



# Dự đoán khả năng rời bỏ dịch vụ của khách hàng trong lĩnh vực viễn thông

Nhóm thực hiện:

13



# Vai trò của AI trong quản trị khách hàng

**Churn prediction** là bài toán quan trọng giúp doanh nghiệp:

- Giảm tỷ lệ rời bỏ dịch vụ
- Tăng doanh thu, giảm chi phí

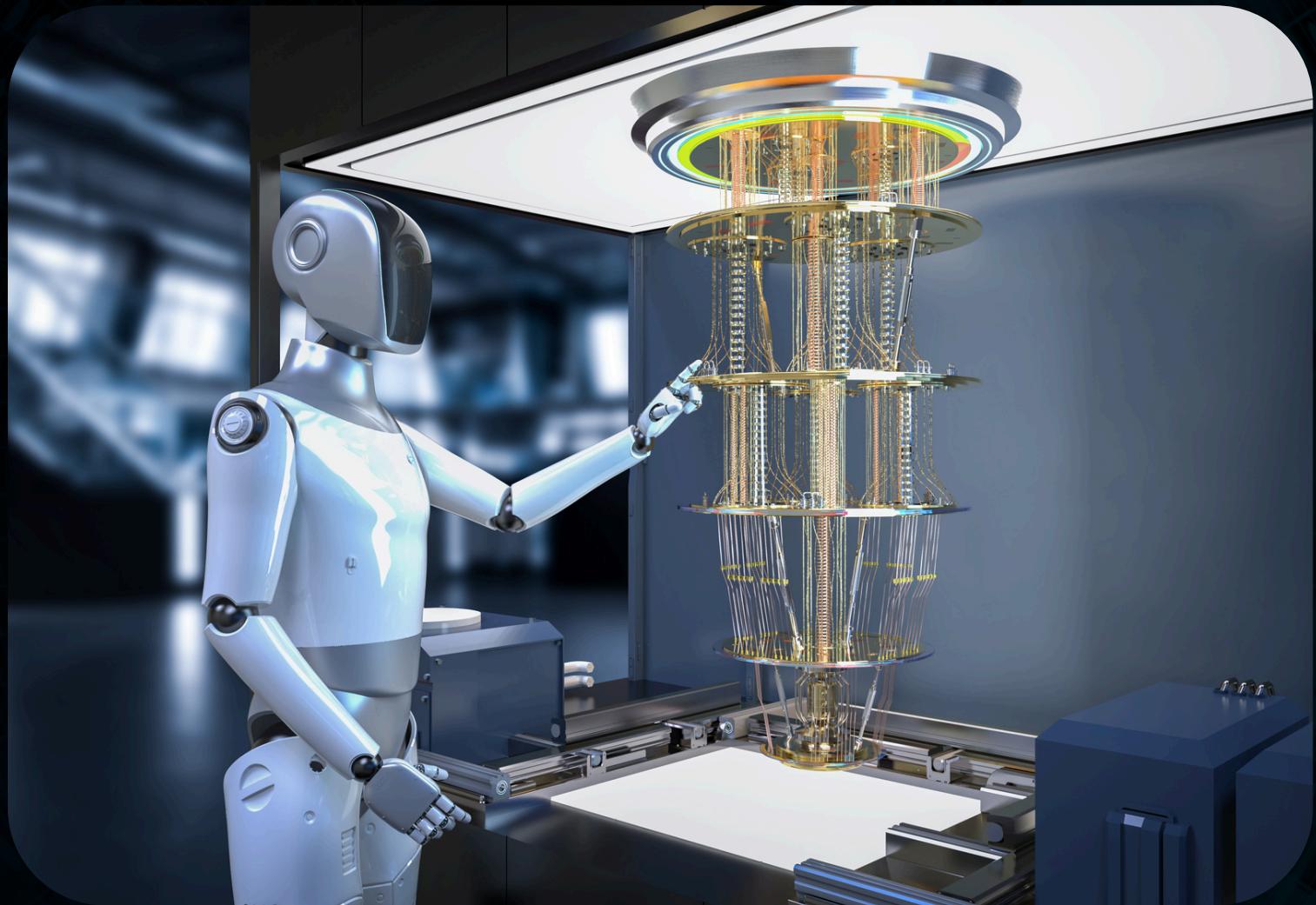
**Mục tiêu đề tài:**

- Xây dựng mô hình học máy dự đoán churn
- Tối ưu hiệu suất mô hình và ứng dụng thực tế



## Tên đề tài:

Ứng dụng học máy trong dự đoán khả năng rời bỏ dịch vụ của khách hàng ngành viễn thông dựa trên dữ liệu hành vi tiêu dùng



