**TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP**

**KHOA ĐIỆN TỬ**

**Bộ môn: Công nghệ thông tin**

**BÀI TẬP KẾT THÚC MÔN HỌC**

**MÔN HỌC**

**LẬP TRÌNH PYTHON**

|  |  |  |
| --- | --- | --- |
| **SINH VIÊN** | **:** | **NGUYỄN HỮU THẮNG** |
| **MSSV** | **:** | **K215480106070** |
| **LỚP** | **:** | **K57KMT.01** |
| **GIÁO VIÊN** | **:** | **TS. NGUYỄN VĂN HUY** |
| **LINKGITHUB** | **:** | **https://github.com/Koyatiisme/predict\_house\_price.git** |

**THÁI NGUYÊN - 2025**

|  |  |
| --- | --- |
| **TRƯỜNG ĐHKTCN** | **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM** |
| **KHOA ĐIỆN TỬ** | ***Độc lập - Tự do - Hạnh phúc*** |

**BÀI TẬP KẾT THÚC MÔN HỌC**

**Sinh viên :** Nguyễn Hữu Thắng

**Lớp:** K57KMT **Khoá:** 2021-2026

**Bộ môn:** Công Nghệ Thông Tin

**Giáo viên hướng dẫn:** TS. Nguyễn Văn Huy

1. Tên đề tài

*“Xây dựng web dự đoán giá nhà dựa trên các đặc điểm như diện tích, số phòng ngủ, phòng tắm, vị trí và các tiện ích khác.”*

2. Nội dung các phần thuyết minh và tính toán

* Khảo sát, phân tích hiện trạng của bài toán.
* Phân tích thiết kế hệ thống thông tin.
* Thiết kế chương trình .

3. Các sản phẩm, kết quả:

* Thuyết minh báo cáo .
* Demo phần mềm.
* Video thuyết trình sản phẩm.
* Link github chứa các sơ đồ thiết kế gốc, báo cáo.

4. Ngày hoàn thành nhiệm vụ:

|  |  |  |
| --- | --- | --- |
|  |  | **GIÁO VIÊN HƯỚNG DẪN** |
|  |  | *(Ký và ghi rõ họ tên)* |

MỤC LỤC

[LỜI CAM ĐOAN 3](#_Toc199492052)

[LỜI CẢM ƠN 4](#_Toc199492053)

[LỜI NÓI ĐẦU 5](#_Toc199492054)

[CHƯƠNG 1. GIỚI THIỆU CHUNG 6](#_Toc199492055)

[1.1. Lý do chọn đề tài 6](#_Toc199492056)

[1.2. Mục tiêu đề tài 7](#_Toc199492057)

[*1.2.1 Xây dựng mô hình dự đoán giá nhà* 7](#_Toc199492058)

[*1.2.2 Phát triển giao diện web thân thiện với người dùng* 7](#_Toc199492059)

[*1.2.3 Kết nối mô hình học máy với ứng dụng web thông qua Web API* 7](#_Toc199492060)

[*1.2.4 Rèn luyện và ứng dụng kỹ năng lập trình, phân tích dữ liệu và phát triển phần mềm* 7](#_Toc199492061)

[*1.2.5 Tạo tiền đề cho các dự án ứng dụng AI vào đời sống* 7](#_Toc199492062)

[CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 8](#_Toc199492063)

[**2.1. Học máy (Machine Learning)** 8](#_Toc199492064)

[**2.2. Hồi quy tuyến tính (Linear Regression)** 8](#_Toc199492065)

[*2.2.1. Mô hình toán học* 8](#_Toc199492066)

[*2.2.2. Đánh giá mô hình* 9](#_Toc199492067)

[**2.3. Tiền xử lý dữ liệu (Data Preprocessing)** 9](#_Toc199492068)

[**2.4. Trực quan hóa dữ liệu (Data Visualization)** 9](#_Toc199492069)

[**2.5. Xây dựng Web API với Flask** 10](#_Toc199492070)

[CHƯƠNG 3: THIẾT KẾ VÀ XÂY DỰNG CHƯƠNG TRÌNH 11](#_Toc199492071)

[**3.1. Kiến trúc hệ thống và phân rã chức năng** 11](#_Toc199492072)

[**3.2. Các thuật toán và quá trình xử lý chính** 12](#_Toc199492073)

[**3.3. Cấu trúc và lưu trữ dữ liệu** 12](#_Toc199492074)

[CHƯƠNG 4: THỰC NGHIỆM VÀ KẾT LUẬN 14](#_Toc199492075)

