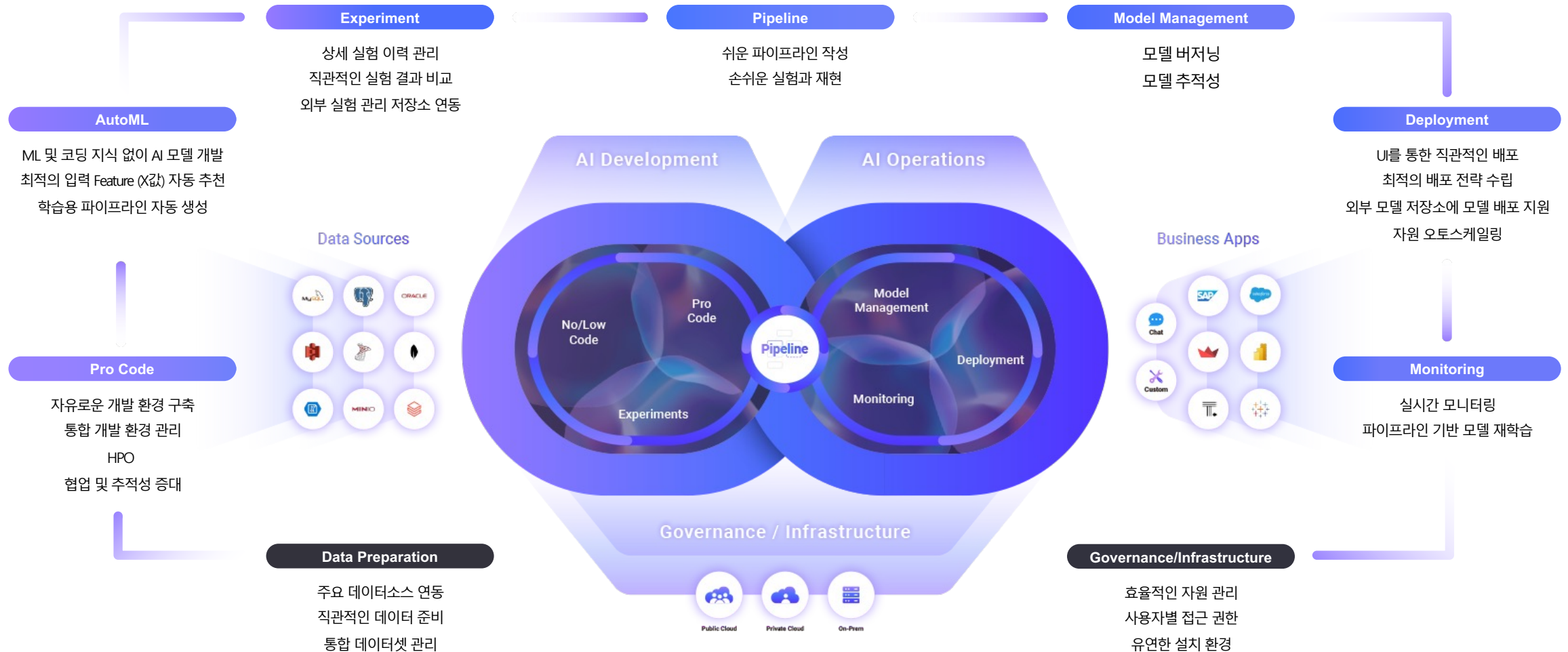




AI 모델 개발부터 자동화된 운영환경까지
AI 운영의 완성, End-to-End AI Platform

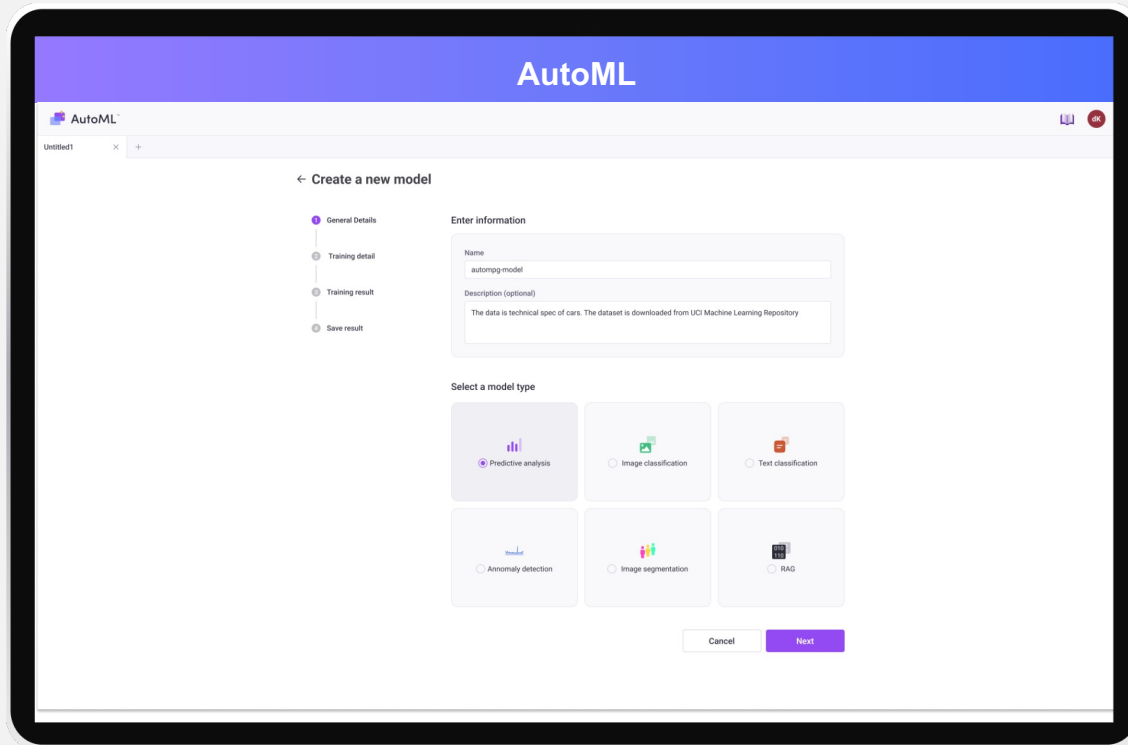
MakinaRocks

What is Runway?



ML 및 코딩 지식 없이 시작하는 AI 모델 개발 : **AutoML**

머신러닝과 친숙하지 않은 실무자도 AutoML을 이용하여 클릭만으로 AI 모델을 쉽게 개발하고 즉시 배포 할 수 있습니다.



