



*자동화된 ML, 나도 해보자*

*Episode 5: 자동화된 ML 모델 배포 활용하기*

한석진  
마이크로소프트

---

## Episode 5

자동화된ML 모델  
배포 활용하기

---

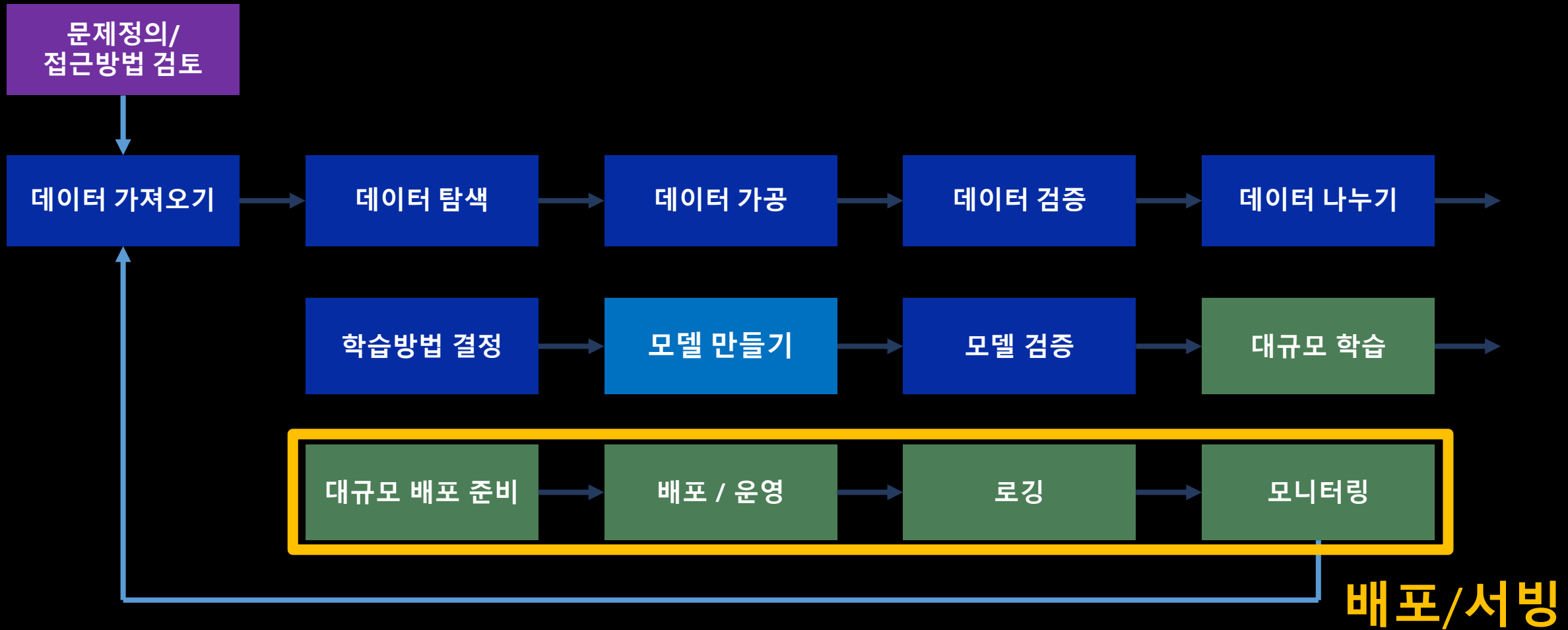
ML 생애주기

코딩 없이 배포하기

배포 결과 확인

애저머신러닝에서 패키징, 배포 (서빙) 개념

# ML 생애주기



# 코딩 없이 배포하기

Microsoft Azure Machine Learning

새로 만들기

홈

작성자가

Notebooks

자동화된 ML

디자이너

자산

데이터 세트

실험

파이프라인

모델

엔드포인트

관리

컴퓨팅

데이터 저장소

데이터 레이블 지정

aml-handson-202001 > 실험 > diabetes-automl

실행 5 Completed

새로 고침

배포

다운로드

모델 설명

취소

로그 스트리밍 사용

자동 줄 바꿈

세부 정보

모델

설명(미리 보기)

메트릭

출력 + 로그

이미지

자식 실행

스냅샷

70\_driver\_log.txt

PY scoring\_file\_v\_1\_0\_0.py

process\_status.json

explanation

logs

outputs

conda\_env\_v\_1\_0\_0.yml

env\_dependencies.json

internal\_cross\_validated\_models.pkl

model.pkl

pipeline\_graph.json

PY

scoring\_file\_v\_1\_0\_0.py

accuracy\_table

automl\_driver.py

confusion\_matrix

모델 배포

이름 ?

설명 ?

컴퓨팅 형식 ?

Azure Kubernetes Service

컴퓨팅 이름 ?

이름 선택 또는 검색

모델: AutoMLEf4af53f30:1

인증 사용

유형

토큰 기반 인증

이 모델은 코드 없는 배포를 지원합니다. 기본 원칙을 읽을 수 있습니다.

