

1. Dự đoán doanh số bán hàng bằng mô hình hồi quy tuyến tính (Linear Regression)

- **Mục tiêu:** Xây dựng một mô hình hồi quy tuyến tính đơn giản để dự đoán doanh số bán hàng dựa trên các yếu tố như chi phí quảng cáo, số lượng sản phẩm tồn kho, và số lượng khách hàng tiềm năng.
- **Dữ liệu:** Có thể sử dụng dữ liệu bán hàng sẵn có từ Kaggle hoặc các tập dữ liệu thương mại điện tử công khai.
- **Thực hiện:** Tập trung vào phân tích dữ liệu đơn giản, làm sạch, và xây dựng mô hình hồi quy tuyến tính để dự đoán.

2. Phân loại email thành spam và không spam bằng Logistic Regression

- **Mục tiêu:** Sử dụng mô hình Logistic Regression để phân loại email thành hai nhóm: spam và không spam.
- **Dữ liệu:** Sử dụng tập dữ liệu phân loại email spam, ví dụ như **Spam Email Dataset** từ UCI Machine Learning Repository.
- **Thực hiện:** Sinh viên chỉ cần làm việc với một mô hình phân loại đơn giản, xử lý dữ liệu văn bản cơ bản và áp dụng Logistic Regression.

3. Phát hiện bệnh tiểu đường bằng Decision Tree

- **Mục tiêu:** Sử dụng cây quyết định (Decision Tree) để phát hiện liệu bệnh nhân có khả năng mắc bệnh tiểu đường dựa trên các chỉ số sức khỏe cơ bản.
- **Dữ liệu:** Tập dữ liệu **Pima Indians Diabetes** từ UCI hoặc Kaggle, có các chỉ số sức khỏe như tuổi, cân nặng, huyết áp.
- **Thực hiện:** Xây dựng một mô hình cây quyết định đơn giản để phân loại dự đoán khả năng mắc bệnh.

4. Dự đoán giá nhà bằng hồi quy đa biến (Multiple Linear Regression)

- **Mục tiêu:** Dự đoán giá nhà dựa trên các yếu tố như diện tích, số phòng ngủ, và vị trí.
- **Dữ liệu:** Sử dụng tập dữ liệu giá nhà như **Boston Housing Dataset** từ Scikit-learn hoặc Kaggle.
- **Thực hiện:** Tập trung vào việc xây dựng mô hình hồi quy đa biến và đánh giá hiệu suất của mô hình.

5. Nhận diện chữ viết tay bằng Mạng Nơ-ron Đơn giản (Simple Neural Network)

- **Mục tiêu:** Sử dụng một mạng nơ-ron đơn giản để phân loại chữ viết tay.
- **Dữ liệu:** Sử dụng tập dữ liệu **MNIST** nổi tiếng, chứa hình ảnh các chữ viết tay.

- **Thực hiện:** Sinh viên sẽ xây dựng một mạng nơ-ron đơn giản với 1-2 lớp ẩn để phân loại các chữ viết tay từ ảnh.