✓ ML 및 코딩 지식 없이 AI모델 개발

AI로 풀고자 하는 문제 유형을 선택 후, 가이드에 따라 클릭만으로 결과를 확인하고 즉시 배포 가능한 AI모델 생성

✓ 다양한 시각화 지원

모델 개발 과정에서 데이터 전처리, 모델 선택을 자동으로 진행하며, 주요 분석 지표를 도표 및 그래프 형태로 제공

✓ 최적의 입력 Feature (X값) 자동 추천

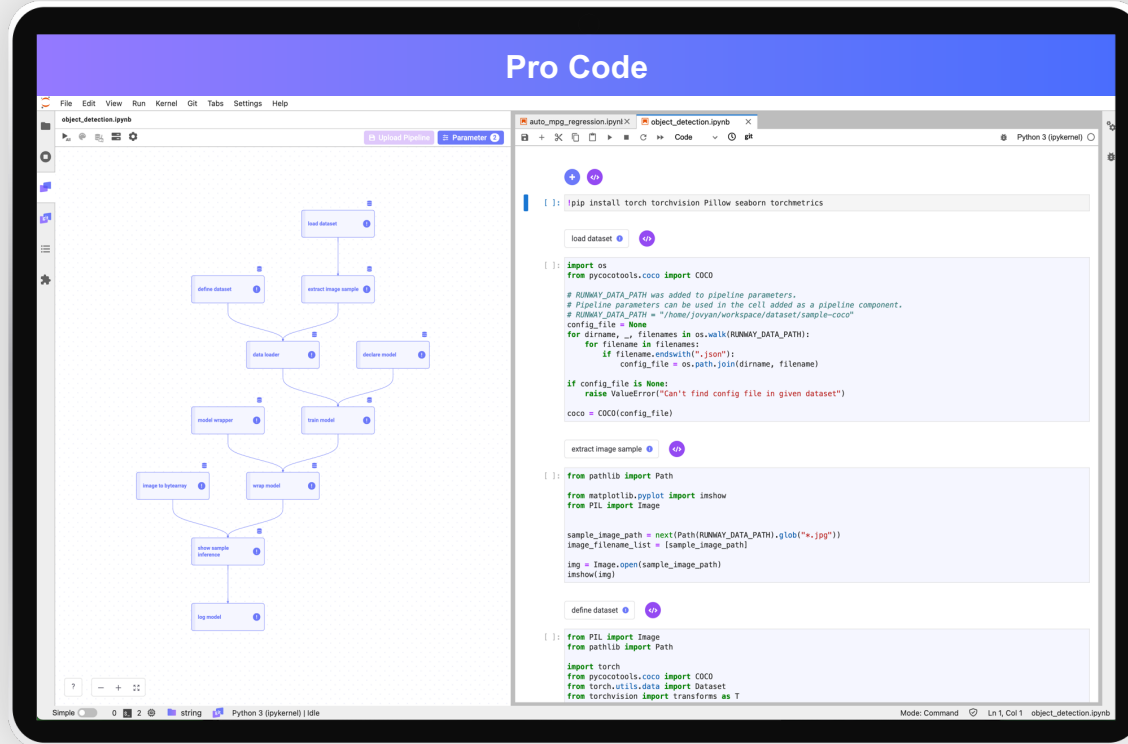
마키나락스의 문제 해결 노하우를 반영한 최적의 입력 Feature 를 자동 추천하여 모델 성능을 보장

✓ 학습용 파이프라인 자동 생성

파이프라인을 자동 생성함에 따라, 전문가들은 쉽게 재현 및 모델 개선을 할 수 있고 운영담당자는 즉시 배포 가능

각 사용자가 원하는 환경에서 작업할 수 있는 통합된 환경 제공 : **Pro Code**

데이터 과학자 및 머신러닝 엔지니어는 데이터를 탐색하고 고성능 AI 모델을 직접 개발하거나 외부에서 개발된 SOTA 모델들을 이용해 다수의 실험들을 수행합니다



✓ 자유로운 개발 환경 구축

다양한 오픈소스 패키지를 설치하거나 자체적으로 패키지 작성 가능

✓ 통합 개발 환경 관리

적용된 개발 환경을 모델 환경에서도 특별한 설정 없이 바로 사용

✓ HPO (Hyper Parameter Optimization)

