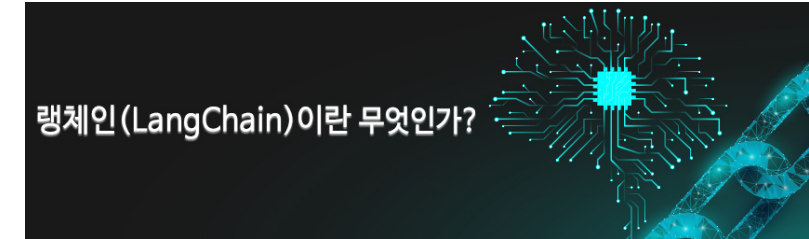


인공지능

# 랭체인(LangChain)이란 무엇인가?

2023-09-18 Janakiram MSV



이 글은 IDG의 기사를 전재하여 제공합니다.  
[원문보기] : <https://www.ciokorea.com/column/305341>

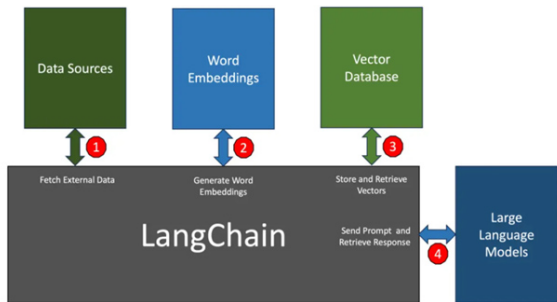
랭체인(LangChain)은 노출하여 대규모 언어 모델과 애플리케이션의 통합을 간소화하는 SDK이다. 대규모 언어 모델의 부상과 함께 중요성이 커지고 있다. 대규모 언어 모델에 대한 최신 트렌드를 쫓는 소프트웨어 개발자라면 연일 쏟아지는 소식이 당황스러울 수 있다. 새로운 오픈소스 모델의 출시나 상용 모델 제공업체의 중요한 새 기능 발표가 매일 등장하는 상황과 비슷하다고 할 수 있다.

어느덧 LLM은 최신 소프트웨어 스택의 한 요소로 빠르게 부상했다. 그러나 오픈AI와 같은 공급업체가 제공하는 모델 API를 사용하든, 오픈소스 모델을 앱에 포함하든 LLM 기반 애플리케이션을 구축하려면 단순히 프롬프트를 보내고 응답을 기다리는 것 이상의 작업이 필요하다. 매개변수 조정부터 프롬프트 보강, 응답 조정에 이르기까지 고려해야 할 요소가 많다. 이를테면 LLM은 상태를 저장하지 않으므로 대화의 이전 메시지를 기억하지 못한다. 기록을 유지하고 LLM에 컨텍스트를 제공하는 것은 개발자의 책임이다. 맥락 정보가 새로운 대화에서 컨텍스트를 다시 가져오기 위해 영구 데이터베이스에 저장되어야 할 수도 있다. 이렇듯 LLM에 단기 및 장기 메모리를 추가하는 것은 개발자의 주요 책임 중 하나이다. 또 다른 문제는 LLM에 대한 일률적인 규칙이 없다는 것이다. 감정 분석, 분류, 질문 답변과 요약 등 다양한 시나리오에 특화된 여러 모델을 사용해야 할 수 있다. 여러 LLM을 처리하는 작업은 복잡하며 상당한 노력을 필요로 한다.

## LLM 앱을 구축하기 위한 통합 API 레이어

랭체인은 LLM과 애플리케이션의 통합을 간소화하도록 설계된 SDK로서 앞서 설명한 대부분의 문제를 해결한다. 랭체인은 ODBC, 또는 표준 SQL 문에 집중하게 함으로써 백엔드 데이터베이스의 구현 세부 정보를 요약하는 JDBC 드라이버와 비슷하다고 할 수 있다. 랭체인은 간단하고 통합된 API를 노출하여 기본 LLM의 구현 세부 사항을 요약하는데, 이 API를 통해 개발자들은 코드를 크게 변경하지 않고 모델을 쉽게 교체하거나 대체할 수 있다.

