

생성형 AI를 위한 파운데이션 모델 플랫폼

주요 이점

- ▶ 사용자가 InstructLab을 통해 대규모 언어 모델(LLM)을 업데이트 및 강화
- ▶ LLM을 상용 데이터와 안전하게 연결하여 모델을 비즈니스 요구 사항에 맞게 조정
- ▶ 생성형 인공지능(생성 AI)을 신속하게 시작하고, 신뢰할 수 있는 보안 중심의 Red Hat® Enterprise Linux® 플랫폼을 통해 결과 제공
- ▶ 설치 및 업데이트를 위해 부팅 가능한 Red Hat Enterprise Linux 컨테이너 이미지로 패키징

제품 개요

Red Hat Enterprise Linux AI는 생성 AI 비즈니스 활용 사례를 위해 LLM을 개발, 테스트 및 배포하기 위한 엔터프라이즈급 생성 AI 파운데이션 모델 플랫폼입니다.

Red Hat Enterprise Linux AI는 다음 요소를 통합합니다.

- ▶ 오픈소스 LLM의 Granite 제품군
- ▶ LLM 미세 조정(fine-tuning)에 대해 커뮤니티 기반 접근 방식을 제공하는 InstructLab 모델 정렬 툴링
- ▶ NVIDIA, Intel, AMD용 PyTorch 및 AI 액셀러레이터 드라이버 소프트웨어 등과 같은 생성 AI 라이브러리 및 종속성, 그리고 Red Hat Enterprise Linux의 부팅 가능한 이미지
- ▶ Red Hat이 제공하는 엔터프라이즈급 기술 지원과 모델 지식재산권 보증
- ▶ Red Hat Enterprise Linux AI는 신뢰할 수 있는 Red Hat Enterprise Linux 플랫폼을 제공하며 생성 AI 여정을 시작하고 결과를 확인하는 데 필요한 구성 요소를 추가합니다.

미래 AI의 특징: 개방성과 투명성

Red Hat Enterprise Linux AI에는 Red Hat이 전적으로 보증하는 오픈소스 Granite 언어 및 코드 모델의 하위 집합이 포함되어 있습니다. 오픈소스 Granite 모델은 광범위한 생성 AI 활용 사례에 맞춰 비용 및 성능에 최적화된 모델을 제공합니다. Granite 모델은 Apache 2.0 라이선스에 따라 릴리스되었습니다. 모델이 오픈소스라는 점 외에도 모델을 학습시키는 데 사용되는 데이터세트도 투명하고 개방적입니다.

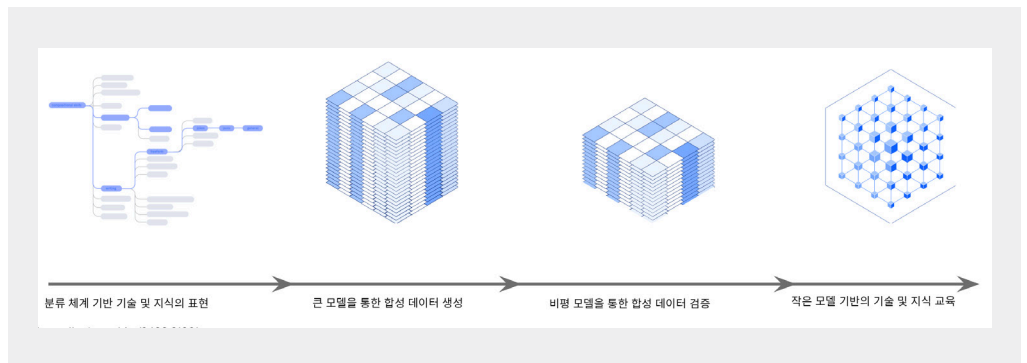
표 1. Red Hat Enterprise Linux AI의 Granite 모델

IBM Granite 언어 모델	Granite-7B-Starter Granite-7B-RedHat-Lab
IBM 코드 모델	Granite-8B-Code-Instruct Granite-8B-Code-Base

생성 AI 모델 학습의 접근성 향상을 통해 가치 창출 시간 단축

오픈소스 Granite 모델 외에도 Red Hat Enterprise Linux AI에는 LAB(Large scale Alignment for chatBots) 기술을 기반으로 하는 InstructLab 모델 정렬 툴링이 포함되어 있습니다. InstructLab을 통해 조직 내 팀들은 기술과 지식으로 LLM에 효율적으로 기여하고 이러한 LLM을 비즈니스의 특정 요구 사항에 맞게 커스터마이징할 수 있습니다.