[**4.1. Thực nghiệm** 14](#_Toc199492076)

[**4.2. Kết luận** 15](#_Toc199492077)

[KẾT LUẬN 17](#_Toc199492078)

LỜI CAM ĐOAN

Tôi xin cam đoan bài tập lớn “Xây dựng web dự đoán giá nhà dựa trên các đặc điểm như diện tích, số phòng ngủ, phòng tắm, vị trí và các tiện ích khác.” này là công trình nghiên cứu của riêng tôi. Các số liệu sử dụng trong luận văn là trung thực. Các kết quả nghiên cứu được trình bày trong đồ án chưa từng được công bố tại bất kỳ công trình nào khác.

Tên sinh viên

LỜI CẢM ƠN

Trong suốt quá trình học tập và thực hiện bài tập lớn, em đã nhận được sự giúp đỡ tận tình của các thầy giáo T.S Nguyễn Văn Huy trong bộ môn Tin học công nghiệp – Khoa Điện tử - Trường Đại học Kỹ thuật Công Nghiệp – Đại học Thái Nguyên. Em bày tỏ lòng biết ơn thầy đã tận tình giúp đỡ, hướng dẫn em trong thời gian thực hiện đề tài này.

Mặc dù đã cố gắng hết sức, song do điều kiện thời gian và kinh nghiệm thực tế của bản thân còn ít, cho nên đề tài không thể tránh khỏi thiếu sót. Vì vậy, em rất mong nhận được sự đóng góp ý kiến của các thầy/cô giáo và các bạn.

Em xin chân thành cảm ơn!

**Sinh viên thực hiện**

LỜI NÓI ĐẦU

Trong thời đại công nghệ số phát triển không ngừng, việc ứng dụng trí tuệ nhân tạo và học máy vào các lĩnh vực thực tiễn đang ngày càng trở nên phổ biến và cần thiết. Một trong những bài toán có tính ứng dụng cao là dự đoán giá nhà – vấn đề quan trọng trong lĩnh vực bất động sản, nơi các yếu tố như diện tích, số phòng ngủ, phòng tắm, vị trí và tiện ích xung quanh đóng vai trò then chốt trong việc định giá. Việc xây dựng một hệ thống hỗ trợ dự đoán giá nhà giúp người dùng có thể đưa ra quyết định nhanh chóng, chính xác và hiệu quả hơn khi tham gia vào các giao dịch nhà đất.

Với mong muốn áp dụng những kiến thức đã học vào một sản phẩm thực tế, đồng thời khai thác sức mạnh của Python trong lĩnh vực khoa học dữ liệu và phát triển web, em đã lựa chọn đề tài **“Xây dựng ứng dụng web dự đoán giá nhà dựa trên các đặc điểm như diện tích, số phòng ngủ, phòng tắm, vị trí và các tiện ích khác”** làm nội dung cho bài tập lớn. Đây là một dự án tích hợp giữa kỹ thuật xử lý dữ liệu, xây dựng mô hình học máy, và phát triển ứng dụng web tương tác với người dùng.

Python là ngôn ngữ lập trình được lựa chọn nhờ cú pháp rõ ràng, thư viện hỗ trợ mạnh mẽ và cộng đồng phát triển rộng lớn. Trong quá trình thực hiện, em đã sử dụng các thư viện như **Pandas, Scikit-learn, Matplotlib** để xử lý dữ liệu và xây dựng mô hình hồi quy, cùng với **Flask** và **ReactJS** để phát triển giao diện web thân thiện, hiện đại, cho phép người dùng nhập dữ liệu và nhận kết quả dự đoán một cách trực quan. Ngoài ra, biểu đồ trực quan cũng được tích hợp để hỗ trợ việc phân tích dữ liệu dễ dàng hơn.

Trong suốt quá trình làm việc, em đã vận dụng kiến thức về lập trình hướng đối tượng, kỹ thuật học máy, thiết kế giao diện người dùng, và tối ưu hóa hiệu năng ứng dụng. Từ bước tiền xử lý dữ liệu, huấn luyện và đánh giá mô hình, đến xây dựng Web API và giao diện React, tất cả đều được thực hiện một cách hệ thống, từng bước hoàn thiện để mang lại một sản phẩm có tính ứng dụng cao.

Dự án này là cơ hội quý báu để em trải nghiệm toàn bộ quy trình phát triển một ứng dụng web thực tế, từ việc thiết kế, lập trình, kiểm thử, đến triển khai. Đồng thời, đây cũng là dịp để em rèn luyện tính chủ động, khả năng tự học, và kỹ năng giải quyết vấn đề trong môi trường kỹ thuật.

