

中 바이두 AI, 성능평가 도구 15개 중 6개 중국어 기반

GPT 보다 나은 성능 주장 '어니4.5' 벤치마크 6개 중 4개 중국어 기반

실시간 코드 작성 평가·추론 평가 GPT와 비교해 성능 크게 뒤쳐져 평가 도구 따라 AI수준 '천차만별'

중국 빅테크 기업 바이두가 새 인공지능 모델 '어니(ERNIE) 4.5'를 발표하면서, 답시크의 'R1'보다 저렴하면서 성능은 미국 오픈AI의 GPT-4.5를 능가한다고 주장했지만, 성능 평가에 사용한 벤치마크의 40%가 중국어 기반인 것으로 나타나 실제 GPT-4.5보다 우수하지 않은 것으로 파악됐다. 어니 4.5는 바이두가 자체 개발한 멀티모달 기반 AI 모델이다.

18일 IT업계에 따르면 바이두는 지난 16일 어니 4.5를 출시하며, 여러 벤



치마크에서 오픈AI의 GPT-4.5보다 나은 성능을 보였다고 주장했다.

바이두가 어니 4.5를 출시하며 공개한 벤치마크 자료를 보면, 성능 평가 도구 15개 중 6개가 중국어 기반이었다.

어니 4.5가 GPT-4.5보다 높은 성적을 기록한 벤치마크는 ▲중국어로 된 다양한 주제의 질문을 통해 AI 모델의 언어 이해 및 문제 해결 능력을 평가하는 'C-이벌(Eval)' ▲중국어로 된 대규모 데이터셋을 활용해 추론 등 다양한 작업의 성능을 평가하는 'CMMMLU' ▲초등학교 수준의 수학 문제로 구성된 데이터셋 'GSM8K' ▲중국어로 된 수학 문제를 통해 AI 모델의 수학·언어 능력을 동시에 평가하는 'C매스' ▲

단락을 읽고 이산적 추론을 통해 질문에 답하는 능력을 평가하는 'DROP' ▲중국어 문맥에서 대명사 해석 능력을 평가하는 'CLUUEWSC'다. 전체 6개 중 4개가 중국어 기반 벤치마크 테스트였던 것.

반면, 어니 4.5는 실시간으로 코드를 작성하고 실행하는 능력을 평가하는 '라이브코드벤치', 의학·법률·공학 등 전문 분야에서의 언어 이해 능력을 평가하는 'MMLU-프로', 복잡하고 다양한 코딩 문제들을 평가하는 '휴먼이벌+', 복잡한 추론과 문제 해결 능력을 평가하는 'BBH' 벤치마크에서는 GPT-4.5와 비교해 성능이 크게 뒤쳐지는 것으로 나타났다.

또 어니 4.5는 대학원 수준의 전문 지식이 필요한 질문에 답하는 능력을 평가하는 'GPQA', AI 모델의 수학적 문제 해결 능력을 평가하는 '매스-500',

프로그래밍 문제에 대한 해결 능력을 평가하는 'MBPP+' 벤치마크에서 GPT-4.5에 밀렸다.

이외에 중국어로 된 간단한 질문에 답변하는 능력을 평가하는 '차이니스 심플QA', 중국 수학 올림피아드 문제를 기반으로 AI 모델의 고급 수학적 사고 능력을 평가하는 'CNMO2024'에서는 GPT-4.5와 비슷한 성능을 보였다.

바이두는 어니 4.5의 벤치마크 점수 평균은 79.6점, GPT-4.5는 79.14점으로, 자사의 AI 모델 성능이 더 우수하다고 주장했다.

바이두가 이번에 공개한 벤치마크 결과는 AI 모델의 능력 수준이 성능 평가 도구로 어떤 것을 사용하느냐에 따라 천차만별로 달라질 수 있다는 점을 시사한다.

소프트웨어정책연구소(SPRI)는 'AI 브리프(2025년 3월호)'에서 "최근 개

발된 AI 모델들이 MMLU와 같은 인기 벤치마크에서 대부분 90% 이상의 정확도를 달성하며 기존 벤치마크로는 모델 성능 평가에 한계가 있다는 여론이 대두되자 AI안전센터(미 AI 안전 비영리 기구)는 지난 1월 고난도 평가를 위한 'HLE(인류 최후의 시험)'를 공개했다"고 설명했다.

HLE는 50개국, 500개 이상의 기관 소속 교수와 연구자 등 천여명의 전문가 협업을 통해 탄생한 벤치마크다. 100개 이상의 주제를 포괄하는 2700개의 까다로운 질문으로 구성됐다.

