

19기 정규세션

ToBig's 18기 강의를  
김성우

# Week 1-(2): Frameworks (and Libraries)

# Contents

Unit 01	Frameworks and Libraries
Unit 02	Types of frameworks and libraries
Unit 03	Sklearn, TensorFlow, Pytorch
Unit 04	Coding Examples

## Unit 01 | Frameworks and Libraries

Framework Definition :

" A framework is a structure that you can build software on. "

프레임워크 위에서 코드를 짠다

Why use frameworks ?

- 이미 검증된 코드와 시스템, 설계 위에서 코드를 할 수 있음
- 시간 단축
- 통일성

Library Definition :

" Collection of codes and modules for optimizing the current task. "

코드에서 라이브러리를 활용한다

Why use libraries ?

- 이미 검증된 코드와 시스템, 설계 위에서 코드를 할 수 있음
- 시간 단축
- 통일성

## Unit 01 | Frameworks and Libraries

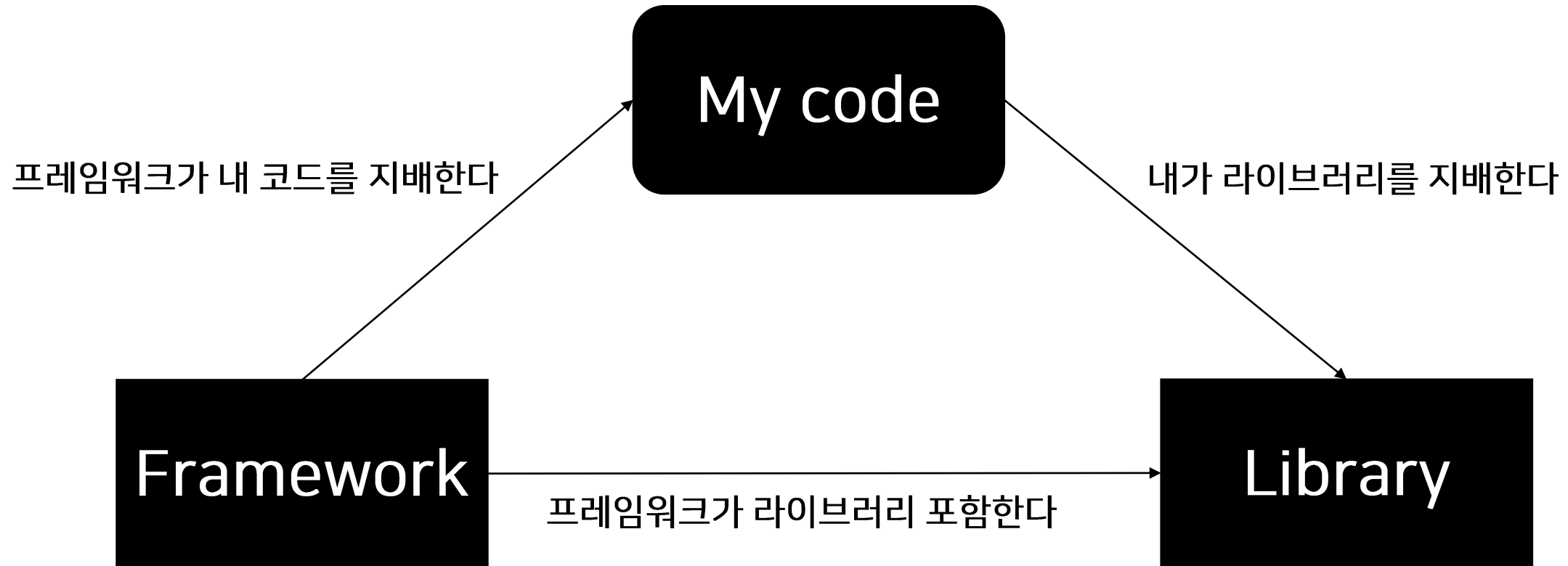
### Framework

- Framework controls the user
- Backbone of code
- 집으로 치자면 집 짓기 전 건설사, 제한된 선택과 건설사의 규정을 따라야함

### Library

- User controls the libraries
- Additional tools for code
- 집으로 치자면 이미 완성된 집에 가구, 장치 등을 추가하는 것, 오로지 내 마음과 내 의지대로 실행 가능

## Unit 01 | Frameworks and Libraries



# Contents

Unit 01	Frameworks and Libraries
Unit 02	Types of frameworks and libraries
Unit 03	Sklearn, TensorFlow, Pytorch
Unit 04	Coding Examples

## Unit 02 | Types of frameworks and libraries

Most popular frameworks in Python

- Django
- Web2Py
- Flask
- Bottle
- CherryPy

The Django logo, featuring the word "django" in a dark green, lowercase, sans-serif font.The Flask logo, which includes a small, detailed illustration of a flask on the left and the word "Flask" in a black, serif font on the right.The Web2Py logo, with "WEB" in black, "2" in a light gray, and "PY" in black, all in a bold, uppercase, sans-serif font.The Bottle logo, featuring the word "Bottle" in a bold, black, sans-serif font, with the letter "o" replaced by a stylized illustration of a flask containing blue liquid and bubbles.The CherryPy logo, which consists of a red cherry with a green leaf on the left, followed by the text "CherryPy" in a black, sans-serif font, with "Py" in red, all contained within a yellow rounded rectangle.

