

**배포일자: 2025. 7. 9(수)**

**[즉시 보도 가능합니다]**

**KT 한국적 AI 믿:음 2.0 Base**

**한국어 AI 성능 평가서 국내 1위 기록**

**▶ KT 자체 개발 ‘믿:음 2.0’, 한국어 LLM 성능 평가에서 국내 모델 최고 순위 달성**

**▶ 국산 기술력 기반의 한국어 AI 모델로서 실질적 경쟁력 입증**

**▶ 다양한 산업 분야에 적용해 ‘한국형 AI’의 시장 확산에 본격적으로 나설 것**

KT가 한국적 AI를 표방하며 자체 개발한 ‘믿:음 2.0’이 한국어 LLM 성능 평가 플랫폼 ‘호랑이(Horangi) 리더보드’에서 국내 1위를 차지하며 가장 우수한 한국어 성능을 입증했다.

KT(대표이사 김영섭)는 지난 4일 공개한 KT의 자체 개발 모델 믿:음 2.0 Base(Mi:dm-2.0-Base-Instruct)가 한국어 LLM 평가 지표 ‘호랑이 리더보드3(Horangi: W&B Korean LLM Leaderboard 3)’에서 파라미터 수 150억 개 미만의 국내 기업 개발 모델 가운데 종합 성능 1위를 기록했다고 8일 밝혔다.

믿:음 2.0 Base의 종합 점수는 0.7197(▲범용 성능 0.7004 ▲응답 정렬(Alignment) 성능 0.739)로 전세계 동급 모델 중에서는 3위에 해당하는 우수한 성과다.

‘호랑이 리더보드’는 글로벌 MLOps 기업 웨이트앤바이어스(Weights&Biases, 이하 W&B)가 주관하는 한국어 특화 LLM 평가 벤치마크이다. 실제 한국어 환경에서의 범용적인 언어 이해력과 응답 안전성을 포함하는 정렬성(Alignment) 등 한국어 LLM의 실용적인 역량을 다면적으로 평가한다. 단편적인 지식을 묻는 질문에 대한 답변 정확성을 평가하는 기존 벤치마크와 달리, 작문·추론·정보 추출 등 실생활 시나리오를 기반으로 평가하여 모델의 실용적 능력을 평가하는 지표로 주목받고 있다.

특히 한국어 고유의 문맥, 표현, 사회적 맥락 등을 반영한 응답인지를 평가한다는 점에서 국내 시장에서의 실효성을 가늠하는 데 중요한 기준으로 꼽힌다. 평가 결과는 W&B가 운영하는 WanDB 플랫폼을 통해 투명하게 공개되어 누구나 직접 검증할 수 있어 신뢰도도 높다고 평가된다.

이번 성과로 믿:음 2.0이 국산 기술력 기반의 한국어 AI 모델로서 실질적 경쟁력을 갖추었음을 객관적으로 입증했다. KT는 믿:음 2.0을 개발하며 한국의 정신과 생활방식, 지식과 사회적 맥락을 반영하여 한국에 가장 잘 맞는 AI로 학습시켰다.

무엇보다 믿:음 2.0은 외산 모델을 기반으로 단순 튜닝하는 방식이 아닌, 아키텍처 설계부터 데이터 구축, 학습까지 전 과정을 KT가 직접 수행하는 ‘프롬 스크래치(From scratch)’ 방식으로 개발한 순수 자체 개발 모델이라는 점에서도 기술적 의미가 크다.

KT Gen AI Lab장(CAIO) 신동훈 상무는 “믿:음 2.0을 국내 다양한 산업 분야에 적용해 ‘한국적 AI’의 시장 확산에 본격적으로 나설 것”이라며, “한국적 AI의 실용화와 기술 발전을 이끄는 핵심 주체로 자리매김하겠다”고 밝혔다.

KT는 믿:음 2.0 기반의 맞춤형 B2B AI 솔루션을 산업 전반에 제공하고, 공공·금융·교육·법률 분야에서 실증을 추진한다. 현재 공개된 버전에서 성능을 더 높인 믿:음 2.0 Pro 공개도 준비하고 있다. 이어 믿:음 2.0의 추론, 멀티모달 기능 확장과 함께 국산 AI 반도체 기업 리벨리온과의 협력을 통한 AI 생태계 강화 등도 계획 중이다.

※ 보충문의: KT 홍보실 변환 차장

※ 사진설명: 호랑이 리더보드 (기준 : 2025. 07. 08)

텍스트, 스크린샷, 번호, 폰트이(가) 표시된 사진

AI가 생성한 콘텐츠는 부정확할 수 있습니다.