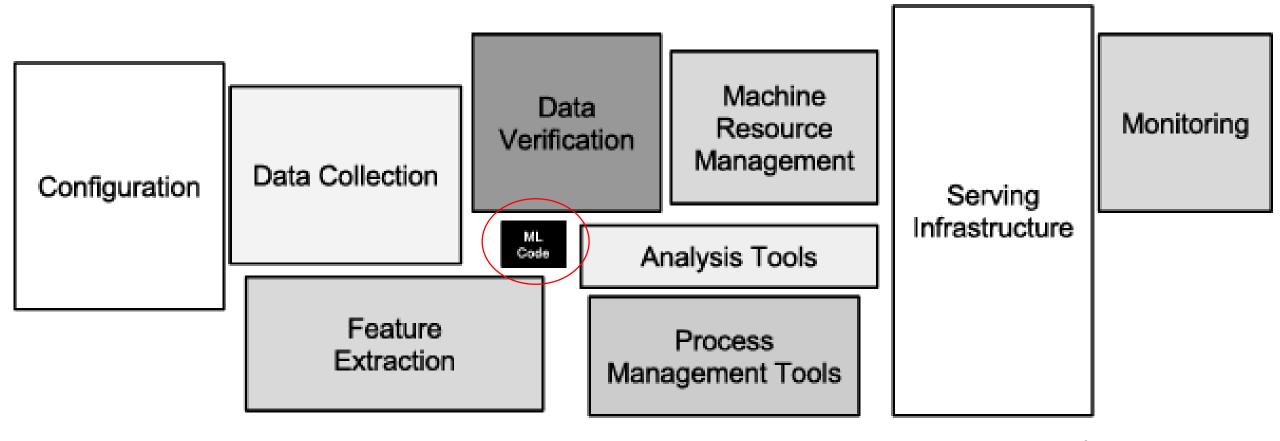
# 로봇을 위한 MLOps 1부

신재익 2025.08.30

#### Reference

- 요즘 우아한 AI 개발
- 15장
- 우아한 기술 블로그
- https://techblog.woowahan.com/17647/
- 우아콘2023 배달 로봇의 주변 환경 인지
- https://www.youtube.com/watch?v=7QFMru0HZto

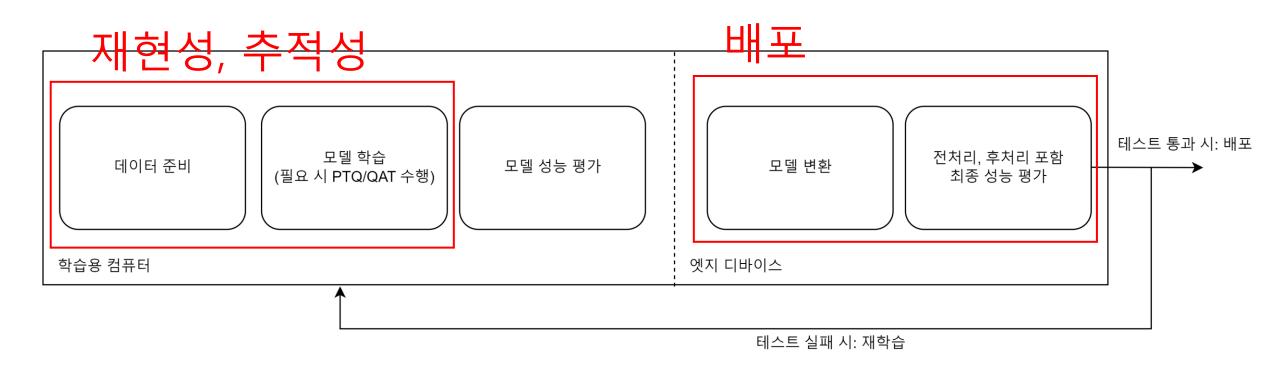
### 기술 부채(technical debt)



Sculley, David, et al. "Hidden technical debt in machine learning systems." Advances in neural information processing systems 28 (2015).

https://zzsza.github.io/data/2018/01/28/hidden-technical-debt-in-maching-learning-systems/

## MLOps: 학습 파이프라인, 서빙 파이프라인



모델 변환 : 양자화 (PTQ, QAT)