랭체인은 챗GPT와 거의 같은 시기에 등장했다. 개발자인 해리스 체이스는 2022년 10월 말, LLM 열풍이 일어나기 시작한 시점에 랭체인을 처음 선보였다. 그 이후로 커뮤니티 구성원들이 적극적으로 기여했으며, 이로 인해 랭체인은 LLM과 상호 작용하는 탁월한 도구로 부상했다. 랭체인은 외부 도구와 통합되어 환경을 조성하는 강력한 프레임워크이다. 이제 이 프레임워크가 어떻게 LLM에서 원하는 결과를 얻기 위한 흐름을 조율하는지 살펴본다.



[그림 1] 랭체인 흐름도

### ① 데이터 소스

애플리케이션이 LLM에 대한 컨텍스트를 구축하기 위해 PDF, 웹 페이지, CSV, 관계형 데이터베이스와 같은 외부 소스에서 데이터를 검색해야 하는 경우가 있다. 랭체인은 서로 다른 소스에서 데이터에 액세스하고 검색할 수 있는 모듈과 원활하게 통합된다.

② 단어 임베딩

일부 외부 소스에서 검색된 데이터는 벡터로 변환되어야 한다. 그래야 텍스트를 LLM과 관련된 단어 임베딩 모델에 전달하게 된다. 실제로 오픈AI의 GPT-3.5 모델은 컨텍스트를 전송하는 데 사용해야 하는 단어 임베딩 모델과 관련되어 있다. 랭체인은 선택한 LLM을 기반으로 최적의 임베딩 모델을 선택한다.

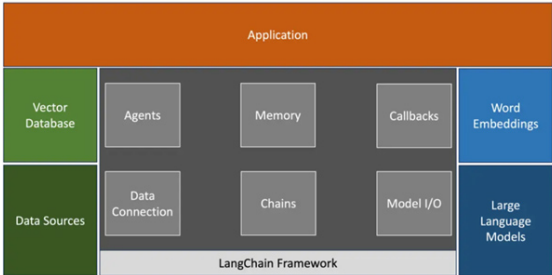
③ 벡터 데이터베이스

생성된 임베딩은 유사성 검색을 위해 벡터 데이터베이스에 저장된다. 랭체인은 메모리 내 배열부터 파인콘(Pinecone)과 같은 호스팅 벡터 데이터베이스에 이르기까지 다양한 소스에서 벡터를 쉽게 저장하고 검색할 수 있도록 지원한다.

④ 언어 모델(LLM)

랭체인은 오픈AI, 코히어(Cohere), AI21에서 제공하는 주류 LLM과 허깅페이스(Hugging Face)에서 제공되는 오픈소스 LLM을 지원한다. 지원되는 모델과 API 엔드포인트 목록은 빠르게 증가하고 있다.

[그림 2]는 랭체인 프레임워크의 핵심을 나타낸다. 스택 상단의 애플리케이션은 파이썬 또는 자바스크립트 SDK를 통해 여러 랭체인 모듈 중 하나와 상호 작용한다. 여러 모듈의 역할을 살펴보면 다음과 같다.



[그림 2] 랭체인 프레임워크

에이전트

에이전트는 랭체인에서 아주 강력한 모듈이다. LLM은 추론과 행동이 가능한데, 이는 ReAct 프롬프트 기법이라고 불린다. 랭체인의 에이전트는 LLM을 사용하여 프롬프트를 행동 계획으로 추출하는 ReAct 프롬프트 제작을 단순화한다. 에이전트는 일종의 동적 체인으로 생각할 수 있다. 에이전트의 기본 일련의 동작을 선택하기 위해 LLM을 사용하는 것이다. 동작의 순서는 체인(코드)으로 하드 코딩된다. 언어 모델은 에이전트 내에서 추론 엔진으로 사용되어 어떤 순서로 어떤 동작을 취할지 결정한다.

메모리

LLM은 상태 비저장형이지만 정확한 응답을 위해 컨텍스트가 필요하다. 랭체인의 메모리 모듈은 모델에 단기 및 장기 메모리를 쉽게 추가할 수 있도록 도와준다. 단기 메모리는 간단한 메커니즘을 통해 대화의 기록을 유지한다. 메시지 기록은 레디스(Redis)와 같은 외부 소스에 저장되어 장기 메모리를 유지할 수 있다.

