

고위험 현장의 안전관리를 위한 AI 클라우드 플랫폼 설계

김기봉
대전보건대학교 컴퓨터정보과 교수

A Design of AI Cloud Platform for Safety Management on High-risk Environment

Ki-Bong Kim
Professor, Dept. of Computer Information, Daejeon Health Institute of Technology

요약 최근 기업과 공공기관에서 안전 이슈는 더는 미룰 수 있는 상황이 아니며, 대형 안전사고가 발생했을 때 직접적인 금전적 손실뿐 아니라 해당 기업 및 공공기관에 대한 사회적 신뢰가 함께 떨어지는 간접적인 손실도 매우 커진다. 특히 사망 사고의 경우는 더욱 피해가 심각하다. 이에 따라 기업 및 공공기관은 산업 안전 교육과 예방에 대한 투자를 확대함에 따라, 고위험 상황이 존재하는 산업현장에서 사용자 행동반경에 영향을 받지 않고 안전관리 서비스가 가능한 개방형 AI 학습모델 생성 기술, 에지단말간 AI협업 기술, 클라우드-에지단말 연동 기술, 멀티모달 위험상황 판단기술, AI 모델 학습 지원 기술을 이용한 시스템 개발이 이루어지고 있다. 특히 인공지능 기술의 발전과 확산으로 안전 이슈에도 해당 기술을 적용하기 위한 연구가 활발해지고 있다. 따라서 본 논문에서는 고위험 현장 안전관리를 위해 AI 모델 학습 지원이 가능한 개방형 클라우드 플랫폼 설계 방안을 제시하였다.

주제어 : 인공지능, AlaaS, 딥러닝, 융합, 4차 산업, 클라우드 플랫폼

Abstract Recently, safety issues in companies and public institutions are no longer a task that can be postponed, and when a major safety accident occurs, not only direct financial loss, but also indirect loss of social trust in the company and public institution is greatly increased. In particular, in the case of a fatal accident, the damage is even more serious. Accordingly, as companies and public institutions expand their investments in industrial safety education and prevention, open AI learning model creation technology that enables safety management services without being affected by user behavior in industrial sites where high-risk situations exist, edge terminals System development using inter-AI collaboration technology, cloud-edge terminal linkage technology, multi-modal risk situation determination technology, and AI model learning support technology is underway. In particular, with the development and spread of artificial intelligence technology, research to apply the technology to safety issues is becoming active. Therefore, in this paper, an open cloud platform design method that can support AI model learning for high-risk site safety management is presented.

Key Words : Artificial Intelligence, AlaaS, Deep Learning, Convergence, the fourth industrial revolution, Cloud Platform

*Corresponding Author : Ki-Bong Kim(kbkim@hit.ac.kr)

Received November 17, 2022
Accepted December 21, 2022

Revised December 1, 2022
Published December 30, 2022

1. 서론

최근 급속하게 발전하고 있는 인공지능 기술은 경제, 사회, 군사 등 다양한 분야에 적용되어 인간의 삶과 산업 패러다임을 변화시키고 있다. 특히, 고위험 현장의 사고 예방 필요한 산업에서도 AI 기술을 도입하여 안정성을 높이고자 하는 움직임이 존재한다. 이에 따라 건설용 인공지능 기술은 용도에 따라 프로젝트 관리, 필드 관리, 리스크 관리, 스케줄 관리, 공급망 관리, 기타로 분류되어 적용되고 있다. 국내 정부 또한 2017년 이후 4차 산업혁명 시대에 대응하고 스마트 건설기술의 활성화를 위한 각종 지원 정책과 방안, 로드맵 등을 제시하면서, 2025년까지 스마트 건설 핵심기술 상용화를 실현하기 위해 현장 근로자 안전용 웨어러블 디바이스에 GPS 추적기, 심박수 측정기, 압력 및 공기 모니터링 센서 등을 부착하는 기술 연구 및 개발을 진행하고 있다[1,2]. 하지만, 산업현장에 안전관리를 위한 서비스가 증가하고 있으나, 단일 센서의 제한적인 감시 범위, 화각 한계를 극복하지 못하고 정해진 시나리오에 따른 안전관리로 인해 이상상황 발생시 제대로 역할을 수행하지 못하는 한계가 있다. 따라서 다양한 예지단말들간에서 서로 정보를 주고받으면서 생체정보까지 감지 가능한 복합 센서 모듈과 인공지능 기반 산업현장 안전관리 솔루션에 대한 수요가 급증하고 있다. 하지만 이러한 건설 분야 뿐만아닌 다양한 산업분야의 AI 서비스 수요 급증에 맞춰 산업현장에 적용가능한 AI 모델을 개발하기에는 시간과 노력, 비용이 소요되는 문제점이 있다[3]. 이에 따라, 해외 기업 중심의 클라우드 플랫폼 시장에서 AI 모델 학습 지원을 위한 플랫폼 서비스 개발이 활발하게 이루어지고 있으며, 국내에서도 과학기술정보통신부와 한국지능정보사회진흥원(NIA)의 연구개발 지원으로 AI를 클라우드에 구현해서 제공하는 서비스형 AI(AIaaS)인 파스-타(PaaS-TA)를 개발하였다[4-7]. 또한, 국내 포털 업체인 네이버는 자체 클라우드 플랫폼 네이버 비즈니스 플랫폼 (NBP)에 사용자가 이용할 수 있도록 여러 AI 서비스를 구축하고, 자체 AI 플랫폼 '클로바'와 AI 기반 번역 '파파고'를 AIaaS 형태로 제공하고 있다[8]. 국외 시장은 대기업 위주로 퍼블릭 클라우드가 융합된 AI 서비스와 개발 환경을 제공하는 AI 플랫폼 등의 형태로 AIaaS(AI as a Service) 솔루션 출시하고 있으며, Microsoft는 자체 클라우드 '애

