▶ 고려대학교 의료정보학과



의료인공지능

고려대학교 의료빅데이터연구소 채민수(minsuchae@korea.ac.kr)

• 참고문헌

- Do it! 점프 투 파이썬,박응용 저, 이지스퍼블리싱
- 파이썬 스킬 업 기초를 넘어서, 파이썬을 파이썬답게 사용하자!, 브라이언 오버랜드, 존 베넷 저/조인석 역, 길벗
- 혼자 공부하는 머신러닝+딥러닝 1:1 과외하듯 배우는 인공지능 자습서, 박해선 저, 한빛미디어
- 핸즈온 머신러닝 사이킷런, 케라스, 텐서플로 2를 활용한 머신러닝, 딥러 닝 완벽 실무, 오렐리앙 제롱 저/박해선 역, 한빛미디어



• 평가방법

- 과제 : 30%

- 중간고사: 30%

- 기말고사 : 30%

- 출석 : 10%



- 텀프로젝트 과제(20%)
 - 주제 : 자유(되도록 의료 관련)
 - 데이터셋:
 - https://aihub.or.kr (AI-Hub)
 - ▶ https://data.go.kr (공공데이터포털)
 - ▶ https://opendata.hira.or.kr (건강보험심사평가원)
 - ▶ https://nhiss.nhis.or.kr (국민건강보험공단)
 - ➤ https://cdw.kumc.or.kr/ (고려대학교 CDW)
 - https://github.com/MinsuChae/MedicalAI/tree/master/dataset
 - ➤ https://github.com/imdatalab/datasets

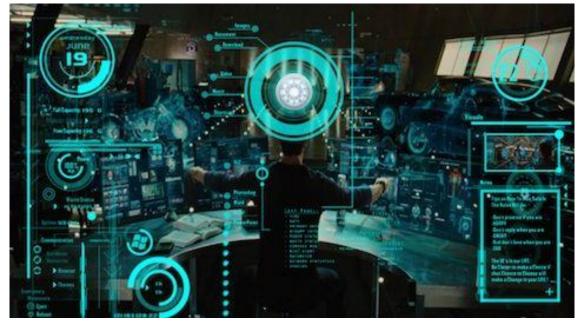




- 텀프로젝트 과제(20%)
 - 발표 시 포함해야 할 항목
 - ▶주제
 - ▶수행한 데이터 전처리
 - ▶사용한 알고리즘 소개
 - ▶수행 결과
 - ▶ 10분 이상 발표



- 인공지능(Artificial Intelligence, AI)
 - 인간의 지능을 모방하여 스스로 컴퓨터가 행동하는 기술



Reference: https://zdnet.co.kr/view/?no=20130702093041



◦ 인공지능(Artificial Intelligence, AI)



Reference: https://ko.gta5-mods.com/scripts/hands-up



Reference:

https://m.post.naver.com/viewer/postView.naver?volumeNo=25823027&memberNo=4704108



• 과거의 인공지능



Reference: https://samulgoongi.com/4517



Reference:

https://www.samsung.com/sec/cookingappliances/MS23K3535AWd2c2/MS23K3535AK/

cs/latest/ko/kakaona - 키워드 : 패턴 인식, 휴리스틱 알고리즘, 전문가 시스템



Reference:

https://developers.kakao.com/do cs/latest/ko/kakaonavi/common

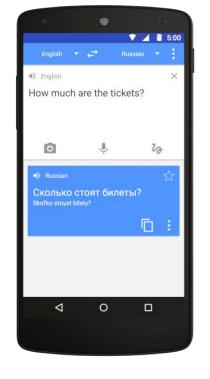


• 현재의 인공지능



1 #!/wsr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 asy
8
9
10
11
12
13
14
15
16
17
Powered by

© OpenAI



Reference: https://clova.ai/

Reference: https://copilot.github.com/

Reference:

