, hình ảnh.
* Học được các quan hệ ngữ cảnh, mối liên hệ ẩn giữa người dùng và sản phẩm mà các mô hình truyền thống khó phát hiện.
* Tăng cường khả năng cá nhân hóa, thích ứng với các tình huống cold-start, dữ liệu mới hoặc hành vi người dùng bất thường.

**2.1.2. Mạng nơ-ron sâu truyền thống (Deep Neural Network - DNN)**

DNN là nền tảng của nhiều ứng dụng học sâu hiện đại, gồm nhiều lớp ẩn (hidden layers) kết nối đầy đủ (fully connected). DNN cho phép mô hình hóa các quan hệ phi tuyến giữa đầu vào và đầu ra, giúp cải thiện độ chính xác và năng lực tổng quát hóa trong các bài toán dự báo, phân loại, ranking sản phẩm, gợi ý sản phẩm.

**Ưu điểm:**

* Tự động trích xuất đặc trưng từ dữ liệu đầu vào.
* Học được các mối quan hệ phức tạp, không tuyến tính.
* Dễ dàng kết hợp nhiều loại đặc trưng (numeric, categorical, text...).

**Hạn chế:**

* Cần lượng dữ liệu lớn để huấn luyện hiệu quả.
* Đôi khi khó giải thích kết quả dự đoán (black box).

**2.1.3. Mạng nơ-ron rộng & sâu (Wide & Deep Neural Network)**

Wide & Deep là kiến trúc kết hợp giữa hai hướng tiếp cận:

* **Nhánh Wide:** Sử dụng các đặc trưng rời rạc, one-hot hoặc cross-feature, học ghi nhớ các quy luật mạnh (memorization) dựa trên dữ liệu lịch sử.
* **Nhánh Deep:** Dùng các lớp embedding và mạng nơ-ron sâu để học tổng quát hóa các quan hệ phi tuyến, trích xuất đặc trưng tự động và phát hiện các tương tác phức tạp giữa các đặc trưng.

**Sự kết hợp này giúp mô hình vừa tận dụng được các đặc trưng phổ biến, vừa phát hiện các mối liên hệ hiếm gặp, từ đó cải thiện đáng kể hiệu quả dự đoán trong hệ thống tư vấn sản phẩm.**

**Ví dụ thực tiễn:**

* Google Play, Youtube, Amazon đều ứng dụng Wide & Deep để gợi ý sản phẩm, video, tối ưu hóa quảng cáo.
* Trên TMĐT Việt Nam, mô hình này phù hợp với các bài toán cá nhân hóa, gợi ý sản phẩm theo hành vi người dùng, tăng tỷ lệ chuyển đổi.

**2.1.4. Mạng nơ-ron hồi quy (RNN) và các biến thể**

RNN và các biến thể như LSTM, GRU là các mô hình phù hợp cho dữ liệu tuần tự, chuỗi thời gian, phiên làm việc (session-based recommendation).

* **Ưu điểm:** Ghi nhớ ngữ cảnh, lịch sử hành vi người dùng, dự đoán hành động tiếp theo hiệu quả hơn.
* **Ứng dụng:** Dự đoán sản phẩm tiếp theo trong chuỗi truy cập, phân tích log truy cập, gợi ý realtime cho user không đăng nhập.

**2.1.5. Transformer và Attention**

Transformer là kiến trúc dựa trên cơ chế attention, giúp học mối quan hệ giữa các sự kiện bất kể vị trí xa gần trong chuỗi. Đặc biệt hiệu quả với dữ liệu lớn, phiên truy cập dài, log hành vi đa chiều.

* **Ưu điểm:**
	+ Khả năng song song hóa khi huấn luyện trên dữ liệu lớn.
	+ Học tốt mối quan hệ phức tạp, cập nhật liên tục.
* **Ứng dụng:** SASRec, BERT4Rec, NARM… giúp gợi ý sản phẩm tiếp theo theo session hoặc trên toàn bộ lịch sử truy cập.

**2.1.6. Autoencoder**

Autoencoder dùng để học biểu diễn ẩn của dữ liệu, phục hồi dữ liệu thiếu, hoàn thiện ma trận user-item, tăng khả năng khái quát hóa mô hình gợi ý.

