



ĐỒ ÁN MÔN HỌC

HỆ THỐNG AI CHẤM ĐIỂM TỰ ĐỘNG

I. Mục tiêu:

- Hiểu cách áp dụng LLM vào bài toán đánh giá/chấm điểm tự động.
- Thành thạo quy trình xây dựng pipeline từ thu thập dữ liệu, tiền xử lý, thiết kế prompt đến đánh giá hiệu quả mô hình.
- Phát triển tư duy phân tích về fairness, explainability và tính ứng dụng thực tiễn của hệ thống AI chấm điểm.

II. Mô tả đồ án:

1. Phạm vi công việc: Sinh viên cần thực hiện các nhiệm vụ sau:

- Chọn một lĩnh vực cụ thể cần chấm điểm (xem gợi ý bên dưới).
- Thu thập dữ liệu đầu vào (câu trả lời, bài thi, bài nói, video,...) và nếu cần thì có cả đáp án chuẩn/mẫu đánh giá để làm dữ liệu tham khảo.
- Thiết kế kiến trúc hệ thống gồm các thành phần:
 - o Tiền xử lý dữ liệu đầu vào (ví dụ: chuyển âm thanh thành văn bản, trích xuất nội dung từ văn bản).
 - o Yêu cầu cơ bản: xây dựng prompt để đánh giá bằng LLM, hoặc thiết kế thêm các bộ đánh giá phụ trợ (rule-based, rubric-based,...).
 - o Triển khai LLM để chấm điểm chuyên sâu bằng cách tăng cường dữ liệu tự động, tiền huấn luyện, tinh chỉnh tham số, knowledge distillation, ...
 - o Trình bày kết quả đánh giá (độ hiệu quả thông qua các độ đo) và khảo sát nhận xét.
- So sánh kết quả đánh giá của LLM với đánh giá của con người (nếu có).
- Đánh giá hệ thống theo các tiêu chí: tính nhất quán, tốc độ, độ chính xác, độ thỏa mãn của người dùng, khả năng mở rộng.



2. Kết quả đầu ra cần đạt:

- Thiết kế một hệ thống AI sử dụng LLM để chấm điểm trong một lĩnh vực cụ thể (do nhóm lựa chọn).
- Trình bày rõ pipeline hệ thống, từ nhập liệu đến đầu ra điểm số.
- Phân tích ưu/nhược điểm của hệ thống.
- Đánh giá độ chính xác bằng cách so sánh với human annotation hoặc chấm điểm thủ công.
- Trình bày demo (có thể qua web app, CLI, hoặc notebook).
- Viết báo cáo đầy đủ các phần: Mở đầu, Phương pháp, Kết quả, Thảo luận, Kết luận.

3. Gợi ý một số chủ đề

Chủ đề	Mô tả cụ thể
Chấm điểm bài thi IELTS Writing/ HSK Speaking	Nhập bài viết Task 1 hoặc Task 2 và yêu cầu LLM chấm theo band score của IELTS. Hoặc transcript/audio của bài thi HSK Speaking và chấm theo band score của HSK
Chấm điểm bài nói phỏng vấn học thuật	Người dùng nói (hoặc nạp transcript), LLM đánh giá độ lưu loát, từ vựng, logic
Chấm điểm bài thi lập trình	Nhập đề bài + lời giải, LLM đánh giá mức độ chính xác, tối ưu, sáng tạo (hoặc kèm test case)
Chấm điểm luận văn tóm tắt / abstract nghiên cứu	Cho đoạn abstract, LLM đánh giá mức độ rõ ràng, đầy đủ, tính mới
Chấm bài thuyết trình PowerPoint	Người dùng nhập nội dung slide (hoặc script), LLM đánh giá theo cấu trúc, nội dung và ngôn ngữ
Chấm điểm nội dung quảng cáo / pitch startup	Dựa vào đoạn pitch, LLM đánh giá sức thuyết phục, rõ ràng, hấp dẫn
Chấm câu trả lời trắc nghiệm mở rộng (short answer)	Dựa vào câu hỏi + câu trả lời + đáp án chuẩn, LLM tính điểm gần đúng và nhận xét
Chấm đối thoại tư vấn khách hàng (CS)	Dùng transcript từ chatbot / nhân viên CS, LLM đánh giá mức độ hài lòng, sự phù hợp câu trả lời
Chấm điểm độ phù hợp của CV ứng viên theo yêu cầu công việc	Hệ thống sử dụng LLM để phân tích CV của ứng viên và mô tả công việc (JD - Job Description), từ đó đưa ra nhận xét và điểm số đánh giá độ phù hợp



4. Gợi ý kiến trúc hệ thống

- Dữ liệu đầu vào: Văn bản, giọng nói, đoạn hội thoại, hoặc video.
- Tiền xử lý: Chuẩn hóa dữ liệu tùy vào hoàn cảnh sử dụng.
- LLM:
 - o Có thể dùng GPT-4, Claude, LLaMA 3, Gemini, Vicuna, Mistral...
- Thiết kế, huấn luyện AI:
 - o Có thể dùng zero-shot, few-shot, instruction tuning, reinforce learning...
- Output:
 - o Điểm số dạng số (0–10), mức độ (A–F), hoặc nhãn phân loại.
 - o Giải thích lý do chấm điểm (nếu có).
- Đánh giá hệ thống:
 - o So sánh với chấm điểm thủ công (human-labeled data).
 - o Dùng Pearson/Spearman correlation, BLEU/ROUGE (với bài luận), hoặc human preference.

