




"풍경 사진 올렸더니 챗봇이 시 짓네...네이버 AI, GPT-4o 못지 않네"

등록 2024.08.22 13:37:01

 사진을 보고 시를 지어줘



 제목 : 해변의 일몰

해변에 앉아 금빛 물결을 바라보네
하늘은 붉은빛으로 물들어가고
구름 사이로 해가 숨바꼭질을 하네

모래 위엔 발자국 하나 둘
파도 소리에 마음을 실어 보내
이 순간만은 세상의 모든 걱정을 잊고

저 멀리 수평선 너머로
해는 서서히 사라져가네
내일을 기약하며 오늘을 마무리하네

해변의 일몰은 언제나 아름다워
내 마음에도 따뜻한 빛을 비추네

뉴스시스

[서울=뉴스시스] 네이버는 오는 27일 대화형 AI 서비스 '클로바X' 업데이트를 통해 시각 정보 처리 능력을 새로 추가한다고 22일 밝혔다. 사진은 클로바X가 지은 시 (사진=네이버 클로바 기술 블로그 캡처)

[서울=뉴스시스]윤정민 기자 = 네이버가 출시 1년을 맞은 거대언어모델(LLM) 하이퍼클로바X에 이미지 분석·추론 기능을 탑재했다. 구글, 오픈AI, 앤트로픽 등 글로벌 빅테크처럼 네이버 LLM도 멀티모달 AI(텍스트, 이미지, 오디오 등 다양한 유형의 데이터를 동시에 처리·이해·생성하는 AI)로 진화했다. 네이버는 한층 고도화된 하이퍼클로바X 일부 성능이 오픈AI LLM 'GPT-4o'에 뒤처지지 않는다고 강조했다.


네이버는 오는 27일 대화형 AI 서비스 '클로바X' 업데이트를 통해 시각 정보 처리 능력을 새로 추가한다고 22일 밝혔다.

클로바X는 그동안 텍스트 기반 질의응답 기능을 제공해 왔다. 네이버 전체 서비스를 넓혀도 이미지 탐색 AI가 적용됐던 곳은 네이버 쇼핑(옴니서치) 등 일부에 불과했다.

하지만 클로바X는 앞으로 사진 속 상황을 추론하고 표, 그래프 등을 분석하고 수학 문제를 풀 수 있게 됐다. 네이버는 자사 블로그를 통해 클로바X에 탑재될 '하이퍼클로바X 비전' 기능(이미지에 대한 상세한 설명, 이미지를 통한 추론, 표·차트 이해 등)을 소개했다.

"채소, 고기, 빵 사진만 올렸는데 내가 햄버거를 먹겠다는 걸 어떻게 알았지?"

사진에 대해서 보이는대로 설명해줘.



사진에는 어린아이가 양에게 먹이를 주고 있는 모습이 담겨 있습니다. 아이는 파란색 바탕에 동물 그림이 그려진 옷을 입고 있으며, 줄무늬 모자를 쓰고 있습니다. 아이의 표정은 집중하고 있는 듯 보이며, 양은 아이가 주는 먹이를 먹으려고 머리를 숙이고 있습니다. 배경에는 다른 양들도 보이며, 이 장소가 양 목장임을 짐작할 수 있습니다.

방금 장봐온 재료들인데 이걸로 뭐 만들 것 같아? 만드는 순서도 알려줘.



사진 속 재료를 보아하니 참깨 햄버거를 만들 것 같습니다. 사진 속 재료로 햄버거를 만드는 순서는 다음과 같습니다.

1. 참깨 햄버거 빵을 반으로 잘라 준비합니다.
2. 소고기 패티를 만들기 위해 소고기를 적당한 크기로 다진 후, 소금, 후추로 간을 하고 잘 치대어 모양을 잡아줍니다.
3. 팬에 기름을 두르고 햄버거 빵을 살짝 구워줍니다.
4. 같은 팬에 패티를 구워줍니다.
5. 토마토와 양파는 얇게 슬라이스하고, 상추는 깨끗이 씻어 물기를 제거합니다.
6. 햄버거 빵에 마요네즈를 바르고, 상추, 토마토, 양파, 패티 순으로 올려줍니다.
7. 기호에 따라 치즈, 베이컨, 피클 등을 추가하고, 햄버거 빵을 덮어줍니다.

이와 같은 순서로 햄버거를 만들면 맛있는 참깨 햄버거를 완성할 수 있습니다.

