

ML 을 Service 화하기 위한 기술, MLOps

1 기존 ML 프로젝트 진행 방식의 한계

기존 ML 프로젝트 진행 방식 돌아켜보기

1

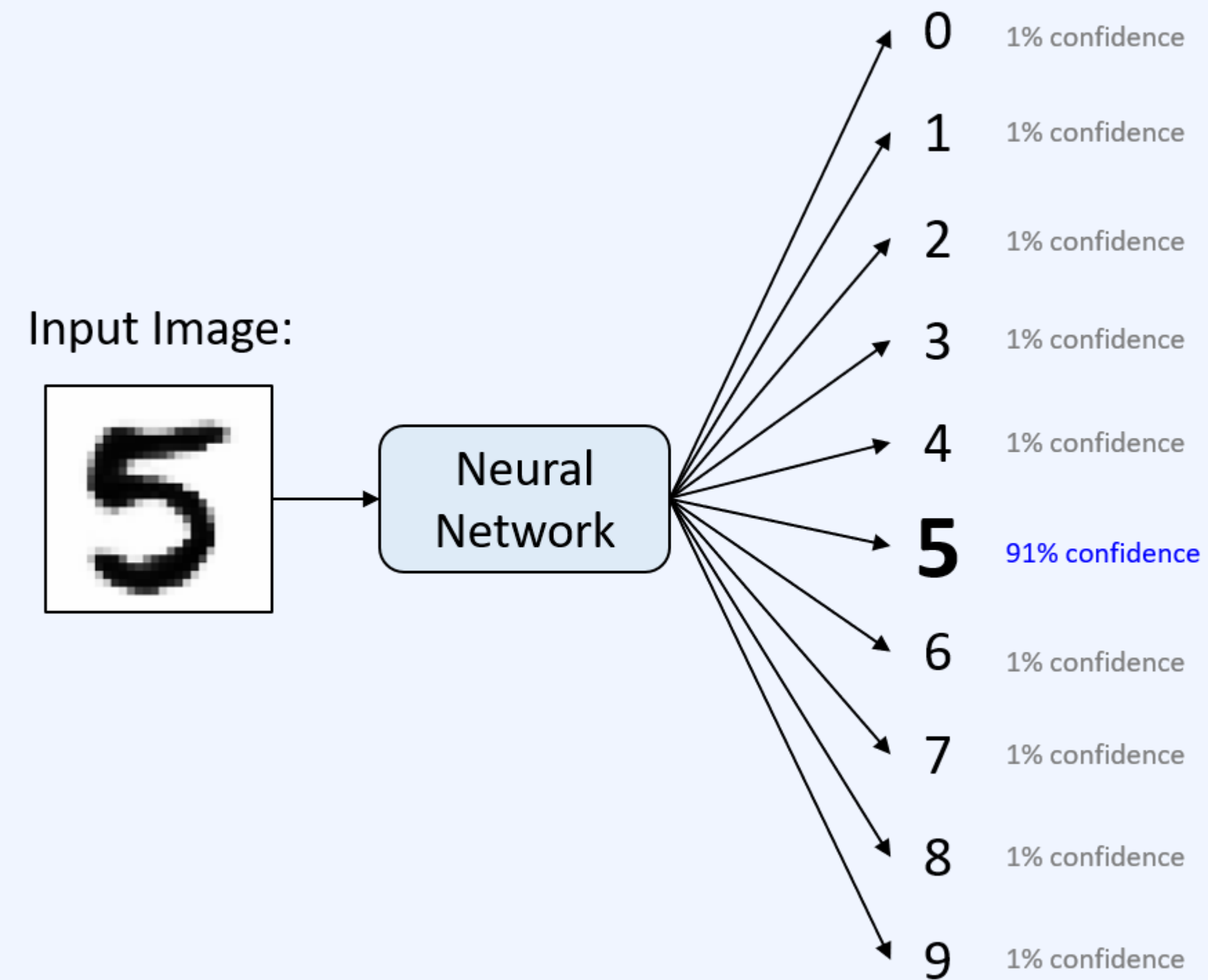
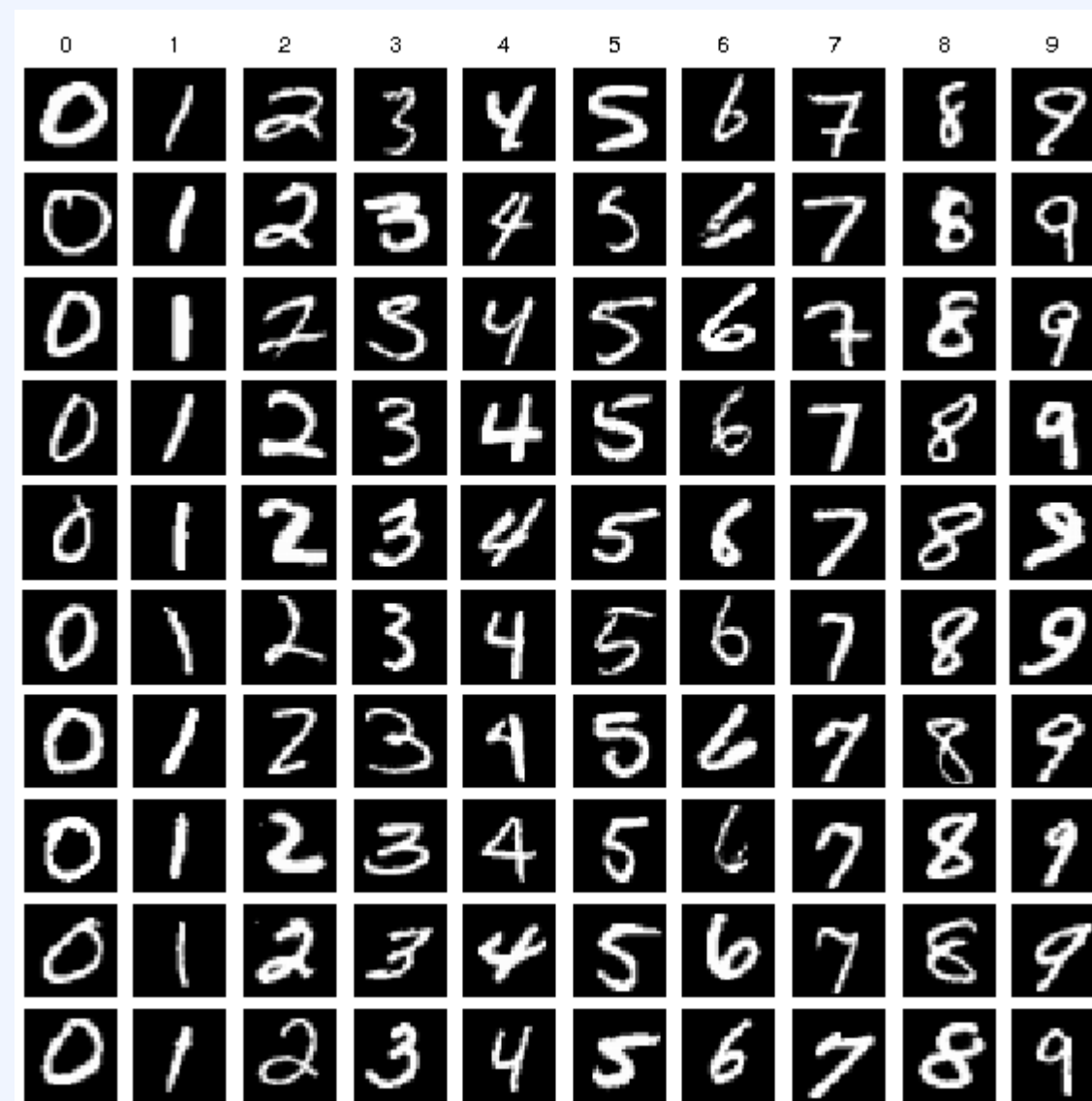
기존 ML 프로젝트
진행 방식의 한계