https://translate.google.com/intl/
ko/about/

- 키워드: 머신러닝, 딥러닝, 자연어 처리, 이미지 분석, 음성 인식, 강화학습



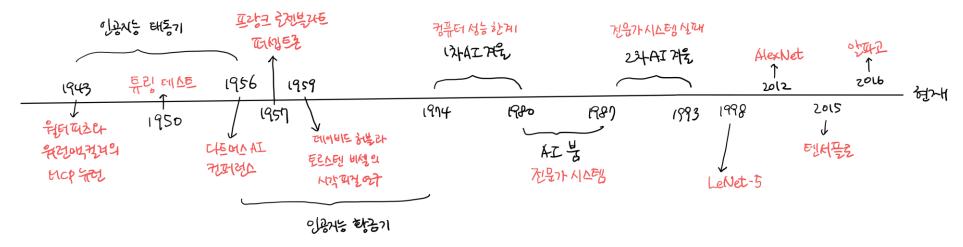
• 근미래의 인공지능





Reference : https://www.tesla.com/ko_KR/autopilot

• 인공지능의 역사



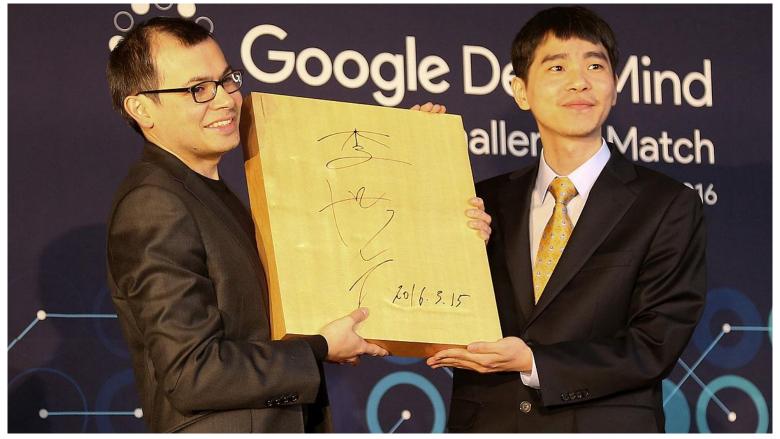
Reference : 혼자 공부하는 머신러닝+딥러닝 1:1 과외하듯 배우는 인공지능 자습서, 박해선 저, 한빛미디어



- AI의 암흑기
 - 1969년 마빈 민스키(Marvin Lee Minsky)와 시모어 페퍼트(Seymour Papert) 가 퍼셉트론(Perceptron)의 이론적 한계점을 증명
 - 퍼셉트론으로 XOR 연산 불가능으로 인해 인공지능 기술에 대한 의구심
- 급격한 인공지능 기술 도래
 - 컴퓨터비전 분야
 - ▶ 1998년 Yann Lecun이 제안한 신경망 모델로 손글씨를 인식함(LeNet-5)
 - ▶2012년 Geoffrey Hinton 팀이 제안한 ImageNet이 기존 알고리즘보다 우수한 성능을 보임
 - 일반인
 - ▶ 2016년 구글의 알파고가 이세돌 9단을 바둑으로 이김



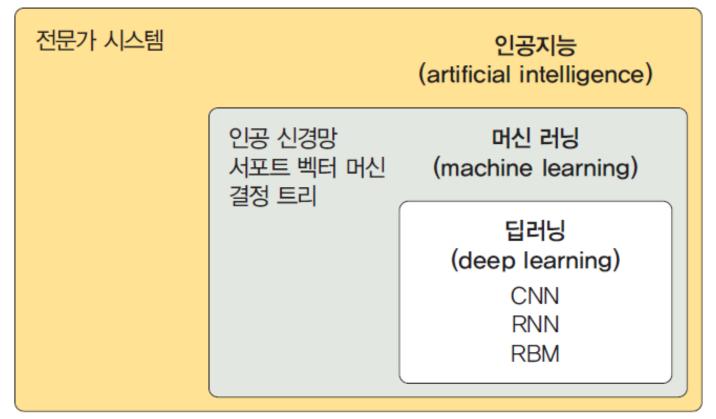
• 급격한 인공지능 기술 도래



Reference: https://deepmind.com/research/case-studies/alphago-the-story-so-far



• 인공지능 범주



Reference : 기초부터 CNN, RNN, 시계열 분석, 성능 최적화, 자연어 처리, 강화 학습, 생성 모델까지! 딥러닝 텐서플로 교과서, 서지영 저, 길벗



• 의료 분야의 인공지능 - MEWS(Modified Early Warning Score)

Table 1. Modified Early Warning Score

Score	3	2	1	0	1	2	3
SBP, mm Hg	≤70	71–80	81–100	101–199		≥200	
HR, beats/min		≤40	41–50	51–100	101–110	111–129	≥130
RR, breaths/min		≤8		9–14	15–20	21–29	≥30
BT, °C		≤35		35.1–38.4		≥38.5	
Mental status				Alert	Reacting to voice	Reacting to pain	Unresponsive

BT indicates body temperature; HR, heart rate; RR, respiratory rate; SBP, systolic blood pressure.

Reference: Joon-myoung, Kwon, et al. "An algorithm based on deep learning for predicting in-hospital cardiac arrest." Journal of the American Heart Association 7.13 (2018): e008678.



