



## NXP, 엣지 AI 기능 위해 eIQ 소프트웨어 확장 지원

- NXP eIQ AI-머신 러닝 개발 소프트웨어에 GenAI 플로우의 RAG 미세 조정 기능과 eIQ 타임 시리즈 스튜디오 추가
- 광범위한 엣지 프로세서에서 보다 쉬운 AI 배포와 사용 지원



NXP 반도체가 eIQ AI-머신 러닝 개발 소프트웨어에 두 가지 새로운 툴을 추가했다. 이로써 소형 MCU부터 크고 강력한 애플리케이션 MPU에 이르기까지 엣지에서 AI 배포와 사용이 보다 용이하도록 지원한다.

**2024년 10월 31일, 대한민국 서울** – 이로 인해 엣지 프로세서 전반에 걸쳐 AI 배포와 활용이 한층 수월해졌다. NXP 반도체가 [eIQ AI-머신 러닝 개발 소프트웨어](#) 에 두 가지 새로운 툴을 추가했다고 발표했다. 이로써 엣지 프로세서 전반에 걸쳐 엣지에서 AI 배포와 사용이 보다 용이해졌다.

eIQ 타임 시리즈 스튜디오(Time Series Studio)는 [MCX MCU 포트폴리오](#)나 [i.MX RT 크로스오버 MCU 포트폴리오](#)와 같은 MCU급 장치에 자동화된 머신 러닝 워크플로우를 제공한다. 이를 통해 시계열 기반 머신 러닝 모델의 개발과 배포를 간소화한다.

GenAI 플로우(GenAI Flow)는 생성형 AI 솔루션을 구동하는 대규모 언어 모델(Large Language Model, LLM)을 위한 빌딩 블록을 제공한다. [NXP의 i.MX 애플리케이션 프로세서 제품군](#)



MPU와 함께 사용하도록 설계된 생성형 AI 솔루션은 특정 맥락 데이터(contextual data)에 대해 LLM을 훈련시켜 엣지에서 인텔리전스를 보다 쉽게 배포할 수 있도록 지원한다. 예를 들어, 사용 설명서를 학습한 LLM이 탑재된 어플라이언스는 사용자와 자연 language로 특정 기능에 액세스하거나 특정 작업 수행, 사용과 유지 관리 최적화 방법에 대해 대화할 수 있다.

엣지에 AI를 배포하면 지연 시간 단축, 사용자 개인정보 보호 개선, 에너지 소비 감소 등 여러 가지 이점을 얻을 수 있다. NXP의 eIQ 툴킷(Toolkit) 확장은 이러한 배포를 보다 쉽고 빠르게 수행할 수 있도록 지원한다. 개발자는 생성형 AI부터 시계열 기반 모델, 비전 기반 모델에 이르기까지 다양한 모델 유형에 액세스할 수 있다. 또한 사용자는 다양한 모델을 보다 광범위한 엣지 프로세서에 배포할 수 있다.

NXP의 산업과 IoT 부문 총괄 매니저 겸 수석 부사장 찰스 닙스(Charles Dachs)는 “AI는 사용자의 요구와 필요에 따라 예측하고 자동화하는 세상의 핵심이며, 실용적인 엣지 배포가 가능하도록 개발돼야 한다. NXP는 즉시 사용 가능한 도구를 통해 모든 범위의 AI 모델과 AI 지원 엣지 프로세서에 걸친 독보적으로 광범위한 옵션을 제공한다. 이 도구는 MCX 포트폴리오와 같은 MCU의 소형 AI 모델, [i.MX RT700](#)과 같은 크로스오버 MCU, [i.MX 95 애플리케이션 프로세서](#)와 같은 보다 강력한 장치에서 실행되는 대형 생성형 AI 모델 전체에 적합하다. NXP는 다양한 시장의 개발자들에게 진정으로 실용적인 엣지 AI를 제공하고 있다”고 말했다.

eIQ 타임 시리즈 스튜디오는 시계열 기반 AI 모델의 개발과 배포에 필요한 시간을 간소화하고 단축한다. 전압, 전류, 온도, 진동, 압력, 소리, ToF(Time of Flight) 등 다양한 입력 신호와 멀티 모달 센서 융합을 위한 신호 조합을 지원한다. 개발자는 자동 머신 러닝 기능을 통해 시간순으로 처리된 원시 데이터(raw data)에서 의미 있는 인사이트를 추출할 수 있다. 또한 성능, 메모리, 플래시 스토리지 크기와 정확도 기준에 따라 맞춤형 AI 모델을 신속하게 구축할 수 있다. 더불어 데이터 큐레이션, 시각화와 분석, 모델 자동 생성, 최적화, 에뮬레이션, 배포를 포함한 포괄적인 개발 환경을 제공한다. 소프트웨어 개발자는 심층적인 데이터 과학이나 AI 전문 지식 없이도 직관적인 인터페이스를 통해 최적화된 이상 징후 탐지, 분류, 회귀 라이브러리를 만들 수 있다.

NXP의 GenAI 플로우를 사용해 엣지 디바이스에서 생성형 AI 애플리케이션에 액세스할 수 있다. 이 소프트웨어 파이프라인은 LLM과 같은 생성형 모델을 최적화할 수 있는 기능을 제공한다. 또한 검색 증강 생성(Retrieval Augmented Generation, RAG)을 통해 모델 또는 프로세서 제공업체에 민감한 정보를 노출하지 않고도 도메인별 지식과 개인 데이터에 대한 모델을 안전하게 미세 조정할 수 있도록 지원한다. 단일 플로우에서 여러 모듈을 연결함으로써



작업에 맞게 LLM을 쉽게 사용자 지정하고 NXP의 i.MX 95 애플리케이션 프로세서와 같은 MPU를 사용한 엣지 배포를 최적화할 수 있다.

[여기](#)에서 추가 기능이 포함된 최신 버전 eIQ 머신 러닝 개발 환경에 액세스하거나 보다 자세한 정보를 확인할 수 있다.

###

### **NXP 반도체 소개**

NXP®반도체(나스닥: NXPI)는 뛰어난 인재들을 모아 더욱 살기 좋고 안전하며 보안 수준이 높은 연결된 세상을 만드는 혁신적인 기술을 개발하고 있다. NXP는 임베디드 애플리케이션용 보안 연결 솔루션의 선도 기업으로서, 자동차, 산업 및 IoT, 모바일, 통신 인프라 시장의 혁신을 주도하고, 보다 지속 가능한 미래로 나아갈 수 있는 솔루션을 제공한다. NXP는 60년 이상의 전문성과 경험을 바탕으로, 전 세계 30개 이상의 국가에서 34,200명의 직원을 고용하고 있다. 2023년 매출은 미화 132억 8천만 달러다. 더 자세한 내용은 <http://www.nxp.com/> 에서 찾아볼 수 있다.