MNIST 데이터 분류 모델 중
성능이 가장 좋았던 모델의 정보가 기억나시나요?

1

기존 ML 프로젝트
진행 방식의 한계



학습 환경과 배포 환경은 같지 않습니다

1

기존 ML 프로젝트
진행 방식의 한계



**Validation
Dataset**

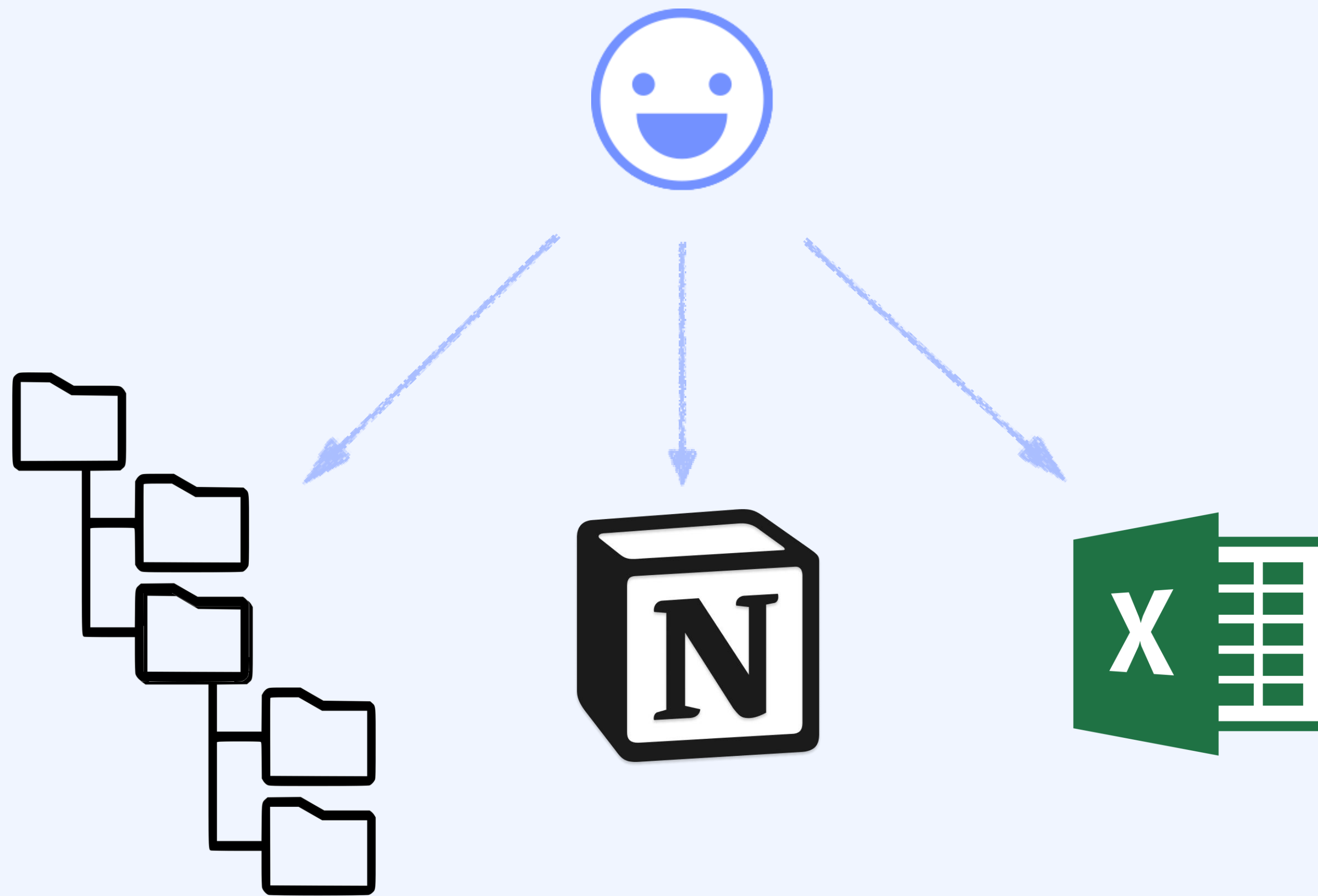


**Test
Dataset**

