**Report week 5**

Contents

[**1.** **Tìm hiểu về LLM** 1](#_Toc203962307)

[**1.2** **Các thành phần cốt lõi của LLM** 2](#_Toc203962308)

[**1.3** **Các loại Mô hình LLM cơ bản** 2](#_Toc203962309)

[**1.4** **Ứng dụng thực tiễn** 2](#_Toc203962310)

[**1.5** **Ưu điểm và hạn chế** 3](#_Toc203962311)

[**1.6** **Kết luận** 3](#_Toc203962312)

[**2.** **Tìm hiểu về Text2SQL** 3](#_Toc203962313)

[**2.1** **Khái niệm cơ bản** 3](#_Toc203962314)

[**2.2** **Kiến trúc và các phương pháp tiếp cận** 3](#_Toc203962315)

[**2.3** **Ứng dụng thực tiễn** 3](#_Toc203962316)

[**2.4** **Các thách thức gặp phải** 4](#_Toc203962317)

[**2.5** **Kết luận** 4](#_Toc203962318)

[**3.** **Tìm hiểu về AI Workflow** 4](#_Toc203962319)

[**3.1** **Khái niệm cơ bản** 4](#_Toc203962320)

[**3.2** **Các bước cơ bản trong AI Workflow** 4](#_Toc203962321)

[**3.3** **Đánh giá tổng quan giữa n8n và Make.com** 4](#_Toc203962322)

[**3.4** **So sánh ưu điểm và nhược điểm** 5](#_Toc203962323)

[**4.** **Hệ thống hỏi đáp giá Coin** 5](#_Toc203962324)

1. **Tìm hiểu về LLM**
   1. **Khái niệm LLM**

LLM(Large Language Model), hay còn được gọi là mô hình ngôn ngữ lớn, là một dạng trí tuệ nhân tạo (AI) tiên tiến được thiết kế để xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP). Đây là những mô hình được huấn luyện trên khối lượng dữ liệu văn bản khổng lồ, bao gồm đa dạng ngôn ngữ, lĩnh vực phong cách viết, và ngữ cảnh sử dụng, từ đó giúp chúng ta hiểu, phân tích và tạo ra ngôn ngữ tự nhiên một cách logic, mạch lạc và gần giống con người

* 1. **Các thành phần cốt lõi của LLM**

Mô hình LLM được cấu thành từ nhiều lớp mạng nơ-tron (neural network) bao gồm:

* Lớp nhúng (Embedding Layer): Đóng vai trò then chốt trong việc tạo ra các biểu diễn nhúng (embedding Layer) từ văn bản đầu vào. Thành phần này của LLM có khả năng nắm bắt cả khía cạnh ngữ nghĩa lẫn cú pháp văn bản, giúp mô hình hiểu rõ ngữ cảnh một cách hiểu quả
* Lớp truyền thẳng (Feedforward – FFN): Bao gồm nhiều lớp được kết nối với nhau có nhiệm vụ biến đổi các biểu diễn nhúng đầu vào. Các lớp này giúp trích xuất các khái niệm trừu tượng ở mức độ cao hơn, hỗ trợ mô hình hiểu được ý định của người dùng từ văn bản đầu vào
* Lớp lặp (Recurrent Layer): Hoạt động bằng cách xử lý tuần tự các từ trong văn bản đầu vào. Nó có khả năng nắm bắt mối quan hệ phức tạp giữa các từ trong câu
* Cơ chế tập trung (Attention Mechanism): Là yếu tố quan trọng giúp mô hình ngôn ngữ tập trung vào các phần cụ thể của văn bản đầu vào có liên quan đến yêu cầu. Lớp này đóng vai trò then chốt trong việc tạo ra kết quả đầu ra có kết quả chính xác
  1. **Các loại Mô hình LLM cơ bản**

Có ba loại mô hình LLM chính:

* Mô hình ngôn ngữ chung (Generic/Raw Language Model): Các mô hình này dự đoán từ tiếp theo dựa trên các mẫu ngôn ngữ đã được học trong quá trình huấn luyện. Chúng thích hợp cho việc truy xuất thông tin
* Mô hình ngôn ngữ điều chỉnh theo hướng dẫn (Instruction – Turned Language Model): Được huấn luyện để dự đoán phản hồi dựa trên các hướng dẫn được cung cấp trong văn bản đầu vào. Khả năng này giúp thực hiện tốt các tác vụ như phân tích cảm xúc, tạo văn bản hoặc mã nguồn
* Mô hình điểu chỉnh theo hội thoại (Dialog-Turned Language Model) : Được thiết kế đặc biệt để tham gia vào các cuộc hội thoại và dự đoán phản hồi tiếp theo trong cuộc trò chuyện. Ví dụ điển hình là chatbot hoặc các hệ thống trò chuyện AI
  1. **Ứng dụng thực tiễn**

LLM được ứng dụng trong nhiều lĩnh vực khác nhau như

* Chatbot và trợ lý ảo
* Tạo nội dung tự động (Content Generation)
* Dịch ngôn ngữ và đa ngôn ngữ hóa nội dung
* Viết và kiếm tra mã lập trình (Code Assistant)
* Phân tích dữ liệu và trích xuất thông tin
* Giáo dục và học tập cá nhân hóa
  1. **Ưu điểm và hạn chế**
* Ưu điểm:
* Khả năng hiểu ngôn ngữ tự nhiên rất tốt
* Linh hoạt trong nhiều tác vụ mà không cần huấn luyện lại
* Có thể sinh nội dung mới, sáng tạo
* Hạn chế:
* Rủi ro về bảo mật
* Độ phức tạp khi triển khai
* Nguy cơ thông tin có thể sai lệch
  1. **Kết luận**

LLM là một bước đột phá trong lĩnh vực trí tuệ nhân tạo, mở ra nhiều cơ hội ứng dụng rộng khắp. Tuy nhiên, cần sử dụng một cách có trách nhiệm, kết hợp kiểm tra đầu ra và cải tiến mô hình để hạn chế các rủi ro đi kèm.

1. **Tìm hiểu về Text2SQL**
   1. **Khái niệm cơ bản**

Text2SQL là bài toán trong lĩnh vực NLP với mục tiêu chuyển đổi câu hỏi/ ngôn ngữ tự nhiên (Natural Language) thành câu truy vấn cơ sở dữ liệu (SQL) tương ứng.

* 1. **Kiến trúc và các phương pháp tiếp cận**

Hệ thống Text2SQL điển hình gồm 3 bước:

* Hiểu ngữ nghĩa câu hỏi (Intent and Entiry Extraction)
* Ánh xạ từ cấu trúc ngôn ngữ sang cấu trúc bảng và cột
* Sinh ra truy vấn SQL hợp lệ tương ứng

Các phương pháp tiếp cận phổ biến như là:

* Rule-Based: Sử dụng luật định nghĩa trước để ánh xạ từ ngôn ngữ sang SQL
* Template-based: Dựa trên mẫu truy vấn có sẵn và điền thông tin phù hợp
* Học sâu (Deep Learning): Dùng mô hình Seq2Seq, Tranformer hoặc LLM để ánh xạ NL sang SQL
* Neural-Symbolic: Kết hợp biểu diễn ngữ nghĩa và kiến trúc logic/cấu trúc
  1. **Ứng dụng thực tiễn**

Các ứng dụng của text2sql được sử dụng trong thực tế như:

* Dashboard tự động: Tạo các biều đồ từ ngôn ngữ
* Trợ lý truy vấn dữ liệu trong CRM/ERP
* Chatbot hỗ trợ phân tích dữ liệu trong ngân hàng, bán lẻ, giáo dục
* Công cụ dạy học giúp sinh viên hiểu cấu lệnh SQL từ câu hỏi
  1. **Các thách thức gặp phải**

**Có nhiều thách thức của text2sql gặp phải như:**

* **Đa nghĩa trong ngôn ngữ tự nhiên: Cùng một câu hỏi có thể có nhiều cách hiểu khác nhau**
* **Cơ sở dữ liệu đa dạng: Tên cột, bảng khác nhau trong mỗi hệ thống**
* **Câu hỏi phức tạp: Có nhiều điều kiện, phép join, nhóm, sắp xếp…**
* **Kiếm tra tính đúng đắn cú pháp và ngữ nghĩa của SQL sinh ra**
  1. **Kết luận**

Text-to-SQL là một lĩnh vực đang phát triển mạnh, đóng vai trò quan trọng trong việc **dân chủ hóa truy vấn dữ liệu**. Với sự kết hợp của các mô hình ngôn ngữ lớn và kỹ thuật học sâu, các hệ thống này đang ngày càng tiến gần đến khả năng hiểu và xử lý câu hỏi như con người.

