

Tóm tắt nội dung tuần 6 của khóa học về LLM

I. Tóm tắt nội dung tuần 6 của khóa học về LLM

Ngày 1: Fine-Tuning Large Language Models - Từ Suy Luận Đến Huấn Luyện

A. Nội dung chính

- Giới thiệu về quá trình chuyển từ suy luận (inference) sang huấn luyện (training).
- Tính quan trọng của dữ liệu trong huấn luyện mô hình.
- Đánh giá sự thành công trong huấn luyện LLM.
- Chuẩn bị dữ liệu: làm sạch, tổ chức, tiền xử lý.
- Tổng quan về Transfer Learning và Fine-Tuning trong huấn luyện mô hình.
- Các bước cài đặt môi trường huấn luyện mô hình LLM.

B. Kỹ năng đạt được

- Hiểu cách đánh giá hiệu quả huấn luyện mô hình.
- Cải thiện quy trình chuẩn bị dữ liệu.
- Nhận biết các phương pháp Fine-Tuning cho mô hình Frontier.
- Cài đặt và sử dụng môi trường lập trình phù hợp với huấn luyện mô hình LLM.

Ngày 2: LLM Deployment Pipeline - Từ Bài Toán Kinh Doanh Đến Giải Pháp Triển Khai

A. Nội dung chính

- Quy trình 5 bước triển khai mô hình LLM trong doanh nghiệp.
- So sánh ba kỹ thuật tối ưu hóa mô hình: Prompting, RAG và Fine-Tuning.
- Xây dựng bản mặt (baseline) cho việc đánh giá hiệu suất mô hình.
- Cách đo lường và theo dõi hiệu suất của mô hình trong môi trường thực tế.
- Thực hành đo lường hiệu suất mô hình bằng cách theo dõi thời gian phản hồi và độ chính xác trong việc hiểu yêu cầu của khách hàng

B. Kỹ năng đạt được

- Lên chiến lược triển khai mô hình LLM trong doanh nghiệp.
- Lựa chọn kỹ thuật tối ưu hóa phù hợp cho bài toán cụ thể.
- Phân tích và điều chỉnh mô hình dựa trên phản hồi từ người dùng thực tế
- Tích hợp LLM vào các hệ thống hiện có

Ngày 3: Feature Engineering và Bag of Words - Xây Dựng Baseline ML Cho NLP

A. Nội dung chính

- Tìm hiểu về feature engineering trong NLP.
- Ứng dụng Bag of Words trong bài toán xây dựng baseline.
- So sánh các mô hình ML truyền thống với mô hình Frontier LLM.
- Thực hành xây dựng các đặc trưng (features) hữu ích cho mô hình NLP
- Chạy thử nghiệm với các mô hình baseline và đánh giá kết quả.

****B. Kỹ năng đạt được****

- Thiết kế pipeline xây dựng baseline cho NLP.
- So sánh hiệu quả ML truyền thống và Frontier LLM.
- Xây dựng các đặc trưng đầu vào tối ưu cho mô hình LLM
- Đánh giá mức độ ảnh hưởng của từng đặc trưng đến kết quả dự đoán

Ngày 4: Đánh Giá Mô Hình Frontier - So Sánh Hiệu Suất

A. Nội dung chính

- Kiểm tra hiệu suất giữa mô hình ML truyền thống và Frontier LLM.
- So sánh hiệu quả giữa AI và con người trong dự đoán giá sản phẩm.
- Các chỉ số đánh giá mô hình như Accuracy, Perplexity, BLEU, ROUGE
- Cách tối ưu hóa mô hình dựa trên kết quả đánh giá

B. Kỹ năng đạt được

- Thiết lập quy trình đánh giá mô hình Frontier.
- Sáng tạo bản mặt kiểm thử (benchmark) cho AI.
- Phân tích kết quả đánh giá để cải thiện mô hình
- Tối ưu hóa mô hình dựa trên phản hồi từ thực nghiệm

Ngày 5: Fine-Tuning LLMs với OpenAI - Chuẩn Bị Dữ Liệu, Huấn Luyện và Đánh Giá

A. Nội dung chính

- Quy trình Fine-Tuning mô hình GPT-4.
- Chuẩn bị dữ liệu để Fine-Tuning.
- Đánh giá hiệu quả Fine-Tuning so với baseline ban đầu.
- Các lỗi phổ biến khi Fine-Tuning mô hình và cách khắc phục

B. Kỹ năng đạt được

- Cách chuẩn bị JSONL data cho OpenAI.
- Thực hành Fine-Tuning mô hình Frontier.
- Tối ưu hóa hyperparameters để đạt hiệu suất tốt nhất
- Hiểu rõ tác động của Fine-Tuning lên hiệu suất mô hình

II. Từ Khóa Quan Trọng Cho Nghiên Cứu LLM

1.Fine-Tuning

- Quá trình tinh chỉnh một mô hình AI có sẵn bằng cách huấn luyện lại trên tập dữ liệu cụ thể. Giúp mô hình thích nghi với một nhiệm vụ riêng biệt thay vì phải huấn luyện từ đầu.

