



PREDICTING CUSTOMER SATISFACTION USING MACHINE LEARNING / DEEP LEARNING

Lê Hồng Sơn
Đỗ Hữu Lộc
Nguyễn Hoàng Thanh Phương
Văn Hoàng Như Ý

DỰ ĐOÁN SỰ HÀI LÒNG CỦA KHÁCH HÀNG BẰNG HỌC MÁY / HỌC SÂU

GIỚI THIỆU

Vấn đề:

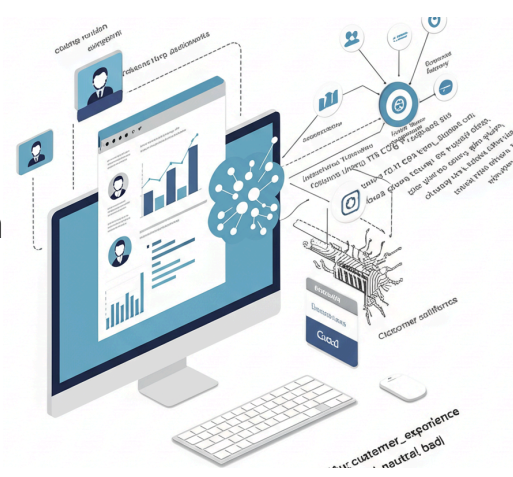
- Input:** Tập dữ liệu được sử dụng trong dự án bao gồm thông tin chi tiết về người dùng, chương trình khách hàng thân thiết, sản phẩm, giao dịch, vận chuyển và trải nghiệm khách hàng.
- Output:** mức độ hài lòng của khách hàng (hài lòng, trung lập, không hài lòng)

Mục tiêu:

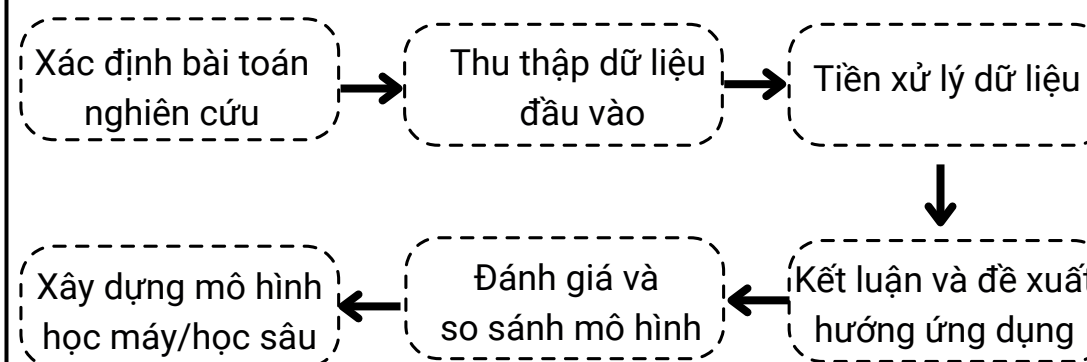
- Hướng đến việc tìm kiếm giải pháp ứng dụng trí tuệ nhân tạo, đặc biệt là các kỹ thuật học máy và học sâu, để phục vụ cho việc phân tích hành vi và đánh giá tâm lý khách hàng trong môi trường thương mại điện tử.

TẬP DỮ LIỆU

- Nguồn dữ liệu: Giả lập môi trường TMĐT (file CSV train_dataset.csv)
- Số lượng thuộc tính: Gồm thông tin người dùng, giao dịch, sản phẩm, chương trình khách hàng thân thiết và trải nghiệm vận chuyển
- Biến mục tiêu: customer_experience (good, neutral, bad)

