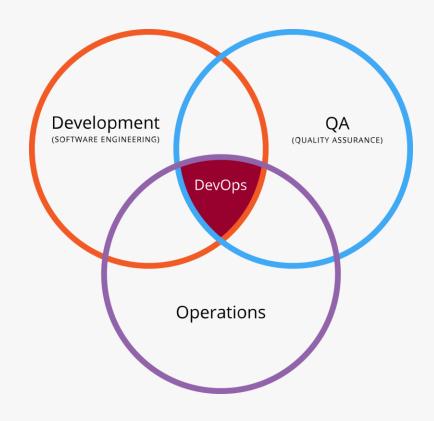
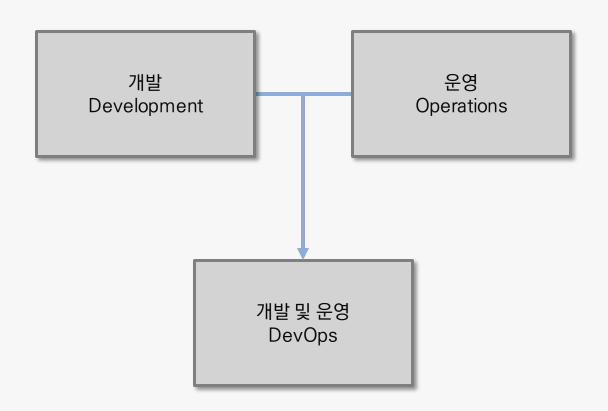
MLOps 公刊 mlf/ow[™]