## **Mục tiêu chính:**

- Tự động hóa dự đoán churn
- Tăng khả năng giữ chân khách hàng
- Cá nhân hóa chiến lược chăm sóc

## **Mục tiêu phụ:**

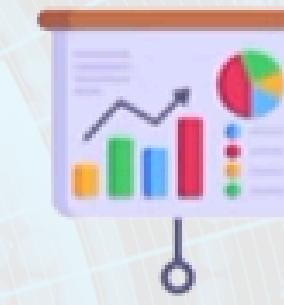
- Khai thác dữ liệu Telco
- Xây dựng mô hình RF, DT, XGBoost
- Tối ưu threshold, SMOTE, đánh giá mô hình

# QUY TRÌNH THỰC HIỆN

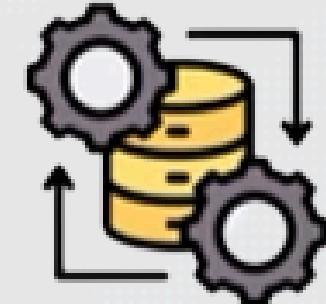
*Workflow*



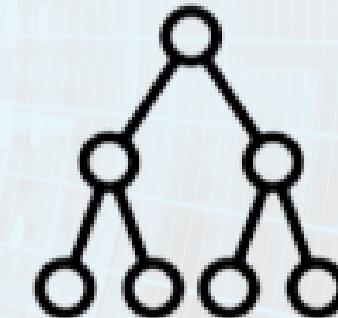
Data Collection



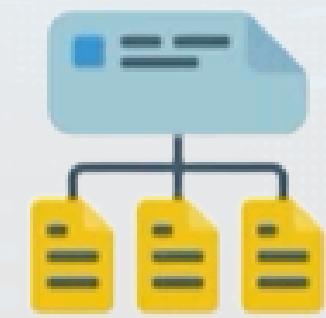
EDA



Data Preprocessing



ML Models



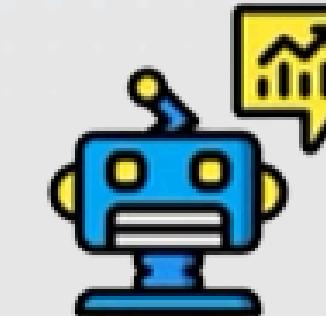
Train Test Split



Unknown data



Best Trained Model



Prediction

Nguồn: Telco Customer Churn (7.043 bản ghi)