혼자서는 할 수 있을 지도 모릅니다

1

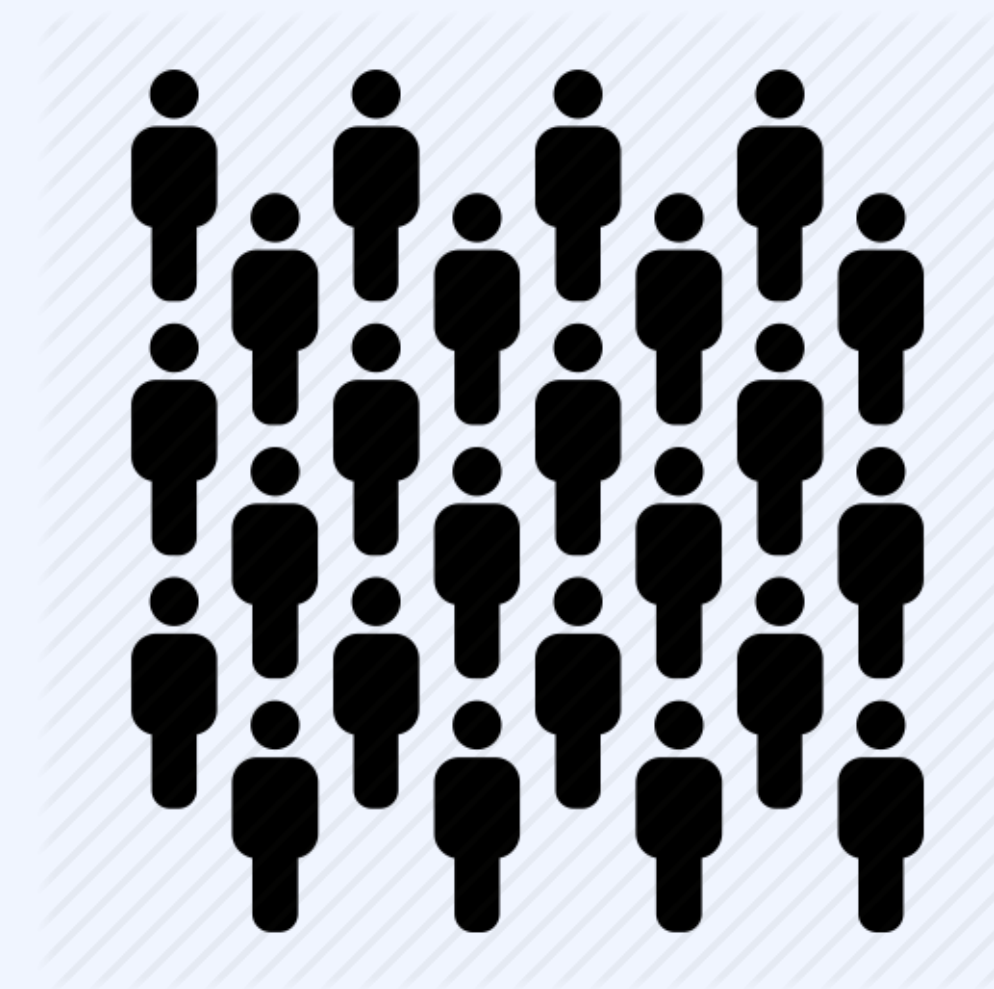
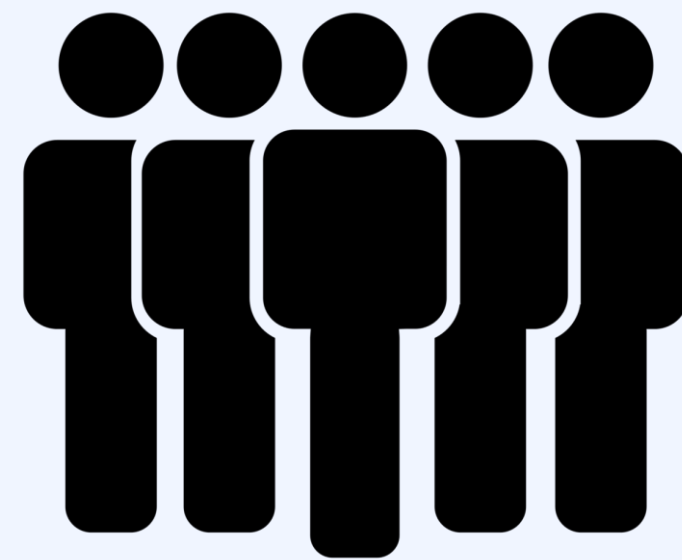
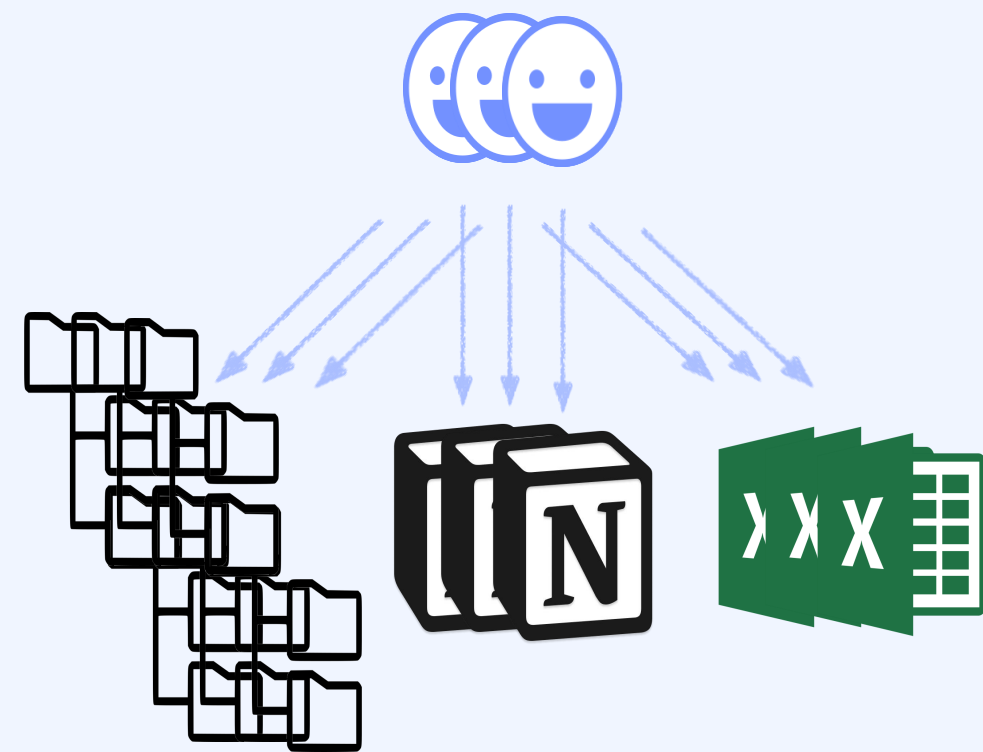
기존 ML 프로젝트
진행 방식의 한계



하지만 여러 명이 함께 협업해야 한다면?