• 의료 분야의 인공지능 - DEWS

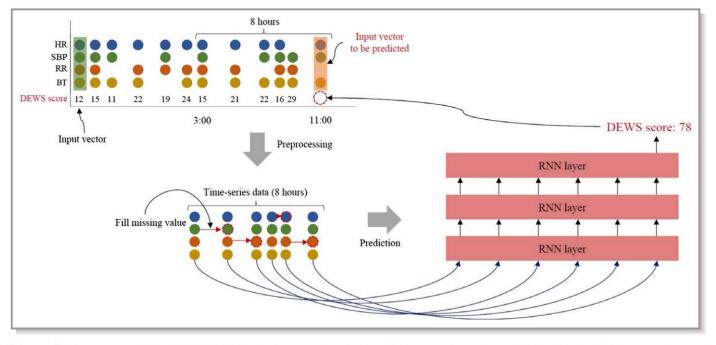


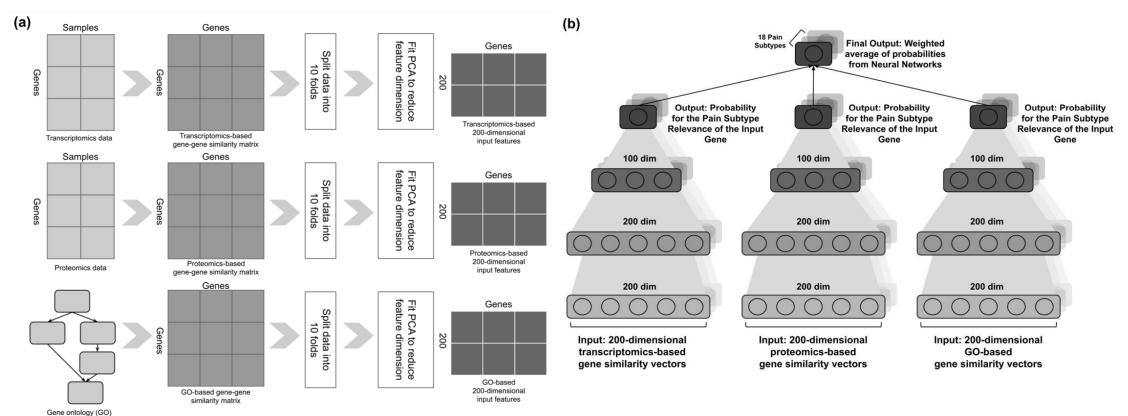
Figure 1. The process of the DEWS. DEWS indicates deep learning—based early warning system; HR, heart rate; RNN, recurrent neural network; RR, respiratory rate; SBP, systolic blood pressure; BT, body temperature.

Reference: Joon-myoung, Kwon, et al. "An algorithm based on deep learning for predicting in-hospital cardiac arrest." Journal of the American Heart Association 7.13 (2018): e008678.



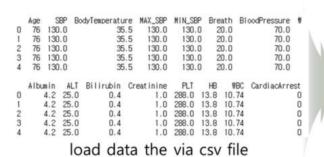
• 의료 분야의 인공지능

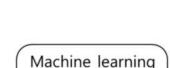
Intelligent Medical Data Lab



Reference: Minji, Jeon, et al. "Prioritizing pain-associated targets with machine learning." Biochemistry 60.18 (2021): 1430-1446.

• 의료 분야의 인공지능





fit and verify.

Test via test

dataset

patient 1 1h data
patient 1 2h data
:
patient 1 72h data
patient 2 1h data
:
patient n 72h data
make
dataframe

patient dataframe

Perform SMOTE algorithm on the training dataset

patient record

patient 1 data

patient 2 data

patient 3 data

patient 4 data

patient 5 data

:

patient n data

group by each

of 72 records

Split Training and test dataset in shallow machine learning. Split Training , validation and test dataset in deep learning

split dataset

data sliced by the values of past patient data and input parameters and target divided

Reference: Minsu, Chae, et al. "Machine Learning-Based Cardiac Arrest Prediction for Early Warning System." Mathematics 10.12 (2022): 2049.

- 프로그래밍 언어란
 - 사용자(Developer)가 컴퓨터(CPU)에게 어느 업무(Instruction)을 해야 할 지 알려주기 위한 언어

- 저수준 언어 : 컴퓨터가 이해하기 용이한 언어

- 고수준 언어: 사용자가 이해하기 용이한 언어

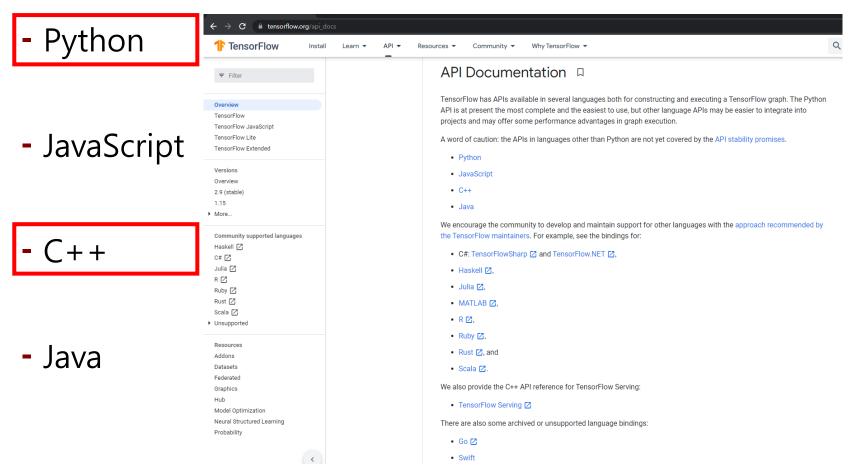


• 프로그래밍 언어 종류

- C, C++, C#
- Java, Scala, Kotlin
- Python
- MATLAB
- R
- Objective-C, Swift
- JavaScript, TypeScript
- Assembler
- Machine Code



• 어느 프로그램을 사용해야 하는가?



