Microsoft

MLOps 101

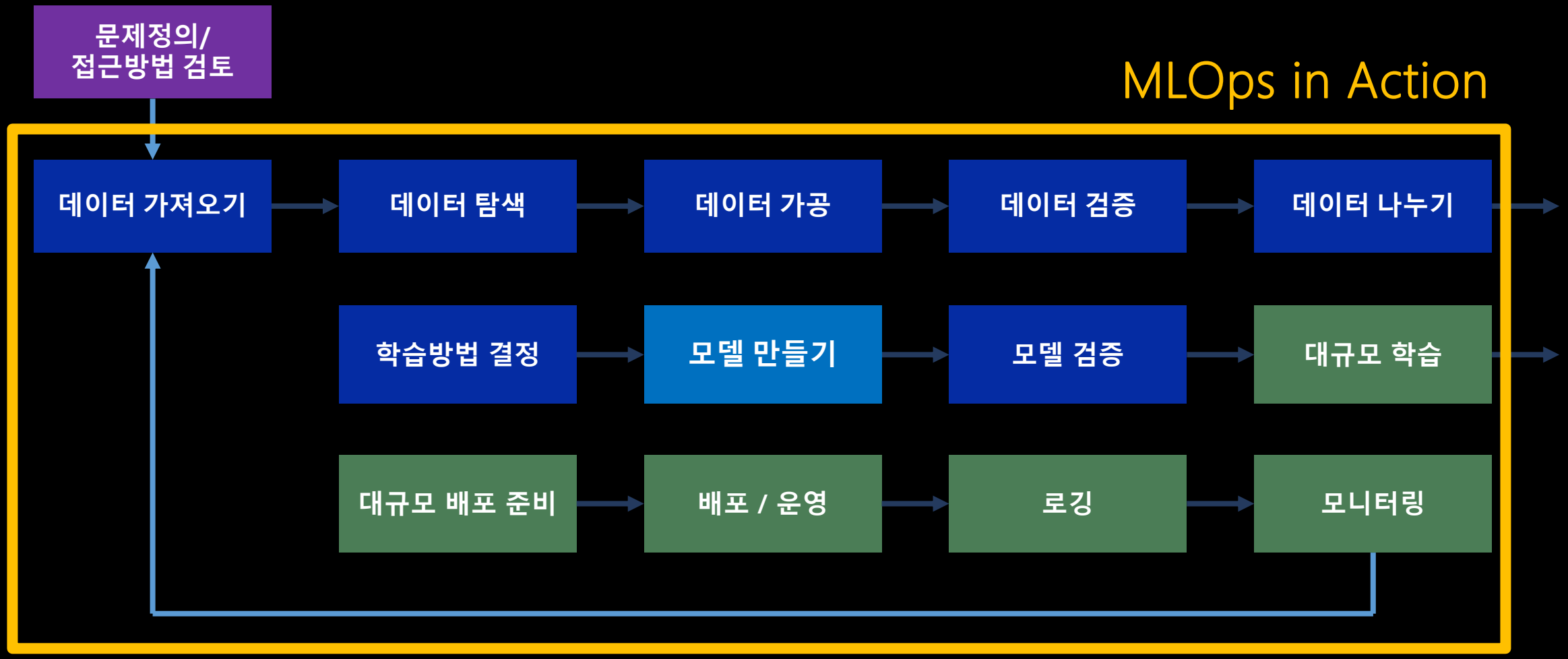
Episode 6: MLOps in Action 엿보기

한석진
마이크로소프트

Episode 6
MLOps in Action
엿보기

- 이제껏 알아본 것들이 실제 어떻게 구현되나
  - 애저머신러닝과 애저데브옵스의 만남 *DEMO*
- MLOps란 무엇이고 어떻게 준비할 수 있는가