1

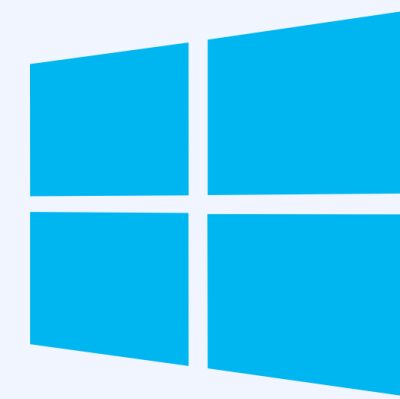
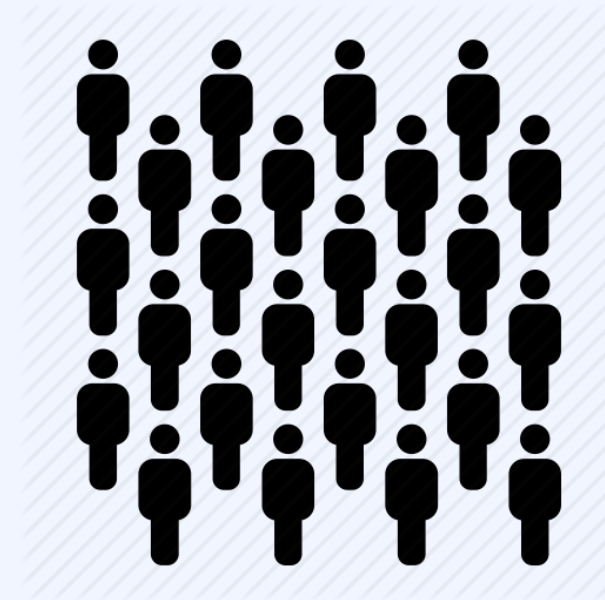
기존 ML 프로젝트
진행 방식의 한계



하지만 여러 명이 함께 협업해야 한다면?

1

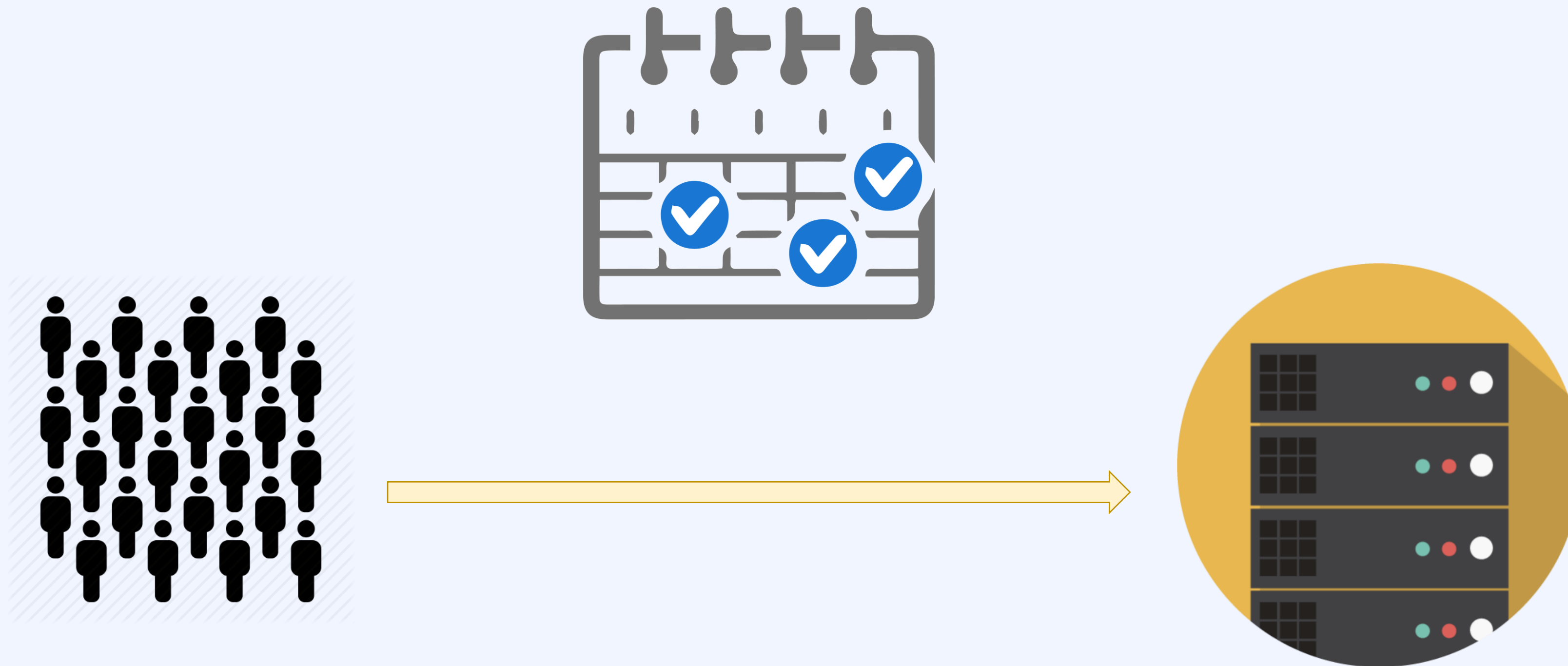
기존 ML 프로젝트
진행 방식의 한계



하나의 컴퓨터를 공유해야 한다면?

1

기존 ML 프로젝트
진행 방식의 한계



소설 속 이야기가 아닙니다

1

기존 ML 프로젝트
진행 방식의 한계



Why 85% of AI projects fail



SEARCH 🔍

SERVICES ▾ CAPABILITIES ▾ INDUSTRIES ▾

[HOME](#) > [ARTICLES](#) > PACTERA WHITE PAPER REVEALS 85 PERCENT OF AI PROJECTS ULTIMATELY FAIL

Pactera White Paper Reveals 85 percent of AI Projects Ultimately Fail

McKinsey
& Company

Gartner®

출처)

- <https://www.pacteraedge.com/pactera-white-paper-reveals-85-percent-ai-projects-ultimately-fail-0>
- <https://www.techrepublic.com/article/why-85-of-ai-projects-fail/>

