

# 강의 소개

머신러닝/딥러닝 기술이 획기적으로 발전하면서 데이터 분석 및 인공지능 관련 연구의 중요성이 사회, 경제, 산업의 거의 모든 분야에 지대한 영향을 미치고 있으며, 앞으로 그런 경향이 더욱 강화될 것으로 기대된다.

본 강의는 비전공자를 위한 머신러닝의 기본 아이디어와 다양한 활용법을 실전 예제와 함께 전달한다. 이와 더불어 파이썬 프로그래밍의 기초도 함께 학습한다.

# 데이터 과학, 인공지능, 머신러닝, 딥러닝

- 데이터 과학: 주어진 데이터로부터 수학과 통계 지식을 활용하여 필요한 정보를 추출하는 학문 분야
- 인공지능: 사고(thinking), 학습(learning) 등 인간의 지적능력을 컴퓨터를 통해 구현하는 기술 또는 해당 연구 분야
- 머신러닝: 컴퓨터가 데이터로부터 스스로 정보를 추출하는 기법 또는 해당 연구 분야.
- 딥러닝: 심층 신경망 이론을 기반으로 복잡한 비선형 문제를 해결하는 머신러닝 기법 또는 해당 연구 분야



그림 출처: 교보문고: 에이지 오브 머신러닝

역사적 관점에서 바라본 인공지능, 머신러닝, 딥러닝의 관계는 다음과 같다.

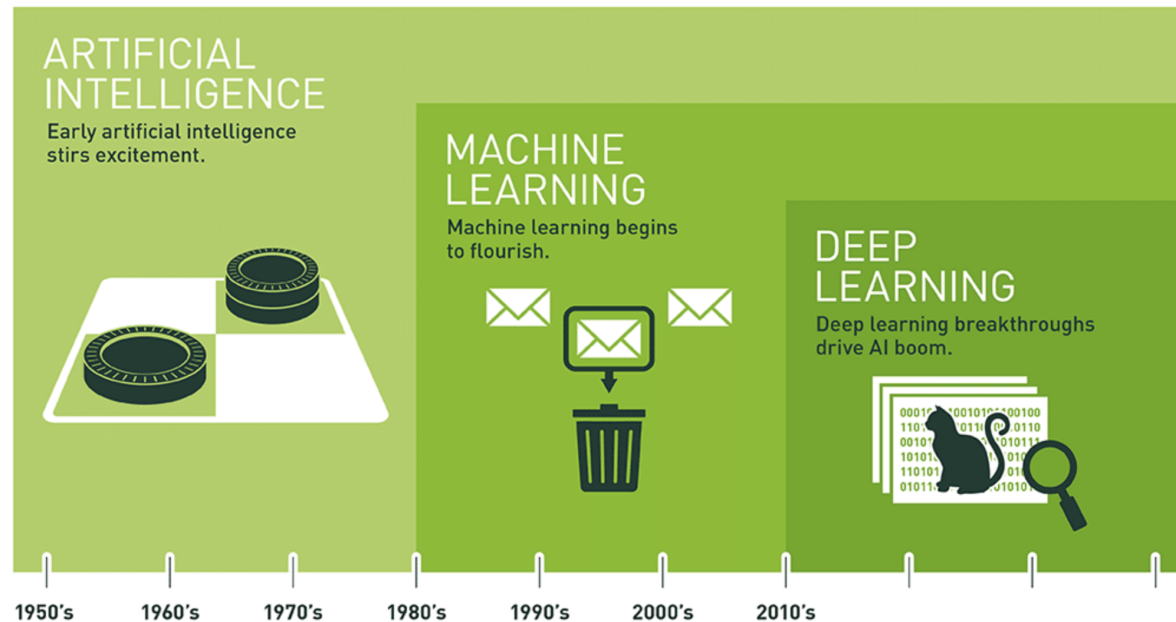


그림 출처: [NVIDIA 블로그](#)

# 파이썬 데이터 분석

파이썬 데이터 분석은 데이터 과학의 주요 연구 도구이며 머신러닝 분야에서 가장 많이 활용된다. 파이썬 데이터 분석 학습을 위해 아래 분야의 기초지식이 요구된다.