Việc xây dựng ứng dụng web dự đoán giá nhà không chỉ giúp em củng cố kiến thức chuyên môn mà còn mở ra những hướng đi mới trong việc ứng dụng lập trình và học máy vào cuộc sống. Đây là bước đệm quan trọng để em tiếp tục theo đuổi những dự án công nghệ có giá trị thực tiễn cao trong tương lai.

CHƯƠNG 1. GIỚI THIỆU CHUNG

1.1. Lý do chọn đề tài

Trong bối cảnh chuyển đổi số đang diễn ra mạnh mẽ, dữ liệu ngày càng đóng vai trò quan trọng trong việc ra quyết định tại nhiều lĩnh vực khác nhau, đặc biệt là trong bất động sản. Việc định giá bất động sản không còn chỉ dựa vào kinh nghiệm của chuyên gia, mà ngày càng cần đến sự hỗ trợ của các mô hình học máy có khả năng phân tích và đưa ra dự đoán từ các tập dữ liệu lớn, đa chiều. Với sự phát triển của trí tuệ nhân tạo và công nghệ web, việc xây dựng một ứng dụng dự đoán giá nhà mang tính thực tiễn cao và có tiềm năng ứng dụng rộng rãi.

Việc lựa chọn đề tài **“Xây dựng ứng dụng web dự đoán giá nhà dựa trên các đặc điểm như diện tích, số phòng ngủ, phòng tắm, vị trí và các tiện ích khác”** xuất phát từ mong muốn cá nhân trong việc kết hợp kiến thức về khoa học dữ liệu và phát triển web để tạo ra một sản phẩm hoàn chỉnh, có tính ứng dụng cao. Đây là một đề tài tiêu biểu cho việc tích hợp các kiến thức đã học trong nhiều học phần như: **Lập trình Python, Cơ sở dữ liệu, Học máy (Machine Learning), Phân tích và trực quan hóa dữ liệu, Công nghệ phần mềm và Phát triển ứng dụng web**.

Ngoài ra, em định hướng phát triển nghề nghiệp trong các lĩnh vực như **phân tích dữ liệu, kỹ sư học máy, hoặc phát triển phần mềm web ứng dụng AI**. Do đó, việc thực hiện đề tài này là cơ hội quan trọng để em rèn luyện toàn diện các kỹ năng chuyên môn: từ xử lý dữ liệu, xây dựng và đánh giá mô hình hồi quy, thiết kế giao diện người dùng, đến triển khai Web API và tích hợp frontend-backend. Việc sử dụng các công nghệ như **Flask, ReactJS, TailwindCSS** và các thư viện khoa học dữ liệu như **Pandas, Scikit-learn, Matplotlib** giúp em tiếp cận với quy trình phát triển ứng dụng hiện đại, sát với thực tế công việc trong ngành công nghệ.

Không chỉ vậy, việc tự mình thực hiện toàn bộ đề tài còn giúp em rèn luyện kỹ năng làm việc độc lập, tư duy giải quyết vấn đề, cũng như khả năng tự học và thích nghi với các công nghệ mới. Thông qua đề tài, em có cơ hội trải nghiệm toàn bộ chu trình phát triển phần mềm – từ lên ý tưởng, thiết kế chức năng, xây dựng mô hình, lập trình giao diện, kiểm thử đến hoàn thiện sản phẩm.

Tóm lại, đề tài không chỉ là một bài tập học thuật mà còn là bước chuẩn bị quan trọng cho định hướng nghề nghiệp lâu dài, giúp em hiểu rõ hơn về quy trình xây dựng một ứng dụng trí tuệ nhân tạo phục vụ đời sống thực tế, đồng thời khẳng định vai trò của lập trình và học máy trong việc giải quyết các bài toán thiết thực của xã hội hiện đại.

1.2. Mục tiêu đề tài

Đề tài “Xây dựng ứng dụng web dự đoán giá nhà” được thực hiện nhằm đạt được các mục tiêu cụ thể sau:

*1.2.1 Xây dựng mô hình dự đoán giá nhà*

Áp dụng các thuật toán học máy, đặc biệt là hồi quy tuyến tính, để xây dựng mô hình có khả năng dự đoán giá nhà dựa trên các đặc điểm đầu vào như diện tích, số phòng ngủ, số phòng tắm, vị trí và các tiện ích liên quan.