저 (Azure)'에 AIaaS를 제공하고 있고, 음성인식, 번역, 이미지 인식 등 많은 AI 서비스를 SaaS 형태로 제공하고 있다[9]. 아마존 회사의 경우 자체 클라우드 서비스 (AWS)에서 여러 AI 서비스를 제공하고 있으며, 직접 코딩할 필요 없이 그대로 가져와서 그 중 일부를 수정할 수 있는SaaS 형태의 서비스를 제공하고 있다. 인공지능 분야에서 가장 많은 영향을 주고 있는 Google은 Cloud AutoML 포트폴리오의 일부인 Cloud AutoML Vision 서비스를 시작하고 있으며, 이 서비스를 통해 사용자는 이미지 인식을 위해 머신러닝(ML) 모델을 더 빠르게 생성할 수 있다는 장점을 가지고 있다[10-13].

따라서, 본 논문에서는 현장 근로자 안전용 웨어러블 디바이스로부터 발생하는 대량의 데이터를 보유하고 있거나 AI 기술 개발을 전문으로 하는 기업이 아니더라도 AI 학습을 통해 고위험 산업 현장에서 안전서비스를 제공할 수 있는 AI 클라우드 플랫폼 설계 방안을 제시하였다.

2. 관련 연구

AI 모델 학습과 생성을 위한 클라우드 플랫폼을 구축하기 위해서는 인프라 구축과 모델 생성 지원 기술 연구가 필요하다. 최근 AI 관련 서비스 개발을 위해 다양한 AI 학습 지원 자동화 도구들이 연구되고 개발되었다. 특히, AutoML 소프트웨어는 머신러닝을 이용하여 AI 모델을 생성하는 과정에 발생하는 다양한 작업을 최대한 자동화하기 위한 소프트웨어이다. 이러한 소프트웨어를 본 플랫폼에 적용하기 위해 현재 제공되고 있는 AutoML 소프트웨어를 아래와 같이 분석하였다[14].

Ludwig은 텐서 플로우 기반으로 코드 작성 없이 딥러닝 모델의 학습과 테스트를 지원하며, 모델 학습을 위해 데이터셋, 입력 컬럼 리스트, 출력 컬럼 리스트가 요구되는 특징을 가지고 있다.

MLBox[15]는 대표적인 자동화 기계학습 파이선 라이브러리로서, 빠른 스캔, 분산 데이터 처리, 포맷, 클리닝 기능을 제공하고 있고 고차원 데이터에서 하이퍼파라미터 최적화를 지원하며, 분류 및 회귀를 위한 미래 예측 모델(Deep Learning, Stacking, LightGBM)과 모델 해석을 통한 예측 기능을 제공한다.

TPOT[16]는 genetic programming을 사용하여 기계 학습 파이프라인을 최적화하는 Python 기반 자동

기계 학습 도구로써 scikit-learn 프레임워크 기반으로 구축되었고, 데이터에 가장 적합한 파이프라인을 찾기 위해 수천 개의 가능한 파이프라인을 지능적으로 탐색하여 기계 학습을 자동화하지만 긴 학습 시간을 갖는 단점이 있다.