**2.2. Phân tích bài toán tư vấn sản phẩm trên sàn TMĐT Postmart**

**2.2.1. Định nghĩa bài toán**

Trên sàn TMĐT Postmart, mỗi ngày phát sinh hàng trăm nghìn lượt truy cập, mua sắm, đánh giá và phản hồi từ phía khách hàng. Dữ liệu ghi nhận tại nhiều bảng khác nhau:

* **ds\_goods:** Thông tin sản phẩm (tên, giá, thương hiệu, số lượt click, số lượt bán, mô tả, thuộc tính, v.v.).
* **ds\_order:** Thông tin đơn hàng (ai mua, thời gian, tổng giá trị, trạng thái…).
* **ds\_ordergoods:** Sản phẩm chi tiết trong từng đơn hàng.
* **ds\_evaluategoods:** Nội dung đánh giá, điểm số sao của sản phẩm từ khách hàng.
* **ds\_member:** Thông tin người dùng (ID, họ tên, vùng miền, hành vi truy cập…).

**Bài toán đặt ra:** Dựa trên các thông tin lịch sử về hành vi người dùng, sản phẩm, đánh giá, xây dựng hệ thống gợi ý danh sách sản phẩm cá nhân hóa nhất cho từng user tại mỗi thời điểm, nhằm:

* Tăng tỷ lệ chuyển đổi, doanh thu trên mỗi khách hàng.
* Giữ chân khách hàng, nâng cao sự hài lòng.
* Phát hiện và gợi ý các sản phẩm mới phù hợp nhu cầu từng user.

**2.2.2. Phân tích luồng dữ liệu và ý nghĩa từng bảng**

* **Bảng ds\_goods:** Cung cấp profile sản phẩm, dữ liệu đầu vào cho embedding, sinh đặc trưng.
* **Bảng ds\_order:** Lịch sử mua sắm, xác định phiên, xu hướng cá nhân hóa.
* **Bảng ds\_ordergoods:** Xác định hành vi đồng mua (co-purchase), xây dựng context cho các thuật toán “bạn thường mua cùng”.
* **Bảng ds\_evaluategoods:** Đánh giá mức độ hài lòng, tạo feature cho ranking.
* **Bảng ds\_member:** Xác định các đặc trưng người dùng (giới, tuổi, vùng…), phục vụ cá nhân hóa sâu hơn.

**2.2.3. Đặc thù dữ liệu Postmart trong bài toán gợi ý**

* Số lượng sản phẩm lớn, nhiều nhóm hàng tương tự nhau, dữ liệu rời rạc, đa dạng.
* Nhiều người dùng không đăng nhập (user anonymous), cần gợi ý theo session.
* Hành vi đa dạng: click, xem, mua, đánh giá, yêu thích.
* Dữ liệu cập nhật liên tục, cần mô hình linh hoạt, dễ cập nhật.

**2.3. Giới thiệu Wide & Deep Neural Network cho Postmart**

**2.3.1. Lý do lựa chọn mô hình Wide & Deep**

Qua phân tích các mô hình hiện đại và dữ liệu thực tế của Postmart, Wide & Deep Neural Network là lựa chọn phù hợp nhất, do:

* **Khả năng kết hợp ghi nhớ các quy luật mạnh (tuyến tính) và học phi tuyến sâu rộng:**
	+ Nhánh wide tận dụng các đặc trưng rời rạc, giúp tận dụng quy luật kinh nghiệm, các mối liên hệ phổ biến (ví dụ: user X từng mua nhóm hàng Y rất nhiều lần…).
	+ Nhánh deep học các quan hệ phi tuyến, tương tác giữa các đặc trưng khó mô tả thủ công (ví dụ: kết hợp hành vi click, đánh giá, giá, mô tả sản phẩm…).
* **Dễ mở rộng và phù hợp dữ liệu TMĐT:**
	+ Có thể bổ sung thêm feature mới mà không cần xây dựng lại toàn bộ mô hình.
	+ Phù hợp với dữ liệu hành vi lớn, rời rạc, đa chiều của TMĐT Postmart.
* **Hiệu quả đã được chứng minh trên thực tế:**
	+ Google, Amazon đã triển khai rất thành công Wide & Deep trong hệ thống gợi ý.
	+ Các thử nghiệm sơ bộ cho thấy Wide & Deep hội tụ nhanh, cho kết quả ổn định và dễ triển khai trên hạ tầng hiện có.

