

DỰ ÁN

- **Customer Segmentation and API recommend**

Dự đoán chi tiêu và gợi ý sản phẩm

- **Data collection:** Tìm kiếm bộ dữ liệu về phân khúc khách hàng trên tất cả các nền tảng (lớn hơn 25 cột và 50000 dòng).
- **Numpy, Pandas:** Làm sạch dữ liệu, xóa trùng lặp, xử lý giá trị trống, kiểm tra và xóa những dữ liệu có giá trị ngoại lai. Đầu ra là một bộ dữ liệu trải dài từ độ tuổi 18 - 80
- **Data visualization:** Sử dụng thư viện matplotlib, seaborn để trực quan hóa dữ liệu. Phân tích các chỉ số từ biểu đồ và trả lời các câu hỏi như : So sánh lượng tiền lương và chi tiêu? khách ở độ tuổi nào? mặt hàng cần được tiếp thị nhiều hơn? loại hàng nào phù hợp với khách hàng nào?...
- **API:** Đưa dataset vào mô hình Random Forest Regressor, huấn luyện và lấy kết quả từ mô hình. Sử dụng Postman để triển khai API mô hình. Kết quả dự đoán là giá tiền mà khách có thể bỏ ra và gợi ý về những sản phẩm phù hợp với khách hàng.

- **Doanh thu của Nike**

Power BI

- **Mục tiêu:** Tìm hiểu tình trạng bán hàng ở các khu vực. Tìm hiểu các yếu tố dẫn đến ảnh hưởng của doanh thu. Độ hài lòng của khách hàng. Chi phí người mua thường là bao nhiêu, giá thành có làm ảnh hưởng tới tần suất mua hàng hay không.
- **Data visualization:** làm sạch dữ liệu đầu vào bằng Power BI, biểu diễn và trực quan hóa từ dữ liệu đã được làm sạch. Tạo báo cáo và giải đáp các câu hỏi, sử dụng diagnostic analytics để tìm ra nguyên nhân vấn đề.

- **Doanh số vé tàu tại Anh**

SQL

- **Data cleanup:** Ngoài các thông tin về vé tàu, các thông tin khác sai lệch khá nhiều, cần định dạng lại, chuẩn hóa, và xử lý các giá trị bị khuyết. Sau khi làm sạch, đầu ra là 1 file csv đã định dạng đúng yêu cầu để có thể tiến hành trực quan hóa dữ liệu.
- **Power BI:** Số lượng nhân sự phục vụ dịch vụ cần tăng trong khoảng thời gian nào? xu hướng sử dụng dịch vụ theo thời gian. Trong tương lai nên nâng cấp các dịch vụ nào phù hợp với nhu cầu khách hàng? Sử dụng kỹ thuật diagnostic analytics và predictive analysis để chuẩn đoán các tình huống đang xảy ra và trong tương lai

HỌC VẤN

- **Đại học Sài Gòn**

Hồ Chí Minh

Công nghệ thông tin - Khoa học máy tính; GPA: 2.7

9 / 2021 - 2026

KĨ NĂNG (sử dụng quen thuộc)

- **Python (Numpy, Pandas, Seaborn, matplotlib)**
- **SQL**
- **Excel**
- **Power BI**