PyCaret[17]은 머신러닝 워크플로를 체계화하는 Python 기반의 로우코드 머신러닝 라이브러리로서 테스트 주기를 단축하고 생산성을 높이는 종단 간 기계 학습 및 모델 관리 도구이다. 다른 오픈 소스 머신 러닝 라이브러리에 비해 매우 긴 코드 라인을 몇 줄 안되는 라인으로 대체할 수 있는 장점을 가지고 있으며, sci-kit-learn, XGBoost, LightGBM, CatBoost, spaCy, Optuna, Hyperopt, Ray 등과 같은 AI 학습 라이브러리 및 프레임워크를 Python 라이브러리로 제공한다. LightAutoML[18]은 이진 분류, 다중 클래스 분류와 같은 자동 분류 및 회귀 모델 생성에 최적화된 프레임워크이다. FLAML[19]은 개발자가 자동으로 정확한 AI 모델을 검색할 수 있는 Python 라이브러리가 포함된 AutoML 프레임워크이다. 따라서, 사용자가 학습을 위한 하이퍼파라미터를 선택하지 않아도 되며, 비교적 빠르게 학습을 시킬 수 있는 장점을 가지고 있다. EvalML[20]은 도메인별 목적 함수를 사용하여 기계 학습 파이프라인을 구축, 최적화 및 평가가 가능하다.

Auto-Sklearn[21]은 머신러닝 개발자에게 알고리즘을 선택하거나 하이퍼파라미터 값을 조정할 수 있는 기능을 제공하며, 베이지안 최적화, 메타 학습, 앙상블 구성에 장점을 갖는다. 아울러 AutoML 시스템 자체를 자동으로 구성할 수 있는 기능이 지원된다. AutoKeras[22]는 Texas A&M 대학의 DATA LAB에서 개발하였으며, Keras를 기반으로 하는 AutoML 시스템으로 여러 작업을 간단한 인터페이스로 처리할 수 있는 기능이 지원된다.

H2O[23]는 확장 가능한 분산 머신러닝 환경을 지원하며, R, Python, Scala, Java, JSON을 지원하며, Hadoop 및 Spark와 같은 빅 데이터 기술과 연계가 가능하다. 또한 GLM(일반화 선형 모델), Gradient Boosting Machine(XGBoost 포함), Random Forests, Deep Neural Networks, Stacked Ensembles, Naive Bayes, Generalized Additive Models(GAM), Cox Proportional Hazards 같은 알고리즘의 구현을 제공한다.

MLOps는 ML의 전체 Lifecycle를 관리하기 위해 ML 시스템 개발과 운영을 통합하고, 데이터 레이블링의 일관성을 유지하면서 자동으로 운영되도록 인프라를 만들 수 있도록 도움을 주는 소프트웨어이다. 따라서, 제안 플랫폼에도 이러한 MLOps 소프트웨어를 적용하여 개발하고자 하며 이를 위해 아래와 같이 관련 소프트웨어들을 분석하였다.

Kubeflow[24]는 머신 러닝 워크플로우를 쉽게 조정하고 배포할 수 있는 오픈소스 기반 도구로써, Jupyter 노트북 연계 기능과 모델 학습, 파이프라인 생성과 같은 머신 러닝 자동화 프로세스 통합 서비스를 제공한다. 또한, TensorFlow나 Istio와 같은 다양한 프레임워크와 통합 및 연계가 가능한 장점을 가지고 있다. 본 논문에서 제안하는 플랫폼은 Kubeflow를 기반으로 설계하였다.

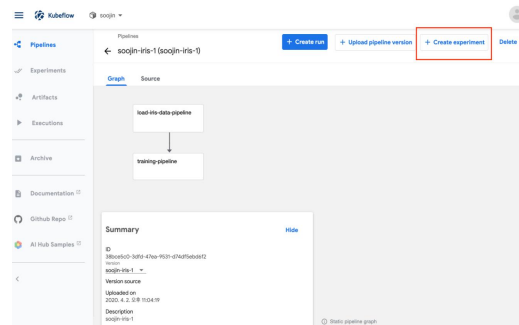


Fig. 1. Kubeflow Pipeline

MLFlow[25]는 학습과정을 추적하거나, 프로젝트 패키징 및 학습 모델 배포 기능 등을 제공하는 오픈소스 플랫폼이다. TensorFlow 및 Pytorch를 비롯한 다양한 기계 학습 라이브러리를 지원하며 개발 프로세스를 간소화한 특징을 가지고 있다. Data Version Control[26]은 AI 학습 데이터 연구 및 머신 러닝 프로젝트를 개발하기 위해 Python으로 개발된 오픈소스 이고, 데이터 셋, 기계 학습 모델, 버전 관리 기능을 제공한다. Pachyderm[27]은 DVC와 마찬가지로 머신 러닝 및 기계 학습 모델을 위해 수집된 데이터의 버전 관리 및 추적이 가능한 도구이다. Docker 및 Kubernetes를 기반으로 구축되어 있다. Metaflow[28]는 Netflix에서 처음 개발한 오픈소스 MLOps 플랫폼이며, 기업용 데이터 학습 프로젝트를 구축하고 관리할 수 있게 해준다. 머신 러닝, 딥 러닝 및 빅 데이터 관련 Python

라이브러리를 통합하여 ML 모델을 효율적으로 학습하고 배포 및 관리하기 위한 기능을 제공한다.

