

SKT, 자체 거대언어모델 기반

멀티모달·범용 문서 해석 기술 공개

**-** **‘A.X 인코더’, 긴 문서도 빠르고 효율적으로 처리해 대규모 LLM 학습에 적합**

**- ‘A.X 4.0 VL Light’, 대규모 멀티모달 한국어 데이터셋 학습된 시각-언어모델**

**- “자체역량 높이고 컨소시엄 협업 박차 가해 글로벌 최고 수준 경쟁력 확보”**

|  |
| --- |
| **엠바고 : 배포 즉시 사용 가능합니다.** |

**[2025. 07. 29.]**

SK텔레콤(대표이사 CEO 유영상, [www.sktelecom.com](http://www.sktelecom.com))은 자사의 LLM인 A.X(에이닷 엑스)를 기반으로 한 시각-언어모델(VLM1))과 LLM 학습을 위한 범용 문서 해석 기술을 선보였다고 29일 밝혔다.

1. Vision Language Model(VLM): 시각 정보와 언어적 정보를 통합 처리하는 인공지능 모델

이날 SKT가 오픈소스 커뮤니티 허깅페이스(Hugging Face)에 공개한 모델은 ‘A.X Encoder(인코더)’와 ‘A.X 4.0 VL Light(비전 랭귀지 라이트)’ 2종이다. 해당 모델들은 학술 연구나 상업 이용 등에 자유롭게 활용 가능하다.

SKT는 7월 한 달 동안 대규모 학습(CPT, Continual Pre-Training)에 기반한 A.X 4.0 모델 2종(표준, 경량)에 이어 프롬 스크래치 방식의 A.X 3.1 모델 2종(표준, 경량)을 순차적으로 선보였다. 이번에 LLM을 산업 영역에 보다 폭넓게 활용하기 위한 기술 2종을 추가하면서 총 6개의 모델을 발표했다.

SKT는 프롬 스크래치2) 방식의 LLM 개발을 꾸준히 이어가는 한편, 향후 발표할 A.X 4.0 추론형 모델 등 지속적으로 개발중인 LLM의 활용도와 성능을 높여갈 계획이다.

2) 프롬 스크래치(From Scratch): 모델의 맨 처음 단계부터 모두 직접 구축

**■ A.X 인코더, 기존 방식 대비 메모리 사용량 적고 고속으로 많은 문서 처리**

자연어처리 기술에서 인코더란, 입력된 문장을 문맥으로 변환하고 이를 바탕으로 다양한 자연어 처리 작업을 수행하도록 돕는 핵심 구성요소다. 문장의 모든 단어들의 상호 관계를 파악, 전체 의미와 맥락을 이해하는 역할을 한다.

SKT는 A.X 모델에 필요한 데이터의 전 과정 처리 프로세스에 적용하기 위해 ‘A.X 인코더’를 개발했다. ‘A.X인코더’는 긴 문서도 빠르고 효율적으로 처리 가능해 대규모 LLM 학습에 적합하다.

‘A.X 인코더’는 1억 4천9백만개(149M)의 매개변수를 바탕으로 작동하며, 자연어 이해 성능지표 평균 85.47점을 달성해 글로벌 최고수준(SOTA)급 성능을 확인했다. 기존 글로벌 오픈소스 모델을 기반으로 KLUE3) 팀에서 공개한 ‘RoBerTa-base’의 성능지표(80.19점)를 상회하는 수준이다.

3) KLUE(Korean Language Understanding Evaluation): 한국어 자연어 벤치마크

‘A.X 인코더’는 16,384개의 토큰까지 처리가 가능해 기존 모델들보다 최대 3배의 추론속도와 2배의 학습속도를 구현할 수 있다. 통상적으로 기존 모델들이 512개의 토큰을 처리해 보통 문장 혹은 문단 정도의 처리가 가능했다면, 그보다 월등히 큰 문맥을 빠르고 효율적으로 처리한다.

이러한 대규모, 고속 문서 처리 기술은 LLM 학습 외에도 AI 기반의 다양한 문서 처리에 효율적으로 적용할 수 있을 것으로 기대된다.

**■ A.X 4.0 VL Light, 가벼운 구조에도 강력한 멀티모달 성능 제공**

‘A.X 4.0 VL Light’는 대규모 멀티모달 한국어 데이터셋이 학습된 시각-언어모델(VLM)이다. 한국어와 관련된 시각정보 및 언어 이해뿐만 아니라, 표·그래프 이해, 제조 도면 이해와 같은 기업용 애플리케이션에서 탁월한 성능을 제공한다.

70억 개(7B) 매개변수의 A.X 4.0 Light 모델을 기반으로 개발되어 사용자 시스템에 쉽게 적용 가능하면서도 중형 모델 수준의 강력한 성능을 자랑한다.

‘A.X 4.0 VL Light’는 한국어 시각 벤치마크에서 평균 79.4점을 기록하며, Qwen2.5-VL32B(73.4점)보다 훨씬 작은 모델 크기에도 불구하고 더 우수한 성능을 보였다. 또한 한국어 텍스트 벤치마크에서는 평균 60.2점을 기록, 경량모델임에도 국내 모델 중에서 최상위권에 포진했다.

한국어 문화 및 맥락적 이해를 평가하기 위해 설계된 멀티모달 벤치마크인 K-Viscuit에서 80.2점을 기록했고, 복잡한 문서 구조와 차트·표를 이해하는데 중점을 둔 KoBizDoc 벤치마크에서는 89.8점을 달성했다. 각각 Qwen2.5-VL32B 모델보다 뛰어나거나(72.3점) 비슷한(88.8점) 수준이다.

이러한 성능에도 높은 효율을 자랑하는 ‘A.X 4.0 VL Light’는 동일한 한국어 데이터입력 시 Qwen2.5-VL32B 대비 약 41% 적은 텍스트 토큰을 사용하여, 사용하는 기업들의 비용을 낮추는데 기여할 수 있다.

김태윤 SK텔레콤 파운데이션 모델 담당은 “독자적인 기술력 확보가 소버린 AI의 핵심인 만큼, 자체 역량을 높이고 컨소시엄 기업들과의 협업에도 박차를 가해 글로벌 최고 수준의 AI 경쟁력을 확보할 것”이라고 밝혔다.

**▶ 관련문의 : SK텔레콤 PR실 테크PR팀 우현섭 매니저 (02-6100-3854)**