- ▶ 기술: 모델에게 특정 태스크를 수행하는 방법을 학습시키는 기능 도메인입니다. 기술은 다음과 같은 두 가지 범주로 분류됩니다.
 - ▶ 구성 가능한 기술
 - ▶ AI 모델이 특정 태스크 또는 기능을 수행할 수 있습니다.
 - ▶ 근거가 있거나(컨텍스트 포함) 근거가 없습니다(컨텍스트 미포함).
 - ▶ 근거가 있는 기술의 예시는 모델에 마크다운 형식의 테이블을 읽을 수 있는 기능을 제공하는 스킬을 추가하는 것입니다.
 - ▶ 근거가 없는 기술의 예시는 모델에게 운율을 맞추는 방법을 가르치는 기술을 추가하는 것입니다.
 - ▶ 기반 기술
 - ▶ 수학, 추론, 코딩과 같은 기술입니다.
- ▶ 지식: 모델에게 더욱 정확하게 질문에 답변할 수 있도록 추가 데이터와 정보를 제공하는 데이터와 팩트입니다.



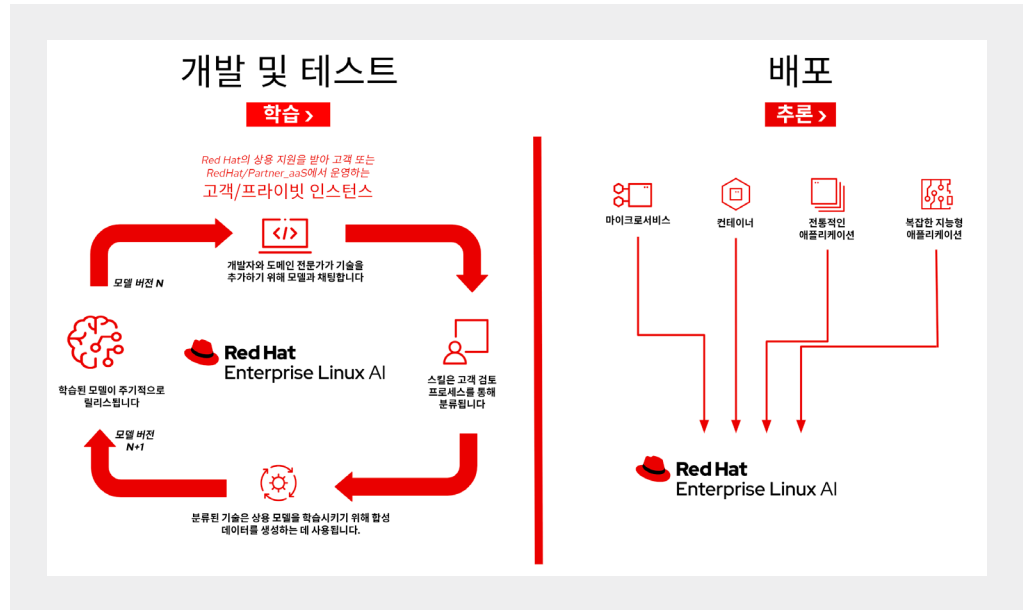
위 이미지는 InstructLab 모델 미세 조정 워크플로우를 개략적으로 나타낸 것입니다.

1. 기술 및 지식 기여는 분류 체계 기반 데이터 리포지토리에 위치합니다.
2. 분류 체계 데이터를 사용하여 방대한 양의 합성 데이터가 생성되어 LLM을 성공적으로 업데이트하고 변경할 만큼의 충분한 대규모 데이터셋을 생산합니다.
3. 비평 모델이 합성 데이터 출력을 검토, 검증, 정리합니다.
4. 모델은 사람이 생성한 수동 입력에 기반한 합성 데이터를 통해 학습합니다.

InstructLab은 일반적으로 LLM을 미세 조정하는 데 필요한 데이터 사이언스 전문 지식이 부족할 수 있는 개발자와 도메인 전문가가 이용하기에 좋습니다. InstructLab 방법론은 팀들이 협업을 통해 데이터나 비즈니스 활용 사례 요구 사항에 적합한 스킬을 학습을 위해 선택한 모델에 추가하는 데 도움이 되므로 가치 창출 시간을 앞당길 수 있습니다.

