

AI 시대를 선도하는 비결

김형섭 / 효성인포메이션시스템 컨설턴트

최근 산업 분야에서의 핫 트렌드는 두말할 것 없이 'AI(인공지능)'다. AI 서비스 시대는 이미 시작되었으며, IT 비즈니스에서 AI는 더욱더 주목받을 전망이다. 산업별로 활발하게 활용 중인 AI의 현황부터 도입 유형별 차별화 전략 등을 알아본다.

AI, 거스를 수 없는 트렌드가 되다

IDC가 올해 초 발표한 자료에 따르면, 국내 AI 시장은 향후 5년간 연평균 15.1%의 성장률을 기록할 것으로 나타났다. 2020년에 8,072억 원이던 국내 AI 시장의 매출 규모는 2025년에 1조 9,074억 원까지 성장할 것이라는 전망이다. AI 시장을 서비스, 소프트웨어, 하드웨어 부문으로 구분할 경우, 그 비율은 30:40:30으로 추정되며, 2022년 AI 하드웨어 인프라 시장 규모만 3,363억 원에 달한다고 한다.

AI를 제외하고는 IT를 논할 수 없는 시대지만, 기업들은 AI를 현업에 적용하기 위한 방안을 놓고 여전히 고민하고 있다. 한 가지 분명한 점은 인적 자원을 최소한으로 투입하면서 미래를 예측해 비즈니스 성과를 이끌어내기 위해서는 AI를 업무에 반드시 적용해야 한다는 것이다.

AI는 이미 다양한 산업 분야에서 활발하게 적용되고 있다. 산업별로 8개의 분야에서 AI를 활용하는 방안에 대해 간략하게 정리하면 다음과 같다.

↓ 산업 분야별 AI 활용 현황

자동차 & 로봇	생산 및 제조	유지 관리 및 품질	유통 및 물류
자율주행 / 위험 및 규정 준수 / 로봇공학	수요예측 / 에너지 소비 최적화 / 스케줄링 최적화 / 인재이탈 예측	예지 정비 / 예측 품질 / 시각적 분석	화물 최적화 / 유통경로 최적화 / 선박운송 최적화
사물인식(고급), 강화학습	강화학습, 추천	사물인식, 추천, 이미지 분할	추천, 이미지 분류
헬스케어	금융 & 서비스	영업 및 마케팅	공공 분야
의료영상 진단 분석 / 맞춤형 진료처방 챗봇 / 치료 연구개발	보험상품 및 고객 분석 / 위조 및 사기 탐지 / 이상거래 감지 / 고객상담 챗봇	고객 분석 / 제품 가격 최적화 / 상품 추천 / 고객상담 챗봇	지능형 CCTV / 지능형 범죄 예방 / 공공 서비스 챗봇 / 연구 프로젝트
이미지 분할, 자동음성인식, 자연어 처리, 사물인식(고급)	강화학습, 자연어 처리, 자동음성인식	자연어 처리, 추천, 이미지 분류, 자동음성인식	사물인식(일반), 추천, 자동음성인식, 강화학습

그림을 보면 다수의 산업 및 업무 분야에서 AI가 활용되고 있고, 산업별 특성에 따라 사용되는 AI 모델 종류가 다르다는 것을 알 수 있다. 먼저, 자동차와 로봇 분야에서는 사물인식과 강화학습 모델, 생산 및 제조 분야에서는 강화학습과 추천 모델, 유통 및 물류 분야에서는 이미지 분류 모델을 주로 사용한다.

즉, 활용 방안에 따라 강화학습처럼 고성능 연산이 필요하기도 하고, 이미지 분류같이 비교적 낮은 사양이 요구되는 부분도 있다. 따라서 자사의 업무 특성과 목표에 맞게 적합한 AI 플랫폼을 도입하는 것이 매우 중요하다.

국내 AI 도입 현황을 분석한 자료에 따르면, AI를 한정된 분야에서 일부만 사용했던 과거와 달리, 지금은 전통적인 업무까지 AI가 적극적으로 활용되고 있다고 한다. 활용 기술 또한 사물인식, 머신러닝(ML), 딥러닝 등 AI 적용 업무 구분이 점차 없어지고 있다. 한 연구기관의 조사를 바탕으로 기업이 AI를 도입할 때의 세 가지 이슈에 대해 간략하게 짚어본다.