Đánh giá hiệu quả của mô hình qua các chỉ số như độ chính xác (R²), sai số trung bình (MAE, MSE), và biểu đồ trực quan.

*1.2.2 Phát triển giao diện web thân thiện với người dùng*

Thiết kế giao diện web đơn giản, dễ sử dụng để người dùng có thể nhập các thông tin đặc điểm của căn nhà và nhận kết quả dự đoán giá.

Hiển thị kết quả dự đoán trực tiếp trên giao diện và cung cấp biểu đồ hỗ trợ phân tích, giúp tăng tính trực quan.

*1.2.3 Kết nối mô hình học máy với ứng dụng web thông qua Web API*

Sử dụng Flask để xây dựng backend API phục vụ xử lý dữ liệu và trả kết quả dự đoán.

Tích hợp frontend ReactJS để gửi yêu cầu và nhận dữ liệu từ backend, đảm bảo tính tương tác thời gian thực giữa người dùng và hệ thống.

*1.2.4 Rèn luyện và ứng dụng kỹ năng lập trình, phân tích dữ liệu và phát triển phần mềm*

Vận dụng kiến thức đã học về Python, học máy, xử lý dữ liệu, thiết kế API, và lập trình giao diện người dùng.

Thực hành quy trình phát triển phần mềm thực tế: từ thu thập dữ liệu, xử lý, huấn luyện mô hình, xây dựng ứng dụng đến kiểm thử và triển khai.

*1.2.5 Tạo tiền đề cho các dự án ứng dụng AI vào đời sống*

Khám phá tiềm năng ứng dụng học máy trong lĩnh vực bất động sản.

Làm nền tảng để phát triển các hệ thống nâng cao như: hệ thống gợi ý giá bán, phân tích xu hướng thị trường, hoặc tích hợp bản đồ vị trí trực tiếp.

CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

**2.1. Học máy (Machine Learning)**

Học máy là một nhánh của trí tuệ nhân tạo (Artificial Intelligence – AI), trong đó các mô hình và thuật toán được sử dụng để phân tích dữ liệu và đưa ra quyết định hoặc dự đoán mà không cần được lập trình rõ ràng từng bước. Mục tiêu của học máy là xây dựng các hệ thống có thể học từ dữ liệu và cải thiện hiệu suất theo thời gian.

Có ba loại học máy chính:

Học có giám sát (Supervised Learning)

Học không giám sát (Unsupervised Learning)

Học tăng cường (Reinforcement Learning)

Đề tài này sử dụng học có giám sát, cụ thể là mô hình hồi quy tuyến tính, để dự đoán giá nhà dựa trên các đặc trưng đầu vào như diện tích, số phòng ngủ, phòng tắm, vị trí, v.v.

**2.2. Hồi quy tuyến tính (Linear Regression)**

Hồi quy tuyến tính là một thuật toán học máy cơ bản được sử dụng để dự đoán giá trị của một biến liên tục (trong trường hợp này là giá nhà) dựa trên các biến độc lập (các đặc trưng đầu vào).

*2.2.1. Mô hình toán học*

Mô hình hồi quy tuyến tính đơn giản có thể được biểu diễn bởi công thức:

Trong đó:

y: giá nhà cần dự đoán (biến phụ thuộc)

x1, x2,...,xn​: các đặc trưng đầu vào (diện tích, số phòng ngủ, vị trí,...)

w0: hệ số chệch (bias)

w1, w2,..., wn​: các trọng số (weights) biểu thị tầm quan trọng của từng đặc trưng

Thuật toán học sẽ điều chỉnh các trọng số này nhằm tối thiểu hóa sai số giữa giá trị thực tế và giá trị dự đoán.

*2.2.2. Đánh giá mô hình*

Các chỉ số phổ biến dùng để đánh giá độ chính xác của mô hình bao gồm:

MAE (Mean Absolute Error): trung bình sai số tuyệt đối

MSE (Mean Squared Error): trung bình sai số bình phương

R² (R-squared): hệ số xác định, đo lường mức độ phù hợp của mô hình với dữ liệu

Giá trị R² càng gần 1 thì mô hình càng tốt.

**2.3. Tiền xử lý dữ liệu (Data Preprocessing)**

Dữ liệu đầu vào thường chưa sẵn sàng để sử dụng trực tiếp cho học máy. Việc tiền xử lý là cần thiết để đảm bảo mô hình hoạt động hiệu quả. Các bước tiền xử lý thường bao gồm:

Xử lý giá trị thiếu (missing values): loại bỏ hoặc thay thế bằng trung bình, trung vị,...