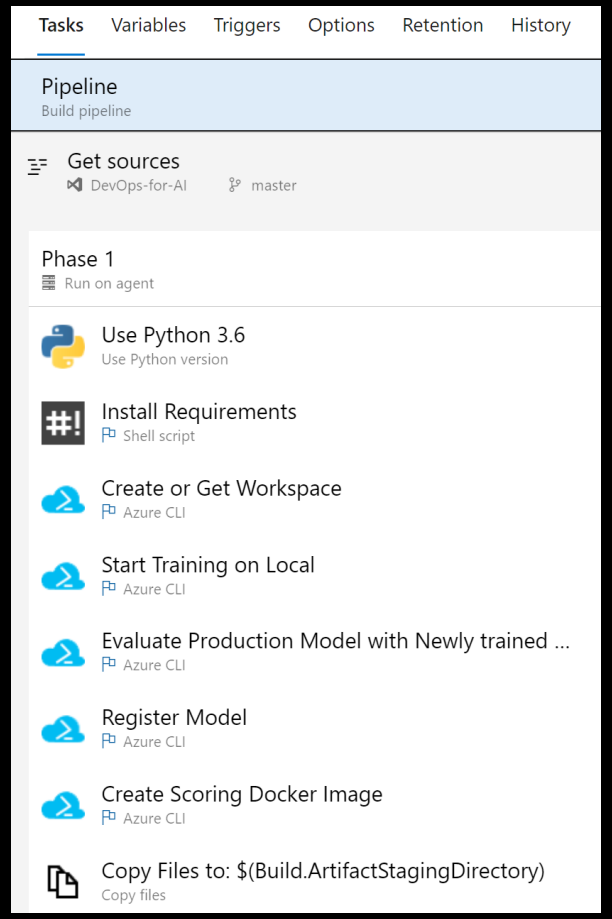
ML 생애주기

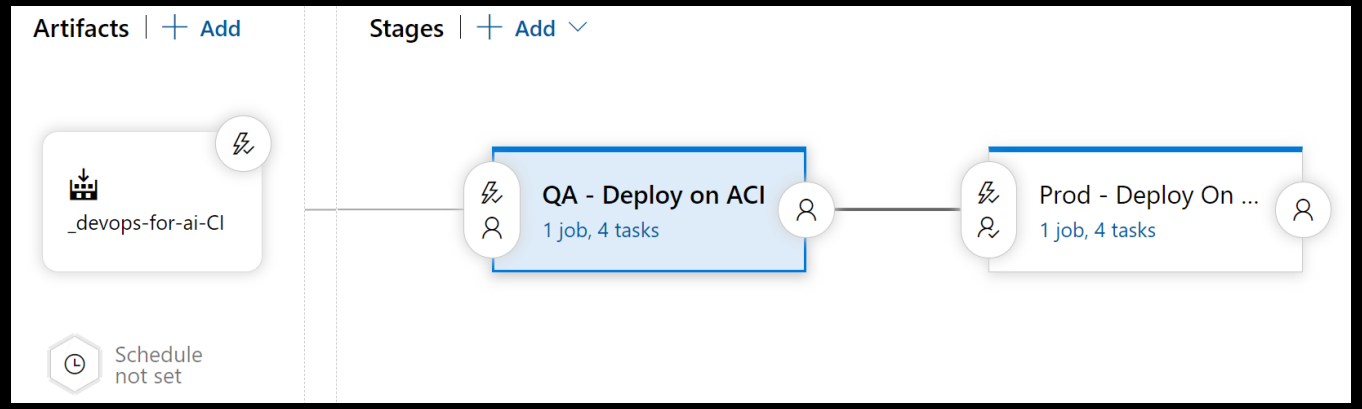

문제정의/
접근방법 검토
MLOps in Action
데이터 가져오기
데이터 탐색
데이터 가공
데이터 검증
데이터 나누기
학습방법 결정
모델 만들기
모델 검증
대규모 학습
대규모 배포 준비
배포 / 운영
로깅
모니터링

DEMO: 애저머신러닝과 애저데브옵스의 만남

실제 어떻게 구현되나 – 초보 버전

MLOps





https://azuredevopsdemogenerator.azurewebsites.net/?name=azure%20machine%20learning

DEMO: 애저머신러닝과
애저데브옵스의 만남

실제 어떻게 구현되나 – 고급 버전

IaC: Infra As Code
(클라우드 자원 생성)