콜백

랭체인은 개발자에게 LLM 애플리케이션의 다양한 단계에 연결할 수 있는 콜백 시스템을 제공한다. 이는 로깅, 모니터링, 스트리밍 및 기타 작업에 유용하다. 파이프라인 내에서 특정 상황이 발생할 때 호출되는 사용자 지정 콜백 핸들러를 작성할 수 있게 해준다. 랭체인의 기본 콜백은 모든 단계의 출력을 콘솔에 간단히 인쇄하는 stdout을 가리킨다.

데이터 연결

데이터 연결 모듈을 LLM 애플리케이션의 ETL 파이프라인에 해당한다. 이 모듈은 PDF 또는 엑셀 파일과 같은 외부 문서를 로드하고, 이를 처리하기 위해 일괄적으로 단어 임베딩으로 변환한 다음, 임베딩을 벡터

데이터베이스에 저장하고, 마지막으로 쿼리를 통해 검색한다. 앞서 설명한 바와 같이, 이는 랭체인의 가장 중요한 구성 요소이다.

체인

LLM과 상호 작용하는 것은 유닉스 파이프라인을 사용하는 것과 많은 면에서 유사하다. 한 모듈의 출력이 다른 모듈에 입력으로 전송된다. 개발자는 종종 원하는 결과를 얻을 때까지 LLM을 사용해 응답을 명확하게 하고 요약해야 한다. 랭체인의 체인은 구성 요소와 LLM을 활용하여 예상되는 응답을 얻는 효율적인 파이프라인을 구축하도록 설계됐다. 간단한 체인에는 프롬프트와 LLM이 포함될 수 있지만 재귀와 같이 LLM을 여러 번 호출하여 결과를 얻는 등 매우 복잡한 체인을 구축할 수도 있다. 예를 들어, 체인은 문서를 요약한 다음 이에 대한 감정 분석을 수행하는 프롬프트가 포함될 수 있다.

모델 I/O

모델 I/O 모듈은 LLM과의 상호 작용을 다룬다. 이는 기본적으로 효과적인 프롬프트를 생성하고, 모델 API를 호출하고, 결과 해석을 돕는다. 생성형 AI의 핵심인 프롬프트 엔지니어링이 랭체인에서 잘 처리된다. 이 모듈은 LLM 제공자가 노출하는 인증, API 매개변수, 엔드포인트를 요약한다. 이 모듈은 마지막으로, 모델에서 보낸 응답을 애플리케이션서 사용할 수 있는 원하는 형식으로 해석하는 작업을 돕는다.

어느덧 랭체인은 생성형 AI 기반 애플리케이션에서 중요한 구성 요소로 빠르게 자리 잡고 있다. 계속해서 발전 중인 환경 덕분에 다양한 구성 요소를 지원할 수 있다. 오픈소스 및 상용 LLM, 벡터 데이터베이스, 데이터 소스 및 임베딩에 대한 지원은 랭체인을 개발자에게 필수적인 도구로 만들고 있다.



- ▶ 해당 콘텐츠는 저작권법에 의하여 보호받는 저작물로 기고자에게 저작권이 있습니다.
- ▶ 해당 콘텐츠는 사전 동의 없이 2차 가공 및 영리적인 이용을 금하고 있습니다.

#인공지능 #AI #랭체인 #LANGCHAIN #생성형AI #GENERATIVEAI #언어모델 #LLM #벡터데이터베이스



이 글이 좋으셨다면 구독&좋아요

여러분의 “구독”과 “좋아요”는 저자에게 큰 힘이 됩니다.

구독하기



Janakiram MSV

InfoWorld

[모든 글 보기](#)

< [목록보기](#)

## 관련 아티클

글로벌 AI 리더들이 전하는 생성형 AI 기술의 전망

'생성형 AI 중심'으로 변화하는 B2B 기업

기업의 업무 자동화 RPA, LLM으로 환골탈태하다

생성형 AI 모델과 대화하는 프롬프트 엔지니어링(Prompt Engineering)

웹, 앱에서 생성형 AI로 인터넷의 진화