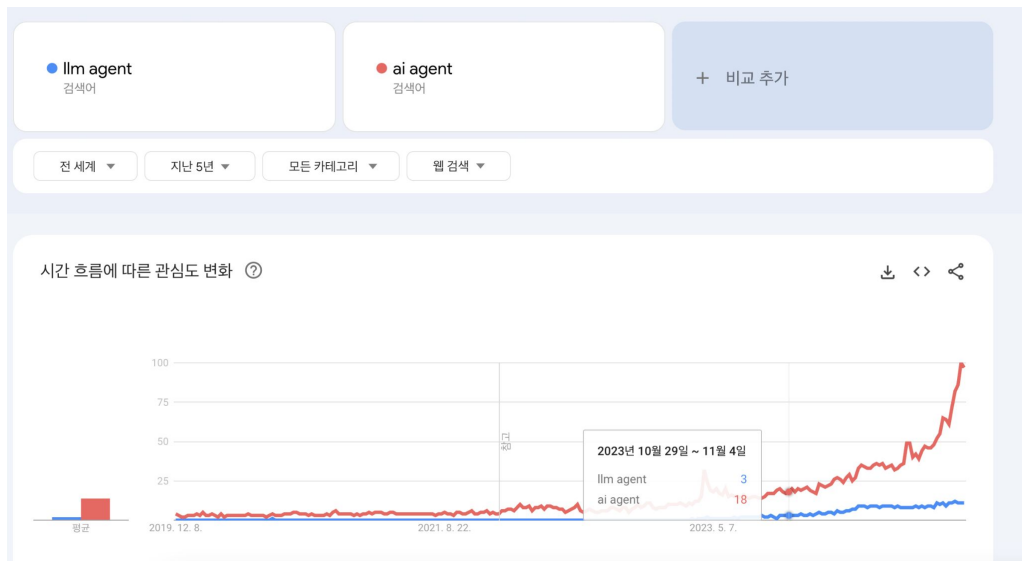




LLM Agent 튜아보기

최유진

Agent 관련 키워드 살펴보기



- 2023년 부터 본격적으로 ai agent에 대한 관심도가 올라감.
- 2024년 llm agent 관련 키워드도 차츰 증가하는 것을 볼 수 있음
- 2022년 말 GPT-30이 출시되었고 GPT-3.5가 2023년 5월에 출시됨.



스터디 책에서 말하는 Agent

- 주변 환경을 감각을 통해 인식하고 의사결정을 내려 행동하는 인공지능 개체
- 에이전트 구성 요소
 - **감각**을 통해 외부 환경과 사용자의 요청을 인식
 - **두뇌**를 통해 가지고 있는 지식이나 지금까지의 기억을 확인해보고 계획을 세우거나 추론을 통해 다음에 어떤 행동을 해야 할지 의사결정
 - 문제를 해결하기 위해 취할 수 있는 적절한 도구를 선택해 **행동**



스터디 책에서 말하는 Agent

- 에이전트의 두뇌
 - 감각을 통해 현재 상황과 사용자의 요청을 인식한 것을 바탕으로
 - 사용자의 요청이 무엇인지 현재 어떤 상황인지 이해하고
 - 목표 달성을 위해 어떤 행동을 취할지 결정한다.
- 에이전트의 감각
 - LLM에서는 기본적으로 텍스트를 입력으로 받고 텍스트를 출력으로 내보낸다.
- 에이전트의 행동
 - 특정 도구를 어떨 때 사용하고 어떻게 사용하는지 설명을 제공하면
 - LLM은 상황에 맞춰 필요한 도구를 선택할 수 있다.



스터디 책에서 말하는 Agent

- 단일 에이전트
 - 모든 과정을 스스로 처리한다.
 - 범용적인 작업을 처리할 수 있는 프롬프트로 동작하기 때문에 다양한 작업에 사용할 수 있다.
 - 단점으로는 작업을 수행하는 과정에서 길을 잃을 가능성이 있다.
 - 목표가 명확하고 작업의 크기가 작은 경우에 선택적으로 활용하는 것이 좋다.
- 멀티 에이전트
 - 여러 에이전트의 협력을 통해 문제를 해결하는 방식
 - 각 에이전트마다 서로 다른 프로필을 주고 작업을 수행하는데, 에이전트 간의 협력으로 작업 품질 향상
 - 구체적이고 세분화된 작업을 잘 수행하고 역할이 부여되면 더 성능이 높아진다.
 - 맞춤형 에이전트: 수평형 대화, 위계형 대화



스터디 책에서 말하는 Agent

- 에이전트 평가하기
 - 튜링 테스트: 에이전트와 사람의 결과물을 구분할 수 있는지 확인하는 평가방식
 - 평가자 LLM을 활용해 평가하는 방식
 - 벤치마크 데이터셋을 이용해 객관적으로 평가하는 방식
 - **유용성: 작업 성공률 같은 기준으로 평가**
 - **사회성:** 언어를 얼마나 숙련되게 사용하는가, 협력*협상 시 얼마나 뛰어난지, 역할에 얼마나 부합하게 행동하는지
 - **가치관:** 신뢰할 수 있는 정확한 정보를 전달하는지, 차별*편향 여부, 사회에 해가 될 수 있는 정보 전달하는지
 - **진화능력:** 변화하는 상황에 맞추어 스스로 발전하는 것, 지속적인 학습, 스스로 목표를 설정하고 달성하는 학습 능력, 환경에 적응하는 능력