**2.3.2. Phân tích ưu điểm, nhược điểm và khuyến nghị**

**Ưu điểm:**

* Khả năng cá nhân hóa tốt.
* Kết hợp giữa thuộc tính rời rạc và liên tục, học được cả quy tắc thủ công lẫn biểu diễn phi tuyến.
* Hiệu suất cao, dễ dàng mở rộng quy mô.

**Nhược điểm:**

* Nếu đặc trưng đầu vào chưa đủ phong phú, hiệu quả sẽ bị hạn chế.
* Phụ thuộc nhiều vào chất lượng dữ liệu đầu vào, đặc biệt là khâu feature engineering.

**Khuyến nghị:**
Wide & Deep là giải pháp phù hợp cho TMĐT Postmart trong giai đoạn hiện tại, vừa đảm bảo hiệu quả, vừa đảm bảo tính mở rộng khi cần phát triển các hệ thống tư vấn phức tạp hơn trong tương lai.

**Kết luận chương 2:**

Chương 2 tập trung phân tích các mô hình mạng nơ-ron học sâu nổi bật trong hệ thống tư vấn, gồm Wide & Deep, RNN, Transformer và Autoencoder. Qua phân tích cấu trúc dữ liệu và nghiệp vụ của Postmart, đề án lựa chọn mô hình Wide & Deep Neural Network là giải pháp trọng tâm, nhờ khả năng tận dụng đồng thời các đặc trưng nghiệp vụ tuyến tính và học sâu các quan hệ phi tuyến trong dữ liệu thực tế. Kiến trúc này chứng minh sự phù hợp, tính mở rộng, dễ triển khai và hiệu quả cá nhân hóa trên quy mô lớn, đáp ứng tốt đặc thù của TMĐT Việt Nam.

**Chương 3: Thực nghiệm, đánh giá và đề xuất**

**3.1. Quy trình chuẩn bị và xử lý dữ liệu**

* Làm sạch, chuẩn hóa dữ liệu, xử lý missing, encoding các trường rời rạc.
* Biểu diễn đặc trưng: one-hot, embedding, vector hóa TF-IDF cho các trường văn bản.
* Chia bộ dữ liệu train – test, chuẩn bị pipeline tiền xử lý tự động.

**3.2. Xây dựng mô hình Wide & Deep**

* Kiến trúc mô hình: nhánh wide sử dụng cross feature, nhánh deep sử dụng embedding và các lớp fully connected.
* Tuning tham số: số lớp, kích thước embedding, batch size, learning rate, số epoch.
* Triển khai trên tập dữ liệu Amazon Reviews for Sentiment Analysis.

**3.3. Huấn luyện và đánh giá mô hình**

* Các chỉ số: accuracy, MAE, MSE, loss.
* Đánh giá ảnh hưởng số epoch (5, 50, 100) đến kết quả thực nghiệm.
* Biểu đồ learning curve, loss, phân tích từng giai đoạn huấn luyện.
* Kết quả: mô hình hội tụ ổn định, accuracy đạt từ 0.82–0.87 tùy epoch, MAE, MSE giảm dần khi tăng số epoch hợp lý, tuy nhiên cần tránh overfitting.

**3.4. Phân tích và so sánh**

* So sánh với baseline (linear, logistic regression, random, SVD…): Wide & Deep thể hiện vượt trội về khả năng cá nhân hóa, hiệu quả dự đoán.
* Đề xuất sử dụng Wide & Deep như giải pháp trọng tâm cho sàn TMĐT Postmart trong giai đoạn đầu.

