**Tiêu đề:** Phát triển Mô hình Dự đoán Điểm Khách Hàng Tiềm Năng (Predictive Lead Scoring) **1. Giới thiệu:**

* **Vấn đề nghiên cứu:**
  + Trong môi trường kinh doanh cạnh tranh, việc xác định và ưu tiên các khách hàng tiềm năng có khả năng chuyển đổi cao là rất quan trọng.
  + Phương pháp chấm điểm khách hàng tiềm năng truyền thống dựa trên các quy tắc tĩnh thường không hiệu quả do bỏ qua các yếu tố động và phức tạp trong hành vi khách hàng.
* **Mục tiêu nghiên cứu:**
  + Xây dựng một mô hình Predictive Lead Scoring chính xác và hiệu quả để đánh giá hiệu suất của mô hình và so sánh với các phương pháp truyền thống.
  + Đề xuất các ứng dụng thực tiễn của mô hình trong quy trình bán hàng.
* **Giả thuyết nghiên cứu:**
  + Mô hình Predictive Lead Scoring sử dụng học máy sẽ cho kết quả dự đoán chính xác hơn so với các phương pháp truyền thống.
  + Việc sử dụng mô hình này sẽ dẫn đến tăng tỷ lệ chuyển đổi và tối ưu hóa chi phí bán hàng.

**Câu hỏi**

* **Xác định mục tiêu:**
  + Tăng tỷ lệ chuyển đổi khách hàng tiềm năng bao nhiêu phần trăm?
  + Tối ưu hóa chi phí marketing và bán hàng như thế nào?
  + Cải thiện trải nghiệm khách hàng ra sao?
* **Xác định phạm vi:**
  + Bạn sẽ sử dụng những nguồn dữ liệu nào?
  + Bạn sẽ tập trung vào những loại khách hàng tiềm năng nào?
  + Bạn sẽ sử dụng những thuật toán học máy nào?

**Dataset:** bank-additional-full.csv: https://github.com/daddydrac/Machine-Learning-For-Predictive-Lead-Scoring/blob/master/data/bank-additional/bank-additional-full.csv

(features) trong dataset:

* **age:** Tuổi của khách hàng.
* **job:** Nghề nghiệp của khách hàng (ví dụ: technician, admin., retired).
* **marital:** Tình trạng hôn nhân của khách hàng (ví dụ: married, divorced, single).
* **education:** Trình độ học vấn của khách hàng (ví dụ: professional.course, high.school, university.degree).
* **default:** Khách hàng có khoản nợ quá hạn hay không (yes/no).
* **housing:** Khách hàng có khoản vay nhà ở hay không (yes/no).
* **loan:** Khách hàng có khoản vay cá nhân hay không (yes/no).
* **contact:** Loại phương thức liên lạc (ví dụ: cellular, telephone).
* **month:** Tháng liên lạc cuối cùng (ví dụ: may, aug, jul).
* **day\_of\_week:** Ngày trong tuần liên lạc cuối cùng (ví dụ: mon, fri, wed).
* **duration:** Thời lượng cuộc gọi liên lạc cuối cùng, tính bằng giây.
* **campaign:** Số lần liên lạc được thực hiện cho khách hàng này trong chiến dịch hiện tại.
* **pdays:** Số ngày đã trôi qua kể từ lần liên lạc cuối cùng với khách hàng từ một chiến dịch trước đó (999 có nghĩa là khách hàng chưa được liên lạc trước đó).
* **previous:** Số lần liên lạc được thực hiện trước chiến dịch hiện tại cho khách hàng này.
* **poutcome:** Kết quả của chiến dịch marketing trước đó (ví dụ: nonexistent, success, failure).
* **emp.var.rate:** Tỷ lệ thay đổi việc làm theo quý.
* **cons.price.idx:** Chỉ số giá tiêu dùng hàng tháng.
* **cons.conf.idx:** Chỉ số niềm tin của người tiêu dùng hàng tháng.
* **euribor3m:** Tỷ lệ euribor 3 tháng hàng ngày.
* **nr.employed:** Số lượng nhân viên theo quý.
* **y:** Biến mục tiêu, cho biết khách hàng có đăng ký tiền gửi có kỳ hạn hay không (yes/no).

Link hướng dẫn: [daddydrac/Machine-Learning-For-Predictive-Lead-Scoring: Predictive Lead Scoring does all the hard work for you by leveraging Machine Learning to provide your sales and marketing team with in-depth customer knowledge and ways to target the hottest and most qualified leads – resulting in saved time and higher revenue streams.](https://github.com/daddydrac/Machine-Learning-For-Predictive-Lead-Scoring/tree/master)

**CoNVO Framework:**

* Context (Bối Cảnh):
  + Ngân hàng X đang chạy chiến dịch telemarketing để bán sản phẩm tiền gửi có kỳ hạn cho khách hàng hiện tại.
  + Người ra quyết định là giám đốc marketing.
* Needs (Nhu Cầu):
  + Tăng tỷ lệ chuyển đổi của chiến dịch hiện tại để giải phóng ngân sách cho chiến dịch sản phẩm thế chấp.
  + Định lượng tác động của nỗ lực mô hình hóa dữ liệu.
* Vision (Tầm Nhìn):
  + Xây dựng mô hình chấm điểm khách hàng tiềm năng để xếp hạng khách hàng theo khả năng đầu tư vào tiền gửi có kỳ hạn.
  + Cung cấp danh sách xếp hạng liên hệ cho đội ngũ marketing.
  + Thiết kế và thực hiện thử nghiệm có kiểm soát để định lượng tác động của chiến dịch.
* Outcome (Kết Quả):
  + Tăng đáng kể ROI thông qua việc triển khai mô hình chấm điểm khách hàng tiềm năng.

2. Quy Trình Phát Triển Mô Hình (CRISP-DM Framework):

* Business Understanding (Hiểu Vấn Đề Kinh Doanh):
  + Xác định mục tiêu kinh doanh và yêu cầu của các bên liên quan.
  + Chuyển đổi các mục tiêu kinh doanh thành các mục tiêu khoa học dữ liệu.
* Data Understanding (Hiểu Dữ Liệu):
  + Thu thập dữ liệu từ UCI Machine Learning Repository (Bank Marketing Dataset).
  + Khám phá dữ liệu (EDA) để hiểu rõ hơn về các biến và mối quan hệ.
* Data Preparation (Chuẩn Bị Dữ Liệu):
  + Làm sạch và chuẩn hóa dữ liệu.
  + Tạo các đặc trưng mới (feature engineering).
  + Chia dữ liệu thành tập huấn luyện và tập kiểm tra.
* Modeling (Mô Hình Hóa):
  + Xây dựng và đào tạo các mô hình học máy (Naive Bayes, Logistic Regression, v.v.).
  + Tinh chỉnh các tham số của mô hình.
  + So sánh kết quả của các mô hình khác nhau.
* Evaluation (Đánh Giá):
  + Đánh giá hiệu suất của mô hình trên tập kiểm tra.
  + Sử dụng các chỉ số đánh giá phù hợp (tỷ lệ chuyển đổi, độ chính xác, v.v.).
  + Đánh giá tác động kinh doanh của mô hình.
* Deployment (Triển Khai):
  + Trong dự án này, bước triển khai bị bỏ qua.
  + Trong các dự án thực tế, triển khai sẽ bao gồm việc tích hợp mô hình vào hệ thống CRM hoặc các công cụ bán hàng.
* Monitoring (Giám Sát)
  + Trong các dự án thực tế, bước giám sát là rất quan trọng để đảm bảo mô hình luôn hoạt động tốt.