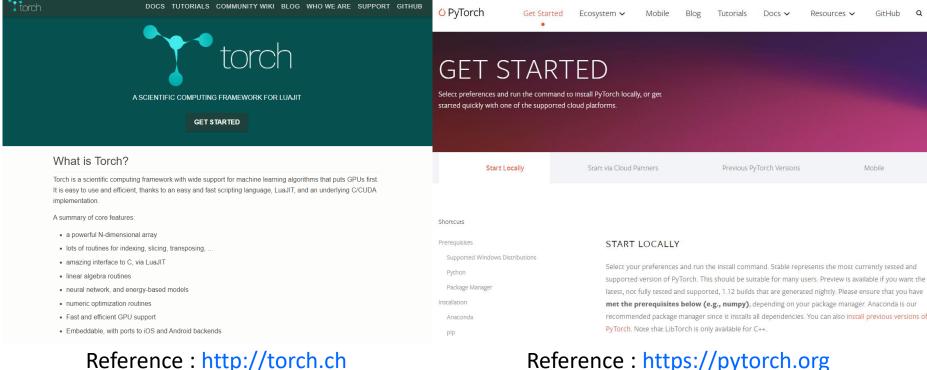


Reference: https://www.tensorflow.org/api_docs

• 어느 프로그램을 사용해야 하는가?

- Rua

Python



Reference : https://pytorch.org

Tutorials

Previous PyTorch Versions

Java



- C++의 장점
 - 일반적인 프로그램은 C++ 언어로 작성한 프로그램 코드(Source)가 빠름
 - 많은 머신러닝 라이브러리의 핵심(OpenCV, CUDA)은 C++로 개발됨

- C++의 단점
 - 환경 설정(라이브러리 설정)의 어려우며 상황에 따라 직접 구현해야 함 example) https://github.com/Neargye/hello tf c api/blob/master/src/create tensor.cpp
 - 일부를 수정하고 그 결과를 확인하기 위해서 전체를 수행해야 함(Compile)



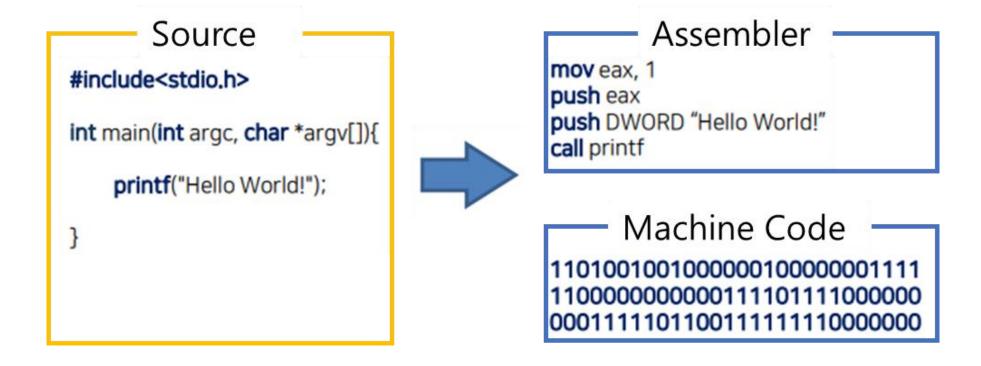
- Python의 장점
 - 라이브러리를 쉽게 설치 및 사용이 가능함

- C++ 기반 핵심 라이브러리 중 호환되는 라이브러리(Wrapper)가 있음
- 일부를 수정하고 그 결과를 확인하기 위해선 그 일부만 다시 수행하면 됨 (Interpreter)

- Python의 단점
 - C++에 비해 상대적으로 느림



• 컴파일(Compile)과 인터프리터(Interpreter)의 차이



Reference: https://velog.io/@chun_gil/%EC%BB%B4%ED%93%A8%ED%84%B0-%EA%B5%AC%EC%A1%B0-2.-%EA%B8%B0%EA%B3%84%EC%96%B4-%EC%96%B4%EC%96%B4%EC%96%B4%EC%96%B4-%EB%B8%94%EB%A6%AC%EC%96%B4-%EA%B3%A0%EA%B8%89%EC%96%B8%EC%96%B4-%EB%B9%84%EA%B5%90



- 프로그래밍 언어 문법이 있는 이유
 - 사용자(Developer)가 컴퓨터(CPU)에게 어느 업무(Instruction)을 해야 할 지 알려주기 위한 언어