# Dữ liệu đầu vào

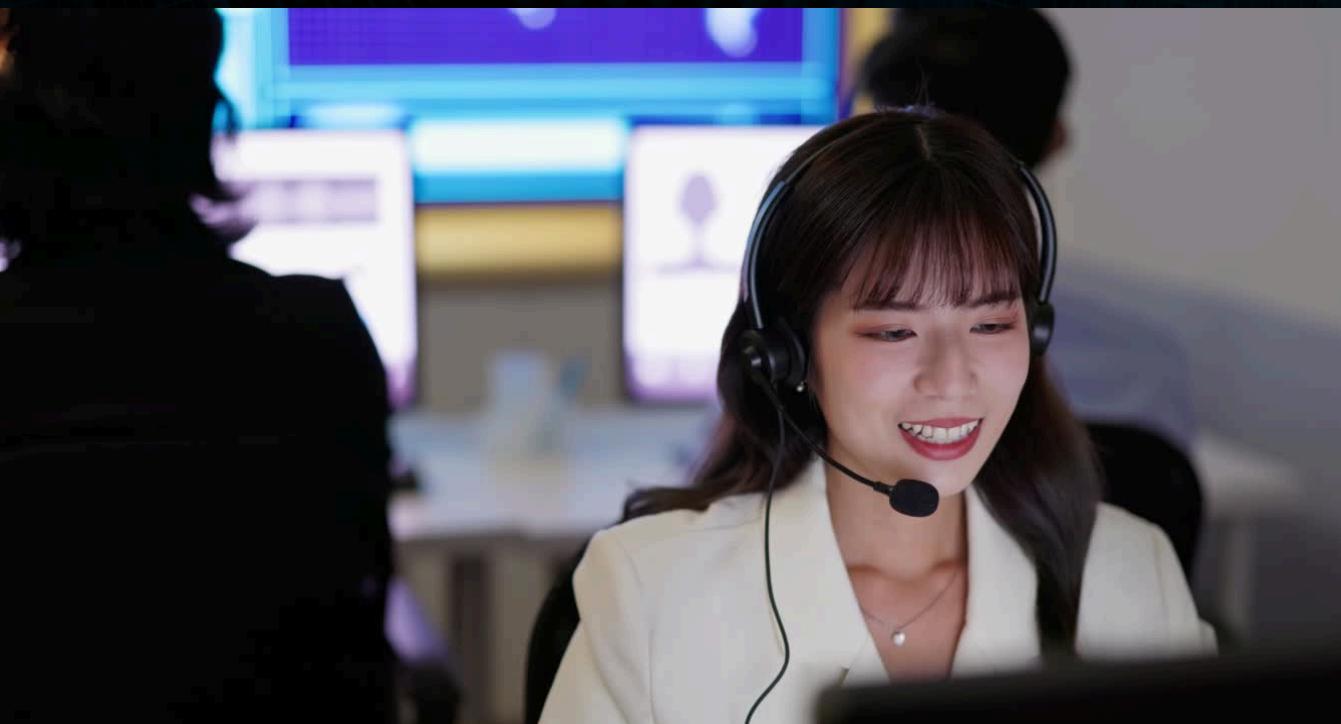
Tên biến	Giải thích ngắn gọn
SeniorCitizen	Người cao tuổi (1 = Có, 0 = Không)
Partner	Có vợ/chồng hay không (Yes/No)
Dependents	Có người phụ thuộc không (Yes/No)
tenure	Thời gian sử dụng dịch vụ (số tháng)
PhoneService	Có dùng điện thoại không (Yes/No)
MultipleLines	Có nhiều đường dây điện thoại (Yes/No/No phone)
InternetService	Loại dịch vụ Internet (DSL/Fiber optic/Không có)
OnlineSecurity	Có dùng bảo mật online (Yes/No/No internet)
OnlineBackup	Có dùng sao lưu online (Yes/No/No internet)
DeviceProtection	Có dùng bảo vệ thiết bị (Yes/No/No internet)
TechSupport	Có hỗ trợ kỹ thuật (Yes/No/No internet)
StreamingTV	Có xem TV trực tuyến (Yes/No/No internet)
StreamingMovies	Có xem phim trực tuyến (Yes/No/No internet)
Contract	Loại hợp đồng (Tháng/1 năm/2 năm)
PaperlessBilling	Có dùng hóa đơn điện tử (Yes/No)
PaymentMethod	Phương thức thanh toán
MonthlyCharges	Phí trả hàng tháng
TotalCharges	Tổng số tiền đã thanh toán
Churn	Có rời bỏ dịch vụ hay không (Yes/No)

Biến quan trọng:  
tenure, MonthlyCharges, TotalCharges  
Contract, InternetService, PaymentMethod...

Biến mục tiêu: Churn (Yes/No)



# Xử lý và chuẩn hóa dữ liệu



- Loại bỏ cột customerID
- Chuyển đổi TotalCharges về float
- Mã hóa các biến phân loại (LabelEncoder)
- Xử lý mất cân bằng bằng SMOTE:
- Churn chiếm 26.5% → xử lý để cân bằng

# Phân tích dữ liệu (EDA)

- Histogram + Boxplot: tenure, charges ổn định
- Heatmap:
- tenure – TotalCharges:  $r = 0.83$  (mạnh)
- MonthlyCharges – TotalCharges:  $r = 0.65$
- Countplot:
- Churn cao nhất ở khách dùng hợp đồng tháng-tháng
- Fiber optic có churn cao hơn DSL

**tenure, TotalCharges, MonthlyCharges phân bố rộng và không đồng đều**

→ cho thấy khách hàng có hành vi sử dụng rất khác nhau.

**Tương quan mạnh giữa tenure và TotalCharges ( $r = 0.83$ ):**

→ Khách hàng càng gắn bó lâu, chi phí tích lũy càng cao.

**Churn cao nhất ở khách dùng hợp đồng “Month-to-month”**

→ Nhóm không cam kết dài hạn, dễ rời bỏ.

**Churn cao hơn ở khách dùng Fiber optic so với DSL**

→ Có thể do chi phí cao hoặc kỳ vọng không được đáp ứng.