SPRI는 "HLE를 통해 과학자와 정책 입안자들이 AI 역량을 평가할 수 있는 공통된 기준을 제시함으로써 개발 방향이나 잠재적 위험, 필요한 정책 조치에 대해 정보에 입각한 논의가 가능해질 것으로 기대된다"고 밝혔다.

/김현정 기자 hjk1@metroseoul.co.kr

KT, 온실가스 줄이는 '친환경 보빈' 도입

폐플라스틱 활용, 10회 이상 사용 목재 사용량 줄여 산림훼손 방지

KT는 국내 통신사 중 처음으로 친환경 '보빈'을 도입한다고 18일 밝혔다. 보빈이란 케이블을 연속적으로 감는 데 사용하는 원통형 구조물을 말한다. 일반적으로 목재로 제품을 만들며, 케이블 보관과 운송에 활용된다.

KT가 이번에 도입하는 친환경 보빈은 산업통상자원부에서 인증한 폐플라스틱(저밀도 폴리에틸렌)으로 만든 제품이다. 10회 이상 다시 사용할 수 있다는 점이 특징이다. 다만, 일반 보빈 대비 가격이 비싸고 수거 작업이 어려워 산업계로 확산되지는 못했다.

그러나 KT는 LS전선·대한광통신·가온전선·머큐리광통신·이에스테크인터내셔널 등 5개 광케이블사 및 친환경



KT 직원들이 오는 4월부터 납품 받는 광케이블에 친환경 보빈을 적용하기 위해 친환경 보빈을 지게차로 옮기고 있다. /KT

보빈 공급사 간 협의를 진행해 렌털 공급 방식으로 친환경 보빈의 가격을 낮췄다고 설명했다. 제품을 원활히 수거하기 위해서는 사물인터넷(IoT) 기반의 관제 서비스를 적용했다.

친환경 보빈은 목재 사용량을 줄여

산림 훼손을 방지하고, 온실가스를 약 1만2500kg 정도 감축한다. 또한 전기 사용량 2만7200kWh 절감과 동일한 효과를 낸다. 이는 4인 가족 기준 78가구가 한 달 동안 사용하는 전기량과 맞먹는다. 아울러 포장 폐기물이 감소하고, 보빈 폐기 비용을 절감하는 효과도 기대할 수 있다.

KT는 4월부터 납품 받는 광케이블에 친환경 보빈을 사용할 계획이다. 연간 약 2500개의 목재 보빈이 친환경 제품으로 대체될 예정이다. KT는 친환경 보빈 사용을 점진적으로 확대해 전체 광케이블에 적용할 계획이다.

KT는 폐플라스틱 재활용 자재를 사용해 친환경 구매를 강화하고, 자원의 선순환을 실현해 광케이블 제조사와 동반 성장을 도모하겠다고 강조했다.

/김서현 기자 seoh@



LG유플러스 직원이 인공지능(AI) 기반의 고객관리(CRM) 시스템에서 AI가 고객 불편을 사전 탐지하는 서비스를 시연해보고 있다. /LG유플러스

LG유플러스, AI가 고객불편 사전 탐지

'고객 관계 관리 시스템' 구축 이상징후 감지엔 사전에 조치

LG유플러스는 인공지능(AI)을 활용해 고객 서비스 불만을 사전에 탐지하고 예방하는 고객 관리 시스템을 구축했다고 18일 밝혔다.

LG유플러스가 도입한 AI 기반 고객 관계 관리(CRM) 시스템은 고객의 개통·AS 정보, 상담 정보, 품질 정보 등 데이터를 통합해 관리하고, AI가 고객의 잠재 불만과 불편을 선제적으로 감지하는 플랫폼이다. 초고속인터넷·인터넷TV(IPTV) 등 홈 서비스를 이용 중인 고객이 불편이나 불만을 표시하기 전에 선제적으로 대응해 서비스에 대한

불만을 낮추는 게 목표다.

기존에는 고객이 서비스센터에 직접 점검을 요청하면 홈 서비스 담당 직원이 파견돼 점검했지만, 이런 방식으로는 고객이 문제를 경험하기 전에 불편 사항을 파악하기 어려운 한계가 존재했다. LG유플러스는 이번 CRM 시스템 구축으로 이상 징후가 감지될 경우 홈 서비스 직원이 사전 조치를 취할 수 있을 것으로 보고 있다. 예를 들어, 초고속인터넷 회선에 이상이 발생해 속도가 저하된 경우 고객이 미처 체감하지 못했더라도 홈 서비스 직원이 CRM이 알려주는 이상 여부를 확인한 뒤 고객에게 연락해 서비스 이상 여부와 조치 사항 등을 안내할 수 있다. /김서현 기자

'1만원대 20GB' 알뜰폰 5G요금제 출시

스마트폰 등 3곳... 상반기 중 더 나올듯

정부가 알뜰폰 도매대가 인하 정책을 내놓은 이후 1만원대로 이용할 수 있는 20GB용량의 5G 요금제가 연이어 출시되고 있다. 이미 출시됐거나 상반기 출시 예정인 요금제가 22개 가량이다.

