



# **ĐÁNH GIÁ HỆ THỐNG MULTI-AGENT KẾT HỢP MÔ HÌNH NGÔN NGỮ LỚN CHO HỌC TẬP CÁ NHÂN HÓA QUA TRÒ CHƠI**

**Trương Nguyễn Yến Nhi - 240101062**

# Tóm tắt

- Lớp: CS2205.FEB2025
- Link Github: <https://github.com/nhitny/CS2205.FEB2025>
- Link YouTube video: <https://youtu.be/MzCDNvw3v5A>



Trương Nguyễn Yến Nhi  
240101062

# Giới thiệu

## Vấn đề nghiên cứu

- Cá nhân hóa học tập giúp đáp ứng năng lực, sở thích, tiến độ người học [1][2].
- Tuy nhiên, việc triển khai quy mô lớn gặp khó khăn do thiếu nhân lực, quá tải giảng viên (25–27 SV/giảng viên) [3].
- Trò chơi hóa (gamification) là phương pháp gia tăng tương tác, tạo hứng thú học tập [8][9].

# Giới thiệu

## Giải pháp

- Hệ thống đa tác tử (Multi-Agent) kết hợp LLM đang nổi lên như một giải pháp triển vọng [1][4][6].
- Framework đã được đề xuất lý thuyết nhưng chưa có thực nghiệm định lượng [1].
- Nghiên cứu này hiện thực hóa hệ thống, đánh giá hiệu quả từng mô hình LLM trong các vai trò Agent.

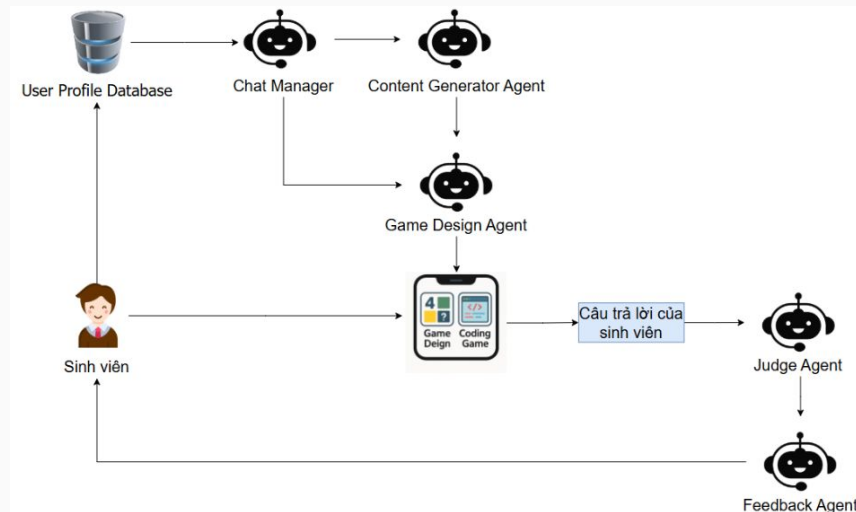
# Mục tiêu

- Hiện thực hóa hệ thống học tập cá nhân hóa gồm các Agent: Content Generator, Game Design, Judge, Feedback [1].
- So sánh hiệu quả các LLM (GPT-3.5, GPT-4 [4], Claude [5], Llama [6]) theo từng vai trò tác tử.
- Đề xuất cấu trúc phân công LLM tối ưu để tăng hiệu quả, khả năng mở rộng.

# Nội dung và Phương pháp

## Kiến trúc hệ thống

- Hệ thống gồm 5 thành phần chính [1]:
  - Content Generator Agent
  - Game Design Agent
  - Judge Agent
  - Feedback Agent
  - Chat Manager
- Mỗi Agent được thử với 4 LLM: GPT-3.5, GPT-4, Claude, Llama.



# Nội dung và Phương pháp

## Thiết kế thực nghiệm

- 50 sinh viên tham gia khảo sát hồ sơ học tập: giới tính, chuyên ngành, sở thích trò chơi,... [1].
- Tạo ra 2 loại trò chơi cá nhân hóa:
  - 4 Pics 1 Word
  - Coding Game
- Các Agent phối hợp để sinh nội dung – thiết kế – đánh giá – phản hồi.

# Nội dung và Phương pháp

## Tiêu chí đánh giá & phương pháp phân tích:

- 5 tiêu chí định lượng:
  - Thời gian phản hồi
  - Độ chính xác (F1-score)
  - Tính phù hợp nội dung (Likert 1–5)
  - Phản hồi đúng ngữ cảnh
  - Mức độ hài lòng người học (Likert 1–5)
- Phân tích kết quả bằng t-test và ANOVA để so sánh hiệu quả các LLM.



# Kết quả dự kiến

- Hệ thống hoạt động ổn định với các Agent và mô hình LLM.
- Tạo được minigame cá nhân hóa sát nội dung học tập.
- Dự đoán GPT-4 và Claude hiệu quả hơn ở vai trò Feedback và Judge [4][5].
- Đề xuất chiến lược gán LLM theo vai trò, mở rộng ứng dụng cho giáo dục đại học Việt Nam.

# Tài liệu tham khảo

- [1] Kiem Hoang, Tin Huynh, Duong Ham Tran, Tri Pham, and Binh Tran. Personalized Learning Through Gamification: Multi-Agent and Large Language Model Approaches. SIU Prize, 2024.
- [2] Atikah Shemshack and Jonathan Michael Spector. A systematic literature review of personalized learning terms. Smart Learning Environments, 7(1):33, 2020.
- [3] James D. Basham, Thomas E. Hall, Richard A. Carter Jr., and William M. Stahl. An operationalized understanding of personalized learning. Journal of Special Education Technology, 31(3):126–136, 2016.
- [4] OpenAI. GPT-4 Technical Report, 2023. <https://openai.com/research/gpt-4>
- [5] Anthropic. Claude 2 Model Overview, 2024. <https://www.anthropic.com>
- [6] Meta AI. LLaMA 3 Model Card, 2024. <https://ai.meta.com/llama>
- [8] Yujie Hong, Nienke Saab, and Wilfried Admiraal. Approaches and game elements used to tailor digital gamification for learning: A systematic literature review. Computers & Education, 105000, 2024.