- 저수준 언어 : 컴퓨터가 이해하기 용이한 언어

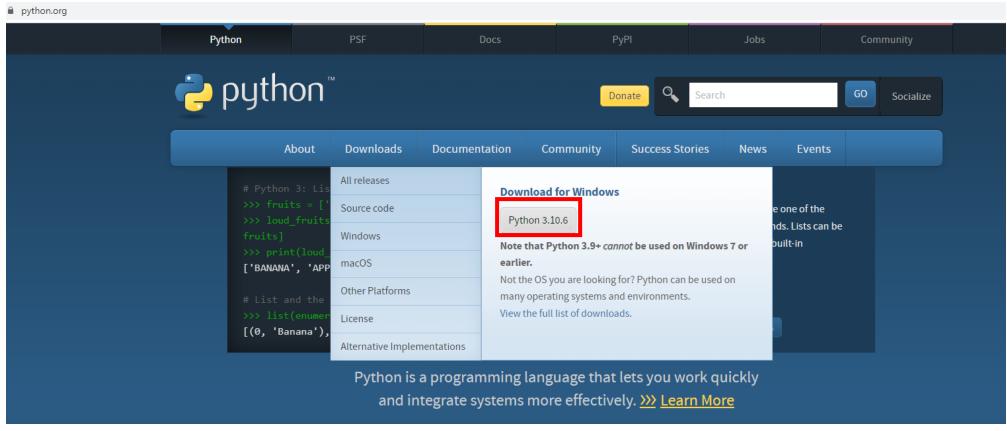
- 고수준 언어: 사용자가 이해하기 용이한 언어



- 윈도우에서의 설치
 - 파이썬 프로그램을 다운받아 설치
 - MS Store 통해 설치
 - Anaconda를 통한 설치 (recommend)
 - WSL 를 통한 설치 (recommend only who experience Linux)
- 리눅스, 맥에서의 설치
 - 프로그램 패키지를 통해 설치
 - Anaconda를 통한 설치 (recommend who use Apple Silicon)