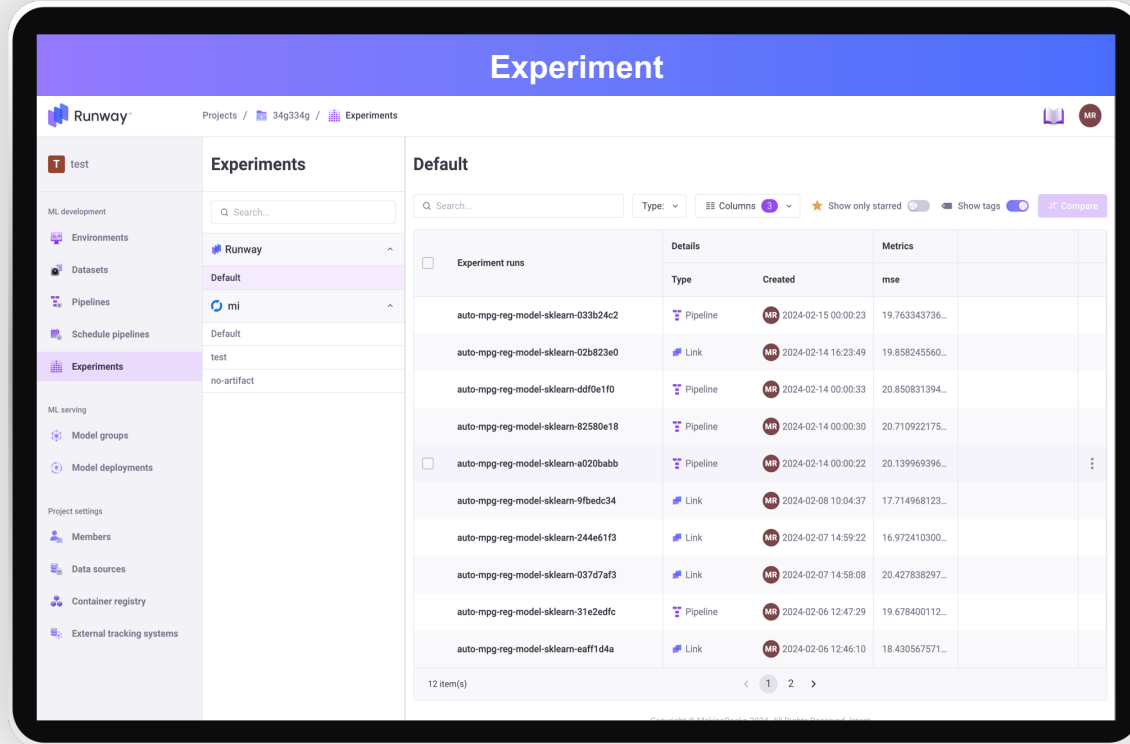
데이터에 맞는 최적의 파라미터를 찾을 수 있는 인터페이스 제공

✓ 협업 및 추적성 증대

코드 레포지토리를 연결하여 파이프라인의 수정을 추적

각 사용자가 원하는 환경에서 작업할 수 있는 통합된 환경 제공 : **Experiment**

개발 단계에서 수행한 다양한 실험 이력을 한 눈에 확인하고
비교 할 수 있고 실험의 결과를 빠르게 재현, 디버깅 합니다



✓ 상세 실험 이력 확인

내부 실험 관리 저장소에 등록된 실험이력 확인

각 실험 별 상세 정보 (Parameter, Metric 등)와 실험에 저장된 모델들의 Artifacts 및 schema 등을 확인

파이프라인을 통해 데이터 전처리, 모델 학습 등의 과정을 손쉽게 수행

✓ 직관적인 실험 결과 비교

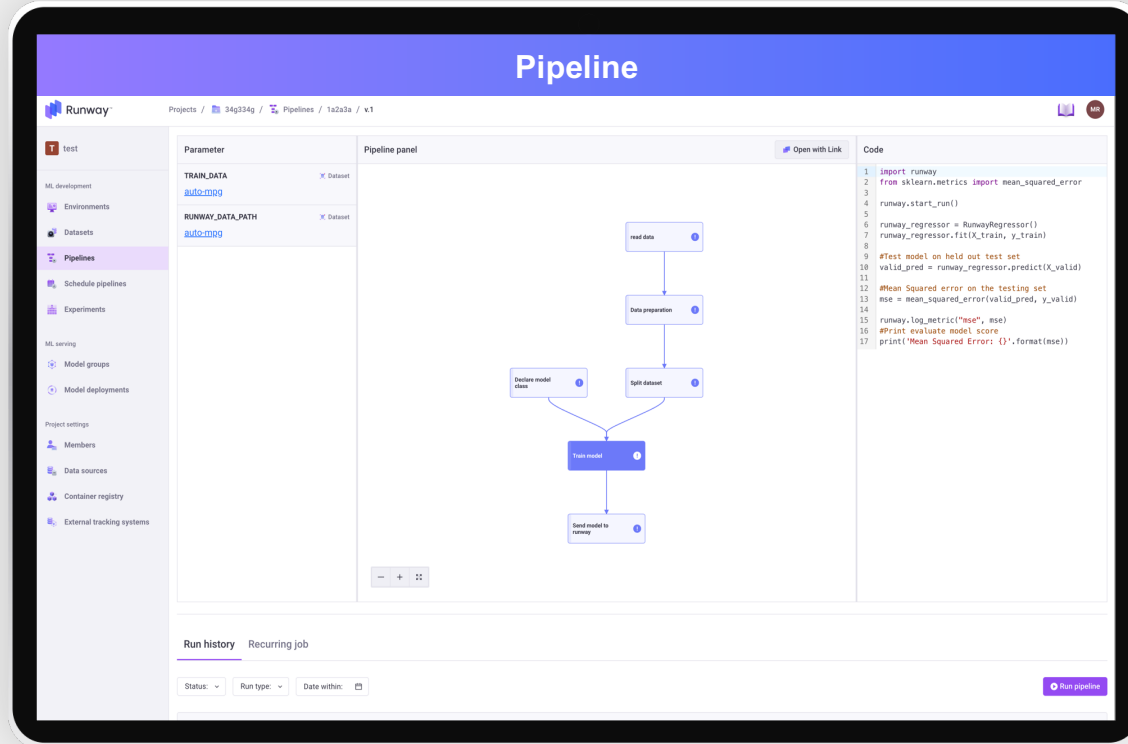
여러 실험의 성능 확인 및 시각화

✓ 외부 실험 관리 저장소 연동

외부 실험 관리 저장소 (WandB, MLFlow 등)을 연결하여 Runway상에서 이력을 확인하고 즉시 배포

각 사용자가 원하는 환경에서 작업할 수 있는 통합된 환경 제공 : **Pipeline**

각 개발 환경에서 작성한 결과물을 재현 가능한 파이프라인으로
손쉽게 작성하여 빠르고 원활한 ML 라이프사이클을 구현합니다



✓ 쉬운 파이프라인 작성

각 작업 환경에서 개발된 코드를 손쉽게 파이프라인으로
변경하는 인터페이스 및 Python SDK 제공

파이프라인을 통해 데이터 전처리, 모델 학습 등의 과정을
손쉽게 수행

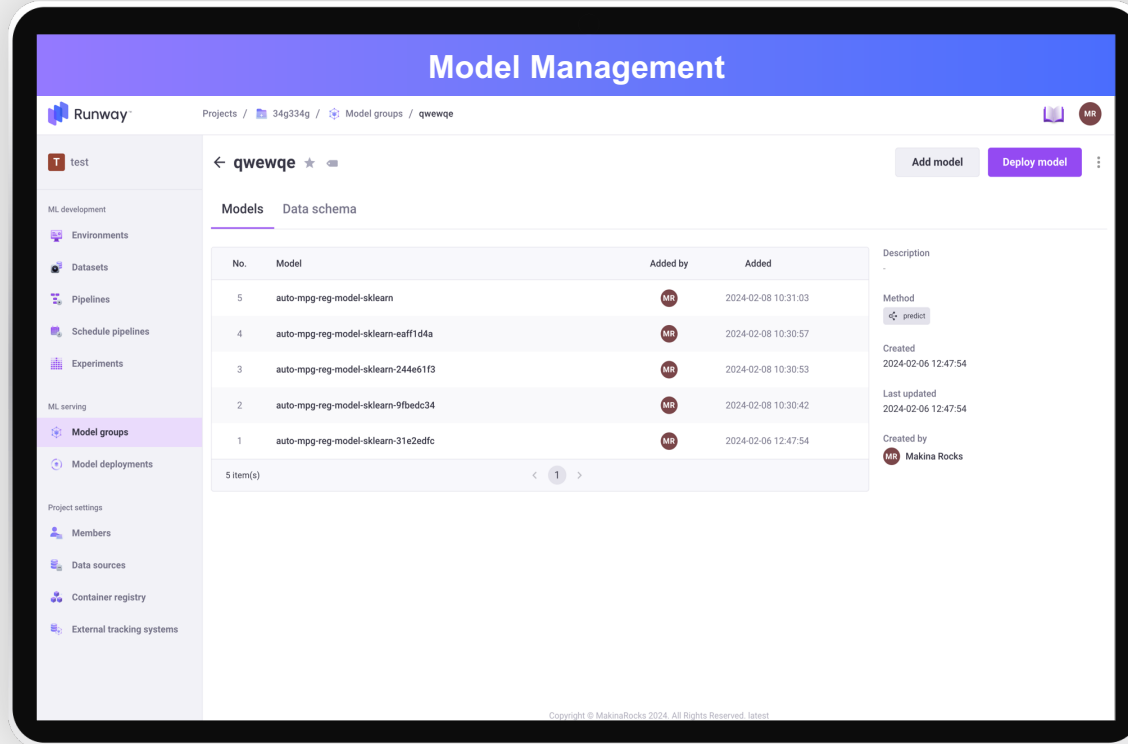
✓ 손쉬운 실험과 재현

파이프라인의 데이터 및 파라미터를 변경해 쉽게 실험

No/Low Code 또는 Pro Code 환경에서 개발한 학습
파이프라인을 저장하고 재현 가능한 형태로 관리

개발과 운영 환경 간 원활한 통합으로 모델 재현성 및 운영 안정성 확보 : **Model Management**

동일한 추론을 하는 모델을 그룹 단위로 관리하여
배포된 모델의 추적성과 재현성을 보장합니다



✓ 모델 버저닝

모델 버전 컨트롤을 통해 AI 서비스 확장 시 필요한 작업을 최소화하고, 배포된 모델에 문제가 발생할 경우 손쉽게 롤백

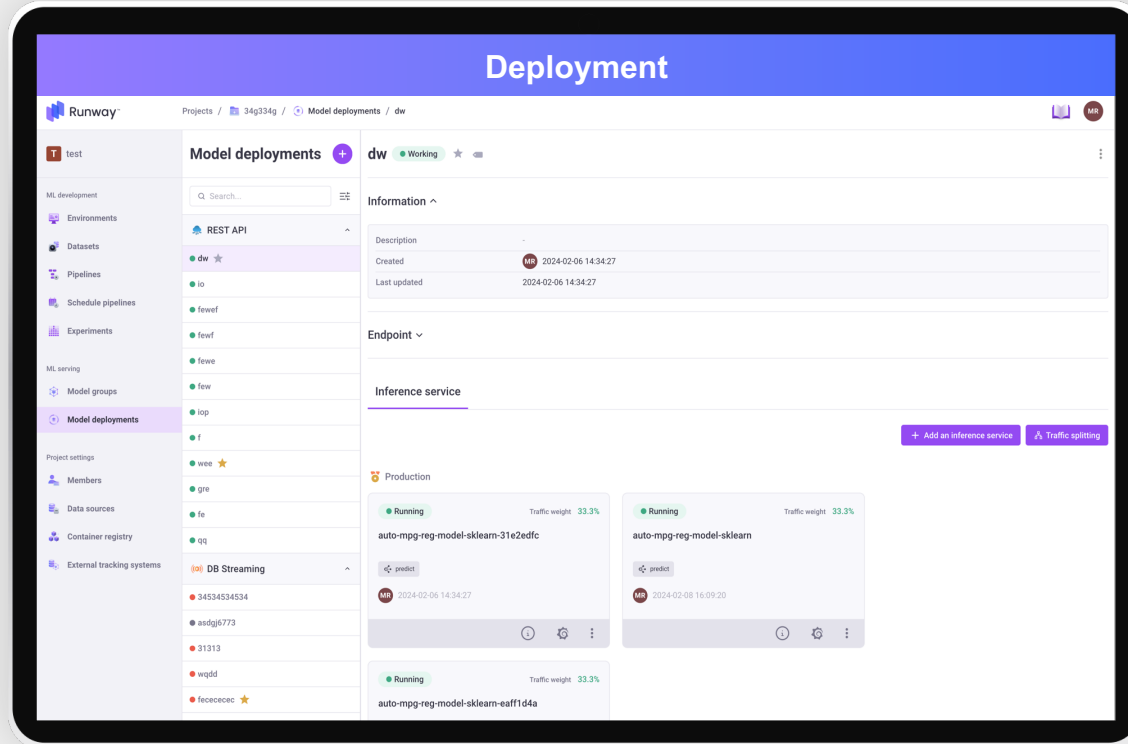
✓ 모델 추적성

모델에서는 등록된 모델들의 소스코드, 데이터 스키마, 라이브러리 및 프레임워크 등을 추적할 수 있으며, 모델의 재현성을 보장

입력 데이터의 변화, 모델 성능 저하 등 AI 서비스에 영향을 미칠 수 있는 상황이 발생할 시, 빠른 재현을 통해 문제 해결

개발과 운영 환경 간 원활한 통합으로 모델 재현성 및 운영 안정성 확보 : **Deployment**

사용자 친화적인 UI를 통해 배포 대상 모델, 배포를 위해 할당할 자원, 서버링 방식 (실시간, API, 배치 서버) 등을 선택하여 쉽고 빠르게 AI 서비스를 구축할 수 있습니다.