2.LoRA (Low-Rank Adaptation)

- Phương pháp Fine-Tuning giúp giảm đáng kể số lượng tham số cần điều chỉnh, tiết kiệm bộ nhớ và tài nguyên tính toán mà vẫn đảm bảo hiệu suất mô hình.

3.Transfer Learning

- Kỹ thuật sử dụng một mô hình đã được huấn luyện trên tập dữ liệu lớn và tiếp tục huấn luyện nó với một tập dữ liệu nhỏ hơn nhưng chuyên biệt hơn, giúp tiết kiệm thời gian và tài nguyên.

4.Prompt Engineering

- Nghệ thuật thiết kế câu lệnh đầu vào cho mô hình ngôn ngữ để có được phản hồi chính xác và tối ưu nhất. Đây là kỹ thuật quan trọng khi sử dụng mô hình LLM mà không cần Fine-Tuning.

5.RAG (Retrieval-Augmented Generation)

- Phương pháp kết hợp giữa tìm kiếm thông tin từ nguồn dữ liệu bên ngoài và sinh văn bản để cải thiện tính chính xác và độ tin cậy của kết quả đầu ra.

6. Dataset Curation

- Quá trình thu thập, làm sạch, xử lý và tổ chức dữ liệu để mô hình AI có thể học hiệu quả hơn. Dữ liệu chất lượng cao quyết định trực tiếp đến hiệu suất của mô hình.

7. Model Evaluation Metrics

- Các chỉ số đánh giá chất lượng mô hình AI như Perplexity (độ khó dự đoán), BLEU (độ chính xác so với dữ liệu tham chiếu), ROUGE (đánh giá tóm tắt văn bản).

8. Hyperparameter Tuning

- Điều chỉnh các tham số quan trọng trong quá trình huấn luyện như learning rate, batch size để tối ưu hóa hiệu suất của mô hình AI.

III. Các Công Nghệ Được Đề Cập Trong Tuần 6

1. OpenAI API

- Được sử dụng trong tuần 6 để Fine-Tuning mô hình GPT-4.
- *Ứng dụng trong tuần 6:* Chạy Fine-Tuning trên mô hình GPT-4 bằng cách tải lên tập dữ liệu JSONL.
- *Vai trò:* Cung cấp nền tảng mạnh mẽ để huấn luyện và triển khai mô hình mà không cần phần cứng riêng.
- *Thực hành:* Huấn luyện một mô hình chuyên biệt cho dự đoán giá sản phẩm dựa trên dữ liệu đầu vào.

2. Hugging Face Transformers

- Sử dụng trong tuần 6 để kiểm thử và đánh giá mô hình ngôn ngữ.
- *Ứng dụng trong tuần 6:* Triển khai baseline NLP với các mô hình BERT và GPT có sẵn.
- *Vai trò:* Cho phép tải các mô hình mã nguồn mở để thực hiện so sánh hiệu suất giữa mô hình truyền thống và Frontier.
- *Thực hành:* Chạy thử nghiệm baseline ML và so sánh với mô hình LLM đã Fine-Tuning.

3.LangChain

- Sử dụng để xây dựng hệ thống truy xuất và tổng hợp thông tin từ nhiều nguồn.
- *Ứng dụng trong tuần 6:* Thực hiện thử nghiệm RAG (Retrieval-Augmented Generation) để nâng cao độ chính xác của mô hình LLM.
- *Vai trò:* Hỗ trợ mô hình trong việc lấy thông tin từ các cơ sở dữ liệu bên ngoài.
- *Thực hành:* Tạo pipeline kết hợp giữa truy xuất dữ liệu và sinh văn bản để nâng cao độ chính xác của mô hình.

4.Scikit-Learn & Pandas

- Được sử dụng để tiền xử lý dữ liệu và xây dựng baseline ML.
- *Ứng dụng trong tuần 6:* Tạo baseline ML để so sánh với mô hình Frontier.
- *Vai trò:* Hỗ trợ tiền xử lý dữ liệu văn bản trước khi huấn luyện mô hình.
- *Thực hành:* Áp dụng feature engineering để trích xuất đặc trưng từ dữ liệu văn bản.