AI 도입 시 발생하는 주요 이슈 세 가지

첫째, 'AI 솔루션에 대한 기술 부족'이다. AI 플랫폼은 단순한 하드웨어 인프라와 운영 솔루션만으로 구성되는 과거의 시스템과 달리, 복잡한 인프라와 다양한 솔루션을 조합해 구성된다. 즉, 모델링 알고리즘, 클라우드, 컨테이너, GPU 서버 가상화, 고성능 스토리지 및 네트워크와 같은 다양한 기술이 필요하다.

둘째는 '부담되는 도입 비용'이다. 고사양 인프라에 AI 솔루션을 도입하거나 시스템을 커스터마이징하기 위한 구축 인건비는 과거의 전통적 시스템 도입 시 투입되는 예산과 비교해 아주 높다. 이는 기업의 부담을 가중시킬 수 있는 사안으로, 비용 문제는 AI 도입에 앞서 해결해야 할 매우 중요한 요소다.

마지막은 '전문 인력과 역량의 부족'이다. 모든 비즈니스와 프로젝트에서 가장 중요한 부분은 바로 '인적 자원'이다. 기술과 예산이 충분하더라도, 역량이 풍부한 전문인력이 없다면 비즈니스를 성공적으로 이끌 수 없다. 따라서 AI 구축 및 운영 역량을 어떻게 확보해야 하는지에 대한 고민도 빼놓을 수 없는 이슈다.

통합 AI 플랫폼으로 완벽 준비

효성인포메이션시스템은 기업들이 AI 도입 이슈를 해결하고, 미래에 발생할 수 있는 문제를 사전 대응하며 자사에 적합한 AI를 도입하고 운영할 수 있는 '통합 AI 플랫폼'을 제안한다.

↓ 효성인포메이션시스템 AI 플랫폼의 네 가지 특징



먼저, 효성인포메이션시스템의 'AI 플랫폼'은 GPU 서버, 초고성능 병렬 파일 스토리지, 네트워크, 컨테이너와 같은 주요 기술 솔루션을 통합 플랫폼 형태로 제공한다. 특히 복잡한 솔루션 조합이 아닌, 슈퍼마이크로의 GPU 서버 인프라에 래블업사의 컨테이너 기반 Backend.AI 솔루션을 조합하여 아키텍처를 단순화하기 때문에 구축과 운영이 쉽다.

두 번째, AI 업무 환경을 위해 필수적인 고성능 저장 자원의 도입과 유지비용을 절감할 수 있는 초고성능 병렬 파일 스토리지 HCSF를 통해 고성능 데이터를 경제적으로 운영할 수 있는 인프라를 제공한다.

세 번째, AI를 업무에 적용하기 위해서는 무엇보다 AI 개발 역량이 필수다. 단계별로 복합적인 연산 자원 요구사항과 IO 가속 기능의 설정 복잡도 증가, 끝없는 호환성과의 충돌은 검증된 역량이 무엇보다 중요함을 방증한다.

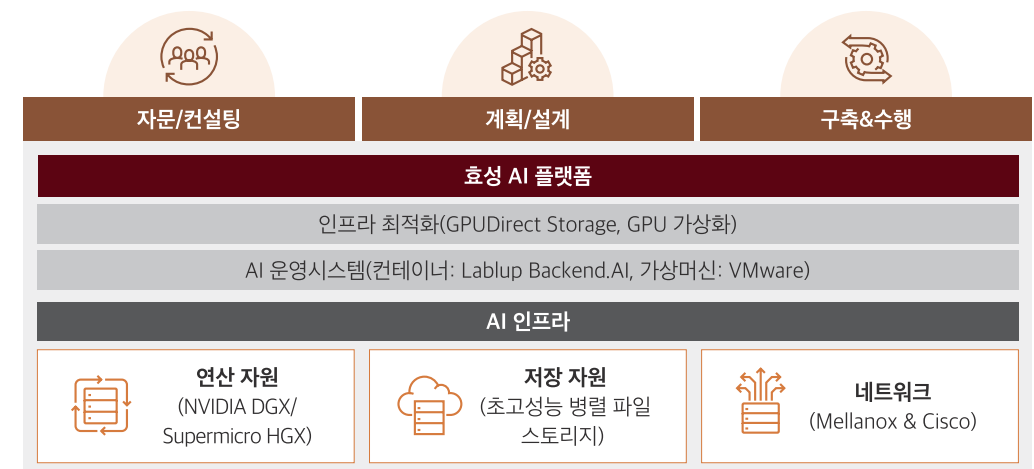
이에 효성인포메이션시스템은 래블업사의 Backend.AI 솔루션을 제안한다. Backend.AI는 AI 개발 및 서비스를 위한 올인원 엔터프라이즈 운영 플랫폼으로 텐서플로우(Tensorflow: 딥러닝과 머신러닝을 위한 오픈소스 라이브러리로 구글에서 개발), 파이토치(Pytorch: 파이썬을 위한 오픈소스 머신러닝 라이브러리)와 같은 사전 정의 개발환경을 제공한다. 이를 통해 데이터 과학자는 연구 환경을 즉시 생성할 수 있다. 또한 직관적인 사용자 UI를 제공하여 효율적인 AI 모델 개발과 용이한 컨테이너의 운영 환경을 지원한다. 현재 Backend.AI는 아태지역 유일의 NVIDIA DGX-레디 소프트웨어로, 국내·외 기업 및 연구소, 대학 등 다수의 레퍼런스를 확보한 전문 솔루션이다.