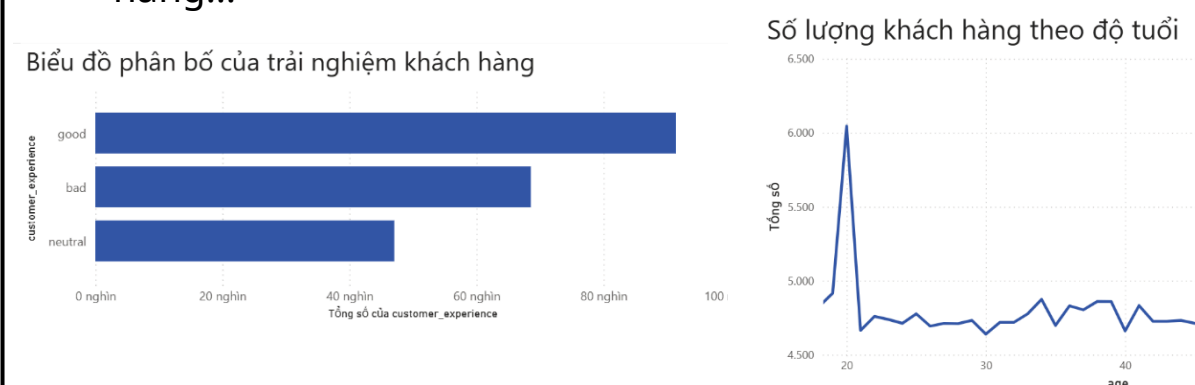


PHƯƠNG PHÁP NGHIÊN CỨU

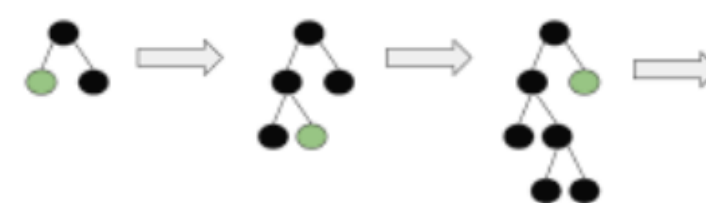


1. Tiền xử lý dữ liệu:

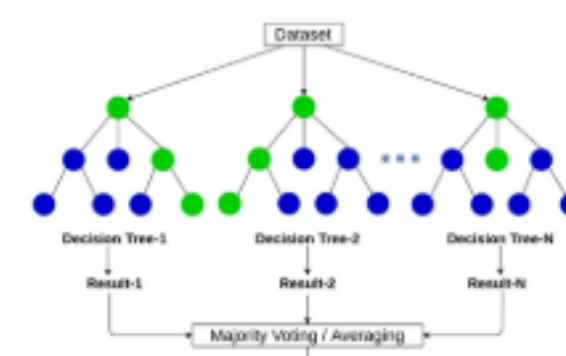
- Xử lý thiếu dữ liệu, chuẩn hóa, mã hóa, loại bỏ ngoại lệ (IQR)
- Feature engineering: nhóm tuổi, loại thẻ, thời gian giao hàng...



2. Mô hình áp dụng



LightGBM



Random Forest

3. Đánh giá mô hình:

- Dùng các chỉ số: Accuracy, F1-Score, Precision, Recall
- K-Fold cross-validation

Lớp	Precision	Recall	F1-Score	Support
0	0.60	0.67	0.63	13,707
1	0.73	0.73	0.73	9,408
2	0.71	0.65	0.68	17,416

Mô hình LightGBM

Lớp	Precision	Recall	F1-Score	Support
0	0.57	0.54	0.55	13,707
1	0.52	0.63	0.57	9,408
2	0.69	0.64	0.67	17,416

Mô hình Random Forest

KẾT QUẢ THỰC NGHIỆM

Chỉ số	LightGBM	Random Forest
Accuracy	0.6731	0.6043
F1-Score	0.6741	0.6054
F1 lớp 0	0.63	0.55
F1 lớp 1	0.73	0.57
F1 lớp 2	0.68	0.67
Tính ổn định	Tốt	Trung bình
Khả năng tổng quát	Tốt	Cần cải thiện

- LightGBM đạt hiệu quả cao nhất cho lớp trung tính (neutral)
- Một số nhầm lẫn xảy ra giữa lớp hài lòng và không hài lòng
- Cần cải thiện thêm bằng kỹ thuật ensemble hoặc tuning hyperparameter

Ý NGHĨA THỰC TIỄN

- Giúp doanh nghiệp phát hiện sớm khách hàng không hài lòng
- Tối ưu chiến lược marketing & chăm sóc khách hàng
- Mô hình giúp giảm chi phí so với các quy trình thu thập đánh giá khách hàng truyền thống

KẾT LUẬN & HƯỚNG PHÁT TRIỂN

- Mô hình LightGBM là ứng viên mạnh cho triển khai thực tế
- Đề xuất mở rộng thêm với các kỹ thuật deep learning như MLP, BERT (phân tích bình luận văn bản)