Jobs in run #20200626.10
IaC-create-resources

Jobs

Job	1m 7s
Initialize job	2s
Checkout mlops-202006@try_c1 t...	1s
Deploy MLOps resources to Azure	50s
Component Detection (auto-inje...	12s
Post-job: Checkout mlops-20200...	<1s
Finalize Job	<1s
Report build status	<1s

MLOps

Jobs in run #20200626.5
mlops-202006

Model CI

Model CI Pipeline	2m 39s
Initialize job	4s
Initialize containers	1m 45s
Checkout mlops-202006@try_c2 t...	1s
Run lint tests	1s
Run unit tests	2s
Publish test results	1s
Publish coverage report	2s
Publish Azure Machine Learning ...	30s
Component Detection (auto-injec...	8s
Post-job: Checkout mlops-20200...	<1s
Stop Containers	1s
Finalize Job	<1s

Train model

Get Pipeline ID for execution	2m 13s
Trigger ML Training Pipeline	7m 44s
Determine if evaluation succee...	2m 12s

Deploy to ACI

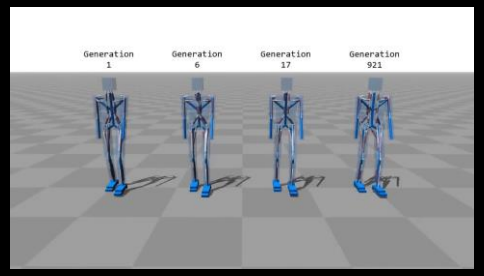
Deploy to ACI	4m 17s

Finalize build

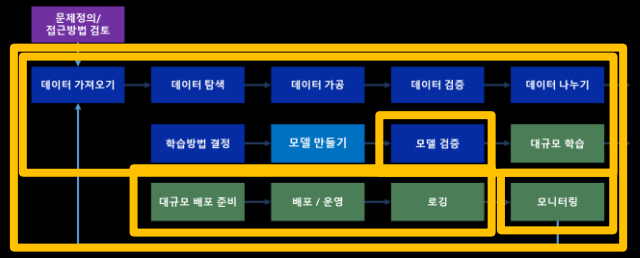
Report build status	<1s

https://github.com/Microsoft/MLOpsPython

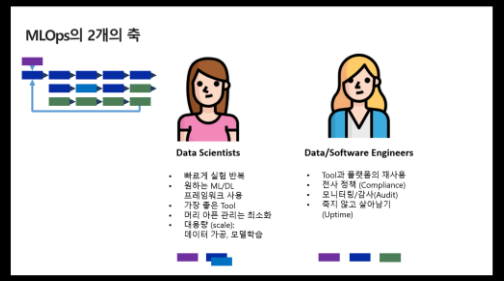
“MLOps란 무엇이고 어떻게 준비할 수 있는가”



ML 생애주기

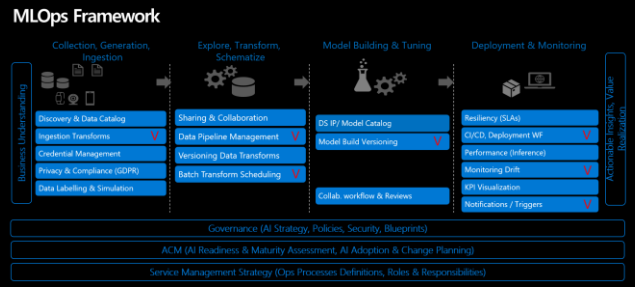


MLOps는 누가 하나: Actors



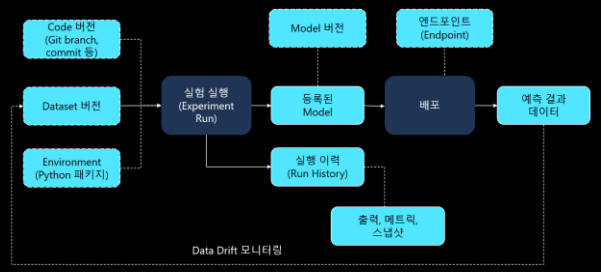
- ep1: MLOps가 뭐길래
- ep2: ML 생애주기 (1) 데이터준비
- ep3: ML 생애주기 (2) 실험/학습
- ep4: ML 생애주기 (3) 모델 해석
- ep5: ML 생애주기 (4) 배포/서빙
- ep6: MLOps in Action 엿보기

MLOps 어디까지 가봤니?