01. 주요 핵심 개념 - 출발점



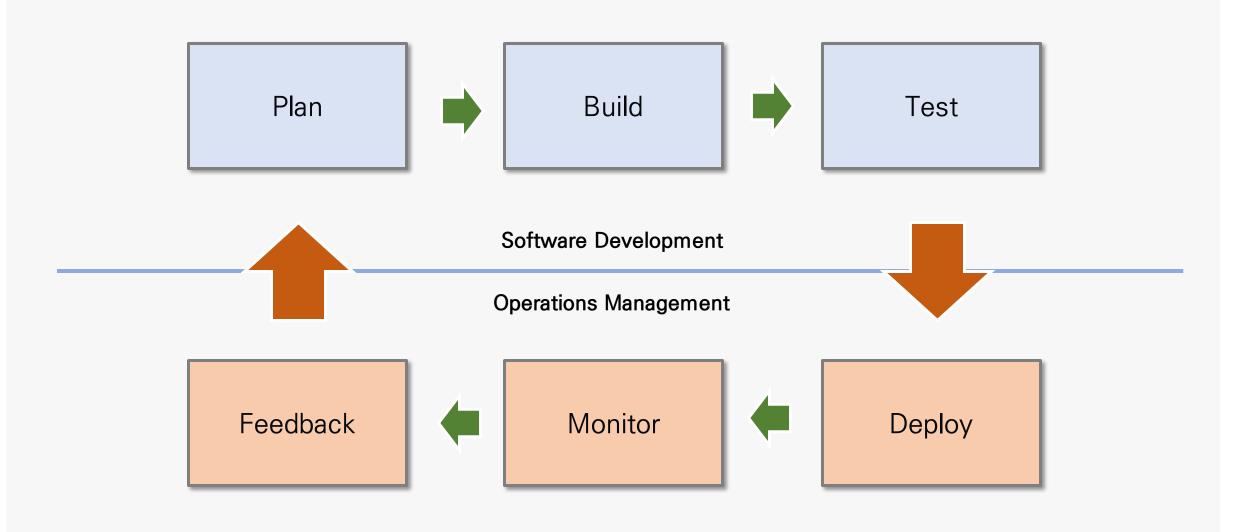


출처: https://en.wiktionary.org/wiki/DevOps



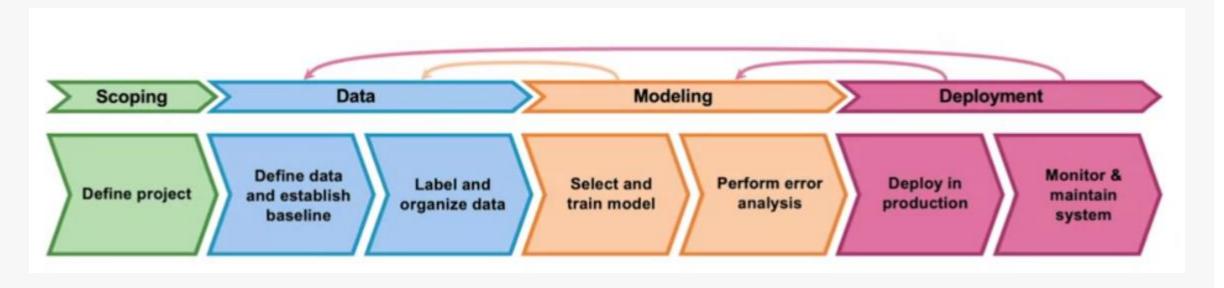
01. 주요 핵심 개념 - DevOps 워크플로우





01. 주요 핵심 개념 - 출발점

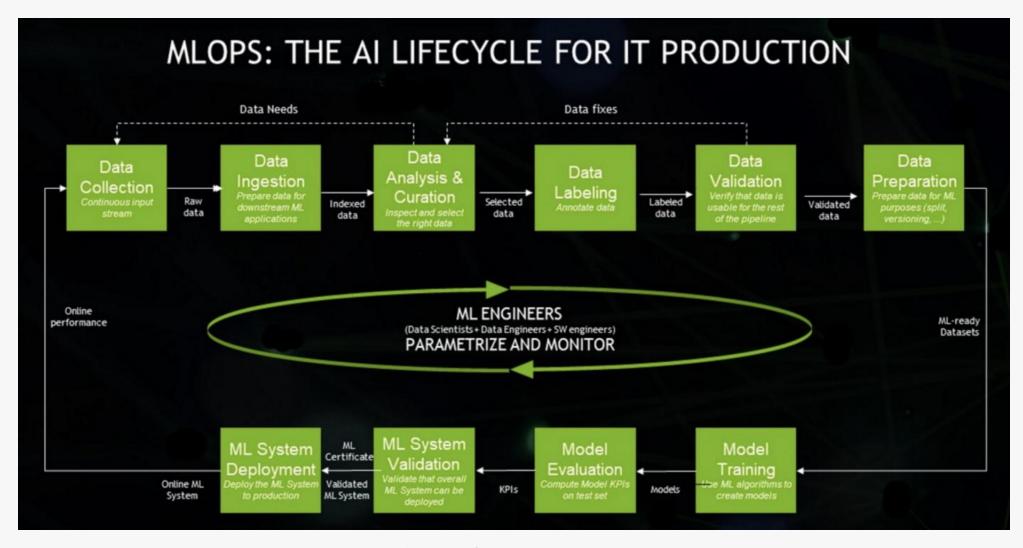




출처: Andrew Ng. Steps of an ML Project in Machine Learning Engineering for Production (MLOps) Specialization