현실 속 ML 서비스화

1

기존 ML 프로젝트
진행 방식의 한계

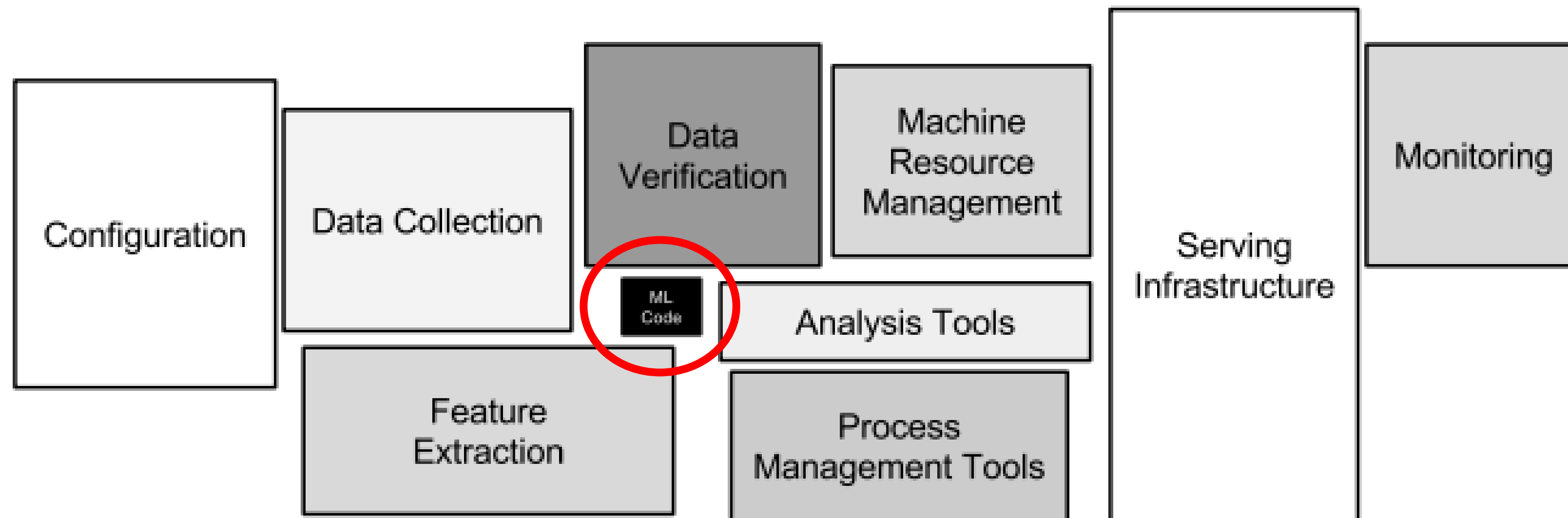


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

이상 속 ML 서비스화

1

기존 ML 프로젝트
진행 방식의 한계

데이터 준비

모델 구현

모델 배포



ML 을 Service 화하기 위한 기술, MLOps

2 ML + DevOps = MLOps

Software 개발 방식

1