## Unit 02 | Types of frameworks and libraries

Popular libraries in Python (data 와 AI 위주로):

- os
- Numpy
- Pandas
- Matplotlib
- Scipy
- sklearn
- Tensorflow
- Pytorch



pandas



NumPy

matplotlib



SciPy



TensorFlow



PyTorch





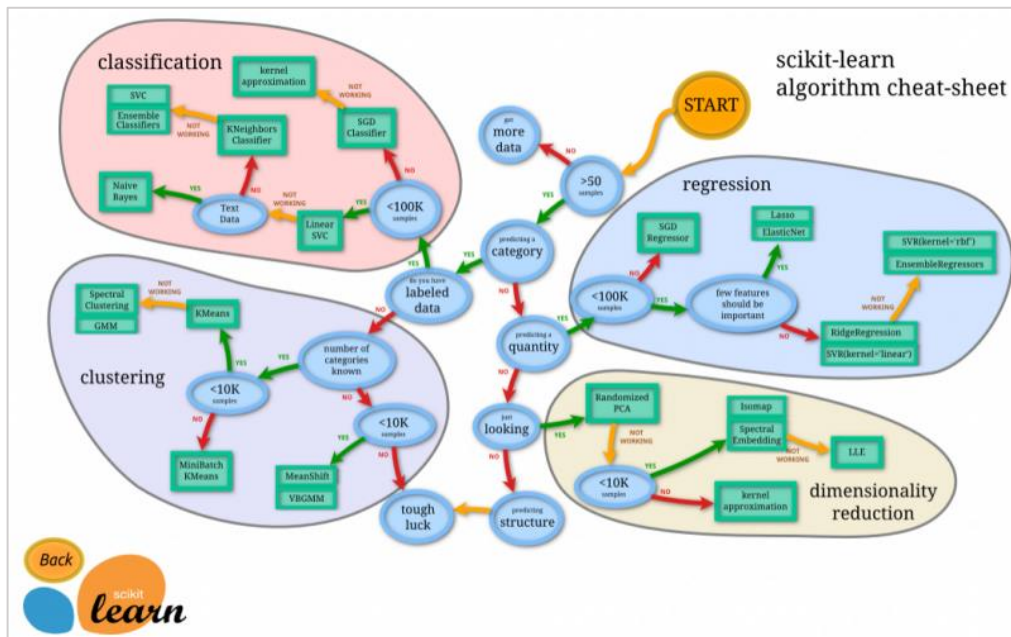
# Contents

Unit 01	Frameworks and Libraries
Unit 02	Types of frameworks and libraries
Unit 03	Sklearn, TensorFlow, Pytorch
Unit 04	Coding Examples

## Unit 03 | Sklearn, Tensorflow, PyTorch

## sklearn

- Scikit-learn(사이킷런) 의 약자 sklearn
- 머신러닝 및 데이터 전처리할 때 자주 사용하는 라이브러리
- 회귀, 분류, 군집화, 차원축소 등 다양하고 검증이 된 기능들을 가지고 있다.
- 머신러닝, 인공지능을 사용할 때 자주 마주치는 라이브러리.



```
X_train, X_test, Y_train, Y_test = train_test_split(X,Y,test_size=0.2, random_state=42)
RF_model = RandomForestRegressor(n_estimators=2000, random_state=42)
RF_model.fit(X_train,Y_train)
```

```
X_train, X_test, Y_train, Y_test = train_test_split(X,Y,test_size=0.2, random_state=41)
XGB_model = xgboost.XGBRegressor(n_estimators=2000, random_state=41)
XGB_model.fit(X_train,Y_train)
```

## Unit 03 | Sklearn, Tensorflow, PyTorch

### Tensorflow

- Tensorflow(텐서플로우) 는 파이썬 딥러닝에 특화된 라이브러리.
- 구글이 2011년에 개발을 시작해 2015년에 오픈 소스로 공개한 딥러닝 학습 라이브러리.



### TensorFlow를 사용 해야 하는 이유

전문가이든 초보자이든, TensorFlow는 ML 모델을 쉽게 빌드 및 배포할 수 있게 해주는 엔드 투 엔드 플랫폼입니다.