✓ UI를 통한 직관적 배포

REST API 엔드포인트를 생성하여 모델을 배포하거나 데이터 소스와 연결하여 실시간으로 모델을 배포

✓ 최적의 배포 전략 수립

모델의 성능 유지를 위해 새롭게 배포하는 모델의 전략 수립 및 결과에 따라 최종 모델 선택

✓ 외부 모델 저장소에 모델 배포 지원

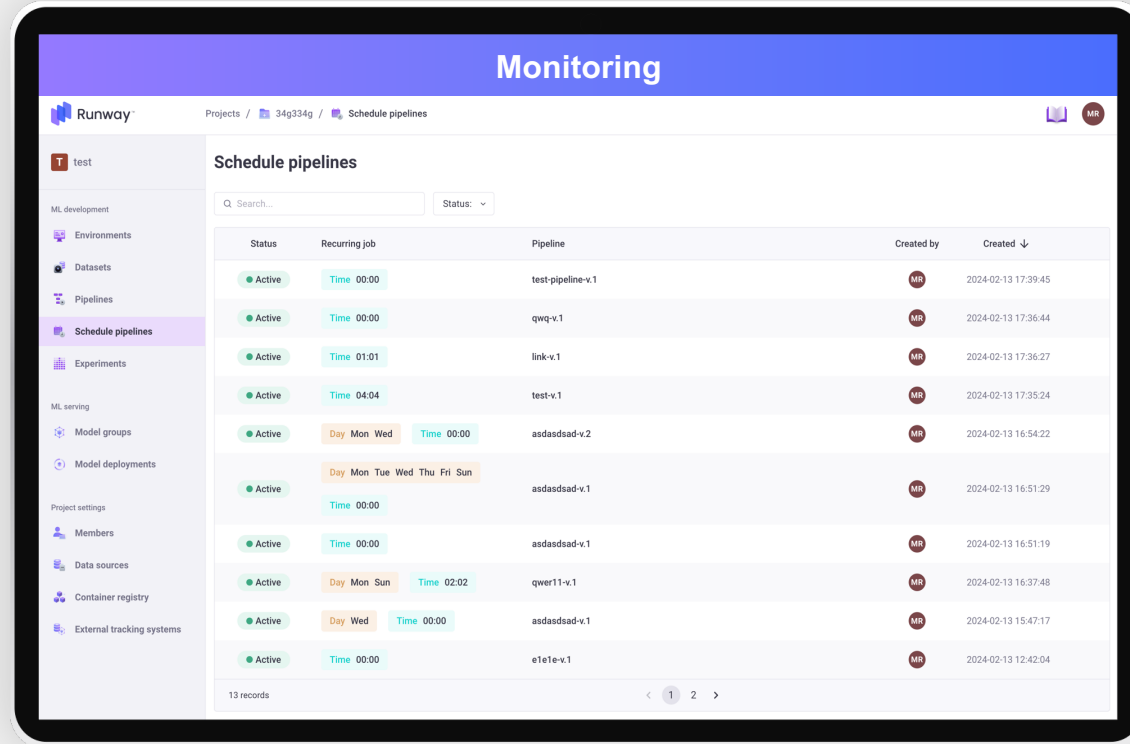
외부실험 저장소(WandB, MLFlow)에 있는 모델을 Runway에서 배포 및 관리

✓ 자원 오토스케일링

API 배포 시 Autoscaling 기능을 통해 컴퓨팅 자원을 효율적으로 관리

컴퓨팅 자원 운영 효율성 보장 및 거버넌스 체계 구축 지원 : **Monitoring**

배포된 모델의 성능을 확인하고 추론 로그를 확인하여
AI 서비스의 운영 안정성을 보장합니다



✓ 실시간 모니터링

실시간으로 모델 성능을 모니터링 하여 구축한
AI 서비스의 성능 유지

파이프라인 실행 로그를 기반으로
성능 저하 원인 분석 및 개선사항 도출

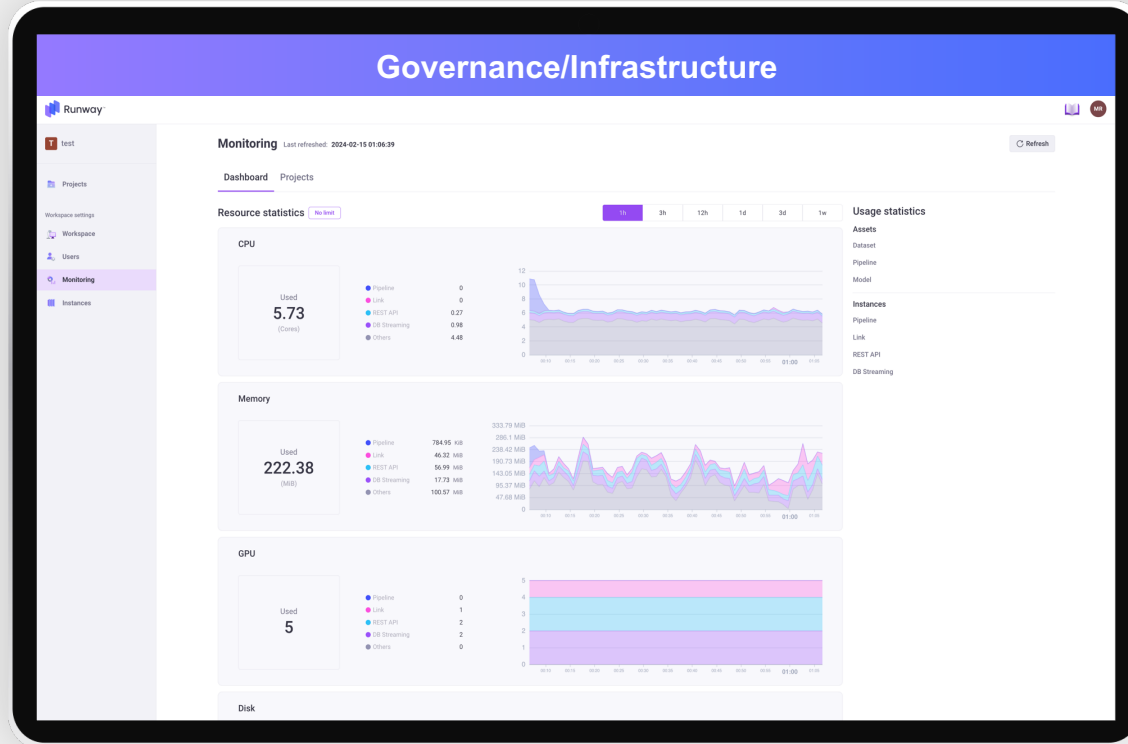
✓ 파이프라인 기반 모델 재학습

배포된 모델은 주기적으로 최신 데이터나
변화하는 조건을 기반으로 업데이트 필요

설정된 주기 또는 특정 이벤트에 따라
자동 재학습 설정 가능

컴퓨팅 자원 운영 효율성 보장 및 거버넌스 체계 구축 지원 : **Governance/Infrastructure**

프로젝트 및 작업 별로 자원을 지정하고 관리할 수 있으며, 다양한 인프라 요구 사항을 충족하여 기존 인프라 환경에 큰 변경 없이 AI 운영 환경을 구현합니다