기존 ML 프로젝트
진행 방식의 한계

전통적인 Software 개발 방식

코드 구현

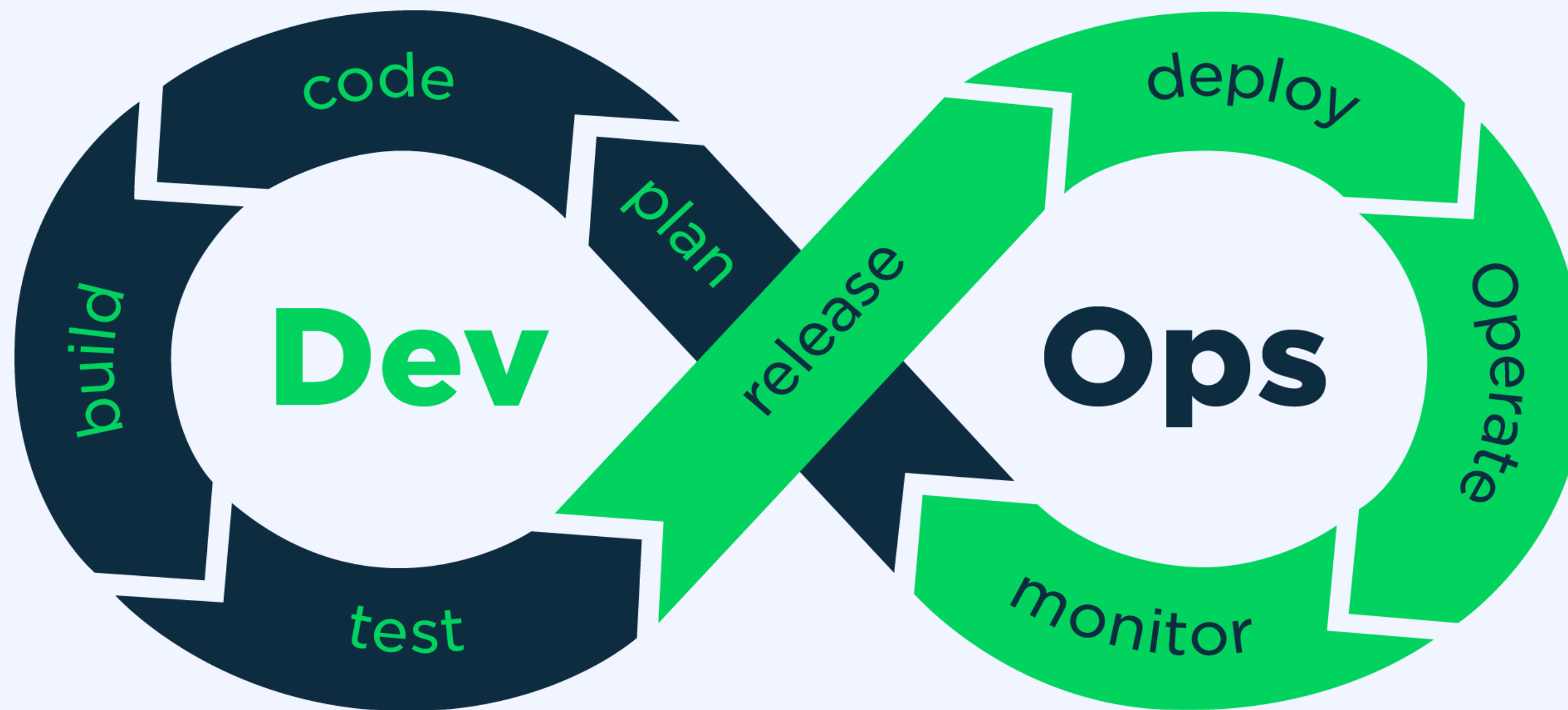
빌드

배포

DevOps 의 탄생

1

기존 ML 프로젝트
진행 방식의 한계



ML 프로젝트와 SW 프로젝트의 유사성

1

기존 ML 프로젝트
진행 방식의 한계

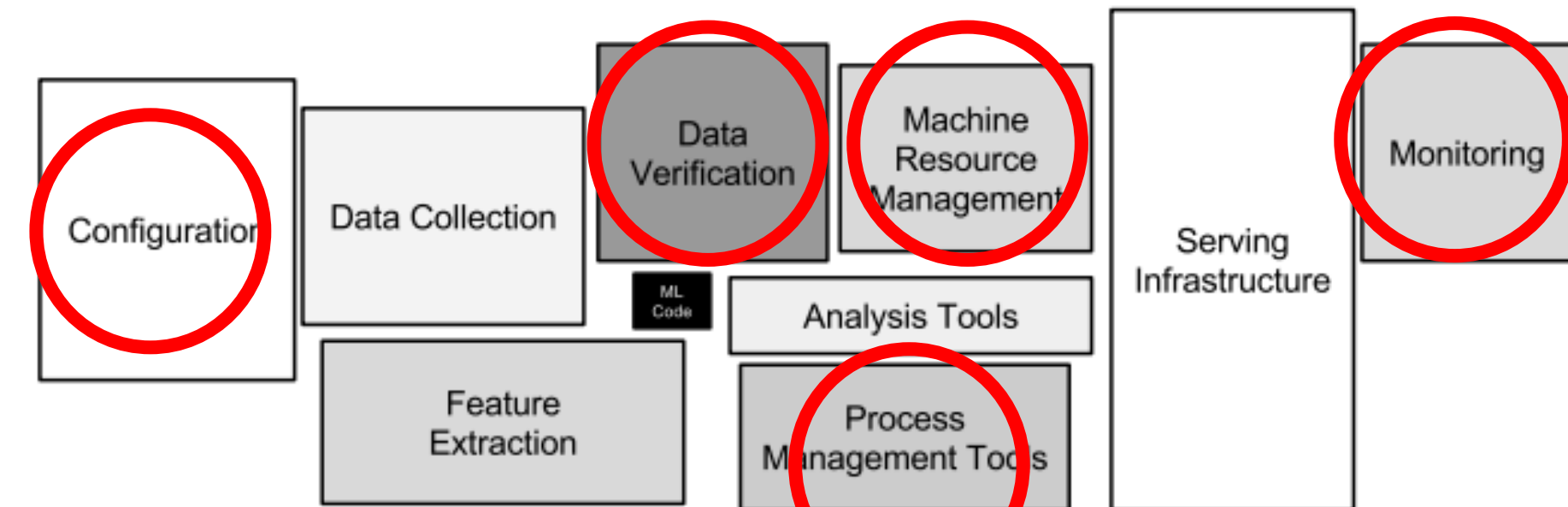
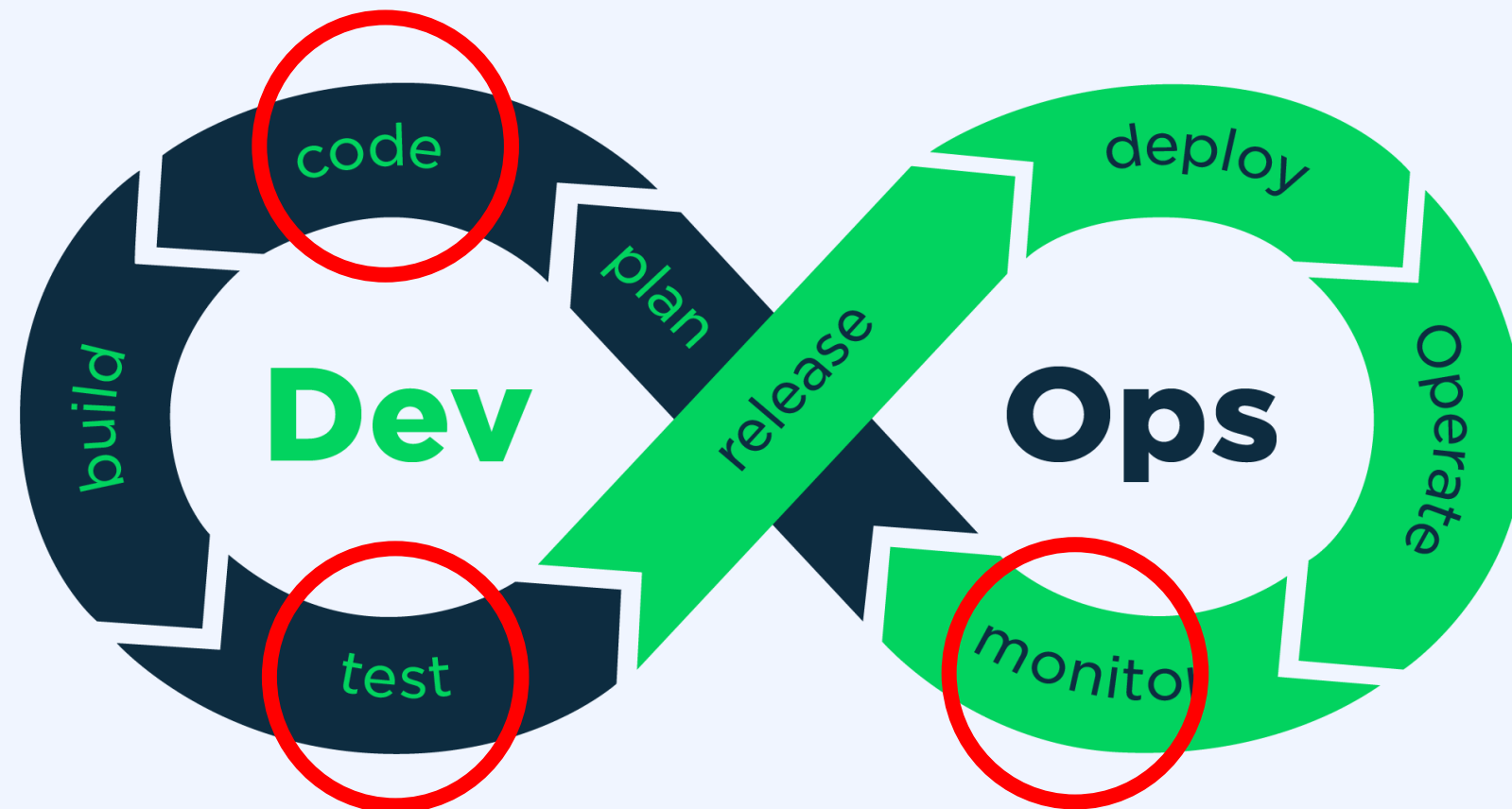


