



IELTS CHATBOT

FINAL PROJECT

CS311 – AI ENGINEERING

Danh sách thành viên:

Huỳnh Đặng Vĩnh Hiền – 21520029

Dương Thị Tú YẾN - 23521846



Mục lục

1. Phân công công việc
2. Tổng quan về đồ án
3. Các tính năng của đồ án
4. Quá trình thu thập dữ liệu
5. Kiến trúc của mô hình
6. Đánh giá hiệu suất
7. Điểm yếu & cải thiện hệ thống
8. Kết luận



Phân công công việc

Công việc	Hiền	Yến
Khảo sát phương pháp	✓	✓
Chuẩn bị dữ liệu	✓	✓
Lập trình frontend		✓
Lập trình backend	✓	
Đánh giá hệ thống	✓	✓
Làm slide	✓	✓



Tổng quan đề án

- Lý do: chứng chỉ IELTS ngày càng trở nên phổ biến và quan trọng. Tuy nhiên, nhiều học viên gặp vấn đề trong việc tìm kiếm tài liệu học phù hợp và hiệu quả.
- Mục tiêu: nhằm phát triển một chatbot hỗ trợ học IELTS, giúp người học dễ dàng tìm kiếm tài liệu học và cung cấp tài liệu PDF ngay trong giao diện chatbot, giúp tiết kiệm thời gian và nâng cao hiệu quả học tập cho người học.



Tổng quan đồ án

- Backend:
 - LlamaIndex: đơn giản, dễ sử dụng, dễ tinh chỉnh
 - LLM: Gemini-flash-1.5
 - Vector search: FAISS
- Frontend:
 - Streamlit: dễ sử dụng, tích hợp trực tiếp với Python, không cần cấu hình phức tạp



Các tính năng của đồ án

- Trò chuyện với chatbot
- Tìm kiếm các bài học theo chủ đề
- Hiển thị tài liệu PDF
- Quản lý hội thoại và lưu trữ