어디서나 가능한 학습 및 배포

Red Hat Enterprise Linux AI는 데이터가 존재하는 곳이라면 어디든지 하이브리드 클라우드 전반에서 모델을 학습시키고, 조정하고, 배포하는 기능과 필요한 모든 툴을 제공하여 조직이 기술 검증(POC)에서부터 프로덕션 서버 기반 배포에 이르는 프로세스를 가속화하도록 돕습니다. 배포된 모델은 회사 내 다양한 서비스와 애플리케이션에 사용될 수 있습니다.



준비가 된 조직에는 Red Hat Enterprise Linux AI가 Red Hat OpenShift® AI를 사용하기 위한 진입점의 역할도 합니다. 이 플랫폼을 사용하면 Red Hat Enterprise Linux AI 배포에 사용되는 것과 동일한 Granite 모델과 InstructLab 접근 방식을 통해 분산형 클러스터 환경 전반에서 모델을 규모에 맞게 학습시키고 조정하고 서빙할 수 있습니다.

특징 및 장점

특징	장점
전체 지원되는 Granite 언어 및 코드 모델, Apache 2.0 라이선스에 따라 오픈소스로 제공	공개적으로 액세스할 수 있는 학습 데이터와 더불어 오픈소스 및 투명한 LLM이 데이터 투명성을 향상하고 데이터 콘텐츠 및 소스에 대한 윤리적 고려 사항을 해결하여 궁극적으로 전체적인 비즈니스 리스크를 낮춥니다.

특징	장점
Granite 모델에 대한 모델 IP 보증	Red Hat Enterprise Linux AI 내의 Granite 모델에 대한 보증은 Red Hat과 IBM이 엄격한 모델 개발 및 테스트에 대해 강한 확신을 갖고 있다는 사실을 반영합니다. 이러한 보장은 고객에게 더욱 강한 확신을 심어주므로 고객은 고객의 성공을 위한 Red Hat의 노력에 대한 강한 신뢰를 바탕으로 Granite 모델을 더욱 안심하고 사용할 수 있습니다.
확장과 액세스 가능한 모델 미세 조정을 위한 InstructLab LLM 정렬 툴링	InstructLab은 LLM을 미세 조정할 수 있는 접근 가능한 방법을 제공하여 심층적인 데이터 사이언스 전문 지식의 필요성을 줄이고 조직 내 다양한 직무가 기여할 수 있도록 지원합니다. 이를 통해 비즈니스는 생성 AI를 도입하여 가치 창출 시간을 앞당기고 ROI(투자수익률)를 극대화할 수 있습니다.
최적화되고 부팅 가능한 모델 런타임 인스턴스	Red Hat Enterprise Linux AI는 Red Hat Enterprise Linux의 이미지 모드라고 불리는 배포 방법인 부팅 가능한 컨테이너 이미지로 제공됩니다. 이 기술은 설치, 구성 및 업데이트 복잡성을 낮추므로 설정 및 변경 관리 프로세스가 간소화됩니다.
AI 하드웨어를 위한 생성 AI 패키지 종속성 및 소프트웨어 드라이버	필수 패키지와 PyTorch, vLLM, NVIDIA 드라이버와 같은 드라이버가 포함된 종합 툴 세트를 사용하여 생성 AI를 즉시 시작할 수 있고 처음부터 생성 AI 비즈니스 활용 사례를 다룰 수 있습니다.

한국레드햇 홈페이지 <https://www.redhat.com/ko>



Red Hat 소개

Red Hat은 세계적인 오픈소스 소프트웨어 솔루션 공급업체로서 커뮤니티 기반의 접근 방식을 통해 신뢰도 높은 고성능 Linux, 하이브리드 클라우드, 컨테이너 및 쿠버네티스 기술을 제공합니다. 또한 Red Hat은 고객이 클라우드 네이티브 애플리케이션을 개발하고, 신규 및 기존 IT 애플리케이션을 통합하고, 복잡한 환경을 자동화하고 관리할 수 있도록 지원합니다. **포춘(Fortune) 선정 500대 기업이 신뢰하는 어드바이저인 Red Hat은 전 세계 고객에게 권위 있는 어워드를 수상한** 지원, 교육 및 컨설팅 서비스를 제공하여 모든 산업 분야에서 개방형 혁신의 이점을 실현할 수 있도록 최선을 다하고 있습니다. Red Hat은 기업, 파트너, 커뮤니티로 구성된 글로벌 네트워크의 허브 역할을 하며 고객들이 성장하고, 확장하고, 디지털 미래에 대비할 수 있도록 지원합니다.

f www.facebook.com/redhatkorea
 구매문의 02-6105-4390
buy-kr@redhat.com