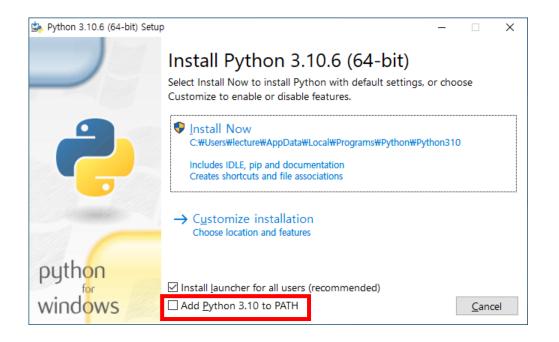


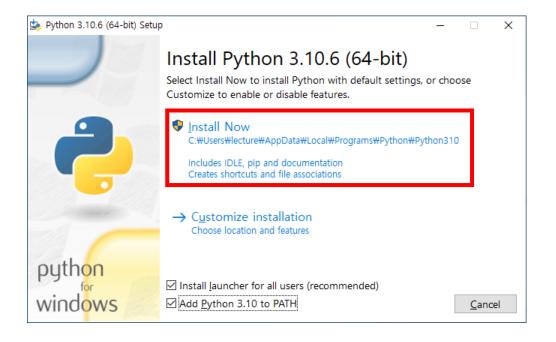
• 파이썬 프로그램을 다운받아 설치



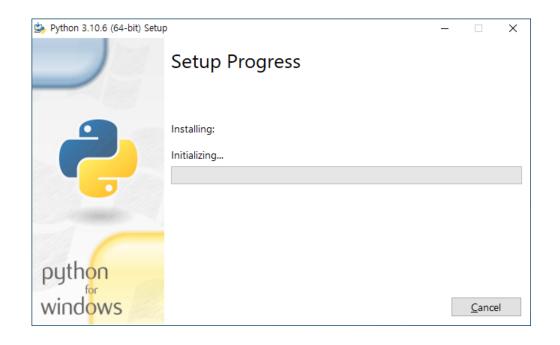
Reference: https://python.org

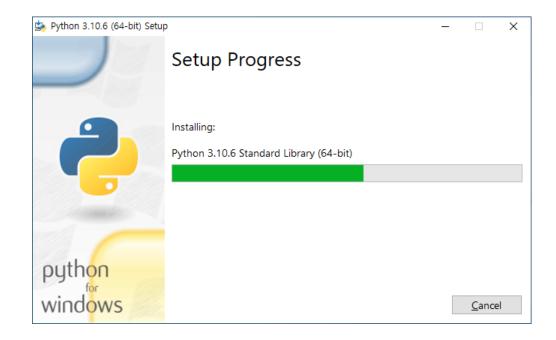




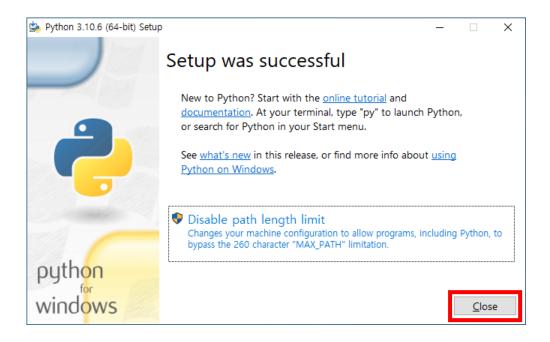




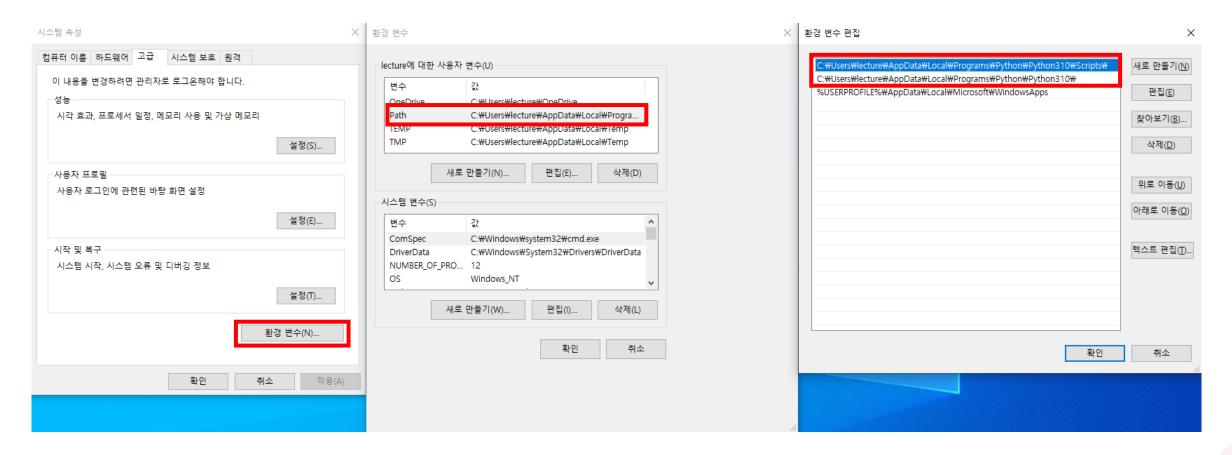






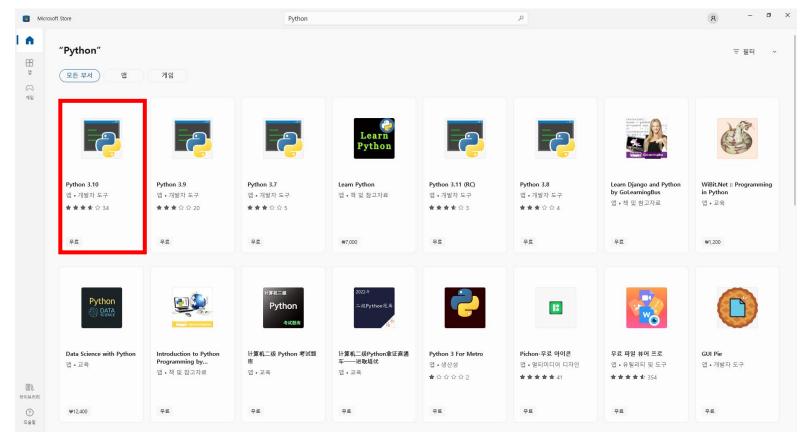






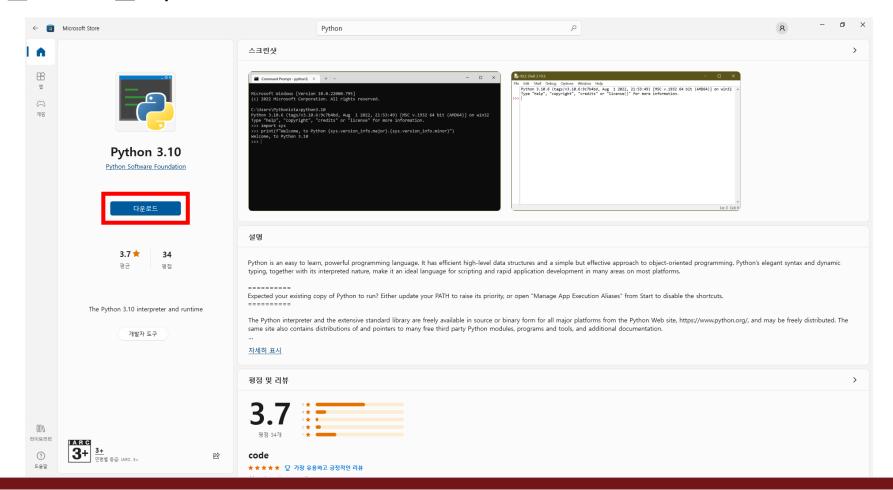


- MS Store 통해 설치
 - Python 검색 및 원하는 버전 클릭



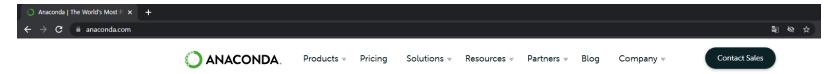


- MS Store 통해 설치
 - 다운로드 클릭





• Anaconda 설치



Data science technology for a better world.

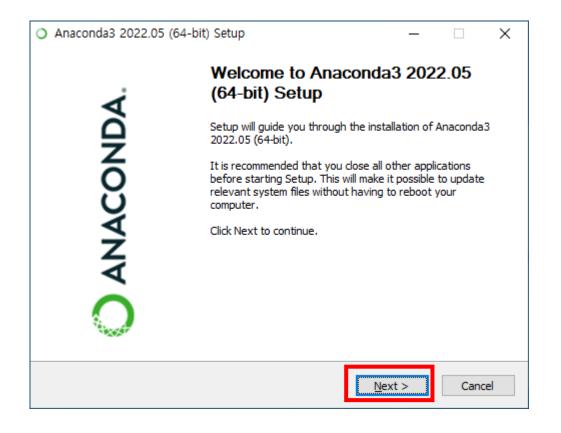
Anaconda offers the easiest way to perform Python/R data science and machine learning on a single machine. Start working with thousands of open-source packages and libraries today.

