**THUYẾT MINH ĐỀ CƯƠNG DỰ ÁN TỐT NGHIỆP**

**Tên** đề tài: Xây dựng Hệ thống **Gợi ý Địa điểm Ăn uống Cá nhân hóa sử dụng Thuật toán Lai kết hợp Phân tích Đánh giá và Yếu tố Vị trí**

**Chương 1: Giới thiệu**

**1.1. Bối cảnh và Đặt vấn đề**

Trong bối cảnh đô thị hóa nhanh chóng, ngành dịch vụ ẩm thực (F&B) tại Việt Nam đã bùng nổ với hàng ngàn nhà hàng, quán cà phê mới ra đời mỗi năm. Các nền tảng đánh giá như Foody.vn hay Google Maps cung cấp một lượng thông tin khổng lồ, tuy nhiên lại đặt người dùng vào tình thế quá tải thông tin. Việc quyết định "hôm nay ăn gì, ở đâu?" trở nên khó khăn hơn bao giờ hết. Các hệ thống gợi ý hiện tại thường chỉ dựa trên mức độ phổ biến chung hoặc quảng cáo, thiếu đi sự cá nhân hóa sâu sắc dựa trên khẩu vị và ngữ cảnh (vị trí, thời gian) của người dùng.

**1.2. Giải pháp Đề xuất và Mục tiêu Sản phẩm**

Dự án này đề xuất xây dựng một **sản phẩm web application hoàn chỉnh** nhằm giải quyết bài toán trên. Hệ thống sẽ là một "trợ lý ẩm thực" thông minh, áp dụng thuật toán lai để đưa ra những gợi ý địa điểm phù hợp nhất, bằng cách kết hợp:

* **Phân tích Nội dung & Thuộc tính (Content-Based):** Phân tích các đặc điểm của địa điểm như loại hình ẩm thực (VD: đồ nướng, lẩu, món chay), mức giá, khu vực. Đặc biệt, sử dụng kỹ thuật Xử lý Ngôn ngữ Tự nhiên (NLP) cơ bản để trích xuất các từ khóa quan trọng từ bình luận của người dùng (VD: "view đẹp", "phục vụ nhanh").
* **Lọc cộng tác (Collaborative Filtering):** "Học hỏi" từ đánh giá của những người dùng có cùng "khẩu vị" để tìm ra những địa điểm mà người dùng có thể sẽ thích.
* **Yếu tố Vị trí (Location-aware):** Ưu tiên gợi ý các địa điểm ở gần vị trí hiện tại của người dùng.

**Sản phẩm cuối cùng** là một website demo tương tác, giúp người dùng dễ dàng khám phá các địa điểm ăn uống phù hợp với nhu cầu tức thời.

**Chương 2: Mục tiêu và Phạm vi**

**2.1. Mục tiêu**

* **Mục tiêu về Sản phẩm:** Xây dựng thành công một ứng dụng web cho phép người dùng tìm kiếm, xem chi tiết, đánh giá địa điểm và nhận được danh sách gợi ý "Hôm nay ăn gì?" được cá nhân hóa và phù hợp ngữ cảnh.
* **Mục tiêu về Học thuật:**
  1. Nghiên cứu và triển khai các mô hình gợi ý: Dựa trên thuộc tính & NLP, Lọc cộng tác, và mô hình Lai có tích hợp yếu tố vị trí.
  2. Đánh giá và so sánh hiệu năng các mô hình thông qua các độ đo thực nghiệm (Precision@k, MAP, NDCG).

**2.2. Phạm vi**

* **Dữ liệu:** Tập trung vào dữ liệu địa điểm ăn uống (thông tin cơ bản và bình luận) tại một khu vực địa lý cụ thể (ví dụ: TP. Hồ Chí Minh), được thu thập từ Foody.vn hoặc Google Maps.
* **Sản phẩm:** Là một sản phẩm Proof-of-Concept (POC) để chứng minh tính khả thi và hiệu quả của giải pháp.

**Chương 3: Phân tích và Thiết kế Hệ thống (Sản phẩm)**

**3.1. Kiến trúc Hệ thống**

Hệ thống được thiết kế theo kiến trúc 3 lớp: Tầng Dữ liệu, Tầng Logic (chứa các model và API), và Tầng Trình diễn (giao diện web).

**3.2. Thiết kế Chức năng và Giao diện Sản phẩm**

Sản phẩm web sẽ bao gồm các chức năng chính:

1. **Trang Khám phá:** Hiển thị bản đồ tương tác với các địa điểm nổi bật. Cung cấp bộ lọc theo quận, loại món, mức giá.
2. **Trang Chi tiết Địa điểm:**
   * Hiển thị thông tin: hình ảnh, địa chỉ, thực đơn, giờ mở cửa.
   * Tóm tắt các từ khóa được nhắc đến nhiều nhất trong bình luận (VD: "không gian rộng", "giá hợp lý").
   * Cho phép người dùng **đánh giá 5 sao** và để lại bình luận.
3. **Trang "Hôm nay ăn gì?" (Core Feature):**
   * Giao diện chính để người dùng nhận gợi ý.
   * Hệ thống tự động đề xuất danh sách các địa điểm dựa trên lịch sử, sở thích và vị trí của người dùng.
   * Mỗi gợi ý có thể kèm giải thích: "Gợi ý cho bạn vì bạn thích 'Đồ nướng' và địa điểm này gần bạn".

**3.3. Công nghệ sử dụng**

* **Thu thập dữ liệu:** Python (Scrapy, BeautifulSoup).
* **Phân tích & Model:** Python (Pandas, Scikit-learn, Surprise, NLTK/PhoNLP).
* **Backend & API:** Python (Flask, FastAPI).
* **Frontend (Sản phẩm Web):** **Streamlit** (Khuyến nghị), có thể tích hợp thư viện bản đồ như Folium.
* **Cơ sở dữ liệu:** SQLite hoặc PostgreSQL (nếu cần xử lý truy vấn không gian).

**Chương 4: Kế hoạch Triển khai**

Dự án được chia thành 4 giai đoạn chính trong 16 tuần:

* **Giai đoạn 1 (Tuần 1-4):** Nghiên cứu & Thu thập Dữ liệu.
  + Xây dựng script cào dữ liệu địa điểm và bình luận.
  + Tiền xử lý và làm sạch dữ liệu.
* **Giai** đoạn **2 (Tuần 5-8):** Xây dựng các mô hình độc lập.
  + Xây dựng mô hình Content-Based (dựa trên thuộc tính và từ khóa NLP).
  + Xây dựng mô hình Collaborative Filtering (dựa trên điểm đánh giá).
* **Giai đoạn 3 (Tuần 9-12):** Xây dựng mô hình Lai và Backend.
  + Kết hợp các mô hình trên, tích hợp logic xử lý vị trí.
  + Xây dựng API để cung cấp kết quả gợi ý.
* **Giai** đoạn **4 (Tuần 13-16):** Hoàn thiện Sản phẩm Web và Báo cáo.
  + Xây dựng giao diện web bằng Streamlit, tích hợp bản đồ.
  + Kết nối giao diện với API backend, kiểm thử toàn diện.
  + Viết báo cáo tốt nghiệp.

**Chương 5: Kết quả Dự kiến**

1. **Sản phẩm Chính:** Một ứng dụng web demo hoạt động, có khả năng gợi ý địa điểm ăn uống thông minh, cá nhân hóa và phù hợp với ngữ cảnh.
2. **Sản** phẩm **Phụ:**
   * Một bộ dữ liệu về các địa điểm F&B tại TP.HCM.
   * Các file mô hình đã