01. 주요 핵심 개념 - 출발점

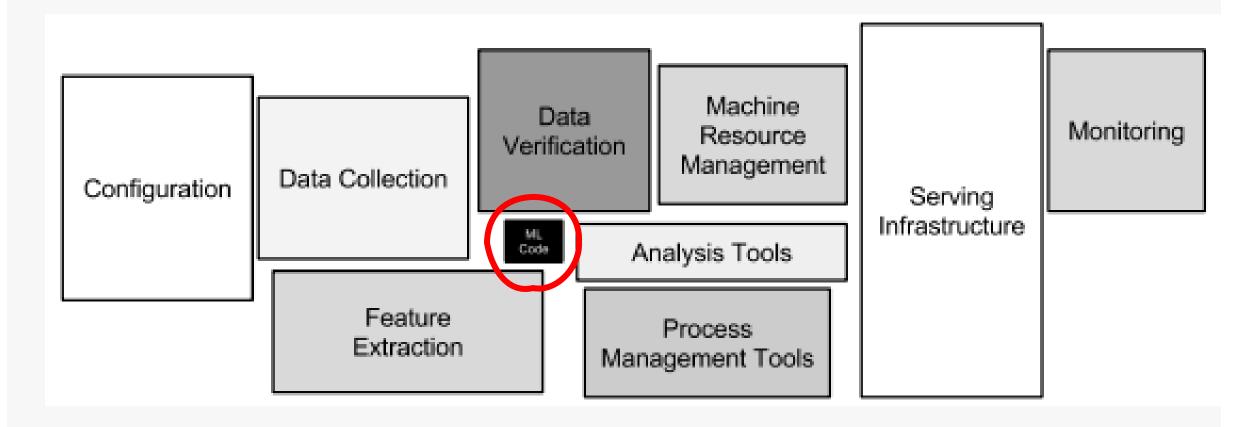




출처: NVIDIA KOREA. AI 개발의 스포트라이트 'MLOps' A to Z 알아보기

01. 주요 핵심 개념 - 현실



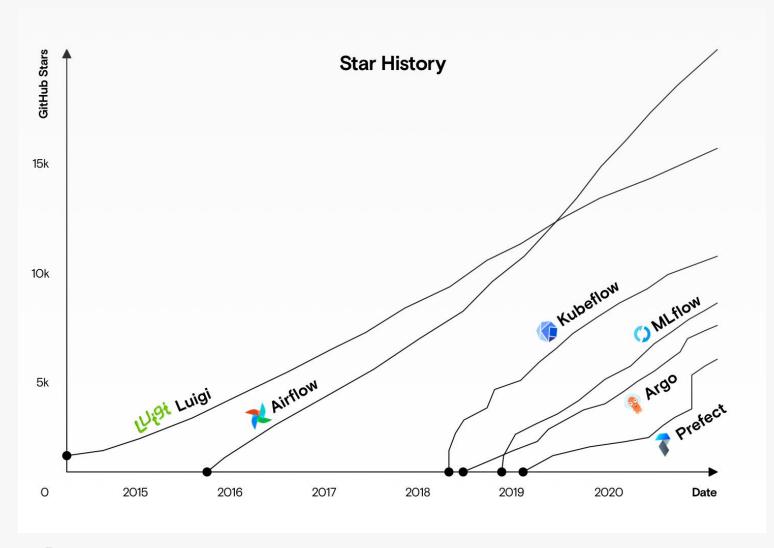


머신러닝 코드는 전체 시스템에 매우 일부분에 지나지 않는다.

출처: Google. (2015). Hidden Technical Data in Machine Learning System

01. 주요 핵심 개념 - 주요 Tools





출처: M. Schmit. Airflow vs. Luigi vs. Argo vs. MLFlow vs. KubeFlow

01. 주요 핵심 개념 – 주요 MLOps Platforms





출처: C. Pachmann (2021). Global MLOps and ML tools landscape.