18일 알뜰통신사업자(KMVNO)협회에 따르면 현재 1만원대에서 이용할 수 있는 '20GB 5G 요금제'를 출시한 알뜰폰 사업자는 스마트·큰사람커넥트·프리텔레콤 등 3곳이다. 이 가운데 프리텔레콤은 프로모션 요금제를 포함해 5개 요금제를 연달아 선보였다.

아이즈비전·유니컴즈·스테이지파이브·KB국민은행·에넥스텔레콤·씨케이커뮤니티 등은 상반기 안으로 1만원대 요금제 출시를 준비 중이다. 가격대는 월 1만4000~1만9000원선이다.

이에 앞서 과기정통부는 지난 1월 알

뜰폰 경쟁력 강화 방안을 발표하고, 도매대가 인하를 통해 사업자들이 경쟁력 있는 요금제를 설계할 수 있도록 지원했다. 도매대가 산정 방식 개정에 따라 제공비용 기반 방식을 도입해 종량 데이터 도매대가는 36.4% 인하됐다.

이 같은 조치로 기존 데이터 1MB당 1.29원에서 0.82원으로 낮아지고, 음성 도매대가는 소폭 낮아졌다. 특히 종량 데이터 도매대가가 처음으로 1원대 밑으로 떨어지면서 알뜰폰 사업자들이 저렴한 요금제를 출시할 수 있는 환경이 조성됐다. 협회는 앞으로 더 많은 사업자들이 참여할 것으로 기대하고 있다.

협회는 정부와 긴밀히 협력해 알뜰폰 산업의 지속적인 성장, 경쟁력 강화를 위해 노력할 계획이라고 밝혔다. 정부는 도매대가 체계를 합리적으로 개선하고 알뜰폰 사업자의 서비스 품질 향상을 지원하기로 했다. /김서현 기자

카카오엔터

라이언 주인공 웹툰 '그래도 라이언' 공개

카카오엔터테인먼트(카카오엔터)가 18일 카카오프렌즈의 대표 인기 캐릭터 라이언을 주인공으로 한 웹툰 '그래도, 라이언'을 카카오페이지와 카카오톡에 공개했다.

웹툰 '그래도, 라이언'은 동등성 왕위 계승자로 태어났으나 갈기가 없어 콤플렉스인 수사자 라이언이 프로도, 네오, 튜브, 어피치, 무지, 제이지 등 카카오프렌즈 친구들을 만나기까지의 우여곡절을 그린다. 라이언 세계관을 집대성한 스토리 기반의 첫 정식 오리지널 웹툰이라고 카카오엔터는 강조했다.

카카오엔터는 웹툰 공개를 기념해 다양한 이벤트를 마련했다. 오는 24일까지 '그래도, 라이언' 감상 독자들을 대상으로 캐시 뽑기권을 제공하고, 이벤트와 카카오토어터 증정 행사도 벌인다. /김현정 기자

SKT, 에이닷에 '제미나이2.0 플래시' 추가

1.5 프로 보다 주요 벤치마크 2배 빨라

SK텔레콤은 인공지능(AI) 에이전트 '에이닷(A.)'에 구글의 자체 거대언어 모델(LLM)인 제미나이 최신 모델인 '제미나이 2.0 플래시'를 추가했다고 18일 밝혔다.

이번 에이닷 AI 모델의 추가로 이용자들은 SK텔레콤의 A.X와 GPT, 클로드, 퍼플렉시티, 제미나이 등 5개 글로벌 대표 AI 서비스 세부 모델들을 자유

롭게 이용할 수 있게 된다.

이번에 새롭게 추가된 '제미나이 2.0 플래시'는 구글이 지난 2월 공개한 최신 AI 모델로 기존 1.5 플래시를 개선한 버전이다. 응답 시간과 성능이 향상되고 주요 벤치마크에서 1.5 프로 모델보다 2배 빠른 속도를 기록했다. 이용자들은 답변과 함께 구글의 실시간 검색 결과를 제공 받고 답변이 어떤 내용을 토대로 생성됐는지 출처 페이지로 이동하는 기능도 쓸 수 있다. /김서현 기자