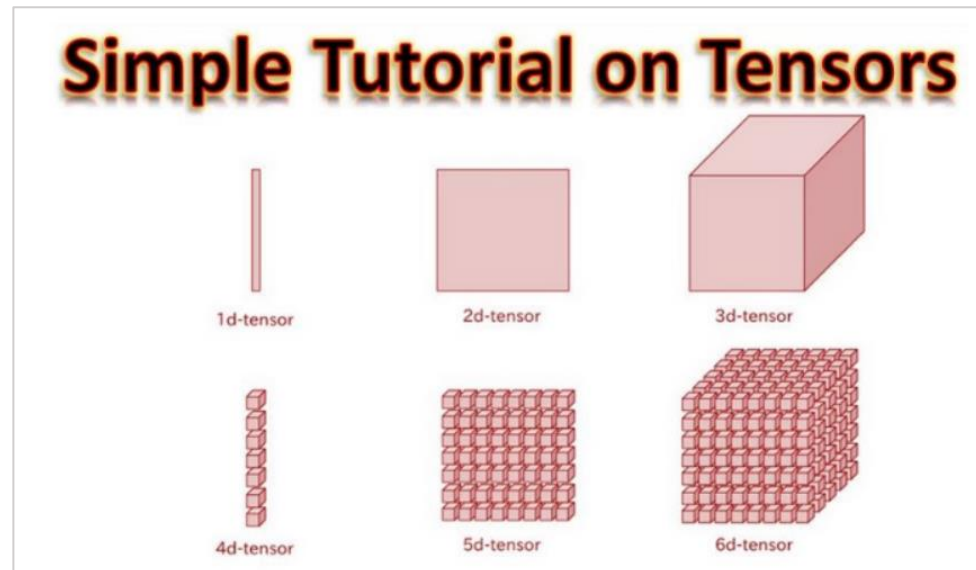
[동영상 보기](#)  [우수사례](#) 



## Unit 03 | Sklearn, Tensorflow, PyTorch

## Tensor?

- Tensor은 수학, 물리 등에서 다양하게 쓰이는 용어.
- 컴퓨터 과학에서는 numpy array (행렬, matrix) 와 많이 유사하지만, tensor 은 GPU 에서 병렬학습이 가능하다는 차이가 있다.
- 효율적인 딥러닝 학습을 위해서는 tensor 을 이용한 병렬학습이 강제되기 때문에, 딥러닝을 한다면 tensor 으로의 변환은 필수적이다.



## Unit 03 | Sklearn, Tensorflow, PyTorch

### PyTorch

- PyTorch 도 Tensorflow 와 같이 딥러닝 학습에 특화된 라이브러리
- 페이스북 인공지능 연구팀이 개발
- Autoencoder, CNN, RNN, LSTM 등 다양하고 폭넓은 딥러닝 패키지 제공



FROM  
RESEARCH TO  
PRODUCTION

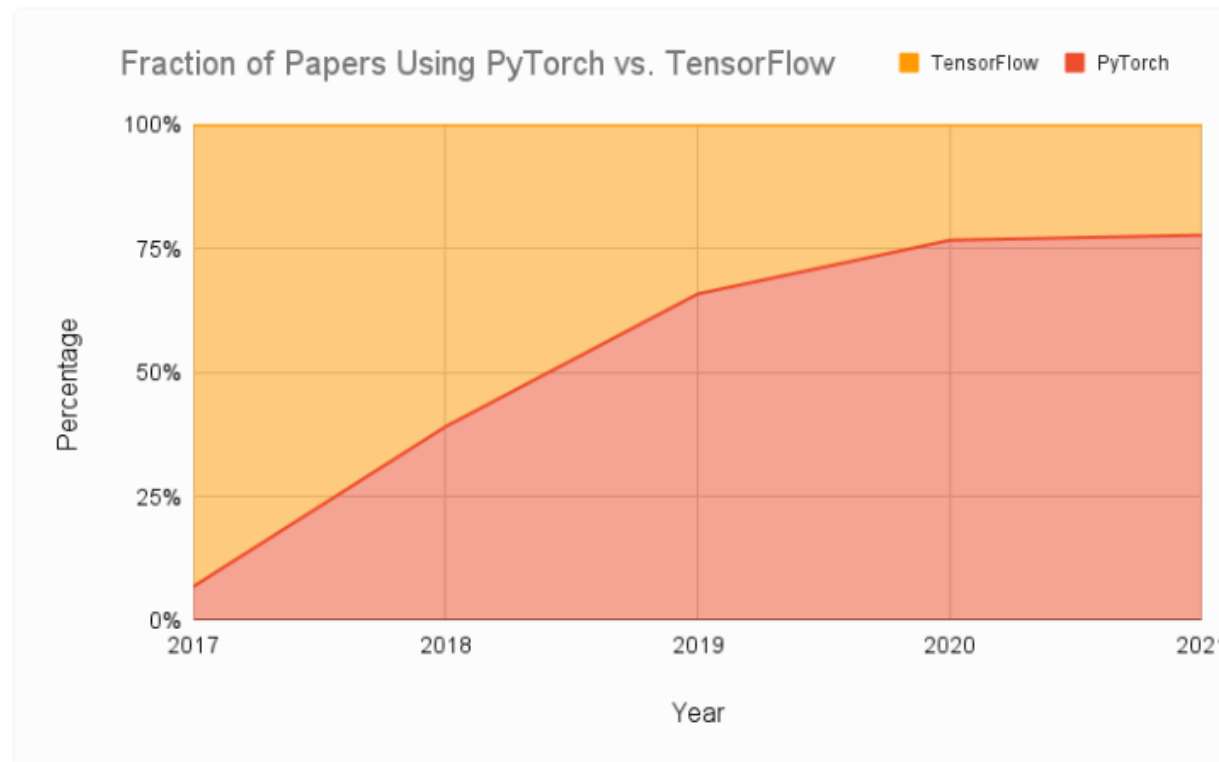
An open source machine learning framework that accelerates the path from research prototyping to production deployment.