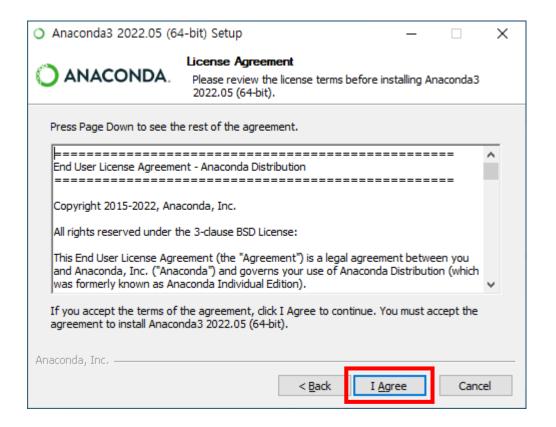


Reference : https://www.anaconda.com/

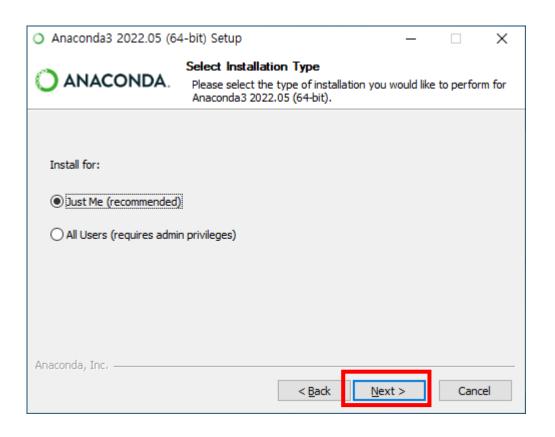


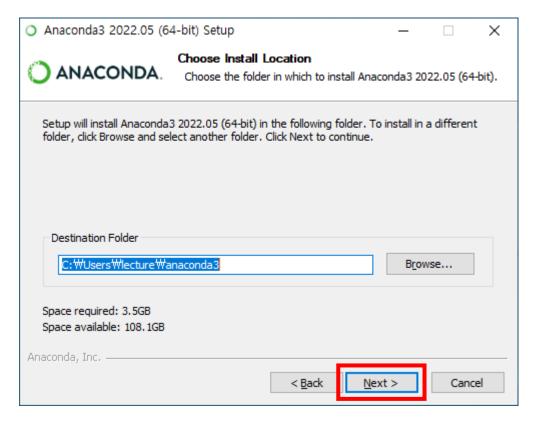
• Anaconda 설치



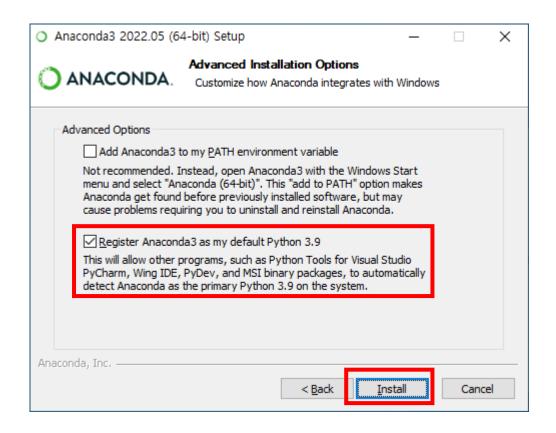


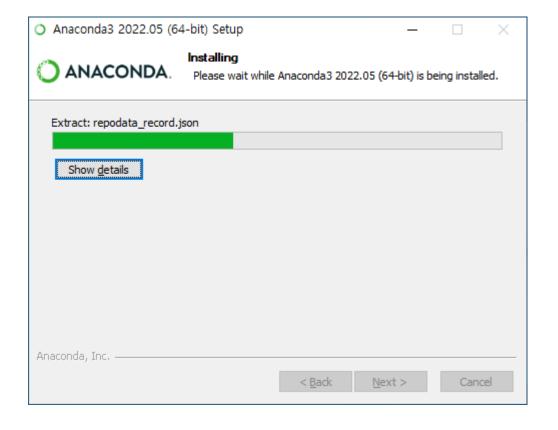




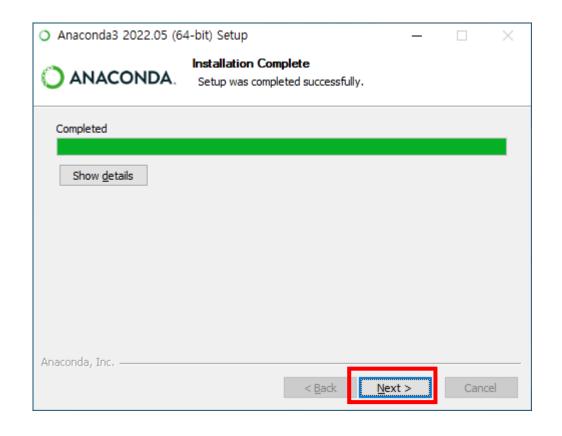


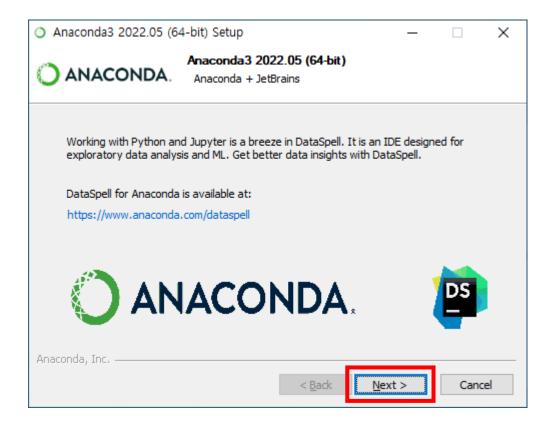




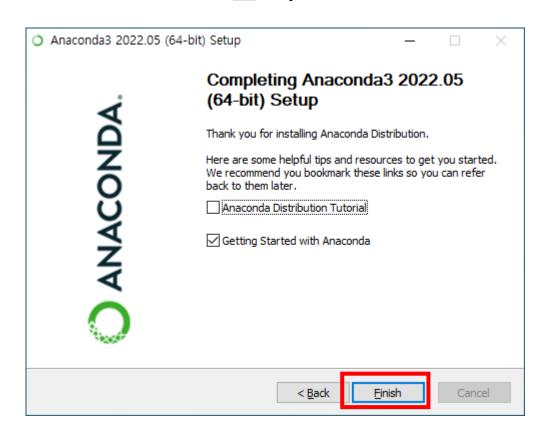


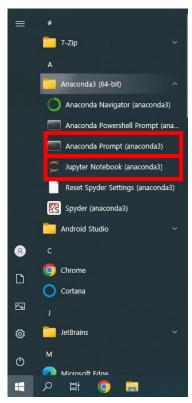


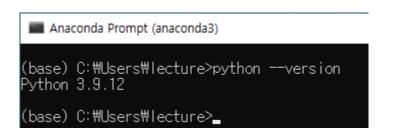














- WSL, 리눅스에서의 설치
 - WSL 설정 (윈도우)
 - python3, python-dev 설치
 - ubuntu sudo apt-get install -y python3 python3-pip

- 맥에서의 설치
 - 파이썬 프로그램을 받아 설치 https://www.python.org/downloads/macos/
 - brew install python3



- 파이썬 개발을 위한 통합개발환경(IDE, Integrated Development Environment)
 - IDLE (파이썬 설치 시 내장 프로그램)

notepad++, sublime text

- vscode, pycharm (recommend for traditional IDE)
- Jupyter Notebook or Jupyter Notebook 기반 IDE (recommend)



- 파이썬의 개요
 - 귀도 반 로섬(Guido van Rossum)이 1991년에 발표한 프로그래밍 언어