3. 시스템 구성

본 논문에서 제안하는 고위험 현장에서 안전관리를 위해 발생하는 대용량 데이터를 이용하여 다양한 서비스가 가능하도록 AI 모델을 학습하고 생성할 수 있는 클라우드 AIaaS 시스템의 구성은 아래 Fig. 2와 같다.

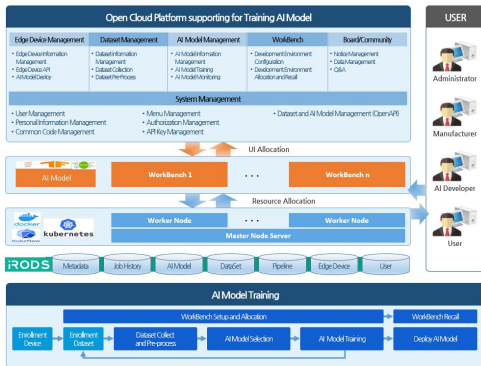


Fig. 2. System Architecture

고위험 현장 안전관리를 위한 AI 클라우드 시스템은 고위험 현장 안전 관리를 위한 지능형 에지 단말기 부분과 개방형 클라우드 플랫폼으로 구분된다.

3.1 AI클라우드 시스템 구분

3.1.1 지능형 에지 단말기

지능형 에지 단말기는 고위험 현장에서 발생할 수 있는 위험 신호를 수집하고, 수집된 신호를 기반으로 안전사고 상황을 인지하여 작업자에게 알림 서비스를 제공하는 기기이다. 따라서 해당 단말기에는 현장 내 작업자 및 위험 대상 위치정보 트랙킹 기능, 현장 내에서 수집되는 데이터 전송 기능, 작업자 위치정보 기반 사고 위험상황 인지, 위험상황 인지/알람/보고와 같은 기능이 필요하다.

3.1.2 개방형 클라우드 플랫폼

본 논문에서 제안하는 플랫폼은 고위험 현장에서 발생하는 데이터를 이용하여 AI 모델을 학습하고 생성하여 단말기에 적용할 수 있는 개발 환경을 클라우드로 서비스하는 솔루션이다. 이에 따라 에지 단말 관리 기

능, 데이터 셋 관리, AI 모델 학습 관리, WorkBench 구성, 데이터 저장 기능이 요구된다.

3.2 에지 단말 관리 기능

AI 학습을 위해 데이터 수집 및 사고 위험상황 알림 역할을 수행하는 단말기의 관리하기 위해 에지 단말 등록, 삭제, 갱신 기능이 필요하다. 또한, 에지 단말 특성에 따라 해당 플랫폼과 통신하기 위한 프로토콜 정의하고 연계할 수 있는 API를 제공해야 한다. 수집되는 데이터는 정의된 데이터 구조에 따라 수집해야 하므로 이를 정의하고 관리하기 위한 기능이 필요하다.

3.3 데이터 셋 관리 기능

지능형 에지 단말기로부터 수집된 데이터를 저장하고 관리하기 위한 메타데이터 및 데이터 정의, 추출, 저장, 검색 기능이 필요하다.

3.3.1 수집 데이터 구조 정의

〈DataType:String〉, 〈항목2:Decimal〉, 〈항목3:Decimal〉와 같이 수집되는 데이터의 종류와 타입을 쌍으로 구성하는 딕셔너리(Dictionary) 형태로 구조를 정의한다.

3.3.2 수집 데이터 추출

정의된 데이터 구조에 따라 수집된 데이터를 추출할 수 있는 기능이 필요하며 아래 Fig. 3은 예시 화면이다.

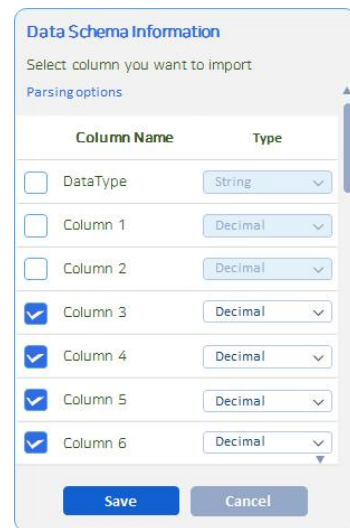


Fig. 3. Definition of Collected Data

3.3.3 데이터셋 전처리(정제 및 라벨링) 기능

아울러, 수집된 데이터를 AI 학습용 데이터셋으로 구성하기 위해 사전의 데이터셋 전처리(정제)와 라벨링이 가능한 아래와 같은 옵션 설정 기능과 데이터 라벨링 및 Annotation 파일 생성 및 업로드(JSON, CSV 파일) 기능이 요구된다.