Install >

## Unit 03 | Sklearn, Tensorflow, PyTorch

Tensorflow  
VS  
PyTorch

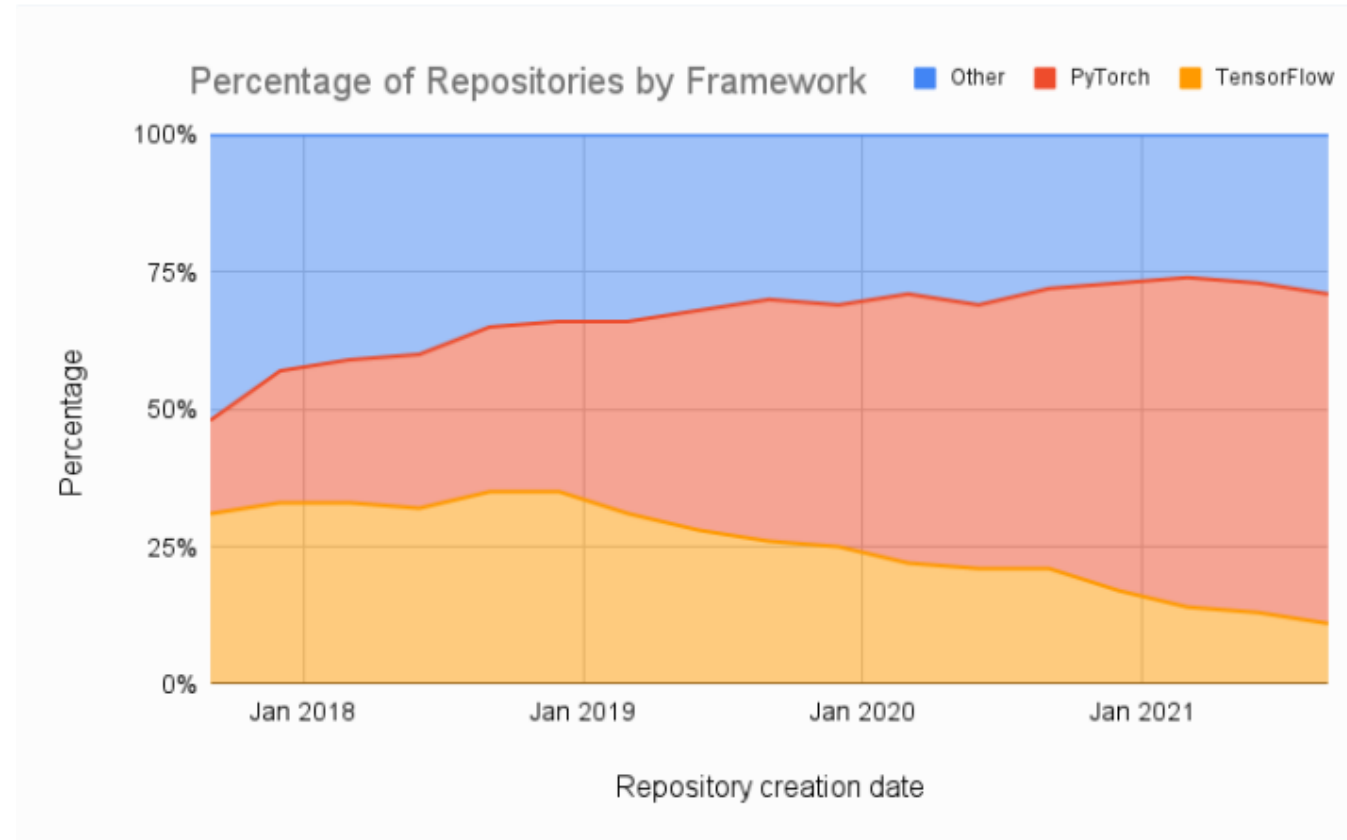
- PyTorch 와 Tensorflow 은 둘 다 대표적인 딥러닝 라이브러리이다.
- 대부분의 연구는 Tensorflow 에서 PyTorch의 이용으로 옮겨가고 있다.
- 아직 많은 산업들은 Tensorflow 사용하고 있다. (연구실보다 적응이 느리기 때문)



## Unit 03 | Sklearn, Tensorflow, PyTorch

### Tensorflow VS PyTorch

- 둘 중에 하나 골라서 사용하면 됨.