✓ 효율적 자원 관리

AI 모델 개발, 학습 및 배포를 위해
할당할 자원 직접 설정

프로젝트 별 사용중인 자원 추적을 통해
자원 관련 비용 최적화

✓ 사용자별 접근 권한

사용자별 접근 권한 관리를 통해 민감한
데이터, 코드, API 등을 안전하게 관리

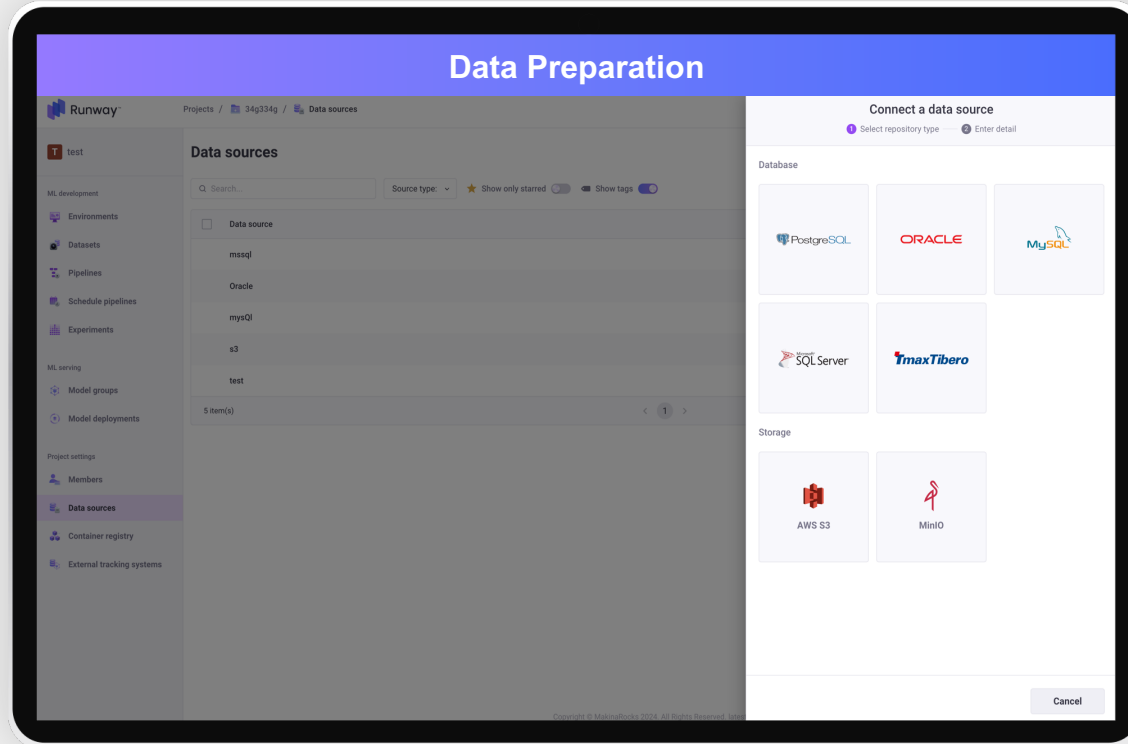
✓ 유연한 설치 환경

온프레미스, 클라우드 (Public, Private) 그리고
하이브리드 환경 설치

필요 시 하이브리드 클러스터 구축

빠른 데이터 연결, 손쉬운 데이터 조합 및 변형으로 데이터 추적성 증대 : **Data Preparation**

주요 데이터 소스를 연동하여 쉽고 빠르게 데이터셋을 생성하고
다양한 참여자가 동일한 데이터셋을 기반으로 모델을 개발하고 소통합니다



✓ 주요 데이터 소스 연동

주요 데이터 소스에 대한 연동을 지원

Amazon S3, Azure Blob Storage, Google Cloud Storage, Snowflake, Databricks Lakehouse, SQL 데이터베이스, NoSQL 데이터베이스 등

✓ 직관적인 데이터 준비

클릭 만을 데이터셋 생성, 정제, 변경할 수 있는
인터페이스 제공

데이터 파이프라인을 시작적으로 표현하여
쉽게 데이터 분석

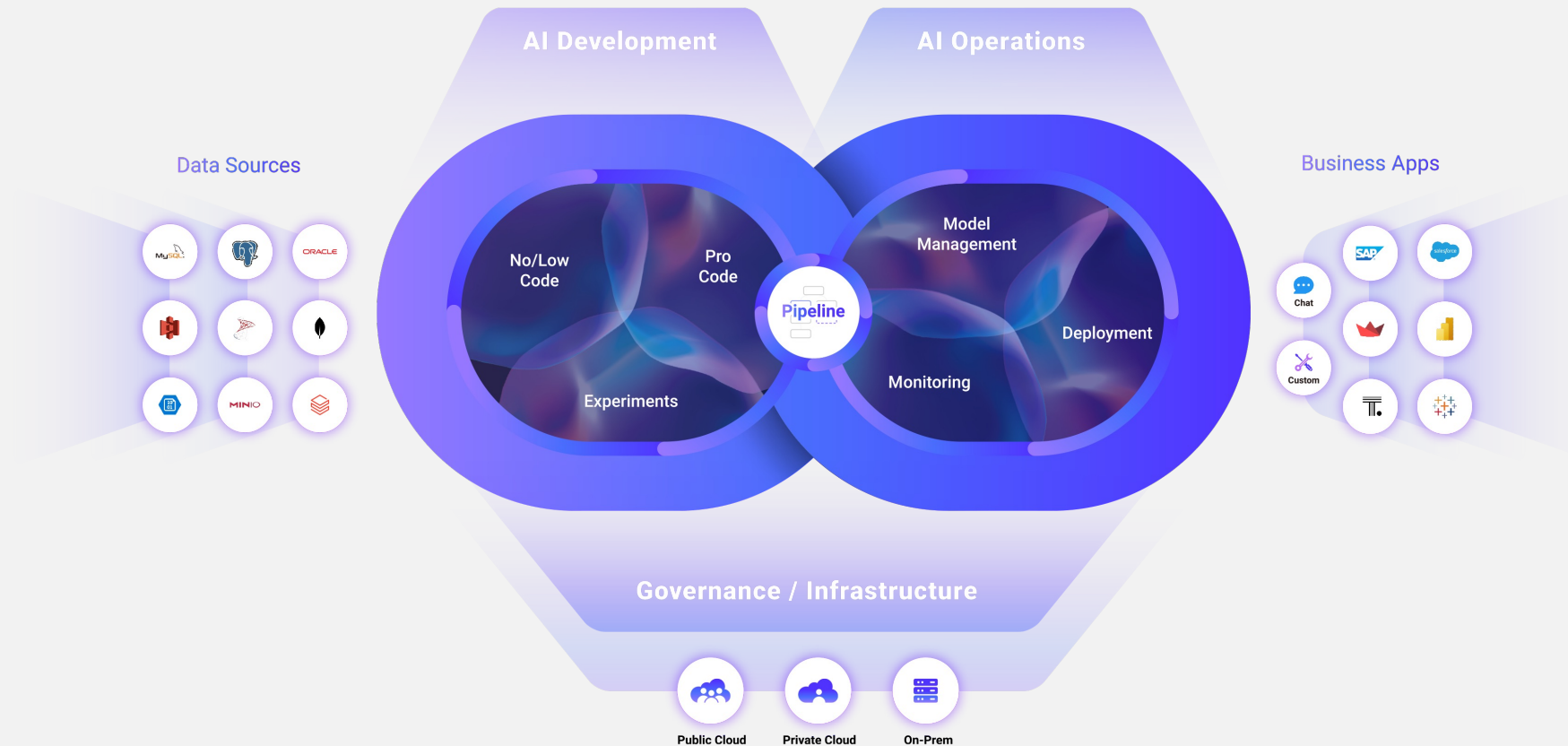
No Code 기반 데이터 전처리 컴포넌트로
데이터 유형별 전처리 함수 제공