- 파이썬 프로그래밍
- 미적분학, 선형대수, 확률과통계

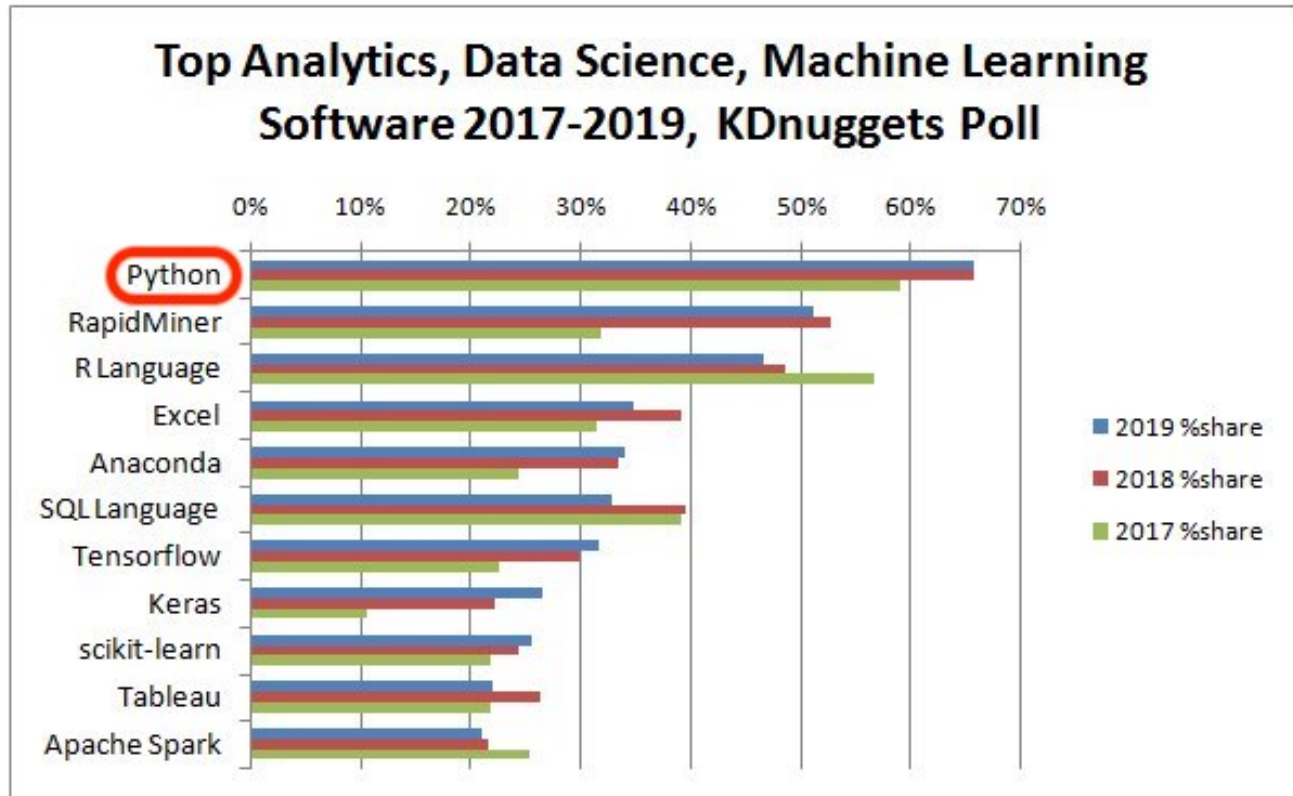
# 파이썬(Python) 언어 인기도

프로그래밍 실습에 사용되는 파이썬(Python)은 현재 데이터 분석 및 머신러닝 분야에서 가장 많이 사용되는 프로그래밍언어이다. TIOBE Index 2021년 2월 기준 세 번째로 가장 많이 사용되는 프로그래밍 언어이며 점유율이 점점 높아지고 있다.

Feb 2021	Feb 2020	Change	Programming Language	Ratings	Change
1	2	▲	C	16.34%	-0.43%
2	1	▼	Java	11.29%	-6.07%
3	3		Python	10.86%	+1.52%
4	4		C++	6.88%	+0.71%
5	5		C#	4.44%	-1.48%
6	6		Visual Basic	4.33%	-1.53%
7	7		JavaScript	2.27%	+0.21%
8	8		PHP	1.75%	-0.27%
9	9		SQL	1.72%	+0.20%
10	12	▲	Assembly language	1.65%	+0.54%

그림출처: [TIOBE Index](#)

데이터 과학 분야로 한정할 경우 파이썬의 인기도가 가장 높다.



그림출처: [KDnuggets](#)

## 언어 특성

파이썬이 데이터 과학 분야에서 인기가 높은 이유는 다음과 같다.

- 범용 프로그래밍언어
- R, 매트랩, SQL, 엑셀 등 특정 분야에서 유용하게 사용되는 언어들의 기능 지원
- 데이터 적재, 시각화, 통계, 자연어 처리, 이미지 처리 등에 필요한 라이브러리 제공
- 머신러닝, 데이터 분석 등에 활용될 수 있는 라이브러리 및 도구의 지속적 개발
  - [Anaconda](#)
  - [TensorFlow](#)
  - [scikit-learn](#)
  - [Keras](#)
  - [PyTorch](#)
  - ...