AI 서비스와 MLOps 도입기 https://techblog.woowahan.com/11582/

데이터 보관, 모델 훈련 및 배포





모델 배포

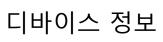


데이터 수집



모델 훈련용 계산 자원

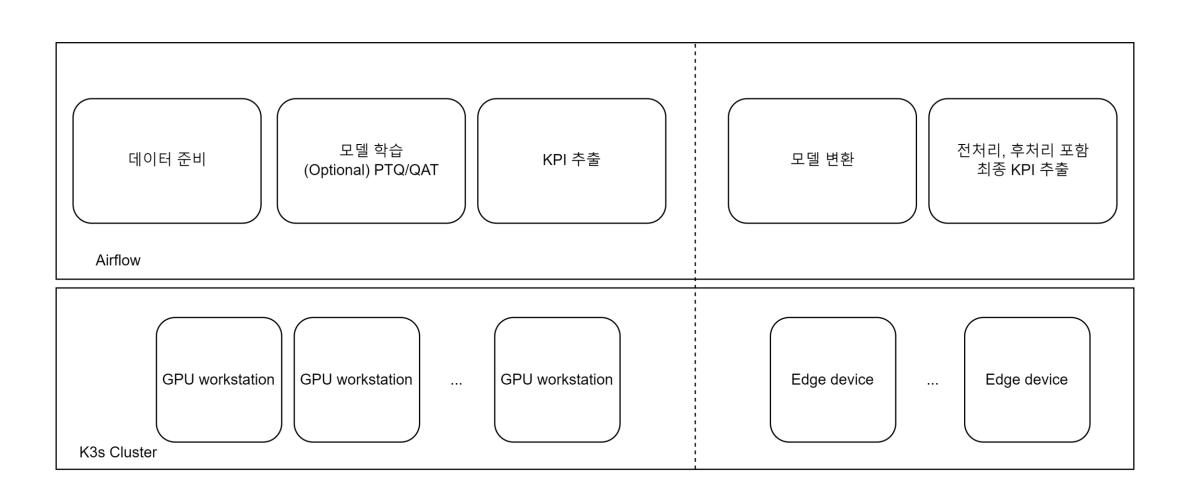




엣지 디바이스 : 로봇



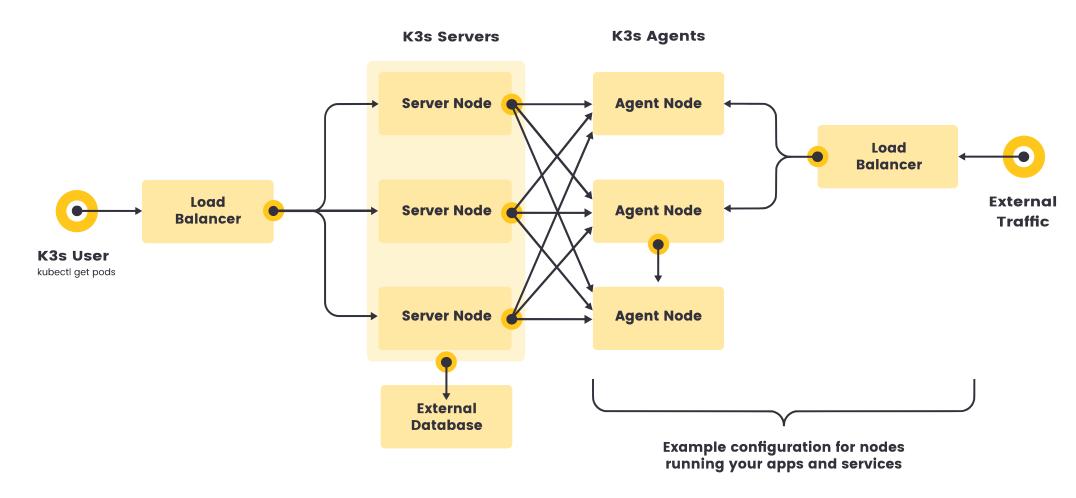
### K3s 과 Airflow 로 구성한 MLOps 시스템



### K3s: lightweight Kubernetes (자원관리)

- K3s
- https://k3s.io/
- https://docs.k3s.io/
- https://github.com/k3s-io/k3s/

#### Example of K3s cluster



### Apache Airflow (파이프라인)

- Airflow
- https://airflow.apache.org/
- https://airflow.apache.org/docs/apacheairflow/stable/index.html
- https://github.com/apache/airflow/
- Kubernetes 의 패키지 매니저 helm 을 이용해서 설치

### DAG (directed acyclic graph)

train\_affinity, edge\_affinity를 이용해 실행하는 장소를 명시 (노드 지정) Resources를 이용해 GPU 자원을 요청 train\_task 로 파이토치를 이용한 모델 훈련 edge\_task 로 nvidia jetson (엣지디바이스) 를 위한 모델 변환