✓ 통합 데이터셋 관리

다양한 참여자가 하나의 통합된 환경에서
동일한 데이터셋을 기반으로 개발

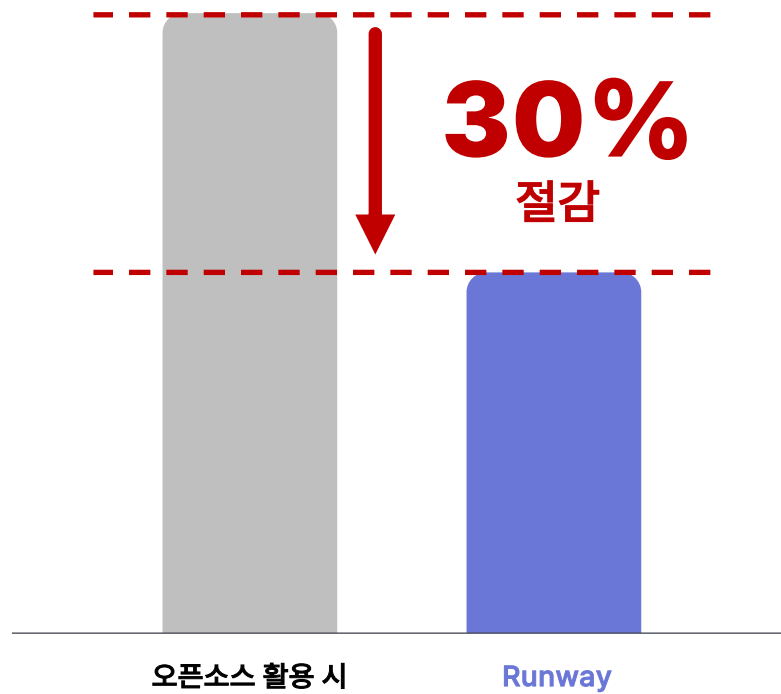
마키나락스 런웨이

런웨이는 AI플랫폼 구축부터 구축 후 운영/활용 측면에서
합리적인 비용으로 자동화된 AI 운영체계를 제공하는 차별화된 솔루션입니다



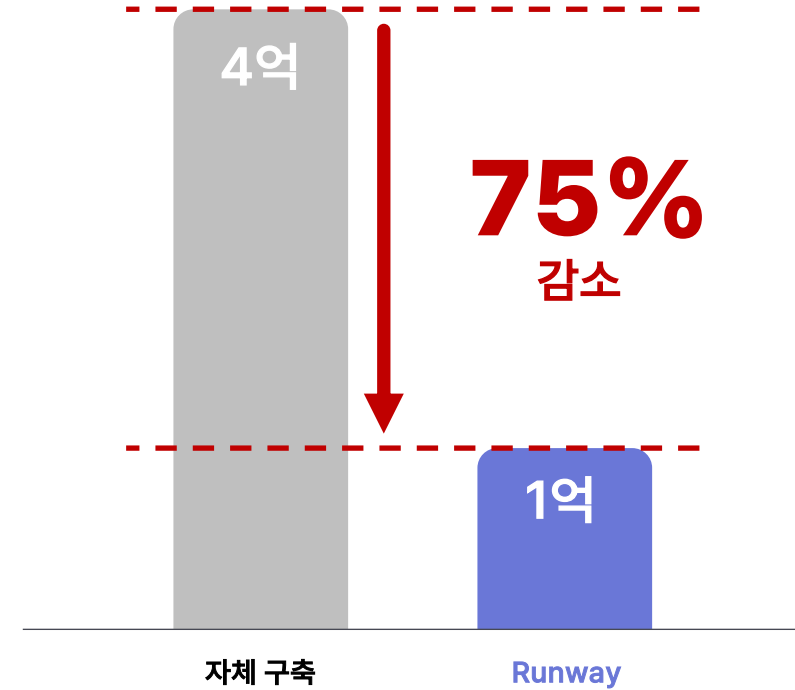
비용 최적화 #1: 초기 도입 비용

구축을 위한 자원비용



* 데이터-개발-학습-배포 기능을 오픈소스로 구현 시 클라우드 비용

1년 내 구축을 위한 인건비



*엔지니어 8명 */연 0.5억

* 백엔드 3명, 인프라 1명, 데이터 엔지니어 2명, ML엔지니어 2명

비용 최적화 #2: 운영 비용 – Onboarding Cost



* 동일 AI 모델 개발 및 배포 시나리오 구현시 스텝 수 비교

Simple

글로벌 솔루션 대비 워크플로우 간소화

Easy

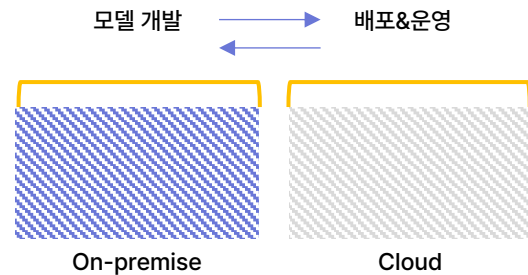
쉬운 GUI 바탕으로 비전문가의
온보딩 시간/비용 절감

기존 IDE 활용 & 기타 SW와의 손쉬운
연동으로 신규 사용자 온보딩 시간 절약

비용 최적화 #2: 운영 비용 – Computing Cost

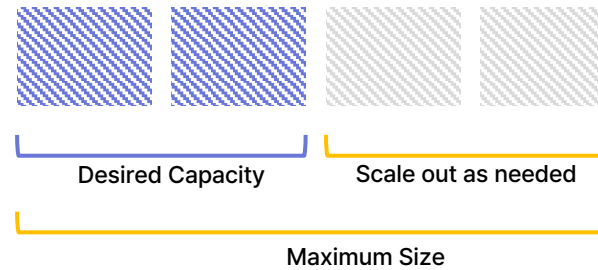
Hybrid

완벽한 하이브리드
MLOps 환경 구축으로
학습- 추론 자원 최적화



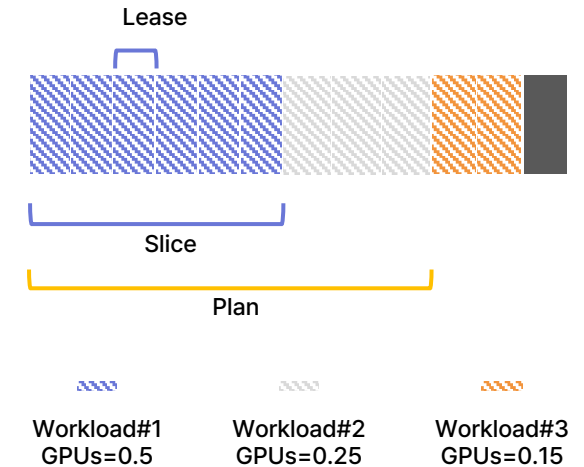
Cloud

클라우드 활용 시
자원 auto-scaling 기능을 통해
클라우드 자원 최적화



On-premise

GPU Slicing을 통한
스케줄링 최적화

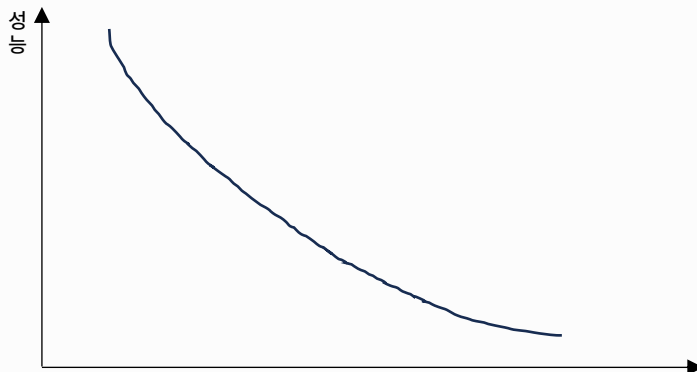


성능 지속성: Continuous Training이란?