- 누락 데이터 처리(Data Imputation) 옵션
 - 데이터 제거
 - 평균 및 중앙값으로 대체
 - 빈번한 값, 0, 특정값으로 대체
 - K-NN(k nearest neighbours) 알고리즘을 이용한 대체
 - DNN을 이용한 누락값 대체
 - 선형회귀 방법을 이용한 Noise Data 처리 방법 최적화
- feature scaling 옵션
 - Min-Max Normalization(최대/최소값 1,0)
 - Standard(평균, 표준편차 사용)
 - MaxAbs(최대, 절대값과 0이 각각 1, 0)
 - Robust(중앙값과 IQR 사용)
- feature selection 옵션
 - Supervised(Intrinsic, Wrapper Method, Filter Method 등)
 - Unsupervised

3.4 AI 모델 학습 관리 기능

본 논문에서 제안하는 AI 클라우드 플랫폼을 이용하여 AI 모델을 생성하는 절차는 아래의 Fig. 4와 같다.

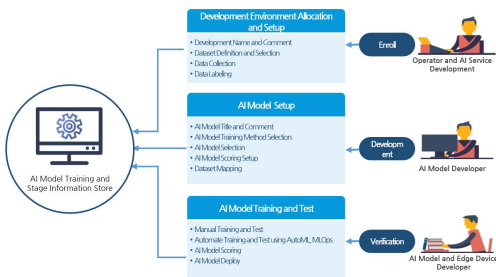


Fig. 4. Training Process of AI Model

AI 서비스를 개발하고자 하는 사용자는 이전에 학습된 AI 모델을 활용할 수 있거나, 새롭게 AI 모델을 학

습할 수 있어야 한다. 이에 따라, AI 모델을 검색하기 위해 개발자의 가치치 부여에 따라 AI 모델을 추천하는 기능을 제공한다. 추천하기 위해 사용되는 모델들은 아래와 같다[13].

- 베이스라인 모델 제공(시계열 데이터 AI 모델 (LSTM, CNN, RNN, GAN, RL 등))
- 가치치(정확도 우선, 성능 우선, 수집 데이터셋 종류 등)에 의한 유사도 평가
- 머신 러닝 알고리즘

Table 1. Comparative of Machine Learning Algorithm

Algorithm	Accuracy	Training time
Two-Class logistic regression	Good	Fast
Two-class decision forest	Excellent	Moderate
Two-class boosted decision tree	Excellent	Moderate
Two-class neural network	Good	Moderate
Two-class averaged perceptron	Good	Moderate
Two-class support vector machine	Good	Fast
Multiclass logistic regression	Good	Fast
Multiclass decision forest	Excellent	Moderate
Multiclass boosted decision tree	Excellent	Moderate
Multiclass neural network	Good	Moderate
Linear regression	Good	Fast
Decision forest regression	Excellent	Moderate
Boosted decision tree regression	Excellent	Moderate
Neural network regression	Good	Moderate
K-means clustering	Excellent	Moderate

선택된 AI 모델을 이용하여 학습을 진행하기 위해, 데이터셋을 구성하는 방법과 평가 결과 정의, 학습을 위한 파이프라인, 코드 등을 같이 확인할 수 있는 기능이 필요하다. 이에 따라, 사용자는 직접 또는 자동화 방법 제공 받아 개발할지를 결정하게 되며, 직접 개발할 경우, 개발 환경을 제공하고, 자동화 방법의 경우 자동화 프로세스를 통한 개발이 가능하도록 해당 환경을 구성하여 제공한다.

- 직접 개발 기능(JupyterLab 지원)
- 아래 Fig. 6과 같은 직접 개발 프로세스에 따라 요구되는 옵션을 설정하는 기능을 제공

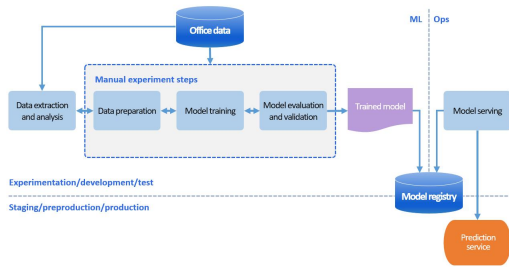


Fig. 6. Manual Development Process

- data science, data visualization, machine learning 개발 환경에 특화된 웹 개발 환경 인터페이스를 제공하는 JupyterLab을 제공

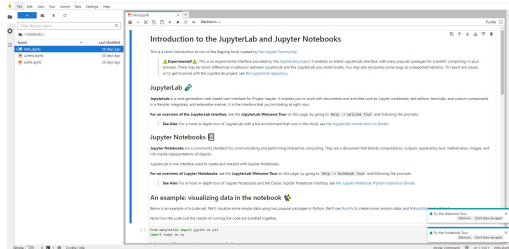


Fig. 7. Web-based JupyterLab Interface

- 자동화 기능
 - AutoML, MLOps 오픈소스 분석 및 자동화 기능으로 아래의 Fig. 8과 같이 자동화 프로세스는 ML 파이프라인을 자동화하여 모델을 지속적으로 학습할 수 있기 때문에, 자동화를 위해 고려해야 할 사항들이 직접 개발 방식에 비해 많음

