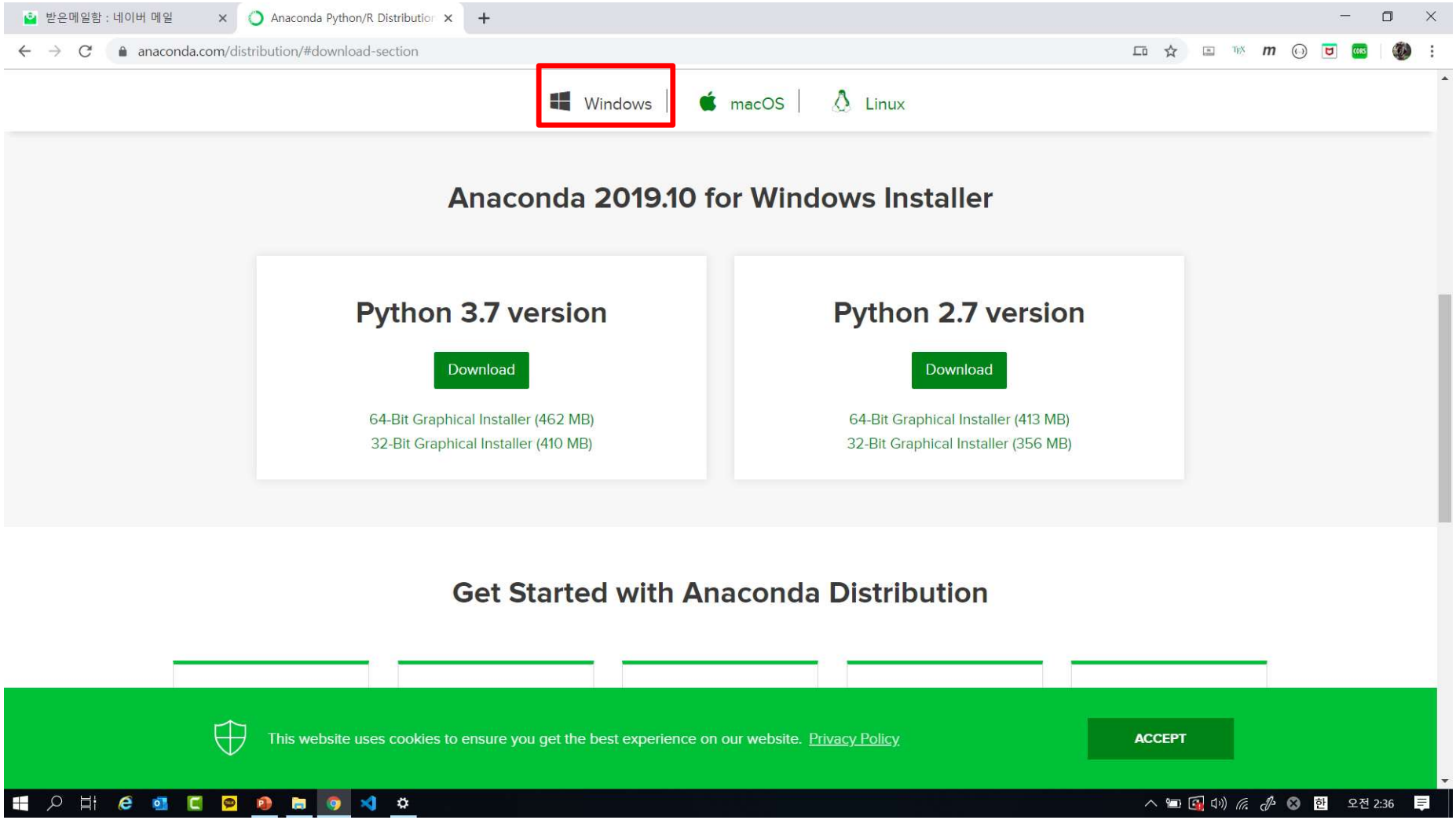
Python 을 활용한 머신러닝 입문

2019. 12

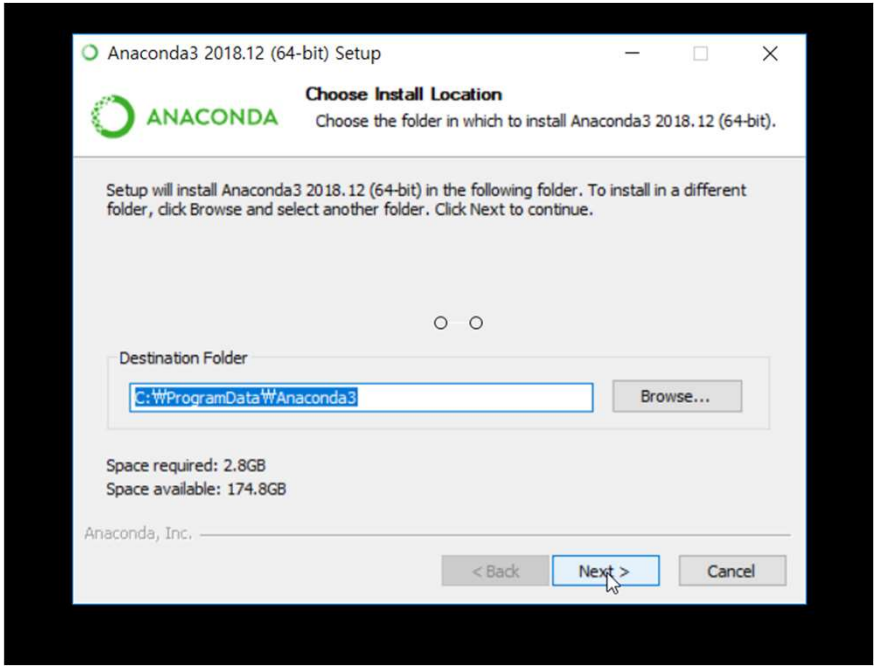
Anaconda Installation

- https://www.anaconda.com/distribution/#download-section



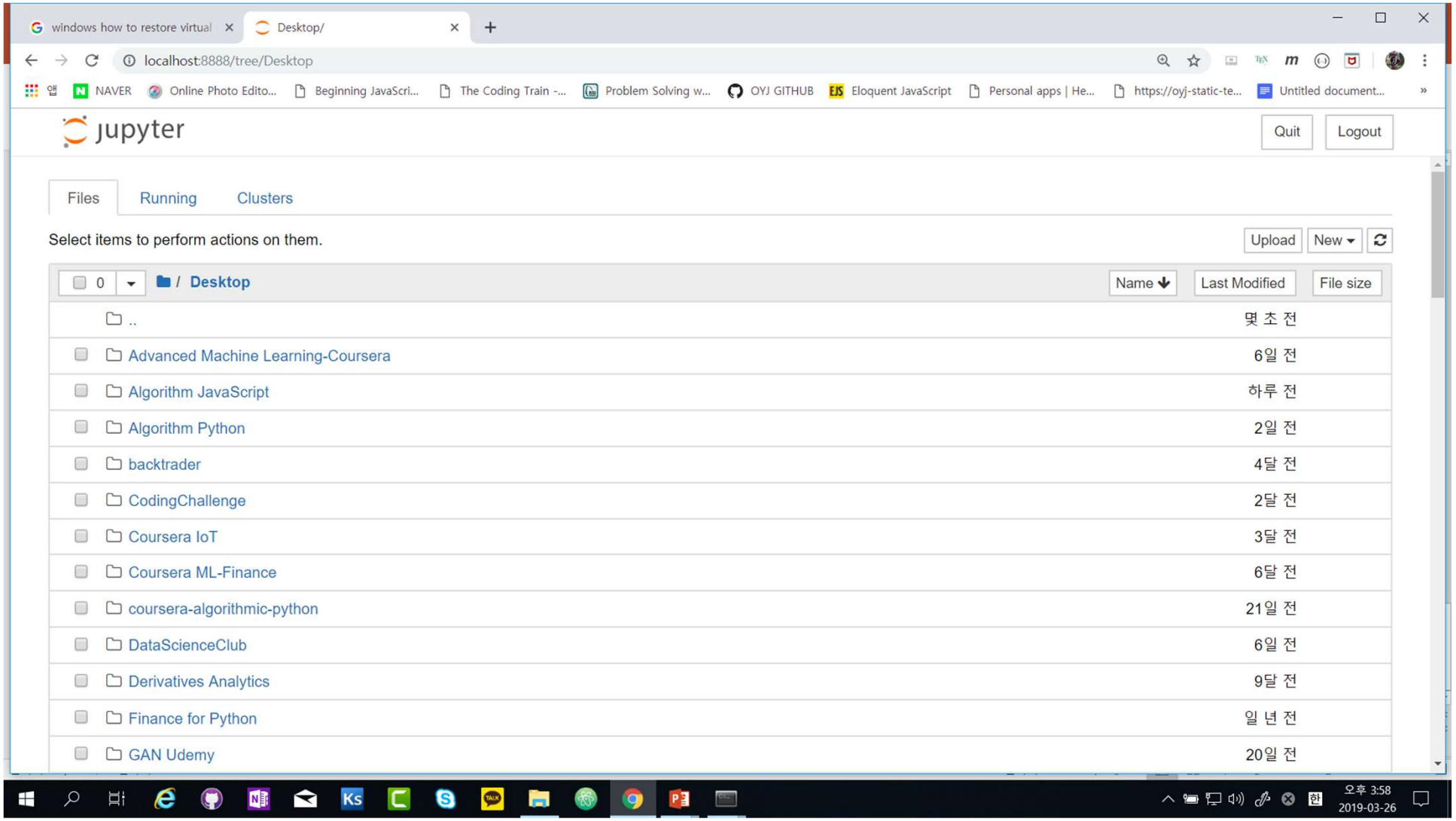
Anaconda 설치


Anaconda3 2018.12 (64-bit) Setup
ANACONDA.
Welcome to Anaconda3 2018.12 (64-bit) Setup
Setup will guide you through the installation of Anaconda3 2018.12 (64-bit).
It is recommended that you close all other applications before starting Setup. This will make it possible to update relevant system files without having to reboot your computer.
Click Next to continue.
Next >
Cancel


Anaconda3 2018.12 (64-bit) Setup
ANACONDA
Choose Install Location
Choose the folder in which to install Anaconda3 2018.12 (64-bit).
Setup will install Anaconda3 2018.12 (64-bit) in the following folder. To install in a different folder, click Browse and select another folder. Click Next to continue.
Destination Folder
C:₩ProgramData₩Anaconda3
Browse...
Space required: 2.8GB
Space available: 174.8GB
Anaconda, Inc.
< Back
Next >
Cancel

- 개인별 작업 directory 생성
- 패키지 설지 : pip install <package name>
- 설치된 Python package 확인 : pip list