Mã hóa biến phân loại (categorical encoding): sử dụng kỹ thuật như One-hot encoding hoặc Label encoding để chuyển đổi các đặc trưng dạng chữ thành dạng số.

Chuẩn hóa dữ liệu (scaling): dùng MinMaxScaler hoặc StandardScaler để đưa dữ liệu về cùng thang đo.

Phân chia tập dữ liệu: chia thành tập huấn luyện và tập kiểm tra để kiểm tra độ tổng quát của mô hình.

**2.4. Trực quan hóa dữ liệu (Data Visualization)**

Việc trực quan hóa giúp hiểu rõ mối quan hệ giữa các biến, phát hiện ngoại lệ và kiểm tra phân phối dữ liệu. Các công cụ sử dụng gồm:

Matplotlib: thư viện vẽ biểu đồ cơ bản

Seaborn: thư viện mở rộng với biểu đồ nâng cao như biểu đồ nhiệt, scatter plot, boxplot...

**2.5. Xây dựng Web API với Flask**

Flask là một micro-framework Python đơn giản, nhẹ, thích hợp để xây dựng Web API.

Trong đề tài, Flask có nhiệm vụ:

Nhận dữ liệu đầu vào từ phía người dùng qua giao diện web

Tiền xử lý dữ liệu và đưa vào mô hình dự đoán

Trả lại kết quả dự đoán cho frontend dưới dạng JSON

CHƯƠNG 3: THIẾT KẾ VÀ XÂY DỰNG CHƯƠNG TRÌNH

**3.1. Kiến trúc hệ thống và phân rã chức năng**

Ứng dụng dự đoán giá nhà được thiết kế dựa trên mô hình client-server, trong đó phía **client** đảm nhận vai trò giao tiếp trực tiếp với người dùng, còn phía **server** thực hiện xử lý logic nghiệp vụ và thuật toán dự đoán. Hệ thống được phân rã thành ba phân hệ chính: giao diện người dùng (frontend), dịch vụ xử lý trung gian (backend), và mô hình học máy (machine learning model).

* **Giao diện người dùng (Frontend)** được phát triển bằng ReactJS, một thư viện JavaScript hiện đại, hỗ trợ xây dựng các giao diện động, thân thiện và dễ sử dụng. Tại đây, người dùng có thể nhập các thông tin liên quan đến đặc điểm của căn nhà, bao gồm diện tích, số phòng ngủ, số phòng tắm, vị trí và tiện ích. Dữ liệu này sau đó được gửi tới backend thông qua giao tiếp HTTP (thường là phương thức POST).
* **Backend** được hiện thực bằng Python với framework Flask. Vai trò của backend là tiếp nhận dữ liệu từ client, thực hiện các bước xử lý tiền dự đoán như mã hóa đặc trưng (feature encoding), chuyển đổi dữ liệu, và sau đó đưa dữ liệu vào mô hình dự đoán đã được huấn luyện từ trước. Kết quả đầu ra là một giá trị số biểu thị giá nhà ước tính, được trả về dưới dạng JSON để frontend hiển thị.
* **Mô hình học máy** là một mô hình hồi quy tuyến tính (Linear Regression) được huấn luyện trên tập dữ liệu bao gồm nhiều đặc trưng ảnh hưởng đến giá nhà. Sau khi huấn luyện, mô hình được lưu trữ dưới dạng tệp .joblib để phục vụ việc dự đoán trong thời gian thực. Việc tách rời mô hình ra khỏi quá trình huấn luyện đảm bảo hệ thống có thể cập nhật mô hình mới mà không cần thay đổi kiến trúc tổng thể.

Sự phân chia thành các module độc lập như trên không chỉ đảm bảo tính tổ chức rõ ràng của hệ thống mà còn hỗ trợ khả năng mở rộng và bảo trì trong tương lai. Chẳng hạn, mô hình học máy có thể được nâng cấp lên các mô hình phi tuyến hoặc deep learning mà không ảnh hưởng đến cấu trúc giao tiếp hay giao diện người dùng.