Continuous Training은 머신러닝 모델을 초기 학습 이후에도 지속적으로 새로운 데이터를 활용하여 모델을 업데이트 하고 개선하는 과정입니다

Level 0 – Improvised learning

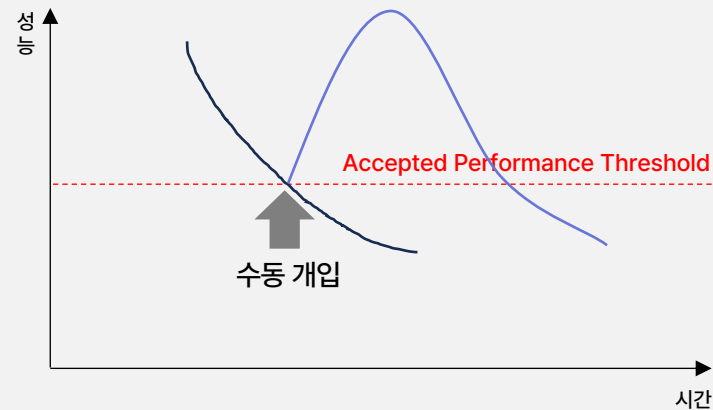
모델 성능 유지를 위해 반복적인 인적비용 발생
새로운 데이터를 활용하고 재배포하는 데 느림



모델 성능 저하 발생 및 AI 실적용의 허들로 작용

Level 1 - Semi-autonomous learning

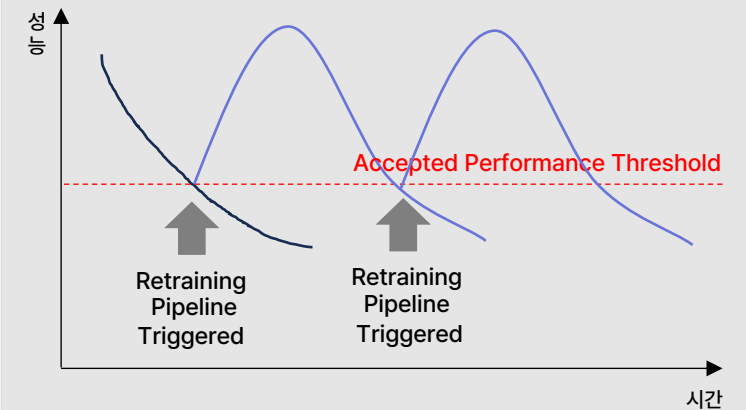
데이터 분포 변화로 인한 성능 저하에 대응해
수동으로 재학습 및 재배포



수동 재학습

Level 2 – Continuous learning

연속적인 Training을 통해
실험부터 배포 운영까지의 AI 라이프사이클 자동화
변화하는 상황에 대응하여 모델 재보정

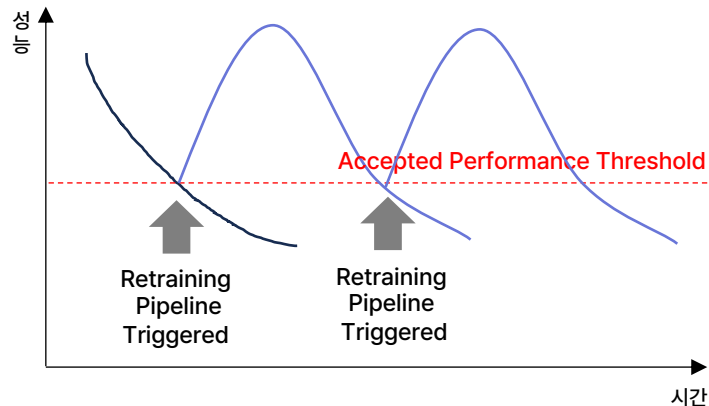


자동 재학습

자동화된 Continuous Training의 요구사항

지속적 학습을 위해 데이터 준비부터 모델 검증, 배포, 피드백까지 전 과정에서의 자동화가 필요합니다

Level 2 – Continuous learning



자동 재학습

10가지 MLOps 기능의 자동화 단계

MLOps 2.0 dimension

	Level of automation		
Data Collection	Manual checks	Data pipeline (batch)	Realtime ingestion
Data Validation	Ad-hoc ingestion	Data validation tests	Realtime data triaging
Feature Engineering	Ad-hoc	Feature store	Automated
Experimentation	Improvised trials	Catalogued trials	Reproducible pipelines
Training	Siloed training	Scalable pipelines	Continuous training
Model Validation	Offline validation	Model validation suite	Automated validation tests
Integration	Manual build & test	Automated build & test	Continuous integration
Deployment	Manual deployment	Automated deployment	Continuous deployment
Model Management	Manual inventory	Model registry	Model monitoring suite
Feedback	Ad-hoc collection	Feedback data stored	Feedback loop to pipeline

Runway™

Runway Use Case

유리 및 세라믹 제조

식품 유통

보험

배터리

Use Cases | 유리 및 세라믹 제조

PLC OT 데이터 기반
AI 플랫폼 구축

Overview

북미 최대의 자동화기업인 로크웰(Rockwell)의
솔루션과 Runway를 결합해 **PLC단에서 수집된
다양한 OT 데이터**(모터, 센서 등)를 활용한 딥러닝 기반
AI 모델을 적용했으며, 이를 빠르게 배포할 수 있는
PLC OT 데이터 기반 AI 플랫폼 구축



Challenge

디스플레이 제조에 사용되는 유리 및 세라믹 공정의
생산성 향상을 위해 AI 모델을 개발했으나, **AI 운영 체계의
부재로 일회성 개발이 지속되어 개발 및 유지보수 비용 증가**

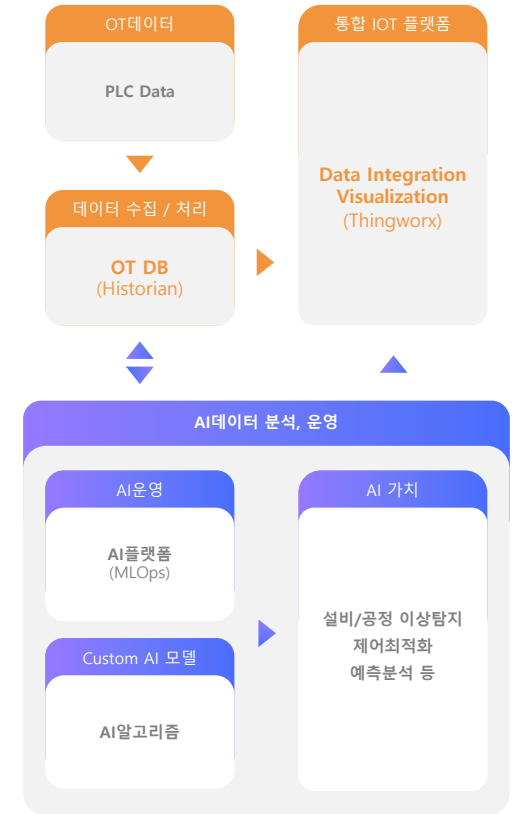
운영 단(OT, Operation Technology) PLC에서 실시간으로 나오는
높은 해상도의 데이터를 AI 모델 개발에 활용하기 어려움

현장의 도메인 전문가의 인사이트를 AI 모델에 반영하지 못해
다양한 Feature가 반영된 고성능 AI 모델 개발의 한계

Approach

OT 운영 단에 구축된 On-Prem 서버와 DB SW를 활용해 현장
데이터를 연동해 높은 해상도의 데이터의 연결성 확보하고,
이를 AI의 빠른 **재학습/재배포가 가능한 AI 운영 체계**(MLOps,
Machine Learning Operations) 와 결합해 제공