머신러닝 주요 흐름





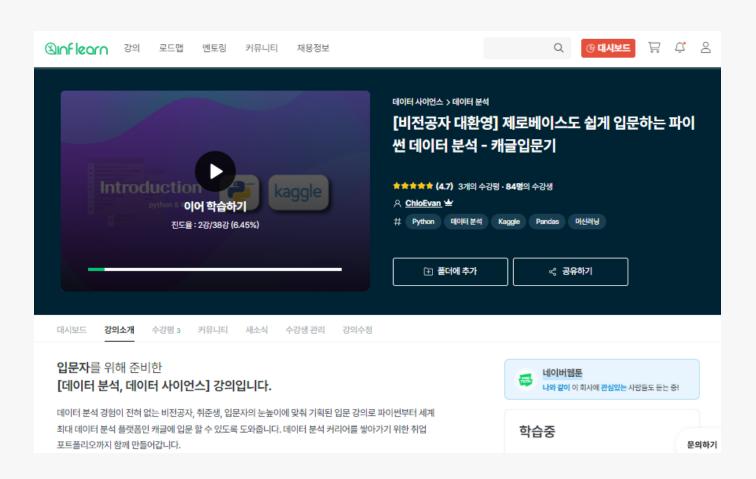
01. 머신러닝 진행 과정 리뷰





기초 강의는 여기에서

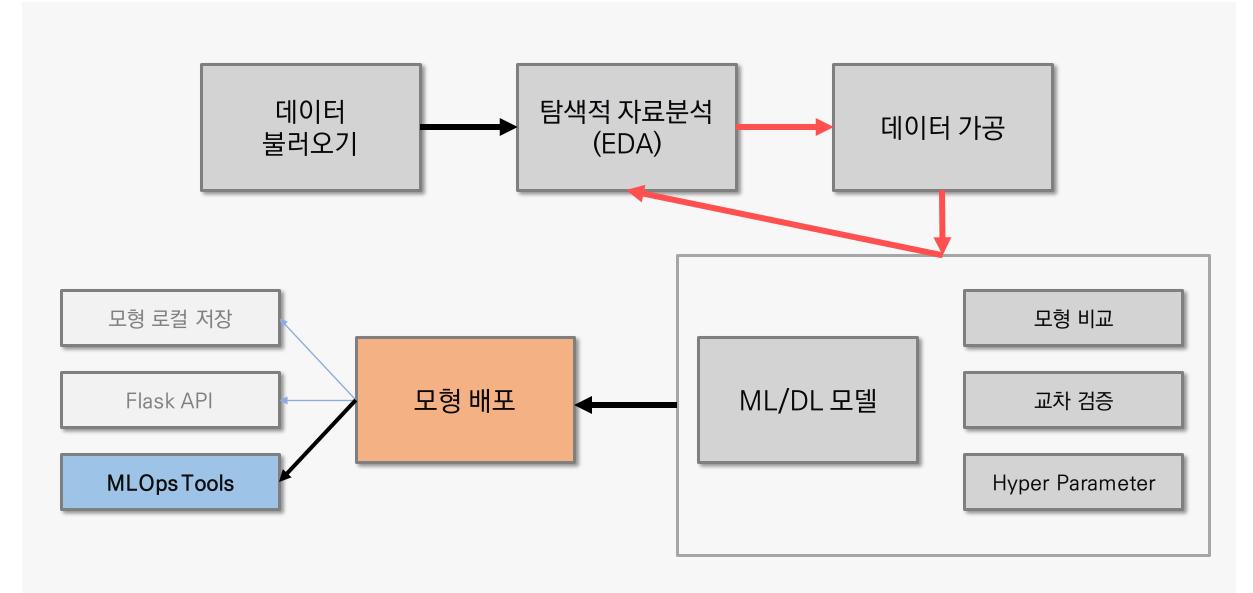




머신러닝 입문: https://www.inflearn.com/course/파이썬-데이터분석-캐글입문?inst=46ae3c78

01. 머신러닝/딥러닝 진행 과정 리뷰





02. 다루지 않는 것



탐색적 자료 분석

모형 선정 과정

하이퍼 파라미터 개념

모형 알고리즘

03. 평가지표 - 회귀



MAE

• *N* = Total Number

• Y_i = Actual Value

• \hat{Y}_i = Predicted Value

 $MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \widehat{y_i}|$

- ✓ Mean Absolute Error
- ✓ 오차는 실젯값과 예측값의 차이
- ✓ 오차에 절댓값 취함
- ✓ 에러의 크기 그대로 반영

MSE

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \widehat{y_i})^2$$

- ✓ Mean Squared Error
- ✓ 오차는 실젯값과 예측값의 차이
- ✓ 오차에 제곱값 취함
- ✓ 이상치 많을수록 MSE값 증가
- ✓ 오차에 대한 가중치 적용

RMSE

$$RMSE = \sqrt{MSE}$$

- ✓ Root Mean Absolute Error
- ✓ 오차는 실젯값과 예측값의 차이
- ✓ MSE값에 루트를 씌움
- ✓ 보편적으로 많이 사용

03. 평가지표 - 회귀



R² Score

- Y_i = Actual Value
- \hat{Y}_i = Predicted Value
- \overline{Y} = Mean of Actual Value

$$R^2 = 1 - \frac{SSE}{SST}$$

$$SSE = \sum_{i=1}^{N} (y_i - \widehat{y}_i)^2$$

$$SST = \sum_{i=1}^{N} (y_i - \bar{y})^2$$

- ✓ Coefficient of Determination
- ✓ SST는 실젯값에서 실젯값 평균을 뺀 결과의 총합
- ✓ SSE는 실젯값에서 예측값 오차의 제곱
- ✓ 1에 가까울수록 좋은 모델
- ✓ 0에 가까울수록 나쁜 모델

