**FORM MÔ TẢ DỰ ÁN DỰ ĐOÁN KỸ NĂNG CỦA HỌC SINH**

1. Mô tả vấn đề:

Tên dự án: Dự đoán kỹ năng học sinh qua biểu hiện

Mục tiêu mô hình: Xác định và dự đoán mức độ thành thạo kỹ năng cụ thể của học sinh dựa trên các biểu hiện.

2. Dữ liệu:

Nguồn dữ liệu: (Ví dụ: Hồ sơ học tập, bài kiểm tra, phản hồi từ giáo viên và bạn bè, video ghi lại lớp học)

Kích thước và định dạng dữ liệu: (Ví dụ: Bảng điểm trong định dạng CSV, video định dạng MP4)

Các thuộc tính dữ liệu: (Ví dụ: Điểm số, nhận xét của giáo viên, tần suất tham gia, biểu cảm khuôn mặt và ngôn ngữ cơ thể trong video)

Cách thu thập dữ liệu: (Mô tả cách thu thập và tổng hợp dữ liệu từ các nguồn khác nhau)

3. Xử lý và chuẩn bị dữ liệu:

Các bước tiền xử lý: (Làm sạch dữ liệu, chuyển đổi video thành dạng có thể xử lý được bằng RNN, mã hóa dữ liệu danh mục)

Chia tập dữ liệu: (Mô tả cách bạn chia tập huấn luyện, validation, và kiểm tra)

4. Thiết kế mô hình:

Kiến trúc mô hình: (Mô tả cấu trúc của RNN, bao gồm số lượng tầng, loại tầng (LSTM, GRU), và số nút)

Hàm mất mát và bộ tối ưu: (Chọn phù hợp với vấn đề dự đoán, ví dụ: cross-entropy)

5. Huấn luyện mô hình:

Thông số huấn luyện: (Số epoch, kích thước batch)

Quy trình validation: (Phương pháp kiểm tra và đánh giá mô hình trong quá trình huấn luyện)

6. Đánh giá và tối ưu:

Phương pháp đánh giá: (Mô tả cách bạn đánh giá hiệu suất của mô hình, ví dụ: độ chính xác, độ mất mát)

Các bước tối ưu: (Điều chỉnh kiến trúc, thử nghiệm với dữ liệu đa dạng hơn, sử dụng kỹ thuật như dropout để ngăn chặn overfitting)

7. Triển khai mô hình:

Cách triển khai: (Ví dụ: Ứng dụng web để giáo viên nhập dữ liệu và nhận dự đoán, hệ thống tự động gửi báo cáo dự đoán tới email của giáo viên)

Các yêu cầu về hiệu suất: (Ví dụ: Thời gian phản hồi nhanh, độ chính xác cao)

8. Bảo mật và quyền riêng tư:

Các biện pháp bảo mật: (Mô tả cách bảo vệ dữ liệu cá nhân và đảm bảo quyền riêng tư của học sinh)

**HƯỚNG ĐI PHÂN TÍCH VÀ XÂY DỰNG MÔ HÌNH:**

Tiền xử lý dữ liệu:

Mã hóa dữ liệu: Bạn cần chuyển đổi dữ liệu phân loại thành dạng số để có thể xử lý bằng mô hình RNN. Có thể sử dụng các phương pháp như mã hóa nhị phân (binary encoding) hoặc mã hóa one-hot.

Xử lý dữ liệu thiếu: Kiểm tra và xử lý các dữ liệu thiếu (nếu có) bằng cách sử dụng các phương pháp như điền giá trị trung bình, median, hoặc sử dụng mô hình để dự đoán giá trị thiếu.

Phân chia dữ liệu:

Chia dữ liệu thành các tập huấn luyện, validation và kiểm tra để đánh giá hiệu suất của mô hình một cách chính xác.

Xây dựng mô hình:

Xác định kiến trúc mô hình phù hợp, có thể bao gồm các tầng LSTM hoặc GRU để xử lý hiệu quả dữ liệu có tính chuỗi thời gian hoặc tuần tự.

Chọn hàm mất mát và bộ tối ưu phù hợp dựa trên bài toán cụ thể (ví dụ, mất mát cross-entropy cho bài toán phân loại).

Huấn luyện và Đánh giá mô hình:

Huấn luyện mô hình sử dụng tập huấn luyện và theo dõi hiệu suất trên tập validation để điều chỉnh tham số và tránh overfitting.

Đánh giá mô hình cuối cùng trên tập kiểm tra để xác định độ chính xác và hiệu suất tổng thể.

Tối ưu và Điều chỉnh mô hình:

Sử dụng kỹ thuật như tuning tham số, kỹ thuật regularization (ví dụ, dropout) để tối ưu mô hình và cải thiện khả năng khái quát hóa.

Triển khai mô hình:

Phát triển một ứng dụng hoặc hệ thống có thể sử dụng mô hình đã được huấn luyện để dự đoán kỹ năng của học sinh mới dựa trên biểu hiện của họ.