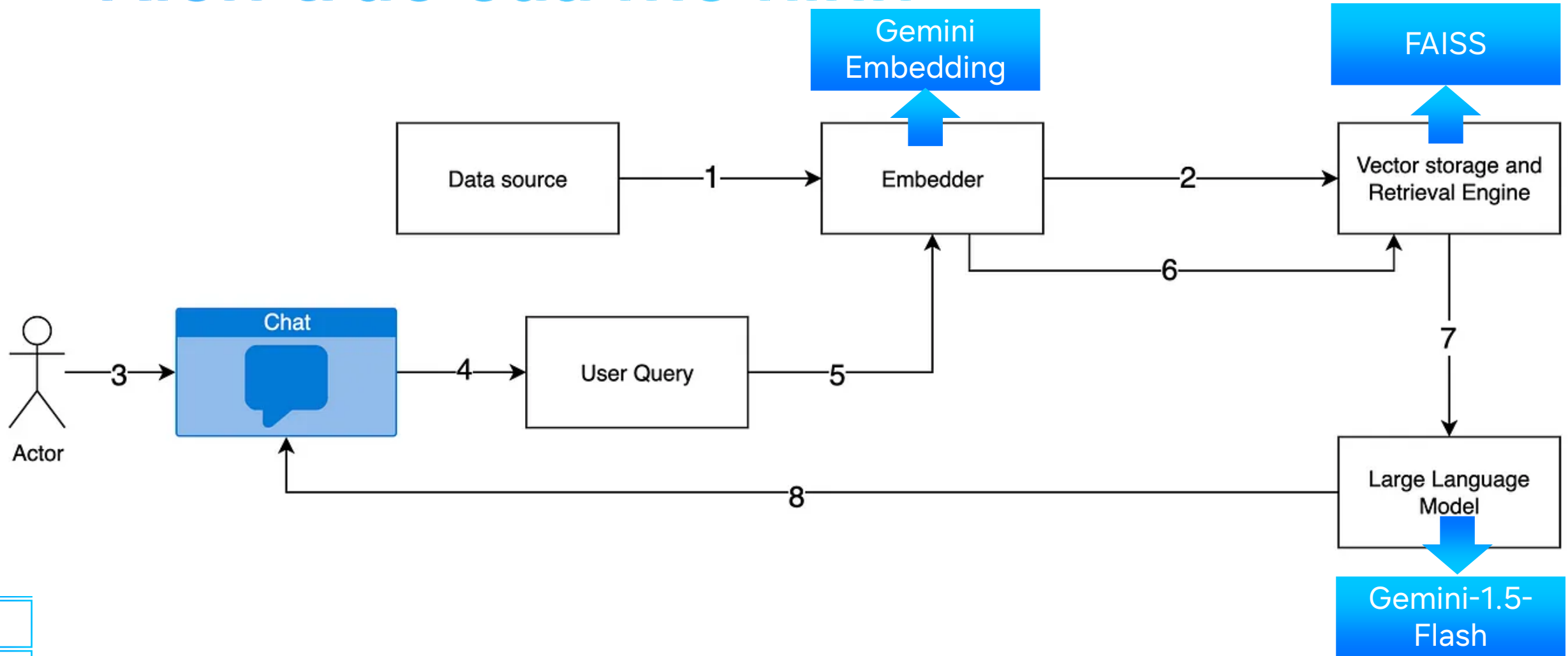
Quá trình thu thập và xử lý dữ liệu

- Bao gồm 2 section là Reading và Speaking
- Nguồn: Cambridge IELTS 16-19 và tự có sẵn
- Sắp xếp & phân loại dữ liệu: theo danh sách topic chuẩn bị trước
- Xử lý PDF
- Tạo dữ liệu JSON

```
{  
  "ID": "1611",  
  "Title": "The Step Pyramid of Djoser",  
  "Pages": "21-22",  
  "Topics": [  
    "History and Culture"  
  ],  
  "Type of Questions": [  
    "Summary Completion",  
    "Multiple Choice Questions"  
  ],  
  "Origin": {  
    "Book": 16,  
    "Test": 1  
  }  
},
```



Kiến trúc của mô hình





Đánh giá hiệu suất

- Do bài toán này có phạm vi rộng, thiếu chặt chẽ, việc đánh giá hiệu suất thông qua các phương pháp truyền thống trở nên thiếu hiệu quả.
- Nhóm thực hiện đề án tiến hành thử nghiệm một số trường hợp phổ biến và tiến hành đánh giá thủ công đề án.
- Kết quả cho thấy mô hình hoạt động tốt ở các trường hợp cơ bản, nhưng dễ lúng túng khi xử lý trường hợp phức tạp.



Đánh giá hiệu suất

Mô hình xử lý tốt

Tôi muốn làm 1 bài IELTS Reading chủ đề Môi trường

Based on the provided text, here's an IELTS Reading passage suggestion on the Environment topic:

1. Title sample: Bats to the rescue
2. Topic: Environment
3. Type of questions: True/False/Not Given, Sentence Completion
4. Origin: Book 17, Test 4, Pages 81-82
5. ID: 1741

READING

READING PASSAGE 1

You should spend about 20 minutes on **Questions 1–13**, which are based on Reading Passage 1 below.

Bats to the rescue

How Madagascar's bats are helping to save the rainforest



Đánh giá hiệu suất

Mô hình xử lý
chưa tốt

Tôi muốn làm 1 bài IELTS Speaking Part 1 chủ đề bất kỳ

The provided text gives example questions for IELTS Speaking Part 1 on several topics. To help you practice, I can offer you a sample Part 1 interview based on one of these topics. Let's choose "Friends." Here's a possible interview, with sample answers to give you an idea of how to respond:

IELTS Speaking Part 1: Friends

Examiner: How important are friends to you?

You: Friends are incredibly important to me. They provide companionship, support, and a sense of belonging. I value the relationships I have with my friends very much, and I feel fortunate to have a strong support network. They help me through difficult times and celebrate with me during happy occasions.



Các vấn đề hiện hữu

- Mô hình chưa nắm bắt được tổng thể các đặt trưng của bài thi IELTS phục vụ cho việc trả lời câu hỏi.
- Mô hình thiếu tính linh động, đôi lúc bị overfit vào dữ liệu được cung cấp ở bước Retrieval.
- Demo chạy còn chậm, thiếu linh hoạt, chưa có tính cạnh tranh.



Hướng phát triển cho mô hình

- Hoàn thiện các phần còn lại của bài thi IELTS, bao gồm Listening và Writing.
- Thêm tính năng chấm bài.
- Cải thiện chất lượng trả về của mô hình.
- Tăng tốc độ khởi động và xử lý của hệ thống.



Kết luận

- Phương pháp RAG thể hiện được tính hiệu quả khi sử dụng trong bối cảnh đồ án.
- Mô hình cần được cải thiện ở nhiều khía cạnh để có khả năng áp dụng trên thực tế.



Hỏi đáp