**3.5. Đề xuất hướng phát triển**

* Mở rộng mô hình với dữ liệu thực tế của Postmart (đa nguồn, đa loại hình: log truy cập, hình ảnh, text review…).
* Triển khai các mô hình tiên tiến hơn: Transformer-based, session-based, multi-modal recommender.
* Kết hợp A/B testing trên nền tảng thực tế, tích hợp phản hồi người dùng để tối ưu liên tục.
* Xây dựng hệ thống tư vấn tự động cập nhật real-time, nâng cao cá nhân hóa và khả năng thích nghi.
* Ứng dụng kiến thức vào các lĩnh vực khác như logistics, giáo dục, y tế…

**III. TỔNG KẾT**

**1. Kết quả đạt được**

Đề án đã hoàn thành các mục tiêu nghiên cứu đặt ra, cụ thể:

* **Hệ thống hóa nền tảng lý thuyết** về các phương pháp xây dựng hệ thống tư vấn sản phẩm trên TMĐT, từ mô hình truyền thống đến các kiến trúc mạng nơ-ron học sâu hiện đại.
* **Phân tích sâu đặc thù dữ liệu và nghiệp vụ** trên sàn Postmart, từ đó định nghĩa rõ ràng bài toán tư vấn, đặc điểm dữ liệu, các yếu tố ảnh hưởng đến chất lượng gợi ý.
* **Lựa chọn và triển khai thành công mô hình Wide & Deep Neural Network** cho hệ thống gợi ý sản phẩm, chứng minh khả năng tận dụng đồng thời các đặc trưng tuyến tính và phi tuyến trong dữ liệu lớn, rời rạc, đa chiều.
* **Tiến hành thử nghiệm thực tế** trên bộ dữ liệu chuẩn Amazon Reviews: mô hình đạt hiệu quả tốt với các chỉ số accuracy, MAE, MSE và learning curve hội tụ ổn định, thể hiện rõ tiềm năng ứng dụng trên quy mô lớn.
* **Đưa ra phân tích, đánh giá kết quả** thực nghiệm một cách khách quan, kết nối sát với thực tiễn TMĐT tại Việt Nam.

**2. Những hạn chế của đề án**

Bên cạnh các kết quả tích cực, đề án còn tồn tại một số hạn chế nhất định:

* **Chưa tích hợp các loại dữ liệu đa phương tiện** (hình ảnh, log hành vi thời gian thực, văn bản đánh giá chi tiết…), mới tập trung vào dữ liệu dạng bảng và text đơn giản.
* **Mới triển khai thử nghiệm với mô hình Wide & Deep**; các mô hình nâng cao như Transformer, session-based, GNN… chưa được thử nghiệm thực tế trong phạm vi đề án.
* **Thiếu các thử nghiệm so sánh trên nhiều bộ dữ liệu TMĐT nội địa**, chưa đánh giá được toàn diện khả năng tổng quát hóa của mô hình trên nhiều sàn khác nhau.
* **Chưa thực hiện đánh giá online A/B test trên nền tảng Postmart thực tế**, do giới hạn về dữ liệu truy cập và thời gian thực hiện.

**3. Hướng phát triển tiếp theo**

Để phát huy hiệu quả và khắc phục các hạn chế nêu trên, các hướng phát triển tiếp theo được đề xuất như sau:

* **Tiến hành thử nghiệm mô hình trên dữ liệu thực tế của Postmart**, kết hợp đa dạng nguồn dữ liệu (log hành vi, lịch sử mua sắm, văn bản đánh giá, hình ảnh sản phẩm...) để tăng khả năng cá nhân hóa và nâng cao chất lượng dự đoán.
* **Khảo sát, thử nghiệm và so sánh với các mô hình học sâu hiện đại khác** như Transformer-based recommendation, session-based, Graph Neural Network (GNN), đồng thời tích hợp thêm các kỹ thuật tăng cường học sâu (regularization, data augmentation, meta-learning...).
* **Mở rộng đánh giá trên nhiều tập dữ liệu TMĐT nội địa và quốc tế**, nhằm kiểm tra tính ổn định, khả năng tổng quát hóa mô hình ở các điều kiện thị trường khác nhau.
* **Tích hợp hệ thống gợi ý vào quy trình vận hành thực tế của Postmart** (A/B test, online serving), thu thập phản hồi người dùng để tối ưu hóa liên tục và nâng cao trải nghiệm thực tiễn.
* **Nghiên cứu hướng kết hợp đa mô hình** (ensemble, hybrid recommender systems), tích hợp dữ liệu thời gian thực, dữ liệu tương tác đa kênh (omnichannel), hướng đến cá nhân hóa sâu hơn và hỗ trợ phát triển TMĐT thông minh, linh hoạt, đáp ứng nhu cầu phát triển lâu dài của Postmart cũng như TMĐT Việt Nam.