# Contents

Unit 01	Frameworks and Libraries
Unit 02	Types of frameworks and libraries
Unit 03	Sklearn, TensorFlow, Pytorch
Unit 04	Coding Examples

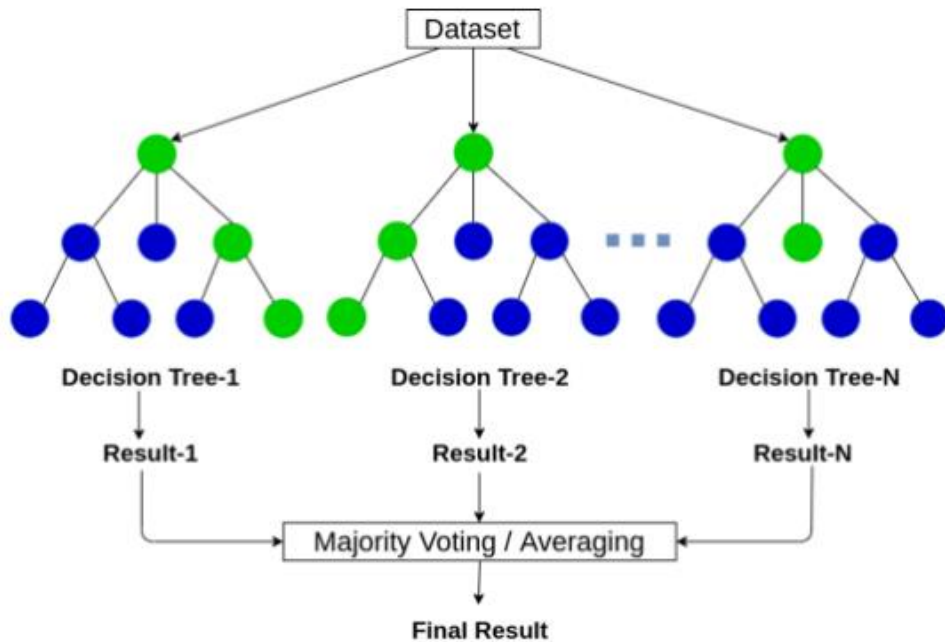


## Unit 04 | Coding Examples

sklearn

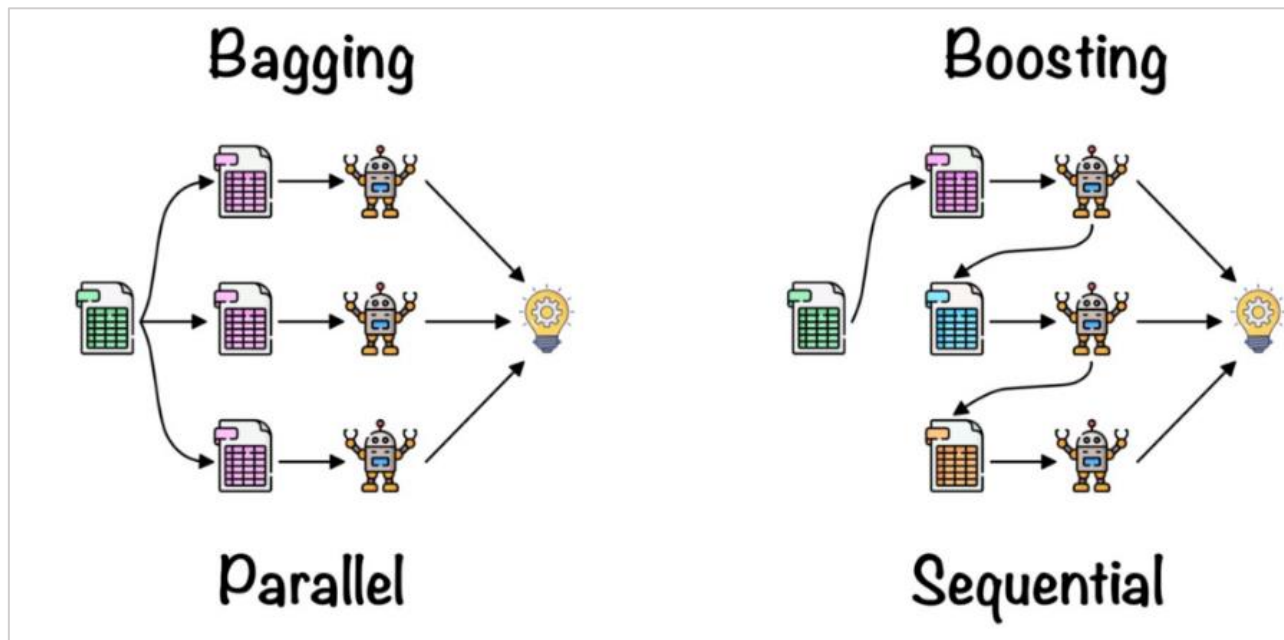
## Random Forest 알고리즘 이용.