↓ Backend.AI의 강점

- ✓ 컨테이너 수준의 GPU 분할 가상화 및 NVIDIA GPU MIG 지원 등 GPU 활용 극대화
- ✓ GUI 기반 컨테이너 운영관리 등 직관적인 관리 및 사용자 경험
- ✓ Tensorflow, Pytorch 등 사전 정의 AI 개발 환경 제공
- ✓ 독자적 엔진으로 최적의 GPU 연산자원 배치 구현 등 AI 및 HPC 성능 최적화

현재 효성인포메이션시스템은 성공적인 AI 프로젝트를 위해 래블업과 MOU를 체결하고 GPU 기반의 AI 플랫폼 시장을 공략한다는 목표 아래 국내 AI, ML, HPC 시장 확대를 위한 공동 영업 및 마케팅 활동을 강화하고 있다.

마지막으로, 가장 중요한 사실은 위의 세 가지 장점이 있는 통합 AI 플랫폼을 효성인포메이션시스템 단일 벤더가 제공한다는 점이다. 연산과 저장을 위한 하드웨어 인프라부터 컨테이너 기반의 GPU 가상화 운영 솔루션까지 통합적으로 제안하고 서비스가 가능하기 때문에, 기업들은 AI 시작 단계부터 구축, 운영까지 단일 벤더와 원활하게 커뮤니케이션하며 프로젝트를 추진할 수 있다. 이 부분은 효성인포메이션시스템이 AI 솔루션을 제공하는 타 벤더와 완벽히 차별화되는 점이다.



AI 도입 유형별 차별화 전략

효성인포메이션시스템은 '통합 AI 플랫폼'을 기반으로 AI를 업무에 적용하고자 하는 고객들을 위해 도입 유형별로 세 가지 전략을 제안한다. AI를 신규로 도입하고자 하는 고객, 기존 사용 고객 그리고 초고성능 운영 효율이 주요 관심사인 대규모 사업 추진 고객이 그 대상이다.

Feature

먼저, 신규로 AI를 도입하려는 고객의 고민은 AI를 도입한 후의 성과 창출에 대한 불확실성과 사업 방향성일 것이다. 대부분의 신규 도입 고객은 주어진 예산 내에서 AI 플랫폼을 테스트하고, 이를 기반으로 본격적인 비즈니스를 추진하기 위한 ROI를 도출하려 한다.

효성인포메이션시스템은 처음 AI를 도입하려는 기업들을 대상으로 최소의 비용으로 AI 플랫폼 환경을 경험하고 느껴볼 수 있도록 AI 테스트베드 환경을 제공한다. 일례로, PCIe 기반의 GPU 서버와 NAS 저장 환경에 Backend.AI 솔루션을 구성해 고객이 직접 GPU 분할 가상화와 컨테이너 기반의 사전 정의된 AI 환경을 경험할 수 있도록 돕는다. 소규모 환경에서 충분히 테스트 해봄으로써 향후 비즈니스와 연계해 성과를 이끌 수 있도록 비용, 기술, 역량 문제를 해결하는 AI 신규 도입 전략이다.