03. 평가지표 - 분류



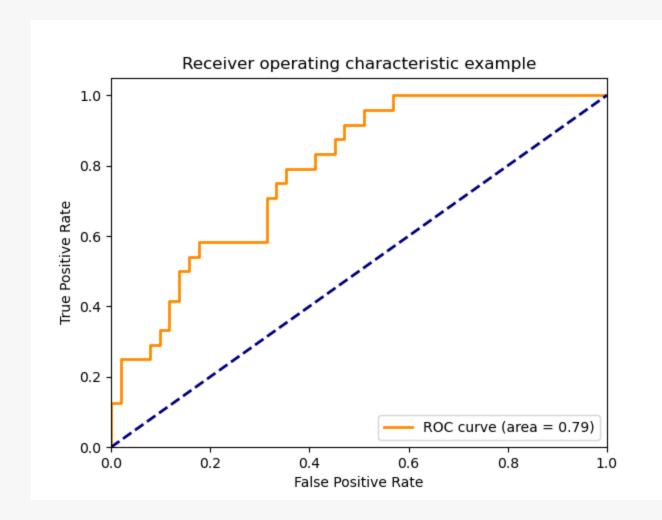
$F1Score = 2 imes rac{Precision imes Recall}{Precision + Recall}$		Predicted Class		
		암 양성(Positive)	암 음성(Negative)	
Actual Class	암 양성(Positive)	True Positive (TP)	False Negative (FN)	재현율 (Sensitivity) $\frac{TP}{(TP+FN)}$
	암 음성(Negative)	False Positive (FP)	True Negative (TN)	특이도 (Specificity) $\frac{TN}{(TN+FP)}$
		정밀도 (Precision)	음성 예측도 (NPV)	정확도 (Accuracy)
		$\frac{TP}{(TP+FP)}$	$\frac{TN}{(TN+FN)}$	$\frac{TN + TN}{(TP + TN + FP + FN)}$

