**BƯỚC 1: XÁC ĐỊNH BÀI TOÁN**

**Tên đề tài**: *Xây dựng chatbot hỗ trợ trả lời tự động trên Discord, cho phép mỗi server có bộ dữ liệu riêng*

**Bài toán**:  
Trong môi trường học tập hoặc cộng đồng online như Discord, việc trả lời các câu hỏi lặp đi lặp lại hoặc ghi nhận những vấn đề của người dùng là một công việc tốn thời gian của admin. Nhóm đề xuất xây dựng một chatbot hoạt động trên Discord giúp:

* Trả lời tự động các câu hỏi theo mẫu (intents).
* Mỗi server có thể tự huấn luyện bộ dữ liệu riêng.
* Ghi lại những câu hỏi chưa trả lời được để admin xử lý.
* Cho phép admin phản hồi và cập nhật dữ liệu bằng lệnh nội bộ.

**BƯỚC 2: GIẢI PHÁP ĐỀ XUẤT**

**✅ Công nghệ sử dụng**

* **Ngôn ngữ lập trình**: Python 3.11
* **Thư viện**:
  + discord.py – để tương tác với Discord API.
  + scikit-learn – để huấn luyện mô hình phân loại câu hỏi (Logistic Regression).
  + pickle, json, os – để lưu trữ mô hình và dữ liệu huấn luyện theo từng server.
* **Mô hình học máy**: Logistic Regression với dữ liệu được vector hóa bằng CountVectorizer.

**✅ Cấu trúc tổng thể hệ thống**

* Mỗi server khi thêm bot vào sẽ được tạo 1 thư mục dữ liệu riêng (data/<server\_id>).
* Người dùng có thể sử dụng /train để huấn luyện mô hình dựa trên intents.json của server.
* Mỗi tin nhắn sẽ được dự đoán intent. Nếu không khớp, tin nhắn sẽ được lưu vào pending\_data.json.
* Admin có thể dùng /reply, /delete, /pending để xử lý các câu chưa trả lời được.

**🔄 Lưu đồ giải thuật (luồng xử lý tin nhắn)**

A diagram of a diagram

AI-generated content may be incorrect.

1. 📥 Nhận tin nhắn từ người dùng trên một server.

2. ❌ Nếu là tin nhắn từ chính bot → bỏ qua.

3. ✅ Nếu là lệnh /train → gọi hàm train\_model() để huấn luyện lại.

4. ✅ Nếu là lệnh /pending, /delete, /reply → thực hiện thao tác tương ứng.

5. ✅ Nếu là tin nhắn thường:

o Tải mô hình theo server\_id.

o Dự đoán intent qua model.predict().

o Nếu độ chính xác cao hơn ngưỡng → gửi phản hồi tương ứng.

o Ngược lại → lưu câu hỏi vào pending.

**Thành viên 1: Xử lý backend & AI model**

* Thiết kế luồng xử lý dữ liệu intents, vectorizer và model.
* Viết file train\_model.py để huấn luyện Logistic Regression.
* Tạo các hàm load\_model, get\_response, save\_pending\_question.
* Đảm nhiệm việc đọc/ghi file intents.json, pending\_data.json.

**👨‍💻 Thành viên 2: Tích hợp bot Discord & quản lý server**

* Viết bot\_multi.py: xử lý on\_message, /train, /reply, /pending, /delete.
* Quản lý logic đa server: mỗi server một thư mục dữ liệu riêng.
* Tạo luồng “bot được thêm vào server mới” và khởi tạo intents mặc định.
* Gắn logic phản hồi theo độ tin cậy và lưu câu hỏi chưa hiểu.

**👨‍💻 Thành viên 3: Lưu đồ + báo cáo + giao diện trực quan**

* Thiết kế **lưu đồ giải thuật** bằng draw.io.
* Viết phần giải thích bài toán, mô hình xử lý, giải pháp kỹ thuật.
* Làm **file README / báo cáo Word / PDF** trình bày rõ ràng.
* Viết mô tả quy trình huấn luyện và phản hồi của bot.