5. Gợi ý cấu trúc báo cáo đồ án

- Giới thiệu: Bài toán và tầm quan trọng; các hướng tiếp cận hiện tại và khoảng trống
- Phương pháp: Thiết kế hệ thống, mô hình đề xuất, chi tiết prompt sử dụng và quá trình tinh chỉnh, ...
- Thực nghiệm
 - o Quy trình thu thập, xây dựng dataset
 - o Cách chấm điểm, đánh giá
 - o Kết quả: bảng điểm benchmark, feedback từ người dùng
- Phân tích và thảo luận
 - o Ưu điểm và hạn chế của hệ thống
 - o Tính công bằng, khả năng mở rộng, explainability
- Kết luận và hướng phát triển

III. Hướng dẫn nộp bài



Nhóm sinh viên cần chuẩn bị đầy đủ các nội dung sau và nộp về trang môn học (Moodle) trước ngày vấn đáp đồ án. Hạn chót sẽ được thông báo sau trên Moodle.

1. **Báo cáo:** Mô tả chi tiết về kiến trúc hệ thống, các thuật toán sử dụng, và quá trình phát triển giải pháp (dạng file nộp: .docx/.zip (nếu viết bằng latex) và .pdf)
2. **Bài thuyết trình:** khoảng 10 phút trình bày nhằm làm rõ các thông tin về kiến trúc hệ thống, thuật toán và quy trình đề xuất, cũng như những đánh giá và phân tích
3. **Video minh họa:** tối đa 5 phút, trình bày các chức năng hệ thống đã xây dựng (chỉ nộp đường dẫn đến video ở chế độ chia sẻ công khai cho mọi người)
4. **Mã nguồn:** Các mã nguồn đã xây dựng và hướng dẫn triển khai (readme). Các trọng số và dữ liệu thực hiện nếu quá nặng thì nhóm sinh viên chỉ nộp kèm đường dẫn
5. **Hệ thống dạng ứng dụng:** Không cần nộp trên trang môn học mà cần chuẩn bị để demo tại buổi vấn đáp

Lưu ý: Sinh viên cần đăng ký nhóm trước khi bắt đầu đồ án, hạn chót đăng ký đến ngày

10/08/2025. Mỗi nhóm có từ **3 - 6** thành viên. Những trường hợp nhóm có số lượng khác cần gửi email đến GVLТ và GVTH để đăng ký và trình bày lý do.

Format nộp bài: Nhóm sinh viên nén toàn bộ thư mục nộp bài gồm các thư mục con có nội dung đã được liệt kê và nén lại với tên theo quy ước sau:

MSHV1_MSHV2_MSHV3_MSHV4_MSHV5.zip

Nếu thư mục cần nộp có kích thước quá lớn thì chuyển thành đường dẫn (có thể down được cho mọi truy cập) và nộp trong file:

MSHV1_MSHV2_MSHV3_MSHV4_MSHV5.txt

IV. Tiêu chí đánh giá

STT	Tiêu chí	Điểm tối đa	Ghi chú
1	Mục tiêu rõ ràng & lựa chọn bài toán phù hợp	1.0	Chọn được lĩnh vực cụ thể, có lý do rõ ràng, và nêu được giá trị thực tiễn
2	Thu thập và xử lý dữ liệu đầu vào	1.5	Quy trình thu thập, gán nhãn, tiền xử lý dữ liệu
3	Thiết kế kiến trúc hệ thống & pipeline hoàn chỉnh	1.0	Kiến trúc hệ thống đánh giá điểm/huấn luyện
4	Khai thác LLM hiệu quả (prompt, fine-tuning, v.v.)	2.5	Các Kỹ thuật và mô hình LLMs sử dụng
5	So sánh mô hình và đánh giá kết quả	1.0	Phân tích ưu điểm/khuyết điểm/mẫu sai, ...
6	Khảo sát đánh giá từ người dùng hoặc so sánh với human judgment	1.0	Quá trình đánh giá thủ công, thu thập mẫu, ...
7	Trình bày báo cáo	1.5	
8	Demo sản phẩm	0.5	
9	Ý tưởng sáng tạo	1.0	Điểm cộng