3.5 자동화 고려 사항

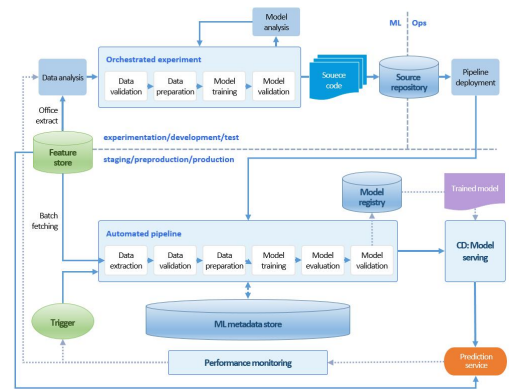


Fig. 8. Automated Development Process

3.5.1 파이프라인(Pipeline)

모델 학습을 위해 어떻게 파이프라인을 저장하고 운영환경에서 사용할 것인지에 대한 기능이 필요하며, 매일, 매주 또는 매월 등의 재학습 빈도는 데이터 패턴의 변경 빈도와 모델 재학습 비용에 따라 선택할 수 있어야 한다. 또한, 모델 성능 저하가 눈에 띄는 경우 모델 재학습 트리거 실행하도록 지원해야 한다.

3.5.2 CT(Continuous Training)

매일 발생하는 새로운 데이터를 사용하여 생성된 모델이 자동으로 학습할 수 있도록 지원해야 한다.

3.5.3 CD(Continuous Delivery)

새로운 데이터로 학습되고 검증된 모델이 지속적으로 배포되어야 하며, 배포 전 모델과 대상 인프라의 호환성(패키지 호환 여부/메모리/컴퓨팅 자원등)을 확인할 수 있는 서비스 API 호출 테스트, QPS 및 지연 시간과 같은 서비스 부하 테스트 기능이 필요하다.

3.5.4 Data validation

수집된 데이터의 검증이 실패하면, 신규 모델의 배포를 중지하기 위해 예상치 못한 데이터가 생성된 경우, 예외 범주를 벗어난 특성이 생성된 경우와 같은 Data schema skews 문제, 데이터의 통계적 속성이 변화되고 있음을 감지하거나 감지된 변화에 따라 모델의 재학습을 수행하는 Data values skews 문제를 고려하여야 한다.

3.5.5 Model validation

모델이 새로운 데이터로 재학습을 마치고, 운영 환경에 반영되기 전에 평가되고 검증하기 위해 테스트 데이터셋으로 평가 메트릭을 생성, 평가 메트릭을 새로운 모델과, 현재 모델 사이에 비교, 새로운 모델이 기존 모델보다 더 나은 성과를 보이는지 검증, 새로운 모델의 성능이 다양한 세그먼트에서 일관된 성과를 보이는지 검증, 인프라 및 서비스 API와 호환성 테스트 기능이 필요하다.

3.5.6 Feature store

학습과 서비스에 사용되는 모든 feature들을 모아둔 저장소 구축이 필요하며, 사용가능한 모든 feature의 저장하고, 항상 최신화된 데이터를 유지할 수 있는 기능이 필요하다.

3.5.7 Metadata management

구축된 파이프라인과 메타데이터 관리를 위해 ML 파이프라인의 실행 정보, 데이터 및 아티팩트의 히스토리 정보 등을 저장하기 위한 기능이 필요하다.

- 실행된 파이프라인 버전, 시작-종료 시간, 소요 시간 등
- 파이프라인의 실행자, 매개변수 인수
- 이전 모델에 대한 포인터(모델의 롤백이 필요한 경우)
- 모델 평가 단계에서 생성된 모델 평가 측정 항목

3.6 AI 모델 모니터링 및 자동 학습 기능

직접개발 또는 자동화 개발에 따라 선택된 AI 모델, 데이터셋, 파이프라인을 이용하여 AI 모델 학습을 실행할 수 있는 기능이다. 또한 배포가 완료된 AI 모델에 대한 지속적인 모니터링을 제공하고, ML 워크플로의 아티팩트, 계보, 실행 추적을 위한 Pipelines의 모든 구성 요소에 대한 입력 및 출력을 자동으로 추적한다. 직접 개발 방식과 자동화 개발 방식은 아래 Table 2와 같은 장단점을 가지고 있다.