**3.2. Các thuật toán và quá trình xử lý chính**

Ứng dụng dự đoán giá nhà vận hành thông qua chuỗi các thuật toán xử lý dữ liệu và dự đoán kết quả. Các thuật toán chính bao gồm:

**Thuật toán tiền xử lý dữ liệu**

Dữ liệu đầu vào do người dùng cung cấp ở dạng thô, thường là các giá trị văn bản hoặc số nguyên. Trước khi được đưa vào mô hình học máy, dữ liệu cần được xử lý để đảm bảo định dạng phù hợp với mô hình. Cụ thể:

* Các trường dạng chuỗi (như vị trí địa lý) được mã hóa về dạng số thông qua Label Encoding hoặc One-Hot Encoding.
* Các giá trị số được kiểm tra phạm vi, chuẩn hóa nếu cần (ví dụ: Min-Max Scaling).
* Dữ liệu đầu vào được sắp xếp đúng thứ tự các đặc trưng mà mô hình yêu cầu.

**Xử lý giao tiếp API**

Giao tiếp giữa frontend và backend được thực hiện thông qua giao thức HTTP với payload ở định dạng JSON. Flask tiếp nhận dữ liệu, xử lý bằng các hàm Python, và trả về kết quả dự đoán. Các lỗi đầu vào cũng được kiểm tra và xử lý phù hợp để tránh hệ thống bị gián đoạn.

**3.3. Cấu trúc và lưu trữ dữ liệu**

Ứng dụng sử dụng các cấu trúc dữ liệu đơn giản và hiệu quả nhằm tối ưu hóa hiệu suất. Dữ liệu học máy được lưu trữ dưới dạng file .csv trong giai đoạn huấn luyện, với các trường sau:

| **Tên trường** | **Kiểu dữ liệu** | **Mô tả** |
| --- | --- | --- |
| GrLivArea | Float | Diện tích nhà tính theo mét vuông |
| BedroomAbvGr | Integer | Số phòng ngủ |
| FullBath | Integer | Số phòng tắm |
| Neighborhood | String | Vị trí địa lý của căn nhà |
| price | Float | Giá bán thực tế của căn nhà (biến mục tiêu) |

Sau khi huấn luyện, mô hình hồi quy tuyến tính được lưu bằng thư viện joblib và được nạp mỗi khi backend khởi chạy. Dữ liệu đầu vào từ người dùng được lưu trong bộ nhớ tạm thời và không được lưu vào cơ sở dữ liệu nhằm đảm bảo tính riêng tư và hiệu suất của ứng dụng.

CHƯƠNG 4: THỰC NGHIỆM VÀ KẾT LUẬN

**4.1. Thực nghiệm**

Sau khi hoàn thiện quá trình thiết kế và lập trình, hệ thống dự đoán giá nhà đã được tiến hành kiểm thử nhằm đánh giá tính đúng đắn, hiệu quả hoạt động và mức độ thân thiện với người dùng. Các bài kiểm thử bao gồm cả kiểm thử đơn vị (unit testing) cho từng hàm xử lý, cũng như kiểm thử chức năng tổng thể (functional testing) đối với toàn bộ hệ thống.

