**SMô hình CRISP-DM (Cross Industry Standard Process for Data Mining)**

CRISP-DM là một phương pháp luận chuẩn được sử dụng rộng rãi trong lĩnh vực khai thác dữ liệu (Data Mining), giúp định hướng và chuẩn hóa quy trình triển khai các dự án khai thác dữ liệu một cách hiệu quả. Mô hình này được phát triển nhằm cung cấp một khuôn khổ chung, có thể áp dụng cho nhiều ngành nghề khác nhau (cross-industry), giúp các nhóm dự án thực hiện từng bước rõ ràng và khoa học từ lúc bắt đầu đến khi hoàn thành.

CRISP-DM gồm 6 giai đoạn chính, được thực hiện tuần tự hoặc lặp đi lặp lại tùy theo tình hình thực tế của dự án:

1. **Business Understanding (Hiểu biết về nghiệp vụ):**  
   Giai đoạn đầu tiên tập trung vào việc hiểu rõ mục tiêu kinh doanh, các yêu cầu và mong đợi của dự án. Đây là bước quan trọng để xác định các vấn đề cần giải quyết và cách thức khai thác dữ liệu sẽ hỗ trợ cho mục tiêu đó.
2. **Data Understanding (Hiểu biết về dữ liệu):**  
   Ở giai đoạn này, nhóm thực hiện thu thập dữ liệu ban đầu, khám phá và kiểm tra dữ liệu nhằm đánh giá chất lượng, tính đầy đủ và tính khả dụng của dữ liệu. Các bước kiểm tra này giúp phát hiện các vấn đề về dữ liệu cần xử lý trước khi phân tích sâu hơn.
3. **Data Preparation (Chuẩn bị dữ liệu):**  
   Đây là giai đoạn tiền xử lý dữ liệu, bao gồm các công việc như làm sạch dữ liệu, lựa chọn dữ liệu liên quan, chuyển đổi và tạo các biến đặc trưng mới. Mục tiêu là xây dựng bộ dữ liệu phù hợp và chuẩn bị tốt cho các mô hình khai thác dữ liệu.
4. **Modeling (Xây dựng mô hình):**  
   Trong bước này, các thuật toán khai thác dữ liệu (machine learning, thống kê, ...) được áp dụng để xây dựng mô hình dự đoán hoặc phân loại dựa trên bộ dữ liệu đã chuẩn bị. Thường cần thử nghiệm và điều chỉnh các tham số để tối ưu hóa kết quả mô hình.
5. **Evaluation (Đánh giá):**  
   Mô hình được đánh giá về độ chính xác, khả năng áp dụng và phù hợp với mục tiêu nghiệp vụ đã đề ra. Nếu mô hình chưa đạt yêu cầu, có thể quay lại các bước trước để điều chỉnh hoặc cải thiện.
6. **Deployment (Triển khai):**  
   Khi mô hình được chấp nhận, nó sẽ được triển khai vào môi trường thực tế để sử dụng trong các hoạt động kinh doanh. Giai đoạn này còn bao gồm việc giám sát và duy trì mô hình trong suốt quá trình vận hành.