03. 평가지표 - 분류



$$TPR = \frac{TP}{TP + FN}$$

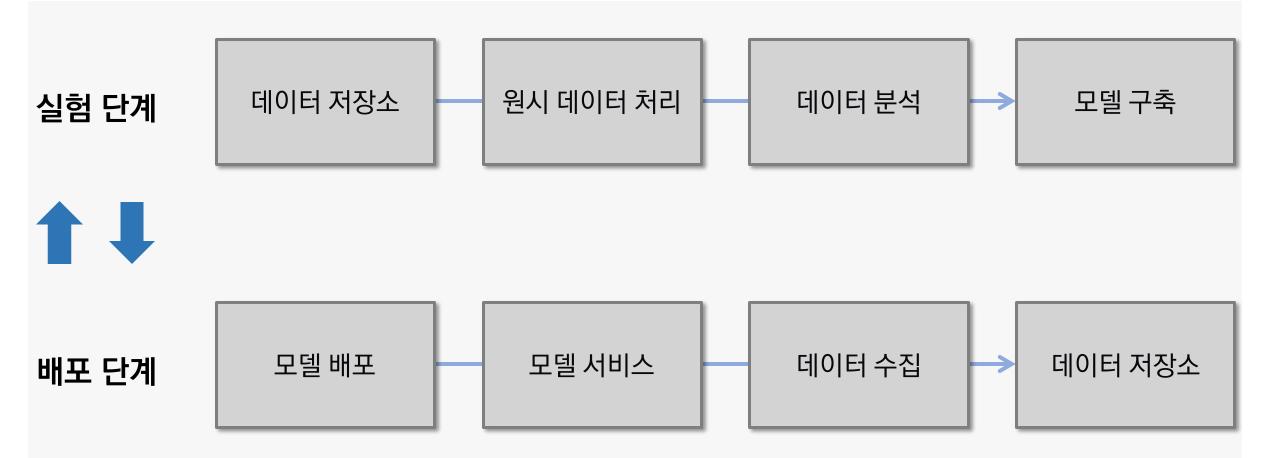
$$FPR = \frac{FP}{FP + TN}$$



MLOps 구축 방안 mlf/www

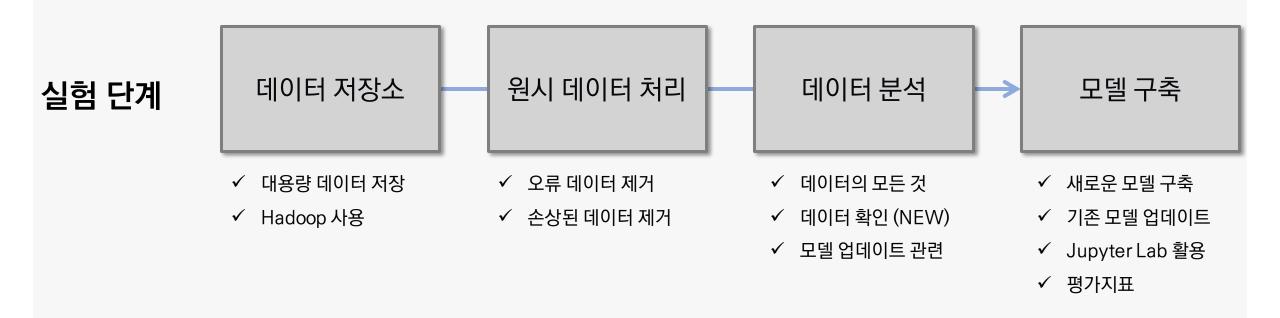
01. 수동 구현





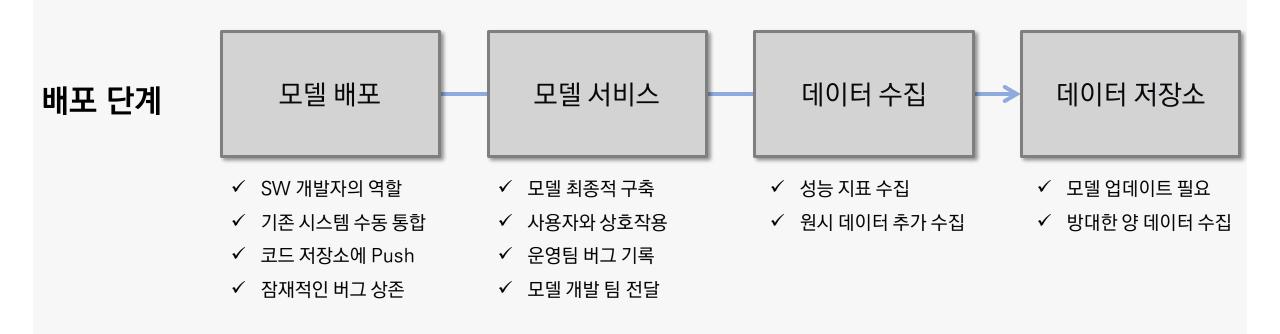
01. 수동 구현: 실험 단계





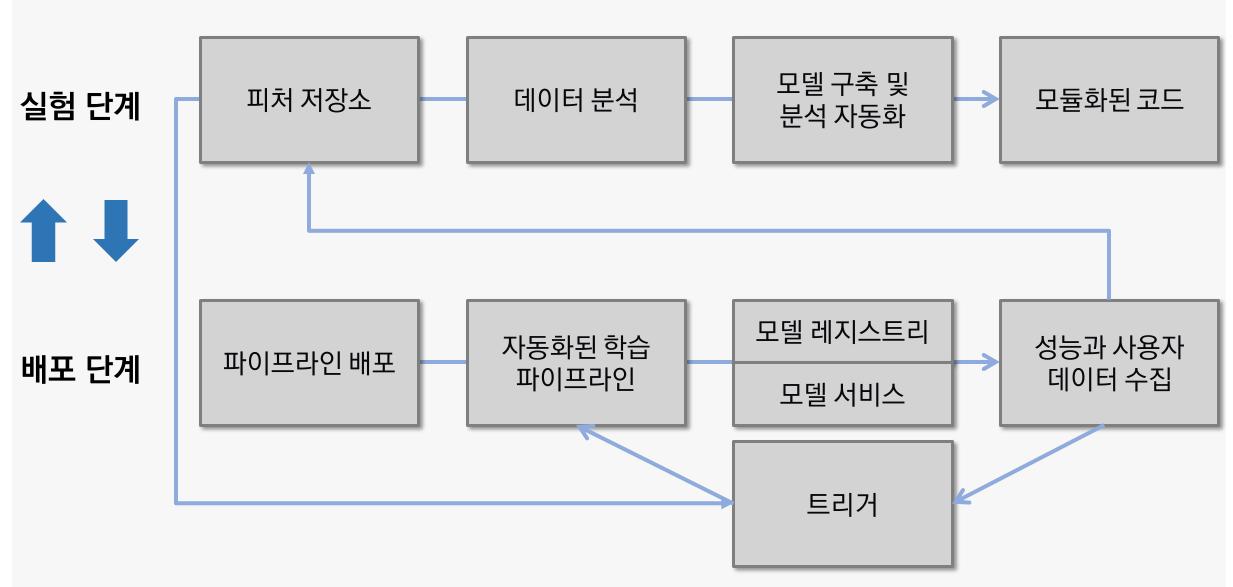
01. 수동 구현: 실험 단계





02. 지속적인 모델 전달