- Random Forest 은 이전의 ML 알고리즘의 주를 이루었던 Least Square Regression 방식에서 벗어난 Tree 를 이용한 앙상블(Ensemble) 기법.
- 머신러닝 알고리즘 중에서 인기 많은 모델이며, 과적합 (Bagging 의 특징) 및 이상치(Tree 의 특징)에 강하다는 특징이 있다.
- 트리를 대량으로 생산하기 때문에 시간이 비교적 오래 걸린다는 단점이 있다. (XGBoost 도 오래 걸리긴 한데)



## Unit 04 | Coding Examples

sklearn



Ensemble Method.  
Bagging and Boosting

Bagging(bootstrap aggregating)  
-> 다수결  
-> Random Forest

Boosting  
-> 다수 적응  
-> AdaBoost, XGBoost

## Unit 04 | Coding Examples

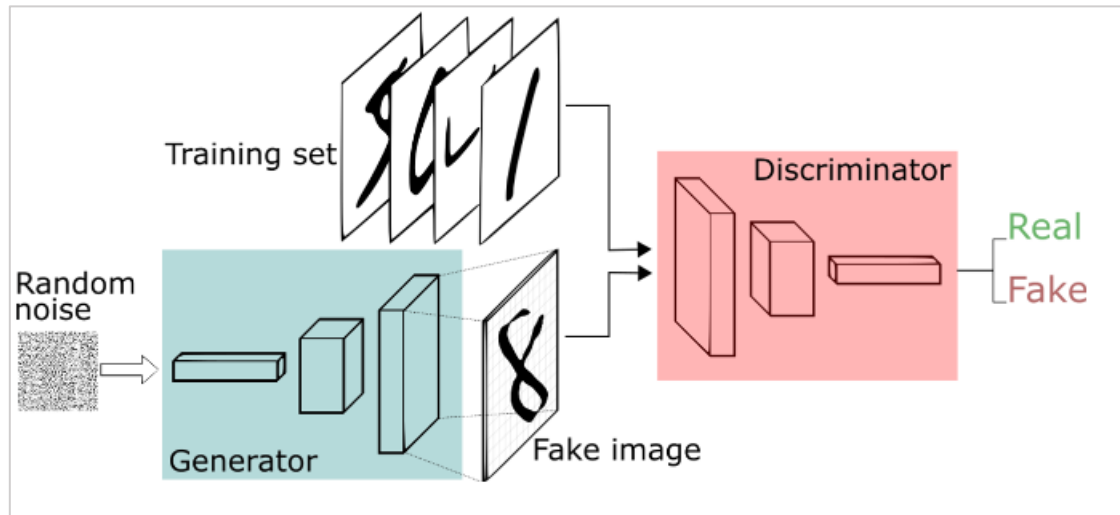
sklearn

구글 코랩 공유

<https://colab.research.google.com/drive/1W92rbalOlhqkrjoxVwKMb5XRkUHgpKiV?usp=sharing>

## Unit 04 | Coding Examples

## PyTorch

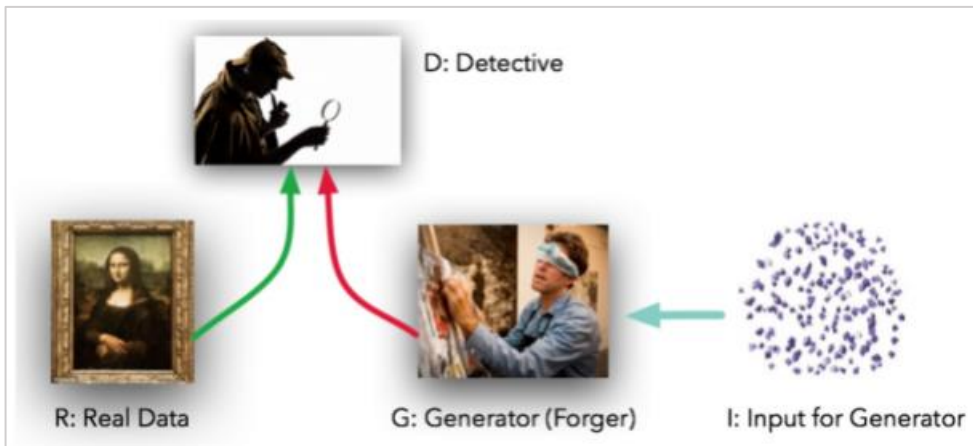
Deep Learning Model  
GAN ( Generative Adversarial Network )