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

버전 관리

데이터 버전 관리
모델 버전 관리

테스트 자동화

모델 학습 자동화
모델 성능 평가 자동화

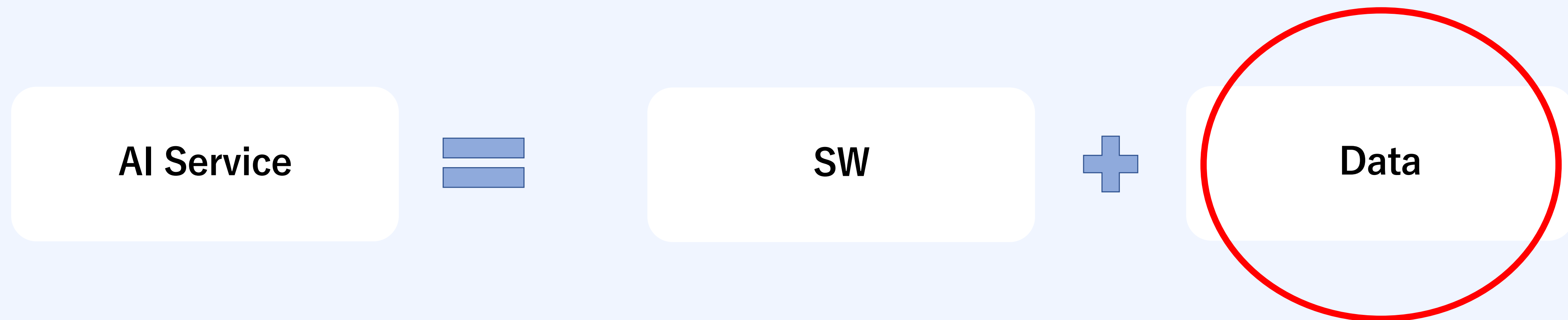
모니터링

서빙 모델 모니터링
데이터 변화 모니터링
시스템 안정성 모니터링

AI 서비스와 일반 SW 의 차이점

1

기존 ML 프로젝트
진행 방식의 한계



ML + DevOps = MLOps

1

기존 ML 프로젝트
진행 방식의 한계

MLOps



ML



DevOps

최신 트렌드, MLOps

1

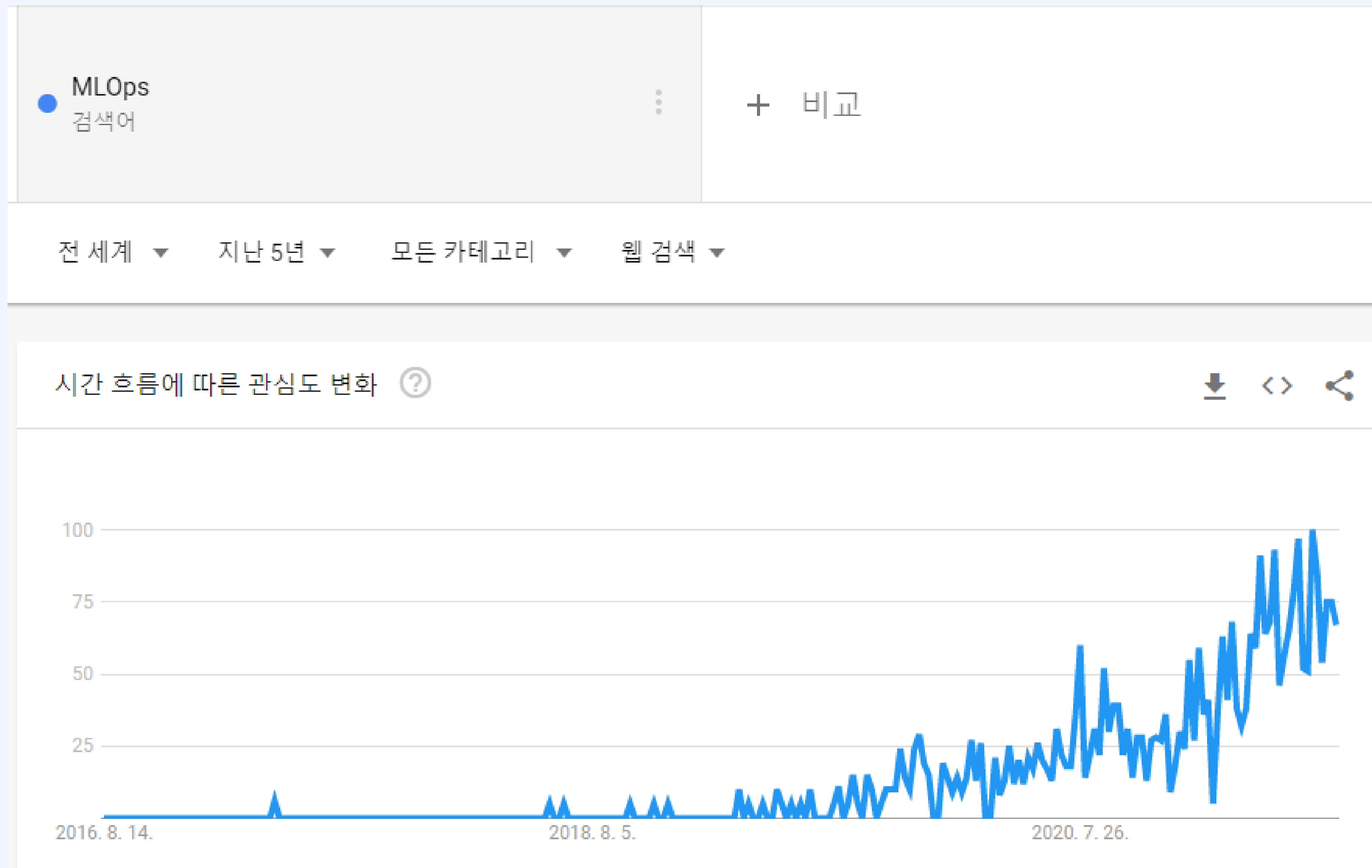
기존 ML 프로젝트
진행 방식의 한계



최신 트렌드, MLOps

1

기존 ML 프로젝트
진행 방식의 한계



최신 트렌드, MLOps

1

기존 ML 프로젝트
진행 방식의 한계

관련 주제 ?	급상승 ▼	다운로드	코드	공유	관련 검색어 ?	급상승 ▼	다운로드	코드	공유
1 기계 학습 - 연구 분야	급등				1 azure mlops	급등			
2 마이크로소프트 애저 - 주제	급등				2 what is mlops	급등			
3 학습 - 주제	급등				3 mlops tools	급등			
4 자료 - 주제	급등				4 mlops course	급등			
5 아마존 웹 서비스 - 웹사이트	급등				5 introducing mlops	급등			

빅테크 기업에서 정의한 MLOps

1

기존 ML 프로젝트
진행 방식의 한계

*MLOps*는 ML 시스템 개발(Dev)과 ML 시스템 운영(Ops)을 통합하는 것을 목표로 하는 ML 엔지니어링 문화 및 방식입니다. *MLOps*을 수행하면 통합, 테스트, 출시, 배포, 인프라 관리를 비롯하여 ML 시스템 구성의 모든 단계에서 자동화 및 모니터링을 지원할 수 있습니다.

- Google

As you move from running individual artificial intelligence and machine learning (AI/ML) projects to using AI/ML to transform your business at scale, the discipline of ML Operations (MLOps) can help. MLOps accounts for the unique aspects of **AI/ML projects in project management, CI/CD, and quality assurance, helping customers improve delivery time, reduce defects, and make data scientists more productive.** MLOps refers to a methodology that is built on applying DevOps practices to machine learning workloads.