Table 2. Difference of Manual Development Process and Automated Development Process

Feature	Manual Development Process	Automated Development Process
Data import and analysis, model training and validation	<ul style="list-style-type: none"> - AI developer directly configures environment and participates in development - Run scripts or personally test on your laptop 	<ul style="list-style-type: none"> - Automatic configuration of development environment and direct development - Automatically performed according to the saved pipeline
AI model deployment	<ul style="list-style-type: none"> - Direct distribution - Distribution of new model versions occurs irregularly 	<ul style="list-style-type: none"> - Direct distribution - Ability to release new model versions on a regular basis - Deploy new model version in case of performance degradation
Active performance monitoring	<ul style="list-style-type: none"> - We do not monitor logs or predictive performance of models - It is not possible to detect that the performance of the model is degraded or that the model is operating abnormally - When a model is deployed to a production environment, it can often fail due to changes in the actual environment and data 	<ul style="list-style-type: none"> - AI model performance improvement through monitoring - Detect performance degradation and model staleness - Retrain the model with the latest data - Continue to try new implementations such as extraction of new features, model architectures, hyperparameters, etc.

3.7 WorkBench 구성 설계

본 논문에서는 아래 Fig. 9와 같이 오픈소스인 Kubernetes를 이용하여 AI 모델 개발 환경의 할당/회수 서비스 제공하도록 설계하였다. 이에 따라, AI 모델 생성, 공유, 활용 서비스와 연계할 수 있는 기능과 Kubernetes를 이용한 자원 할당 및 회수, 자원 모니터링 기능을 설계하였다. 또한, Kubernetes의 컨트롤 플레인 컴포넌트, 노드 컴포넌트, 애드온, 컨테이너 리소스 모니터링, 클러스터-레벨 로깅 컴포넌트와 AI 모델 생성 서비스와 연계가 가능하도록 기능을 설계하였다.

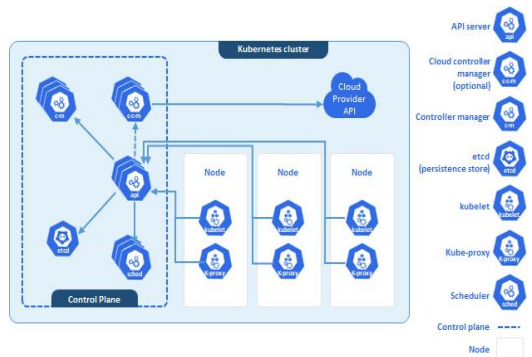


Fig. 9. Kubernetes Component

3.8 데이터 저장 기능 설계

수집된 데이터는 아래 Fig. 10과 같이 iRods 기반 데이터 분산 저장 기능을 통해 이기종 단말의 학습모델, 서비스, 메타데이터, 코드, 파이프라인을 저장하도록 하였다[29-31]. 이를 위해, 데이터 및 메타데이터, 코드, AI 모델, 파이프라인 저장 스키마 설계하였고, 메타데이터 분류 및 검색 기능을 구현하였다. 본 논문에서 구성한 메타데이터들은 수집된 데이터의 종류, 파이프라인 종류, AI 모델 버전 및 변수항목, 업로드한 코드 버전, 변수항목 등이 있다.

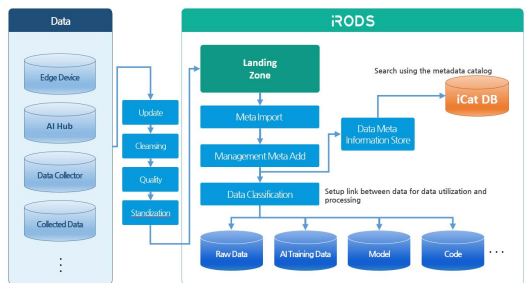


Fig. 10. Structure of Data Store based on iRODS

본 논문에서 제안하는 플랫폼은 웹 환경에서 개발되므로 이와 연계하기 위해서는 iRods에서 제공하는 iCat과 Jargon-core API를 활용해야 하므로 해당 API를 이용한 데이터, 메타데이터 저장 기능을 설계하였다.

4. 결론

최근 급속하게 발전하고 있는 인공지능 기술은 경제, 사회, 군사 등 다양한 분야에 적용되어 인간의 삶과 산업 패러다임을 변화시키고 있다. 특히, 고위험 현장의 사고 예방 필요한 산업에서도 AI 기술을 도입하여 안정성을 높이고자 하는 움직임이 존재한다. 따라서, 본 논문에서는 현장 근로자 안전용 웨어러블 디바이스로부터 발생하는 대량의 데이터를 보유하고 있거나 AI 기술 개발을 전문으로 하는 기업이 아니라도 AI 학습을 통해 고위험 산업 현장에서 안전서비스를 제공할 수 있는 AI 클라우드 플랫폼 설계 방안을 제시하였다. 본 논문에서 제안한 플랫폼은 직접 개발 방식과 자동 개발 방식을 모두 지원하며, 수집된 데이터 정의부터 학습된 모델 배포까지 가능한 장점을 갖는다. 향후에는 본 논문에서 제시하는 설계 기능을 기반으로 실제 구현 및 개발할 예정이며, AI 모델 학습 및 배포, 주기적 모델 개선을 자동화하여 효율적으로 AI 서비스가 가능하도록 할 예정이다.