- 적대적 생성 네트워크로, 딥러닝 비전에서 시작해 Sound Generation, Text Generation 등으로 넓게 응용되고 있는 모델
- Generator 와 Discriminator 의 min-max loss 으로 학습을 진행
- Diffusion model 이라고 업그레이드 모델도 hot함

## Unit 04 | Coding Examples

## PyTorch

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$



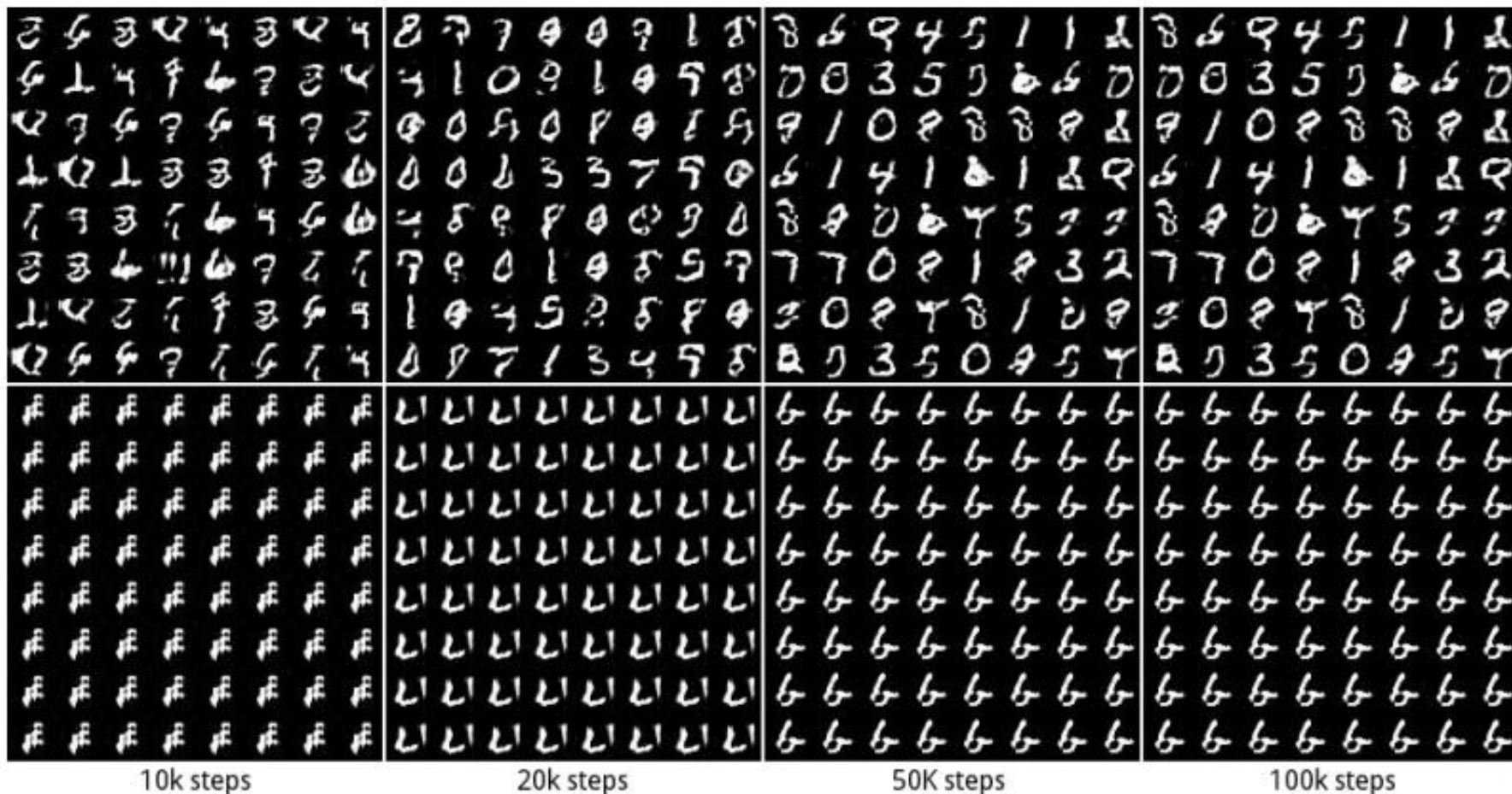
- 딥러닝 모델에서 Loss Function 이 매우 중요한 역할을 함
- GAN 에서 Generator 은 loss 을 최대화 하려고 하고, Discriminator 은 loss 을 최소화 하려고 한다.
- G와 D 가 상호보완적으로 학습을 해야 한다.

## Unit 04 | Coding Examples

## PyTorch

Mode Collapse 예시

Discriminator 이 제한적인 값들만 내는데  
Generator 은 이를 짚고 있어서 더 이상 학습이 진행되지 않음.