실험 추적관리

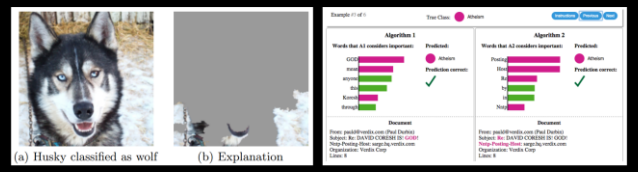
DEMO: 실험 추적관리: 데이터셋, 코드, 환경, 모델, 서빙



모델 해석이 왜 중요한가?

- “단순하게 설명할 수 없다면, 충분히 제대로 이해한 것이 아니다” — Albert Einstein
- 정확도만으로는 더 이상 충분하지 않다
- 내 모델이 예측을 잘 하는 것 같지만 사실은 완전 착각하고 있지는 않는가

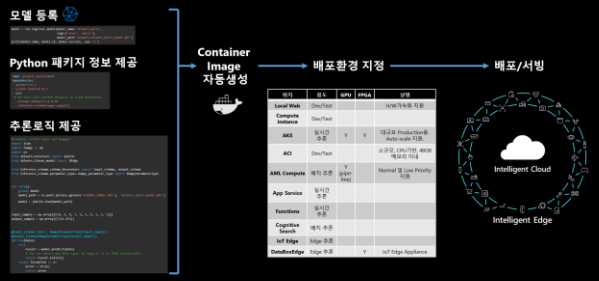
해석이 더 어렵다고 포기?



출처: https://arxiv.org/abs/1602.04938, “Why Should I Trust You?”: Explaining the Predictions of Any Classifier

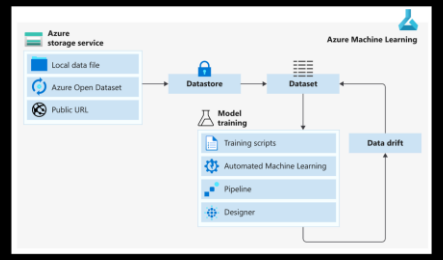
애저머신러닝에서 패키징, 배포 (서빙) 개념

DEMO: 자동화된 ML에서의 No code 배포



모델의 모니터링: 데이터 드리프트(Data Drift)

DEMO: 애저머신러닝에서 데이터 드리프트 확인

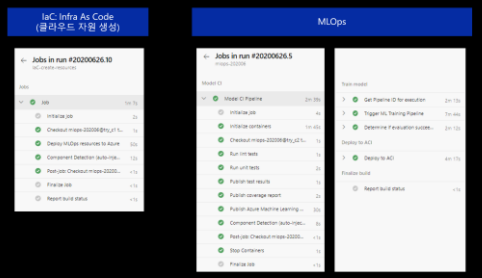


Data Drift

- 데이터의 패턴 특성이 시간이 지나며 변하는 현상
- 학습에 활용된 당시의 데이터와 차이가 커질수록 예측 성능이 저하될 가능성
- 따라서 Data Drift 여부를 지속적으로 모니터링하면 모델의 재학습 시점을 판단하는데 간접적인 방법이 될 수 있음

이제껏 알아본 것들이 실제 어떻게 구현되나

DEMO: 애저머신러닝과 애저데브옵스의 만남



참고자료

- 애저머신러닝 공식 온라인문서: https://aka.ms/azpls/azureml-docs
- 애저머신러닝 MLOps 개념: https://aka.ms/azpls/azureml-docs-mlops
- 애저머신러닝의 자동화된 ML 따라하기 실습 가이드: https://aka.ms/azpls/azureml-learn-automl
- 페이스북 광화문 AI: https://www.facebook.com/groups/GwangAI
- 애저머신러닝 관련 유용한 자료 모음: https://aka.ms/azpls/awesome-azureml

Episode 6
MLOps in Action
엿보기

- 이제껏 알아본 것들이 실제 어떻게 구현되나
  - 애저머신러닝과 애저데브옵스의 만남 *DEMO*
- MLOps란 무엇이고 어떻게 준비할 수 있는가