REFERENCES

- [1] J. Seo. (2018). Smart Wearable Technologies in Construction, *Construction engineering and management*, 19(4), 58-63.
- [2] R. Lee, M. Park & S. H. Lee. (2019). An Advanced IoT Data Collection Service for Data-centric Smart Cities, *International Journal of Innovative Technology and Exploring Engineering*, 8(8S2), 323-327.
- [3] R. Lee, R. Y. Jang, M. W. Park, S.-H. Lee & M. S. Choi. (2020). AIaaS Technology Trends and Activation Plan, *Communications of the Korean Institute of Information Scientists and Engineers*, 38(8), 49-57
- [4] PaaS-TA. (n. d.). <https://paas-ta.kr/>
- [5] Y. H. Kim. (2021), *Open Cloud Platform PaaS-TA Trend and Ecosystem Activation*, *Software Policy & Research Institute*, Monthly Software Oriented Society.
- [6] R. Y. Jang, R. Lee, M. W. Park & S. H. Lee. (2020) Development of an AI Analysis Service System based on OpenFaaS, *The Journal of the Korea Contents Association*, 20(7), 97-106
DOI : 10.5392/JKCA.2020.20.07.097
- [7] V. Ishakian, V. Muthusamy & A. Slominski. (2018). Serving Deep Learning Models in a Serverless Platform, *2018 IEEE International Conference on Cloud Engineering (IC2E)*. IEEE,
- [8] Naver Cloud Platform. (n. d.) <https://www.ncloud.com>
- [9] Azure AIaaS. (n. d.). <https://azure.microsoft.com/ko-kr/resources/cloud-computing-dictionary/what-is-azure/azure-iaas/#overview>
- [10] Google Cloud Platform. (n. d.). <https://cloud.google.com/>
- [11] E. Bisong. (2019). *An Overview of Google Cloud Platform Services. Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Apress, Berkeley, CA, pp.7-10.
- [12] Google Cloud AutoML. (n. d.). <https://cloud.google.com/automl?hl=ko>
- [13] Azure Machine Learning. (n. d.). <https://learn.microsoft.com/ko-kr/azure/machine-learning/how-to-select-algorithms>
- [14] Y.-H. Moon, I. H. Shin, Y. J. Lee & O. G. Min. (2019). Recent Research & Development Trends in Automated Machine Learning, *Electronics and telecommunications trends*, 34(4), 32-42
DOI : 10.22648/ETRI.2019.J.340404
- [15] MLBox. (n. d.). <https://mlbox.readthedocs.io/en/latest/>
- [16] TPOT. (n. d.). <http://automl.info/tpot/>
- [17] PyCaret. (n. d.). <https://pycaret.org/>
- [18] LightAutoML. (n. d.). <https://lightautoml.readthedocs.io/>
- [19] FLAML. (n. d.). <https://microsoft.github.io/FLAML/>
- [20] EvalML. (n. d.). <https://evalml.alteryx.com/en/stable/>
- [21] Auto-Sklearn. (n. d.). <https://automl.github.io/auto-sklearn/master/>
- [22] AutoKeras. (n. d.). <https://autokeras.com/>
- [23] H2O. (n. d.). <https://h2o.ai/>
- [24] Kubeflow. (n. d.). <https://www.kubeflow.org/>
- [25] MLFlow. (n. d.). <https://mlflow.org/>

- [26] Data Version Control(DVC). (n. d.). <https://dvc.org/>
- [27] Pachyderm (n. d.). <https://www.pachyderm.com/>
- [28] Metaflow. (n. d.). <https://metaflow.org/>
- [29] iRODS. (n. d.). <https://irods.org/>
- [30] J. Kim, Y. Han, D. Kim & Y.-K. Lee. (2015). Climate Data Support System using iRODS, *Journal of KIISE*, 31(2), 3-15.
- [31] Y. K. Han, J. Kim, S. Lee & Y. K. Lee. (2010). Massive Electronic Record Management System using iRODS, *Journal of KIISE*, 16(8), 825-836.

김 기 봉(Ki-Bong Kim)

[중신회원]



- 1993년 2월 : 충남대학교 대학원 전산학과 (이학석사)
- 1998년 8월 : 충남대학교 대학원 전산학과 (이학박사)
- 1995년 9월 ~ 1997년 2월 : 혜전대학교 전임강사

- 1997년 3월 ~ 현재 : 대전보건대학교 컴퓨터정보과 교수
- 관심분야 : 데이터베이스, 의료정보시스템
- E-mail : kbkim@hit.ac.kr