

# 쿠버네티스

≡ 태그

## 쿠버네티스 사용 이유

| 인프라의 모든 기능들이 코드로 관리될 수 있다.

→ 도커 오케스트레이션

## 쿠버네티스

| 머신러닝 파이프라인 구현 오픈소스

▼ 머신러닝 프로젝트에서 머신러닝 코드는 전체의 5%에 불과함

- data collection
- configuration
- serving Infrastructure
- feature extraction
- data verification
- analysis
- process management
- monitoring

→ 추상화 + 재사용성 : 쿠베플로우

## 쿠베플로우 특징

## 1. 조합 가능성

- a. 머신러닝 실무자들에 익숙한 데이터 과학 도구
- b. 기계 학습을 용이하게 하기 위해 파이프라인 및 독립적 사용 가능

## 2. 이식성

- a. 컨테이너 기반 설계+클라우드 네이티브 아키텍처 → 특정 개발환경 종속 X: 랩톱, 온프레미스
- b. 실험 및 프로토타입 작업 수행, 프로덕션 배포 용이

## 3. 유동성

- a. 클러스터 요구에 맞게 기본 컨테이너의 수와 크기를 동적으로 변경

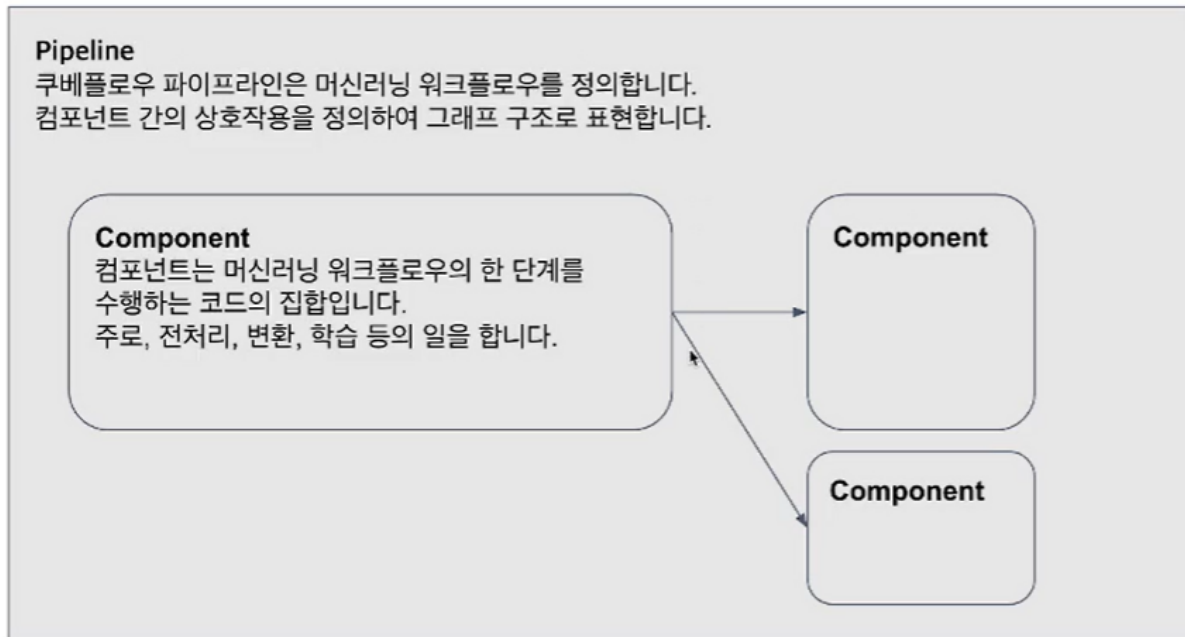
머신러닝 프로젝트의 구성요소를 프레임워크 독립적으로 사용할 수 있다.

- 어느정도 실험이 종료되면 머신러닝 파이프라인을 표준화
- YAML 파일을 이용해 전체 클러스터를 정의 가능

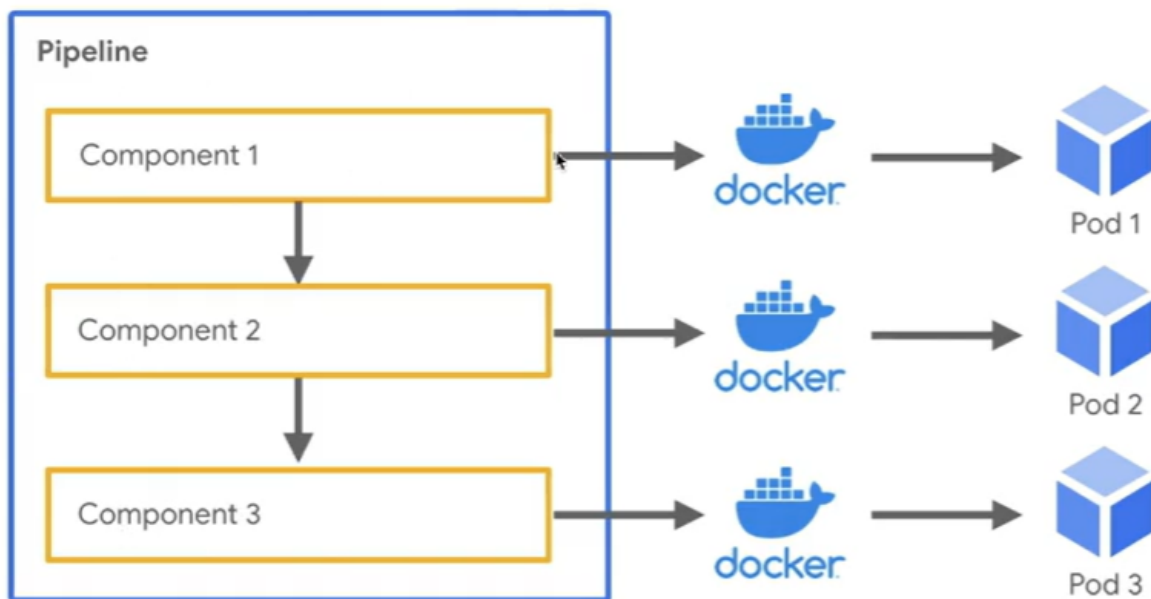
10:00 한 번 더 보기

# 쿠버플로우 파이프라인

- 머신러닝 파이프라인 오케스트레이션 단순화
- 실험, 재현, 공유
- 컴포넌트를 빠르게 재사용하고 연결



파이프라인은 컴포넌트(코드의 집합이므로 인풋과 함수처럼 아웃풋이 있음) 간의 상호작용을 그래프 구조로 표현하여 머신러닝 워크플로우를 정의한다.



## 재사용 가능한 컴포넌트의 원리

1. 프로그램은 출력 데이터의 경로를 컴포넌트의 파라미터로 받아야 함
- 2.