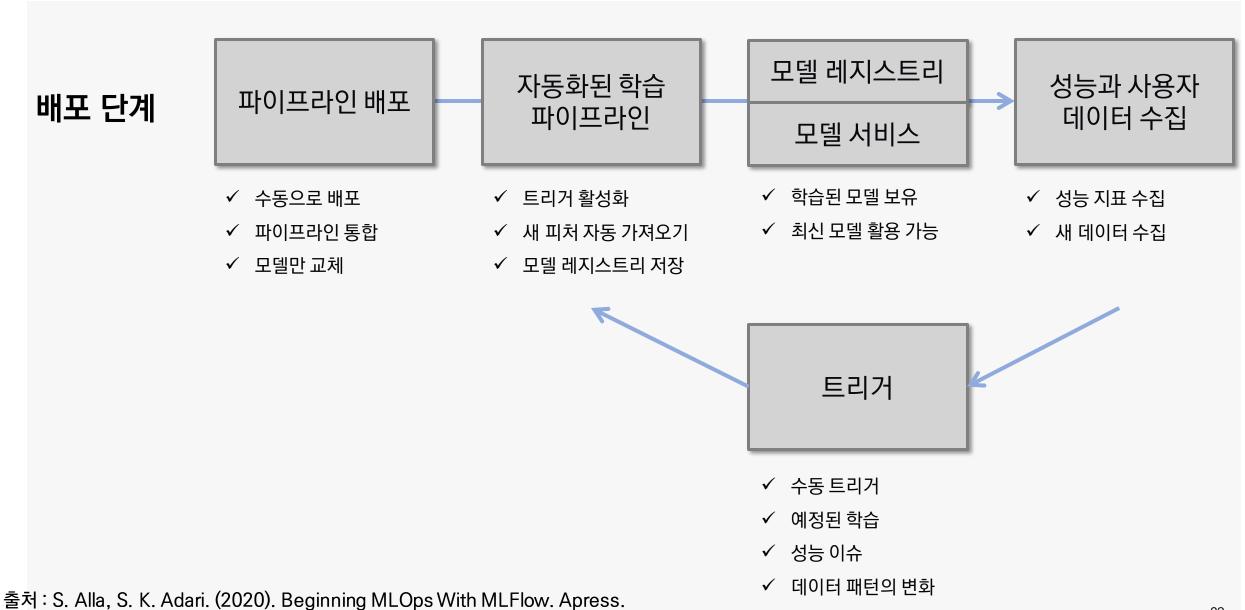
02. 지속적인 모델 전달: 실험 단계





02. 지속적인 모델 전달: 배포 단계

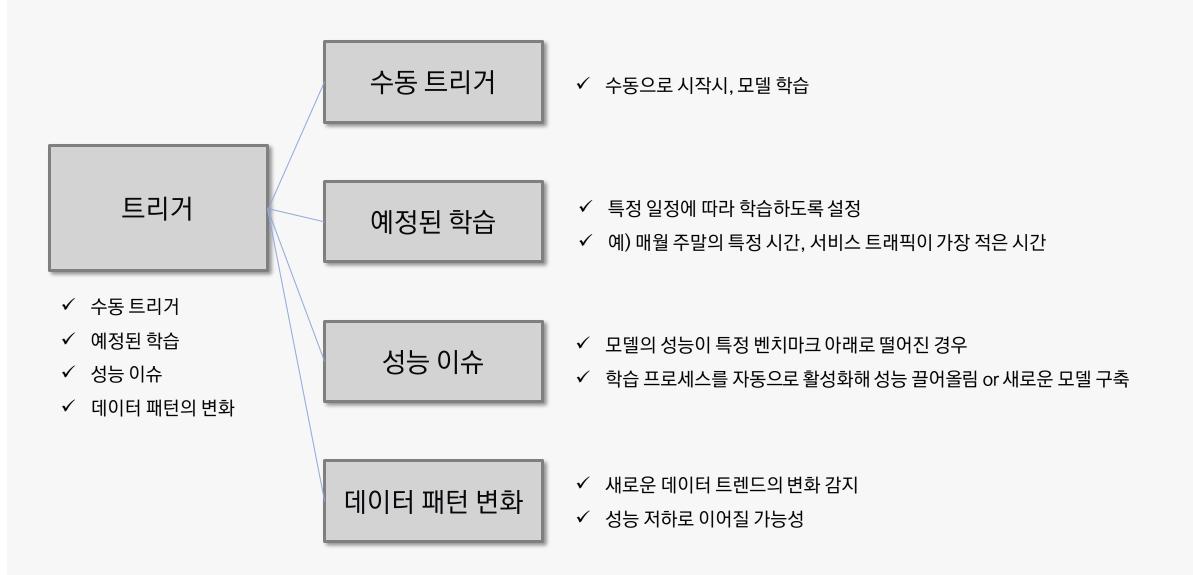




강사: EVAN Copyright ⓒ DSCHLOE. All Rights Reserved. 무단 전재 및 배포 금지

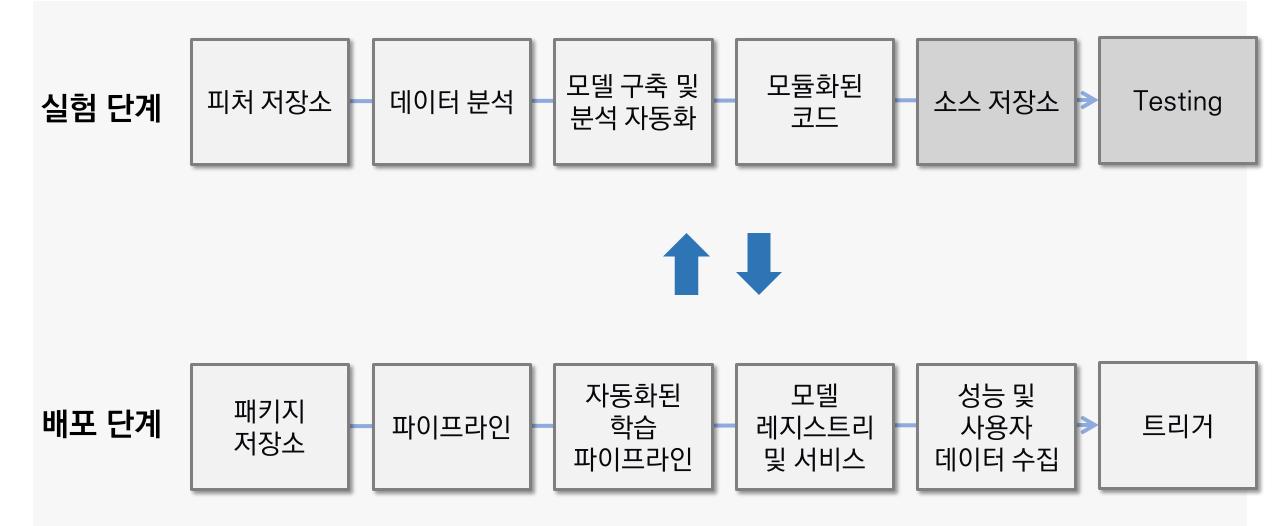
02. 지속적인 모델 전달: 트리거 상세 내용





03. 파이프라인의 지속적인 통합과 전달





03. 파이프라인의 지속적인 통합과 전달: 실험 단계



소스 저장소