사용자 지정 배포 자산 사용

사용자 지정 배포 자산 사용

속성

상태

Completed

만들어짐

Sep 22, 2020 12:14 AM

시작됨

Sep 22, 2020 12:18 AM

기간

37.79초

컴퓨팅 대상

cpu-cluster

실행 ID

AutoML\_ef4af53f-3abe-479a-9ed6-b819360b

실행 번호

5

스크립트 이름

automl\_driver.py

만든 사람

Seokjin Han

입력 데이터 세트

없음

출력 데이터 세트

# 배포 결과 확인

Microsoft Azure Machine Learning

새로 만들기

홈

작성자

Notebooks

자동화된 ML

디자이너

자산

데이터 세트

실험

파이프라인

모델

엔드포인트

관리

컴퓨팅

bankmarketing

기기인 간행

false

Swagger URI

http://74f9f...

ral.azurecon

CPU

1

메모리

2 GB

Application In

true

Application In

https://port...

ae93-4e43-

-aml-handso

mlhandson2

Swagger Editor

Supported by SMARTBEAR

```
1 swagger: '2.0'
2 info:
3   title: bankma
4   description:
5     for the Azu
6     service bar
7   version: '1.0
8   schemes:
9     - https
10  consumes:
11    - applicator
12  produces:
13    - applicator
14  securityDefinitions:
15    Bearer:
16      type: apiKey
17      name: Authc
18      in: header
19      description:
20        Bearer at
21  paths:
22    /:
23      get:
24        operation
25        Service
26        descripti
27        check e
28        the ser
29        given p
30      responses:
31        '200':
32          descr
33          is
34          res
35          ret
36          cor
37          schen
38          typ
39          exam
40          app
```

amlhandson20205583227653 | 성능

Application Insights | 디렉터리: Microsoft

검색(Ctrl+/)

액세스 제어(IAM)

태그

문제 진단 및 해결

조사

애플리케이션 맵

스마트 검색

라이브 메트릭

트랜잭션 검색

가용성

실패

성능

문제 해결 가이드(미리 보기)

모니터링

경고

메트릭

로그

통합 문서

Application Insight로 서비스 모니터링

현재 시간: 지난 24시간

역할 = 모두

서버

브라우저

작업 종속성 역할

평균 50번째 95번째 99번째

작업 시간: 범위 확대

요청 수

오전 03:00 오전 06:00 오전 09:00 오후 12:00 오후 03:00 오후 06:00 오후 09:00 수 25

02:01 AM

02:01 AM

작업 선택

검색하여 항목 필터링...

작업 이름

지속 시간(평균)

개수

고정

전체 0.0669 ms 1.27k

# 애저머신러닝에서 패키징, 배포 (서빙) 개념

## 모델 등록

```
model = run.register_model(model_name='sklearn_mnist',
                           tags={'area': 'mnist'},
                           model_path='outputs/sklearn_mnist_model.pkl')
print(model.name, model.id, model.version, sep='\t')
```

## Python 패키지 정보 제공

```
name: project_environment
dependencies:
- python=3.6.2
- scikit-learn=0.22.1
- pip:
  # You must list azureml-defaults as a pip dependency
  - azureml-defaults>=1.0.45
  - inference-schema[numpy-support]
```

## 추론로직 제공

```
#Example: scikit-learn and Swagger
import json
import numpy as np
import os
from sklearn.externals import joblib
from sklearn.linear_model import Ridge

from inference_schema.schema_decorators import input_schema, output_schema
from inference_schema.parameter_types.numpy_parameter_type import NumpyParameterType

def init():
    global model
    model_path = os.path.join(os.getenv('AZUREML_MODEL_DIR'), 'sklearn_mnist_model.pkl')
    model = joblib.load(model_path)

input_sample = np.array([[10, 9, 8, 7, 6, 5, 4, 3, 2, 1]])
output_sample = np.array([3726.995])

@input_schema('data', NumpyParameterType(input_sample))
@output_schema(NumpyParameterType(output_sample))
def run(data):
    try:
        result = model.predict(data)
        # You can return any data type, as long as it is JSON serializable.
        return result.tolist()
    except Exception as e:
        error = str(e)
        return error
```

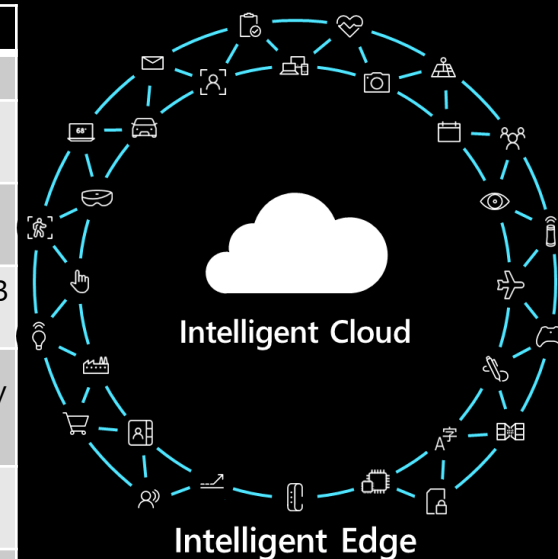
## Container Image 자동 생성



## 배포환경 지정

## 배포/서빙

위치	용도	GPU	FPGA	설명
Local Web	Dev/Test			H/W가속화 지원
Compute Instance	Dev/Test			
AKS	실시간 추론	Y	Y	대규모 Production용. Auto-scale 지원.
ACI	Dev/Test			소규모, CPU기반, 48GB 메모리 이내
AML Compute	배치 추론	Y (pipe-line)		Normal 및 Low Priority 지원
App Service	실시간 추론			
Functions	실시간 추론			
Cognitive Search	배치 추론			
IoT Edge	Edge 추론			
Stack Edge	Edge 추론	Y	Y	IoT Edge Appliance



\* 자동화된 ML에서는 생략

---

## Episode 5

### 자동화된ML 모델 배포 활용하기

---

ML 생애주기

코딩 없이 배포하기

배포 결과 확인

애저머신러닝에서 패키징, 배포 (서빙) 개념

# {다음 시간에는}

---

**Episode 6**  
자동화된ML  
코드로 돌려보기

---

입맛대로 샘플 고르기  
컴퓨팅 인스턴스 만들기  
샘플 가져오기 / 샘플 실행하기  
결과 확인하기 (학습결과, 모델설명, 배포까지)