 - 파이썬2와 파이썬3이 존재함

- 다이나믹 타입

- 인터프리터

- 독특한 문법



◦ 머신러닝을 위한 IDE

- Colaboratory(Colab) - https://colab.research.google.com/

- Jupyter notebook
 - ▶ pip install jupyter



• 프로그래밍 언어에서의 문법의 필요성

- 저는 인간입니다.

- 나는 인간입니다.

- 인간입니다 나는.

- 인간입니다.



• 프로그래밍 언어에서의 문법의 필요성

$$- a = b$$

$$- b = a$$







• 파이썬에서 지원하는 머신러닝 프레임워크 및 라이브러리

- 텐서플로우(Tensorflow)

- 케라스(Keras)

- 토치(Pytorch)

- 사이킷런(Scikit-learn)



• 프레임워크와 라이브러리를 사용하는 이유

- 연산 속도(GPGPU, CUDA)

- 학습의 용이함(Learning Curve)

- 목표(Purpose)



• GPU 연산의 필요성





- 코랩의 장점
 - 무료(혹은 유료 구독서비스)
 - 13GB 메모리(유료 시 23GB(Pro), 51GB(Pro+))
 - K80 혹은 T4 혹은 P100 할당
 - Scikit-learn, Tensorflow 및 시각화 패키지 설치되어 있음
- 。단점
 - 라이브러리 버전의 변경 가능성
 - 최대 보장 시간 제한(무료 12시간, 유료 24시간)



。실습

print("Hello world!")

[실행결과] Hello world!





• 실습

예	최근 사용	Google Drive	GitHub	업로드
GitHub URL을 입력	하거나 조직 또는 사용자로 검		비공개 저장소 포함	
https://g	ithub.com/Mir	alAI	Q	
저장소: 🖸 MinsuChae/Medic	calAl 🔻	브랜치: 🗹 master 🕶		
경로				
2022-2/1주	차 실습.ipynb			Q Z
				새 노트 취소