- ✓ 서로 다른 파이프라인 및 모델에 대한 모든 패키지 된 코드 저장
- ✓ 테스트를 통해 모든 컴포넌트가 올바르게 작동하는지 확인

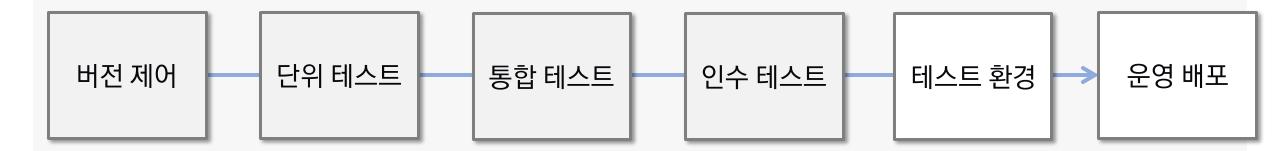
테스팅

- ✓ 가장 중요한 단계, 새로운 환경에 지속적으로 설계, 구축 및 배포되는 자동화의 결과
- ✓ 자동화 테스트 예시
 - 검증 테스트 절차 통해 하이퍼파라미터의 정확한 튜닝이 이루어지는가?
 - 각 파이프라인 구성 요소가 올바르게 작동하는가?
 - 데이터 처리가 올바르게 수행되고 있는가?
 - 모델 분석이 올바르게 작동하는가?

04. 파이프라인 및 자동화



파이프라인의 예시



MLFlow 실습 mlf/ow