고객사의 현장 데이터와 공정에 높은 이해를 가진
도메인 전문 기업(Rockwell)의 솔루션과 연계해 지속가능한
높은 성능의 AI 모델 개발



Value Delivered

데이터 기반 사전예방적 접근으로 유지보수 비용 20-30% 절감

비효율적인 유지보수 및 점검 계획을 최소화 해 작업자 생산성 50-60% 증대

인공지능 기반 예측 분석으로 다운타임 5-8% 감소

Use Cases | 식품 유통

신선 식품 할인 정책 최적화 신속 학습/배포 AI플랫폼 구축

Overview

당일입고 신선식품의 **요일/시간대별 할인율 최적화 AI 서비스 구축**을 위한
실시간 데이터 반영이 가능한
신속 학습/배포 AI 플랫폼 구축



Challenge

점포별 마감 담당자의 **경험적 판단에 의존한 무분별한 할인율 적용**(조기할인, 과다할인 등)으로 인한 판매가 인하 발생 및 소비자의 할인 기대감 고착

AI를 활용한 분석 및 적정가격 산출을 통해 **이익 증대 및 폐기비용 감소** 달성

효과성 검증 후 **타 점포로의 수평 전개**를 통한 전체 이익 신장을 궁극적 목표로 함

Approach

할인라벨 발생 이력을 통해 개선 및 절감 효과가 클 것으로 예상되는 주요 분석 대상을 선정하여, 할인상품/할인시간대/할인율 분석을 통한 **AI 모델링 작업**

데이터 및 분석 결과에 따른 '할인-판매 추이' 관계를 통해 다양한 상황과 조건을 고려한 **최적의 할인율 도출**

Runway 기반 파이프라인 기능을 통해 일일 온라인 데이터 분석 후, 현장 실 적용을 위해 기존 시스템과 Direct Interface, **안정적으로 모델 배포 서비스 운용**

Runway 기반 신속 학습/배포 사이클을 이용하여 실시간으로 전달되는 **현장의 피드백을 빠르게 적용**

Value Delivered

실 적용 후 현장 업무로드 절감 및 데이터에 기반한 의사결정을 통해, 신선식품의 이익 증가와 폐기비용 감소 확인

AI 모델의 확장성 입증, AI 플랫폼 기반으로 타 점포 확대 적용 추진

Use Cases | 보험

자동차 사고 사진정보 식별/관리 AI 시스템 구축

Overview

차량 사고 이미지 분석에 사용되는
다양한 AI 모델을 복합적으로 적용하고,
데이터 관리부터 신규 차종 출시에 따른
모델 재학습 및 배포가 빠르게 가능한
자동차 사고 사진정보 식별/관리 시스템 구축



Challenge

기존 차량사고 이미지 분석 시스템이 monolithic 아키텍처(단일 AI 모델만을 배포할 수 있는 환경)로 구성되어 시스템 업데이트/유지보수가 어렵고, **모듈의 일부분만 변경해도 전체 서비스에 영향을 미침**

AI 모델 개발-운영 환경이 분리되어 모델 성능 개선을 위한 **재학습 관련 다수 제약 발생**

모델 간 성능 비교를 위한 모니터링 과정이 시각화되지 않아 **모델의 성능 저하를 식별하기 어렵고, 서비스의 품질 관리가 어려움**

Approach

병렬 마이크로서비스 아키텍처 (다중 모델 배포가 가능한 환경)로의 전환을 통해
모듈 변경에 따른 운영 서비스 영향도를 최소화하고 새로운 서비스로 확장이 가능한 배포 환경 구축

개발 - 운영 환경을 통합하고 재학습 파이프라인을 구축하여 일원화된 플랫폼에서
지속 가능한 ML 서비스 제공

모델 성능 비교 GUI를 제공하여 재학습 모델과 서비스 모델간 A/B 테스트 모니터링 지원

Value Delivered

기존에는 단일 모델을 배포하여 서비스를 제공했으나, 현재 8개의 다양한 모델을
복합적으로 사용하여 서비스 운영 중

연간 1회 진행됐던 재학습 프로세스를 Runway를 통해 수시로 재학습이 가능하여
신차 출시 주기에 맞춰 신속하게 모델 업데이트 가능

Site License 정책을 활용, 기존 영구 라이선스 대비 소프트웨어 유지보수 비용 40% 감소

Use Cases | 배터리

배터리 극판 공정 회전체 이상탐지 솔루션

Overview

생산성과 비용에 직결되는 연결 공정 내 존재하는 다수의 회전체 장비들의 상태를 한눈에 파악하고 관리하여 예측 가능한 생산 공정 운영 시스템 구축



Challenge

생산 공정 핵심 장비의 예기치 않은 가동 중단이 주 생산 라인 폐쇄로 이어져 **생산성 저하 및 대안 실행(긴급 생산 가동 등)**을 위한 **인건비 상승** 요인 발생

수백-수천대에 달하는 장비 중 단 하나의 장비 비가동 발생 시 전체 공정이 중단되기에 **전체 장비를 유기적으로 관리하는 시스템** 필요

장비의 현재 상태와 관계 없이 유지보수를 위해 **일정 수준의 인력을 지속적으로 투입**해야하는 비용 부담 존재

Approach

인공지능 기반 이상탐지 솔루션 및 **실시간 데이터 수집 파이프라인**을 활용하여 소형/중형/대형 회전체 장비에 모두 적용 가능한 **안정적인 데이터 수집 프로세스 구축**(100,000,000/day)

각각의 회전체 장비에 대한 이상탐지 모델을 운영하고 추론을 지원하며, 추론된 각 장비의 상태를 **한눈에 모니터링 가능하도록 원페이지 대시보드** 제공

향후 다양한 공정과 새로 추가되는 장비를 모니터링 하고, 인공지능 기반 솔루션을 통합할 수 있는 확장 가능한 **AI 운영 플랫폼** 구축

Value Delivered

12대 장비 대상 이상탐지 솔루션 적용, 이상 발생 시 사전에 고장을 예측하고 작업자에게 고장 예측 알람을 제공하여 장비 다운타임 최소화

매주 수작업으로 진행되는 이상탐지 활동을 자동화, 작업 인력의 운용 효율화 및 인건비 절감

Use Cases | 전자 부품 제조

전기전자 산업 AI 워크플로우 최적화를 위한 All-in-one AI 플랫폼 구축

Overview

데이터 구축부터 자원 최적화까지 갖춘 All-in-one AI 플랫폼을 구축하여 제조 공정의 머신 비전 AI 워크플로우를 효율화



Background

비전 모델을 재학습-재배포를 할 수 있는 운영 환경의 부재로 지속적인 모델의 성능 유지가 어려움

신규 이미지 데이터로 재학습이 필요한 경우, 라벨링 부터 재학습까지의 긴 프로세스 반복으로 업무 비효율 초래

반복 학습에 필요한 HPC 자원 활용 최적화 필요성 대두

Solution

학습할 이미지 데이터의 라벨링을 자동화하는 오토 어노테이션툴을 ML옵스 환경과 연동

HPC(High Performance computer) 스케줄러를 연동하여 컴퓨팅 자원의 워크로드 관리, 실시간 모니터링 등 고도화된 자원 관리 지원

데이터셋 준비-모델 학습-재학습-재배포를 자동화하는 지속적 학습(CT, Continuous Training) 환경 구축

Outcome

데이터 라벨링과 재학습-재배포 사이클을 자동화 하여 기존 프로세스 대비 시간을 약 80% 절감

하나의 AI 플랫폼 상에서 데이터 라벨링-모델 학습-모니터링-에러 데이터 재라벨링-재학습 및 재배포 사이클 운영

HPC 자원 관리 스케줄러 연동으로 컴퓨팅 자원의 최적화를 통한 안정성과 비용절감 실현