1. **Tìm hiểu về AI Workflow** 
   1. **Khái niệm cơ bản**

AI Workflow là quy trình làm việc để phát triển và triển khai mô hình AI, từ khâu thu thập dữ liệu đến khi đưa mô hình vào thực tế. Nó bao gồm nhiều bước chặt chẽ với nhau nhằm đảm bảo rằng AI hoạt động hiểu quả, đáp ứng mục tiêu đề ra

* 1. **Các bước cơ bản trong AI Workflow**

Các bước cơ bản trong AI Workflow:

* Xác định vấn đề: xác định rõ vấn đề cần giải quyết và mục đích của mô hình AI
* Thu thập và xử lý dữ liệu: Thu thập, làm sạch và gán nhãn cho dữ liệu
* Xây dựng mô hình (Model Building): Thiết kế mô hình học máy dựa trên dữ liệu đã xử lý
* Huấn luyện mô hình (Training): Bước quan trọng để mô hình học từ dữ liệu
* Đánh giá và cải thiện : Đảm bảo mô hình đạt hiểu suất cao với dữ liệu thực tế
* Triển khai: Hoàn thiện và triển khai trong thực tế
* Giám sát và bảo trì : Theo dõi hiệu suất của mô hình sau khi triển khai
  1. **Đánh giá tổng quan giữa n8n và Make.com**

Make.com và n8n là những công cụ tự động hóa no-code hiểu quả giúp các doanh nghiệp tiết kiệm thời gian để tập trung vào các hoạt động chiến lược

N8n được đánh giá cao vì mã nguồn mở, mang đến sự linh hoạt và khả năng tùy chỉnh cao. Trong khi đó, Make.com tập trung vào các trải nghiệm của người dùng nên ưu tiên sự tiện lợi, dễ sử dụng và giao diện bắt mắt

* 1. **So sánh ưu điểm và nhược điểm**

|  |  |  |
| --- | --- | --- |
|  | **Make.com** | **n8n** |
| **Ưu điểm** | * Nền tảng no-code có giao diện trực quan * Thao tác kéo và thả dễ sử dụng * Mô hình hóa workflow bằng sơ đồ tư duy * Tích hợp nhiều dịch vụ đơn giản, nhanh chóng * Giá cả phải chăng có gói miễn phí * Có cộng đồng người sử dụng đông đảo, nhiệt tình | * Mã nguồn mở, hỗ trở self-host * Khả năng tùy chỉnh đa dạng, linh hoạt * Cho phép nhúng code * Xử lý tốt các work flow phức tạp * Độ bảo mật cao |
| **Nhược điểm** | * Một số thông báo lỗi có thể gây khó hiểu * Nếu Workflow bị lỗi bất cứ điểm nào thì sẽ dừng cả workflow * Tốc độ đôi khi chậm khi xử lý các file lớn hoặc quy trình làm việc phức tạp * Dịch vụ hỗ trợ khách hàng chưa ấn tưởng | * Khó làm chủ hơn các công cụ khác * Cách cài bản self-host hơi phức tạp * Bản cloud không miễn phí * Yêu cầu đội ngũ kỹ thuật có chuyên môn và hiểu code |

Kết luận :

* Make.com dễ sử dụng, phù hợp với những người chưa có kinh nghiệm muốn tìm giải pháp tự động hóa no-code có giao diện trực quan, đơn giản mà hiểu quả
* n8n phù hợp với đội ngũ có chuyên môn kỹ thuật và thích làm việc với các thiết lập phức tạp

1. **Hệ thống hỏi đáp giá Coin**

Hệ thống sử dụng bao gồm:

* Ngôn ngữ python
* Thư viện Flask: Để xây dựng ứng dụng cho web nhẹ
* Html/css: Xây dựng giao diện web
* Tập dữ liệu: chứa dữ liệu giá lịch sử của 15 đồng coin từ 2020 đến nay

Demo:

A screenshot of a computer

AI-generated content may be incorrect.