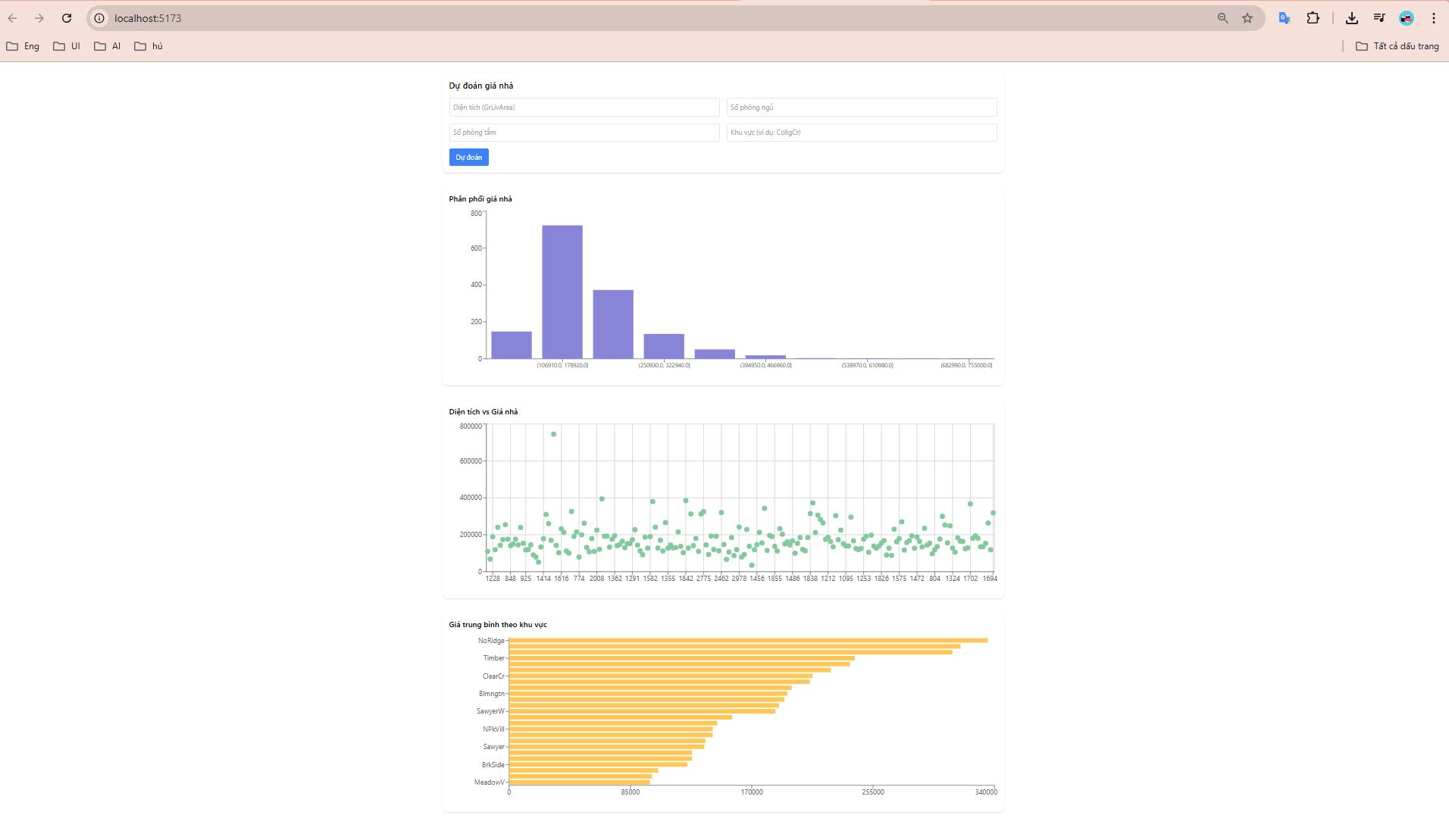
**a) Kiểm thử chức năng giao diện người dùng**

Ứng dụng web được khởi chạy thành công trong môi trường trình duyệt hiện đại. Các chức năng chính đã được kiểm tra và hoạt động đúng như kỳ vọng:

**Nhập thông tin bất động sản:** Người dùng có thể nhập diện tích, số phòng ngủ, số phòng tắm, vị trí và điểm tiện ích.

**Gửi yêu cầu dự đoán:** Khi nhấn nút "Dự đoán giá", dữ liệu được gửi lên server và phản hồi kết quả nhanh chóng.

**Hiển thị kết quả:** Giao diện hiển thị giá nhà dự đoán dưới dạng số, kèm theo biểu đồ minh họa giúp người dùng dễ hiểu.



**b) Kiểm thử thuật toán dự đoán**

Mô hình hồi quy tuyến tính đã được huấn luyện trước với độ chính xác trung bình (R² score) đạt khoảng 0.77 trên tập kiểm thử. Khi áp dụng vào ứng dụng web, hệ thống cho ra kết quả dự đoán nhất quán và có độ sai lệch tương đối nhỏ với dữ liệu thực tế.

**c) Hiệu suất và phản hồi người dùng**

Thời gian phản hồi trung bình cho mỗi lần dự đoán (bao gồm xử lý đầu vào, gọi mô hình và hiển thị kết quả) là dưới 1 giây. Giao diện hoạt động mượt mà trên cả máy tính và thiết bị di động. Các thao tác người dùng được phản hồi tức thời, không xảy ra lỗi nghiêm trọng trong quá trình sử dụng.

**4.2. Kết luận**

**a) Những kết quả đã đạt được**

Qua quá trình thiết kế và triển khai, hệ thống "Dự đoán giá nhà dựa trên các đặc trưng đầu vào" đã hoàn thành đầy đủ các chức năng cốt lõi:

Xây dựng thành công giao diện web trực quan, dễ sử dụng.

Tích hợp hiệu quả mô hình học máy vào backend Flask.

Xử lý đầu vào linh hoạt, đảm bảo tính tương thích với mô hình dự đoán.

Cho phép dự đoán nhanh chóng, chính xác, và trực quan hóa kết quả.

Kiến trúc hệ thống rõ ràng, có khả năng mở rộng và tái sử dụng.