두 번째는 무분별하게 도입한 기존의 하드웨어 인프라의 성능 및 운영 방안, 그리고 향후에 도입하는 인프라의 체계적인 도입 방안에 대해 고민하는 고객 유형이다. 이러한 고객들에게 효성인포메이션시스템은 고객이 이미 보유하고 있는 GPU를 종류별로 그룹화하고, 추가로 사용자와 프로젝트별로 리소스를 할당함으로써 시간 무분별하게 사용되던 자원의 체계적인 활용 방안을 제안한다. 같은 종류의 GPU별로 리소스 그룹을 통합할 때 연산 자원이 가진 최대 성능치를 발휘할 수 있다. 즉, GPU가 최대 성능을 발휘하도록 구현함으로써 리소스 낭비를 최소화할 수 있다.

통합 AI 플랫폼이라면, 고성능 학습이 필요한 부분은 NVLink를 지원하는 GPU 그룹으로 묶고, 추가로 HCSF와 인피니밴드¹⁾ 네트워크를 연계해서 성능 최적화된 구성이 가능하다. 기본적인 연산 자원만 필요한 업무를 위해서는 기존에 도입한 PCIe와 NAS 스토리지를 묶어서 요구 성능별로 리소스 그룹을 만들고, 이후에 연구과제별로 프로젝트 그룹의 사용자를 할당해서 프로젝트 성격에 맞는 적절한 리소스 그룹을 매핑하면, 성능 수준은 유지하면서 GPU 리소스 낭비는 최소화할 수 있다. 현재 AI를 업무에 적용해서 사용하고 있는 고객이라면, 이러한 전략이 최적의 AI 플랫폼 구성 방안이 될 수 있다.

마지막은 이미 AI에 대한 업무 이해도가 높은 기업들을 위한 도입 전략이다. 대규모 GPU 팜(Farm) 구축을 원하는 고객들은 '성능 최적화'가 AI를 도입하는 최우선 목적이다. 학습 시간의 최소화를 원하는 이러한 고객들을 위해 효성인포메이션시스템은 '통합 AI 플랫폼' 풀 패키지를 제안한다.

1) 인피니밴드(InfiniBand): 고성능 컴퓨팅과 기업용 데이터센터에서 사용되는 스위치 방식의 통신 연결 방식. 컴퓨팅 노드와 스토리지 장비와 같은 고성능 I/O 장비를 연결할 때 주로 사용된다.

슈퍼마이크로사의 NVLink 지원 GPU 서버와 초고성능 병렬 파일 스토리지 HCSF, Backend.AI 컨테이너 환경과 함께 효성인포메이션시스템의 GPUDirect 스토리지와 같은 IO 성능 최적화 기술을 통합한다면, 고객들이 원하는 최대의 성능을 구현할 수 있다. 특히, 운영과 비용, 성능 효율을 모두 보장하기 때문에 향후 AI 업무 확대를 위한 초석을 마련할 수 있다는 점에서 매력적이다.

↓ 대규모 활용 고객을 위한 AI 통합 솔루션



이 외에도 연산 자원과 저장 자원의 인프라 공급과 함께 이를 운영할 수 있는 AIOps, MLOps 솔루션, 추가적인 VM 환경이 필요할 경우 VM까지 함께 공급한다. 이처럼 효성인포메이션시스템의 통합 AI 플랫폼은 전체적인 구성에 인프라 최적화 기술까지 더해 자문, 컨설팅부터 계획 및 설계, 구축 및 수행까지 단일 벤더에서 통합적으로 제안하는 최고의 솔루션이다.

현재 효성인포메이션시스템의 DX센터에는 고객들이 AI를 업무에 적용하는 과정을 직접 경험할 수 있도록 'AI 업무 프로세스 체험' 과정이 마련되어 있다. AI 모델 개발 환경을 경험하고 AI 모델 학습 및 추론에 대해 시연해볼 수 있기 때문에 AI 도입을 원하는 고객들에게 최적의 환경이 될 것이다.

AI 도입을 여전히 고민 중이라면 바로 지금이 기회다. 기업의 비즈니스 요구사항을 명확히 파악하고 그에 맞는 로드맵을 수립하는 효성인포메이션시스템과 함께 AI 시대를 선도할 수 있는 최적의 솔루션을 경험해보자.