- Aws

MLOps is modeled on the existing discipline **of DevOps, the modern practice of efficiently writing, deploying and running enterprise applications**

- NVIDIA

출처)

- https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning?hl=ko#devops_versus_mlops
- <https://docs.aws.amazon.com/sagemaker/latest/dg/sagemaker-projects-why.html>
- <https://blogs.nvidia.com/blog/2020/09/03/what-is-mlops/>

MLOps 란

1

기존 ML 프로젝트
진행 방식의 한계

MLOps 란
ML 을 효율적으로 개발하고
성공적으로 서비스화하고 운영할 때
필요한 모든 것을 다루는 분야

ML 을 Service 화하기 위한 기술, MLOps

3 MLOps 의 구성요소

데이터, 모델, 서빙

1

기존 ML 프로젝트
진행 방식의 한계

데이터

모델

서빙

데이터

1

기존 ML 프로젝트
진행 방식의 한계

데이터 수집
파이프라인

Sqoop, Flume, Kafka, Flink,
Spark Streaming, Airflow

데이터 저장

MySQL, Hadoop, Amazon S3, MinIO

데이터 관리

TFDV, DVC, Feast, Amundsen

모델

1

기존 ML 프로젝트
진행 방식의 한계

모델 개발

Jupyter Hub, Docker, Kubeflow,
Optuna, Ray, katib

모델 버전 관리

Git, MLflow, Github Action, Jenkins

모델 학습
스케줄링 관리

Grafana, Kubernetes

서빙

1

기존 ML 프로젝트
진행 방식의 한계

모델 패키징

Docker, Flask, FastAPI, BentoML,
Kubeflow, TFServing, seldon-core

서빙
모니터링

Prometheus, Grafana, Thanos

파이프라인
매니징

Kubeflow, argo workflows, Airflow

Cloud Vendor

1

기존 ML 프로젝트
진행 방식의 한계

Aws SageMaker



Amazon SageMaker

GCP Vertex AI



vertex.ai

**Azure Machine
Learning**



Azure Machine Learning

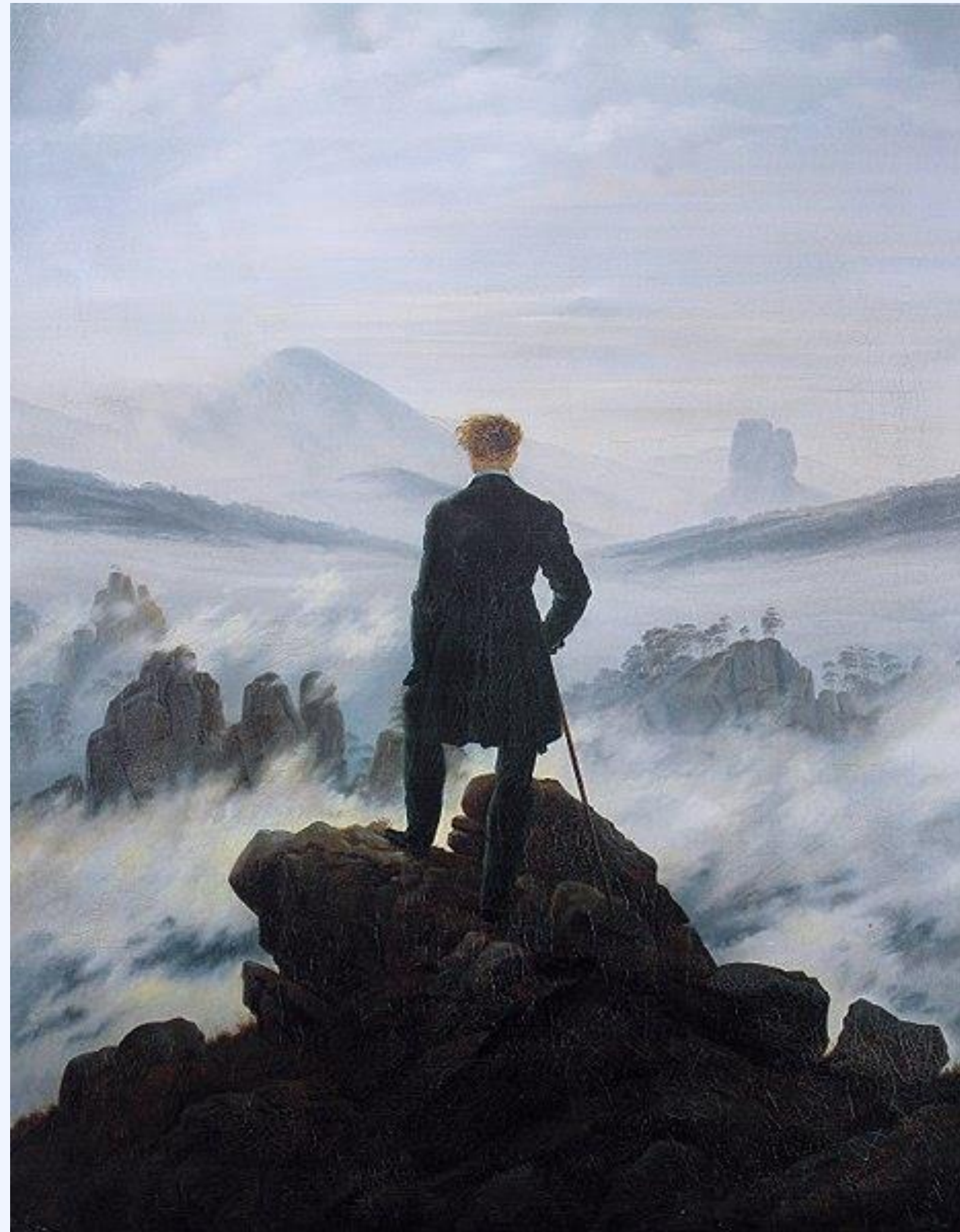
ML 을 Service 화하기 위한 기술, MLOps

4 앞으로의 강의에 대하여

MLOps 라는 미지의 분야

1

기존 ML 프로젝트
진행 방식의 한계



출처)

- https://en.wikipedia.org/wiki/Wanderer_above_the_Sea_of_Fog

MLOps 의 기본

1

기존 ML 프로젝트
진행 방식의 한계

ML 이론

SW 구현 능력

클라우드 지식

협업 능력

앞으로의 강의

1

기존 ML 프로젝트
진행 방식의 한계

도커와 쿠버네티스

MLOps 를
대표하는 오픈소스

Aws SageMaker