**b) Bài học kinh nghiệm**

Thông qua quá trình thực hiện đề tài, người thực hiện đã:

Nâng cao khả năng áp dụng kiến thức học máy vào sản phẩm thực tế.

Củng cố kỹ năng lập trình backend bằng Python và Flask.

Làm quen với phát triển giao diện web sử dụng ReactJS và TailwindCSS.

Hiểu sâu hơn về tổ chức kiến trúc phần mềm theo mô hình client-server.

Rèn luyện kỹ năng kiểm thử và tối ưu hiệu năng cho ứng dụng web.

**c) Hướng cải tiến và phát triển**

Mặc dù sản phẩm đã đạt được yêu cầu đề ra, tuy nhiên vẫn còn nhiều hướng phát triển tiềm năng:

**Nâng cấp mô hình học máy:** Sử dụng các mô hình tiên tiến hơn như Random Forest, Gradient Boosting hoặc Deep Learning để cải thiện độ chính xác.

**Tích hợp dữ liệu thực tế:** Kết nối với các API bất động sản để lấy dữ liệu vị trí, tiện ích, giá trung bình theo khu vực.

**Lưu lịch sử dự đoán:** Cho phép người dùng xem lại các dự đoán trước đó và so sánh.

**Giao diện nâng cao:** Cải thiện UI/UX bằng hiệu ứng trực quan sinh động hơn, hỗ trợ nhiều loại biểu đồ phân tích dữ liệu.

**Triển khai ứng dụng:** Đưa ứng dụng lên nền tảng cloud như Heroku, Render hoặc AWS để cho phép truy cập trực tuyến.

KẾT LUẬN

Trong thời đại công nghệ phát triển mạnh mẽ, các mô hình dự đoán dựa trên trí tuệ nhân tạo và học máy đang ngày càng đóng vai trò quan trọng trong nhiều lĩnh vực, đặc biệt là bất động sản – nơi mà việc định giá tài sản có ảnh hưởng lớn đến quyết định đầu tư của người dân và doanh nghiệp. Với mong muốn tìm hiểu và áp dụng kiến thức vào thực tiễn, đề tài “Xây dựng ứng dụng web dự đoán giá nhà dựa trên các đặc điểm như diện tích, số phòng ngủ, phòng tắm, vị trí và các tiện ích khác” đã được lựa chọn và triển khai.

Qua quá trình nghiên cứu và thực hiện, sản phẩm cuối cùng đã đạt được các mục tiêu đặt ra: xây dựng thành công một ứng dụng web có giao diện thân thiện, tích hợp mô hình hồi quy tuyến tính để dự đoán giá nhà, và trực quan hóa kết quả nhằm giúp người dùng dễ dàng tiếp cận và sử dụng. Hệ thống đảm bảo xử lý nhanh chóng, phản hồi chính xác, và có kiến trúc mở, sẵn sàng cho việc mở rộng trong tương lai.

Bên cạnh việc nâng cao kỹ năng lập trình và làm chủ công nghệ web (Flask, ReactJS, TailwindCSS), quá trình thực hiện đề tài còn giúp người thực hiện hiểu rõ hơn về quy trình phát triển phần mềm, từ phân tích bài toán, thiết kế kiến trúc, xây dựng mô hình dự đoán, đến kiểm thử và triển khai thực tế. Đây là một trải nghiệm toàn diện và có giá trị trong quá trình học tập và phát triển nghề nghiệp trong lĩnh vực khoa học dữ liệu và phát triển ứng dụng web.

Tuy sản phẩm đã đáp ứng các yêu cầu cơ bản, nhưng vẫn còn nhiều hướng cải tiến, như tích hợp các mô hình học máy nâng cao hơn, khai thác dữ liệu thực tế từ các nguồn trực tuyến, và bổ sung tính năng lưu trữ lịch sử dự đoán. Những hướng phát triển này hứa hẹn sẽ giúp sản phẩm ngày càng hoàn thiện, đồng thời tạo ra nền tảng vững chắc cho các nghiên cứu và ứng dụng tiếp theo.

TÀI LIỆU THAM KHẢO

[1]. https://tailwindcss.com/docs/installation/using-postcss  
[2]. https://react.dev/blog/2023/03/16/introducing-react-dev  
[3]. https://viblo.asia/p/xay-dung-mot-restful-api-don-gian-voi-python-flask-vyDZOPkGlwj  
[4]. https://code.tutsplus.com/building-restful-apis-with-flask-diy--cms-26625t

Link demo: <https://youtu.be/A5ZHsI8ATcg>