

Data Science(Machine Learning, Deep Learning)

<Deep Learning - Anaconda3 Navigator, Jupyter Notebook>

<Jupyter Notebook>

Google Colab

계산과 속도가 중요한 DataScience

Colab 시작 페이지

https://colab.research.google.com/?hl=ko#scrollTo=Nma_JWh-W-IF

.ipynb로 저장됨

일반 Python .py로 저장

[데이터 사이언스에 대한 접근]

$$f(x) = x + 1$$

f(x) = 변수

function = 기능이라는 의미

변수에 따라 값이 달라짐

$$f(x) = ax + b$$

x = 변수

a와 b에 대해서 알아내는 것 = 데이터 사이언스

미지수를 알아내는데 어떻게 알아내는가?

a와 b 상수

$$y = f(x)$$

인공지능(artificial intelligence, AI)은 컴퓨터 시스템을 사용하여 인간 지능과 유사한 작업을 수행하도록 설계된 기술과 시스템

DL, AI는 미지의 세계

데이터 시각화 시키는 것이 중요함.

Machine Learning의 WorkFlow

문제 정의 - 데이터 수집 - 데이터 전처리/분석 - 알고리즘 - 평가 - 배포

<파이토치(PyTorch)>

텐서(Tensor) - 선형대수

Tensor - 딥러닝에서 가장 기본이 되는 단위 중 하나

0차원 Scalar

1차원 Vector(벡터)

2차원 Matrix(행렬)

3차원 이상의 것을 Tensor로 의미

<파이토치(PyTorch)와 텐서플로(TensorFlow), 케라스(Keras)>

딥러닝에서 사용하는 프로그램(딥러닝 프레임워크)

딥러닝 프레임워크의 종류도 텐서플로(TensorFlow), 케라스(Keras), 테아노(Theano),

파이토치(Pytorch), CNTK 등 다양한 프레임워크가 있음.

메타 연구팀이 개발한 파이썬 기반 오픈 라이브러리 딥러닝 프레임워크 - 파이토치
구글이 개발한 오픈 딥러닝 프레임워크 라이브러리 - 텐서플로

AI > ML > DL

<데이터 전처리 - 캐글 (Kaggle)>