# LỰA CHỌN MÔ HÌNH

So sánh 3 mô hình học máy:

- Decision Tree: Accuracy 78% – thiếu ổn định
- XGBoost: Accuracy 83% – tốt nhưng chậm
- Random Forest: Accuracy 84% – ổn định nhất → được chọn

## HIỆU CHỈNH NGƯỠNG DỰ ĐOÁN

Tiêu chí	0.5	0.4
Accuracy	0.78	0.75
Recall (churn)	0.58	0.70
F1-score	0.58	0.60
TP tăng	–	+47 khách



# KẾT QUẢ VẬN HÀNH

- Mô hình dự đoán nhanh, dễ áp dụng
- Hiển thị kết quả rõ ràng cho CSKH
- Có thể mở rộng thành giao diện web / tích hợp vào CRM

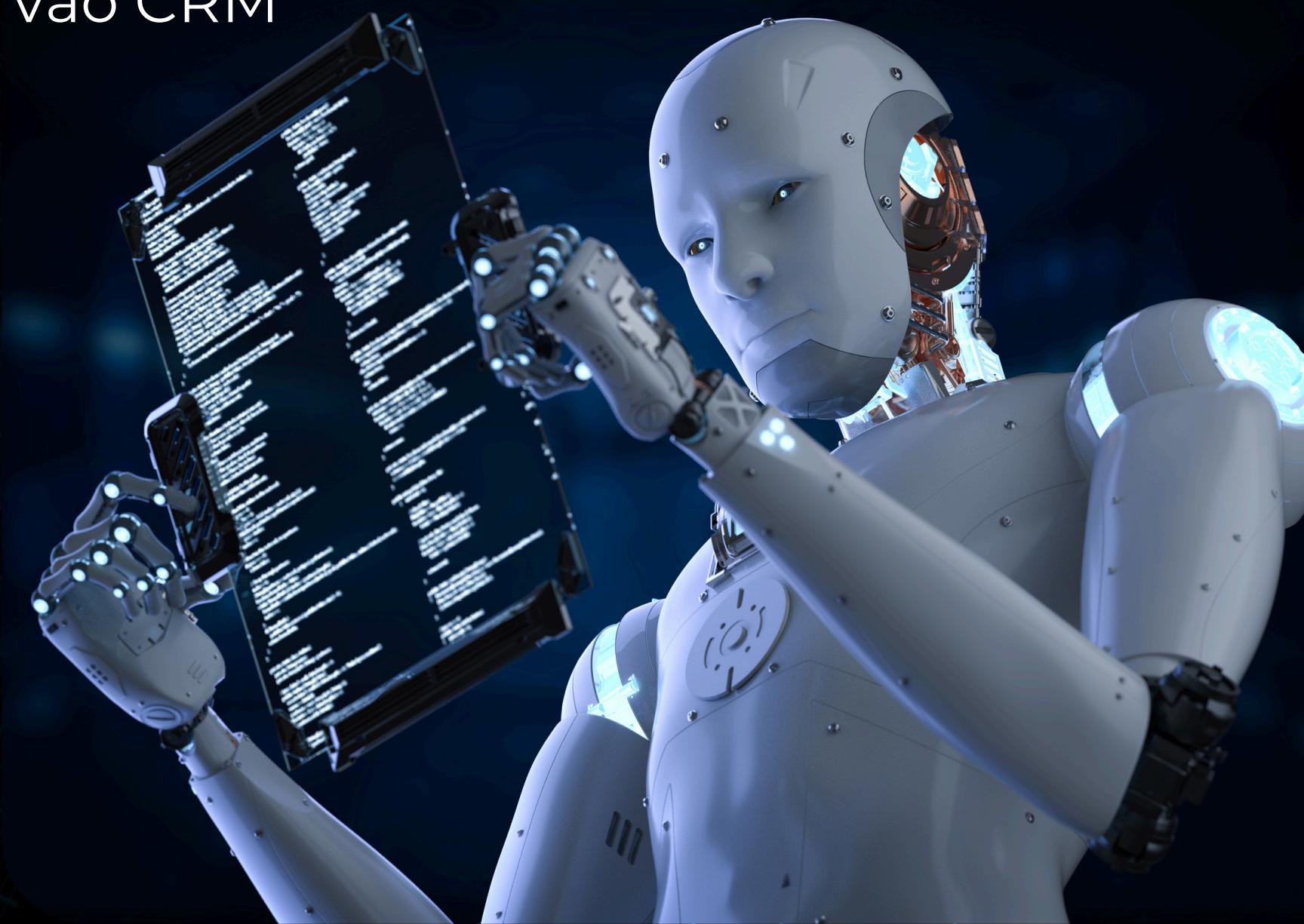
## HƯỚNG PHÁT TRIỂN

Tích hợp mô hình vào hệ thống doanh nghiệp

Kết hợp với Customer Lifetime Value (CLV) để ưu tiên giữ chân

Triển khai giao diện web / ứng dụng

Sử dụng Explainable AI (SHAP) cho giải thích chi tiết hơn



CẢM ƠN THẦY  
VÀ CÁC BẠN ĐÃ  
LẮNG NGHE