[서울=뉴시스] 네이버는 오는 27일 대화형 AI 서비스 '클로바X' 업데이트를 통해 시각 정보 처리 능력을 새로 추가한다고 22일 밝혔다. 사진은 클로바X가 지은 시 (사진=네이버 클로바 기술 블로그 캡처)


예를 들어 아이가 양에게 먹이를 주고 있는 모습이 담긴 사진을 게재하면 클로바X가 "아이가 양에게 먹이를 주고 있습니다. 아이는 파란색 바탕에 동물 그림이 그려진 옷을 입고 있으며 줄무늬 모자를 쓰고 있습니다. 배경에는 다른 양들도 보이며 이 장소가 양 목장임을 짐작할 수 있습니다"라고 답한다.

양파, 토마토, 소고기, 햄버거 빵 등 사진을 함께 올린 뒤 "방금 장 봐온 재료들인데 이걸로 뭐 만들 것 같아? 만드는 순서도 알려줘"라고 말하면 클로바X가 햄버거 조리법을 알려준다.


키위새 인형 등이 나뭇잎, 열매를 든 사진을 주고 "이 듀오를 위한 밴드 이름을 지어 줘"라고 말하면 클로바X가 '키위 팝스',

'리프 앤 베리 밴드', '숲속 친구들' 등의 이름을 제시하는 기능도 선보였다.


하이퍼클로바X 비전 vs GPT-4o, 韓 검정고시 승자는 네이버

 어느 회사 점유율이 제일 높아?



 이 그래프에 따르면, 'Company A'의 시장 점유율이 50%로 가장 높습니다. 그다음으로는 'Company B'가 25%, 'Company C'가 15%, 'Company D'가 10%의 시장 점유율을 가지고 있습니다. 따라서 'Company A'가 가장 큰 시장 점유율을 차지하고 있습니다.



[다른 답변 보기](#) 

[서울=뉴시스] 네이버는 오는 27일 클로바X 서비스 업데이트를 통해 시각 정보 처리 능력을 새로 추가한다고 22일 밝혔다. 하이퍼클로바X 이미지 이해 '코드 생성' 부분 (사진=네이버 제공) *재판매 및 DB 금지

네이버는 AI 챗봇 고도화로 다른 글로벌 빅테크가 개발한 AI 챗봇과의 경쟁력도 갖추게 됐다. 이미 구글, 오픈AI, 앤트로픽 등은 이미지 분석, 음성 대화 등의 기능을 갖춘 LLM을 자사 챗봇에 탑재한 바 있다.

구글은 LLM '제미나이'에 고급 음성 기능을 더한 '제미나이 라이브'를 출시했다. 풍부한 표현력으로 사용자와 대화할 수 있으며 챗봇이 답변 중 후속 질문을 하더라도 바로 대응할 수 있다는 게 구글 측 설명이다.

오픈AI도 사용자와 실시간으로 음성 대화를 나눌 수 있는 LLM 'GPT-4o'를 공개했다. 스칼렛 요한슨이 GPT-4o 음성 중 하나인 '스카이'가 자신의 목소리를 무단 도용했다고 주장해 화제가 되기도 했다. 현재는 스카이 음성을 삭제한 모드로 제공되고 있다.

앤트로픽도 이미지 인식 능력을 갖춘 LLM '클로드 3.5 소넬트'를 보유하고 있다. 이미지를 분석해 코드를 생성하고 편집하는 기능을 갖췄다.

네이버는 이들 기업보다 늦게 멀티모달 LLM을 공개했으나 성능은 뒤처지지 않는다고 강조했다. 네이버가 국내 초·중·고등학교 검정고시 총 1480개 문항을 AI 모델에 이미지 형태로 입력하고 문제를 풀게 한 결과 클로바X가 약 84%의 정답률을 기록했다고 전했다. 오픈AI GPT-4o의 78%보다 높은 정답률을 보였다는 게 네이버 측 설명이다.

네이버는 클로바X가 논리적 글쓰기, 코드 작성, 번역 등의 작업에 활발히 사용된 것에서 더 나아가 이미지 이해 능력을 기반으로 개인의 생산성 향상 도구로서 활용 범위가 한층 넓어질 것으로 기대한다고 전했다.

©공감언론 뉴시스 alpaca@newsis.com

Copyright © NEWSIS.COM, 무단 전재 및 재배포 금지