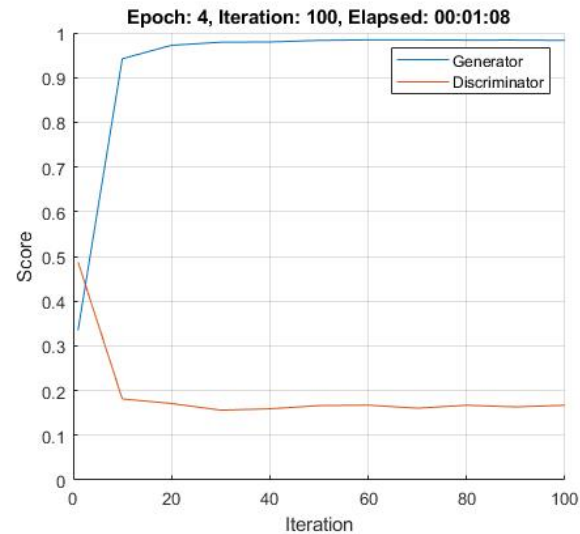
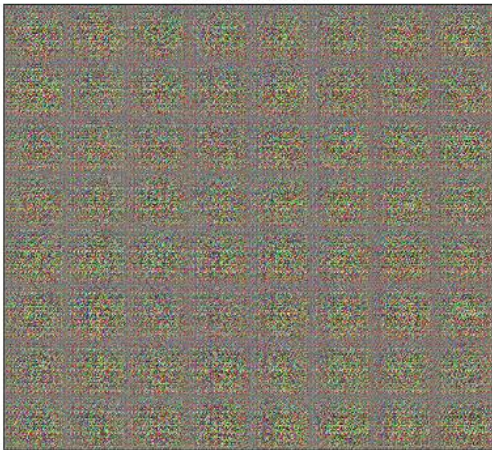
## Unit 04 | Coding Examples

## PyTorch

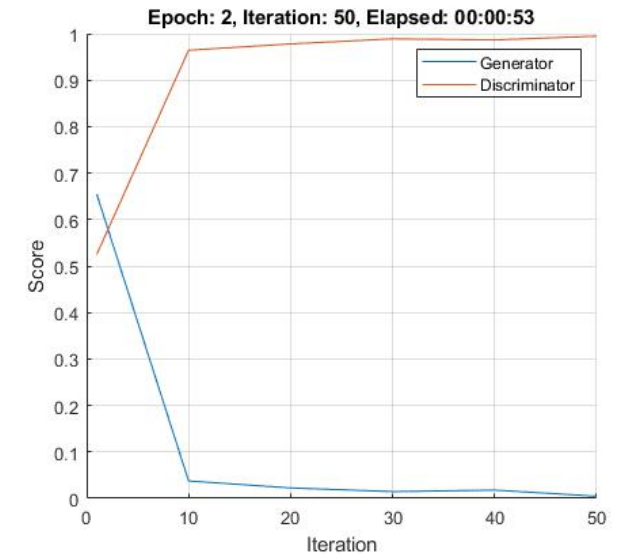
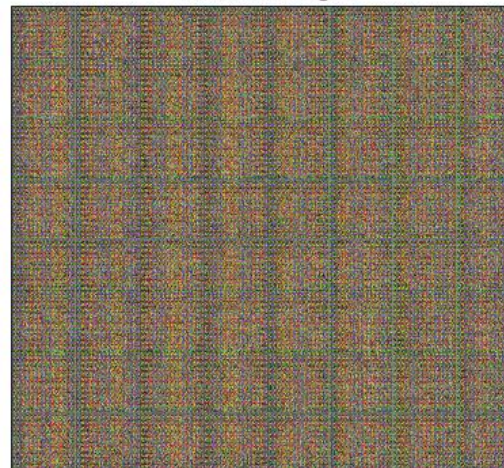
## Random Collapse 예시

Discriminator 또는 Generator 이 상대를 초반에 압도해서 fail 이 되는 경우. 둘 다 학습이 더 이상 되지 않기 때문에 초반의 랜덤한 값들이 최종 값으로 생성된다.

Generated Images



Generated Images



## Unit 04 | Coding Examples

PyTorch

구글 코랩 공유

<https://colab.research.google.com/drive/1IYVIUIGsUzWWoBoLPhzC9VbnWGA70tcG?usp=sharing>



## Homework

'Week1\_Framework.ipynb' 채우기.

1. Library 와 Framework 의 차이 간단하게 서술하기. (100자 내외)
2. 딥러닝과 머신러닝의 관계 및 특징, 차이 간단하게 서술하기. (200자 내외)
3. ipynb 코드 보고 이해하고 주석 달기.

참고

<https://volcano-screen-ada.notion.site/5de439afd31f41509bf2a0fcd3322544>

감사합니다

강의 관련 문의, 과제 관련 문의, ToBig's 관련 문의, 그냥 심심하다,  
010-8746-7447(김성우) 또는 카톡으로 해주세요!

ppt 제작 : 김성우  
ppt 테마 : 투빅스 정규세션