LLM Agent 관련 사례

- 에이닷
 - 뮤직 에이전트
 - 사용자가 텍스트 또는 음성으로 쉽게 음악을 찾을 수 있고 음악을 고르기 귀찮다면 취향에 맞게 플레이리스트를 편집할 수 있음
 - Log Data 기반으로 범용 LLM으로 자동화하여 지속적으로 업데이트를 시킴.
 - <https://devocean.sk.com/blog/techBoardDetail.do?ID=166281&boardType=techBlog>
 - 영화 예매 에이전트
 - 대화로 쉽게 편리하게 영화를 예매할 수 있도록 도와줌
 - <https://devocean.sk.com/blog/techBoardDetail.do?ID=166979&boardType=techBlog>
- 코르카의 Multi Agent architecture로 Customer Service 문제 해결
 - 사용자 질의를 분석하여 카테고리화 질문 추출해 QA 페어를 Search를 해서 해당 지식 기반으로 응답
 - [관련 링크](#)



LLM 에이전트 오픈소스

- AutoGen - Microsoft
 - 원하는 작업을 성공적으로 실행하기 위해 서로 대화할 수 있는 여러 에이전트를 사용하여 LLM의 기능을 활용하는 응용 프로그램을 만들 수 있는 기능을 개발자에게 제공하는 오픈소스 프레임워크
 - <https://github.com/microsoft/autogen>
- Auto-GPT
 - GPT-4와 같은 언어 모델을 기반으로 하는 자동화된 에이전트 시스템
 - <https://github.com/Significant-Gravitas/AutoGPT>



LLM Agent 관련 연구

- LG AI 연구원 LLM 에이전트 연구
 - [MultiPrompter: Cooperative Prompt Optimization with Multi-Agent Reinforcement Learning](#)
 - 파운데이션 모델을 사용할 때마다 사람이 매번 프롬프트 최적화를 하는 것이 아니라 자동으로 최적의 프롬프트를 찾아줌.
 - [Prospector: Improving LLM Agents with Self-Asking and Trajectory Ranking](#)
 - 기존 LLM 에이전트는 주어진 과제를 완수하기 위한 **Action** 선택과 평가가 제한적인데 이러한 문제를 해결하기 위해 LLM Actor와 LLM Critic 두 개의 상호 보완적인 LLM으로 구성된 에이전트를 제안
- 마이크로소프트리서치 아시아
 - [Alympics: Language Agents Meet Game Theory](#)
 - 게임이론 시나리오를 통해 LLM 에이전트들의 전략적 결정을 탐구할 수 있는 시뮬레이션



개인 프로젝트 소개

- AutoDraft
 - 블로그 작성을 도와주는 에이전트
- 선정배경
 - 최근 개발 분야에서도 자신을 PR 해야 할 필요성이 높아지고 있다. 이를 위해서는 다양한 SNS를 통해 활동하게 되는데 지속적으로 콘텐츠를 새롭게 작성하는 것이 쉽지 않다.
 - 이를 위해 초안을 넣어주면 내용을 완성해주고 작성에 도움이 되는 추가 자료 역시 제공하고자 한다.
- 완성 기한(~16-17일로 예상)
 - 12.17 Berkeley LLM Agent Hackathon 마감일
- 공유 코드: <https://github.com/yujin37/AutoDraft>



AutoDraft 기능

- 글 생성
- 해시태그 추천
- 제목 추천
- 요약 생성



개발 상황

- 글 생성
 - 사용자의 블로그 글을 분석하여 주요 특징들을 정리
 - 생성하고자 하는 글의 초안을 넣어주면 이전 작성 글들을 바탕으로 작성해줌
- 제목 추천(영어 **only**)
 - 완성된 글을 바탕으로 제목을 생성해줌
 - 블로그나 그외 작성 시 제목으로 활용할 수 있게 함.
 - 한국어 추가 학습을 시킬 계획



개발 상황

- 해시태그 추천
 - 관련 논문, 코드 탐색 중
 - 텍스트를 넣으면 특징을 뽑아내거나 혹은 인기있는 해시태그를 활용하는 방식으로 생각 중
- 요약 생성
 - AI HUB에 있는 요약 생성 데이터를 활용할 계획
- 웹 페이지(모델 구현을 위한...)
 - 현재는 API 형태로 모델들을 호출할 수 있게 하려고 하고 있음
 - 이를 보여주기 위한 웹페이지를 직접 React로 구현하거나 혹은 간단하게 만들 수 있는 사이트를 활용하려고 함.



기대 효과

- 사용자가 콘텐츠 생성 시 블로그화 하는데 시간을 줄일 수 있음
- 제목, 해시태그 등 콘텐츠화하기 위해 트렌드 분석하는 시간을 줄이게 됨
- 개인별 작성 특징을 기반으로 작성되기 때문에 좀 더 맞춤형으로 작성 가능