>jupyter notebook



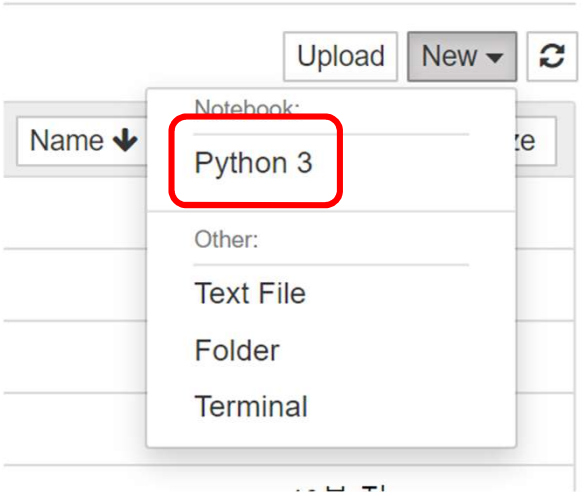
Anaconda 로 설치한 경우 시작 icon 자동 생성

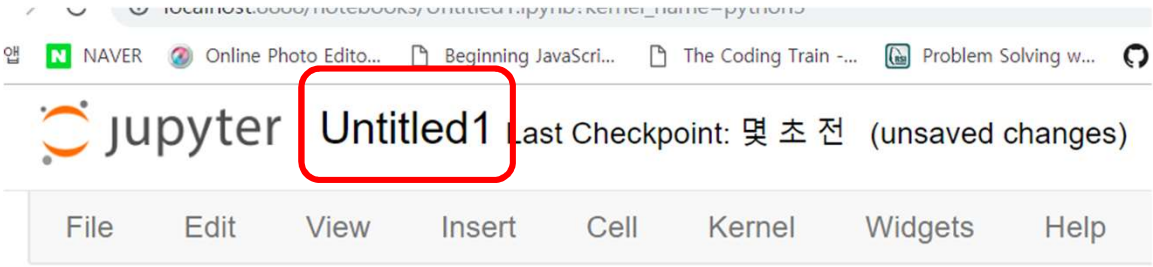
명령 prompt > jupyter notebook

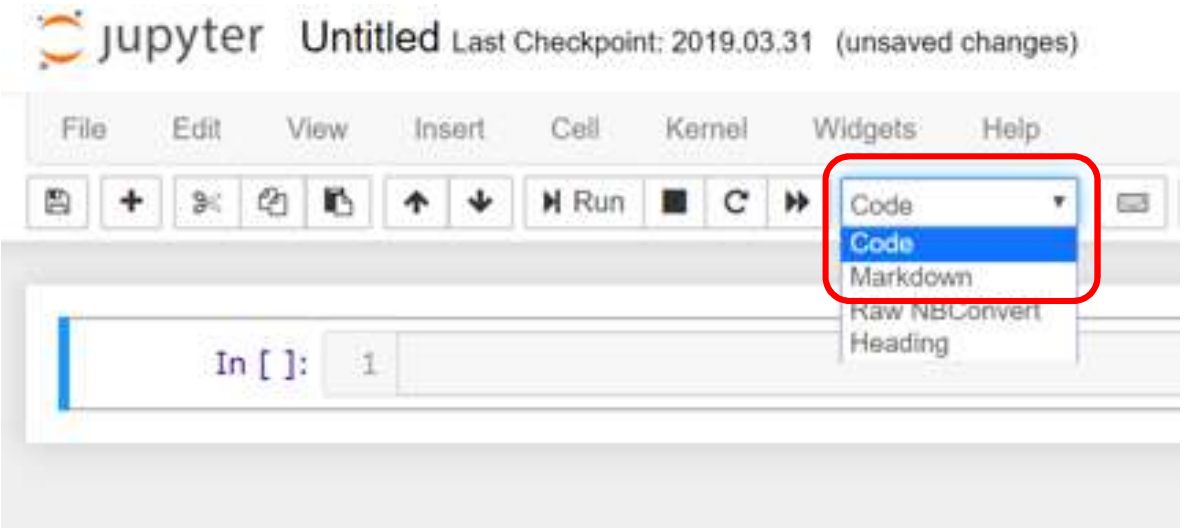
```
(py37) C:₩Users₩trimu>jupyter notebook
[I 09:19:03.913 NotebookApp] JupyterLab extension loaded from C:₩Users₩trimu₩Miniconda3₩envs₩py37₩lib₩site-packages₩jupy
terlab
[I 09:19:03.913 NotebookApp] JupyterLab application directory is C:₩Users₩trimu₩Miniconda3₩envs₩py37₩share₩jupyter₩lab
[I 09:19:03.915 NotebookApp] Serving notebooks from local directory: C:₩Users₩trimu
[I 09:19:03.916 NotebookApp] The Jupyter Notebook is running at:
[I 09:19:03.916 NotebookApp] http://localhost:8888/?token=cde0891905a6a5459965881382c153b4f9f5593f1cab1221
[I 09:19:03.916 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 09:19:04.085 NotebookApp]

    To access the notebook, open this file in a browser:
        file:///C:/Users/trimu/AppData/Roaming/jupyter/runtime/nbserver-27684-open.html
    Or copy and paste one of these URLs:
        http://localhost:8888/?token=cde0891905a6a5459965881382c153b4f9f5593f1cab1221
```

```
C:\Users\whoami>jupyter notebook list
Currently running servers:
http://localhost:8888/?token=9381e1ec86f4ace207780ac683ab440fde7d7a4ecf65af0c ::
```

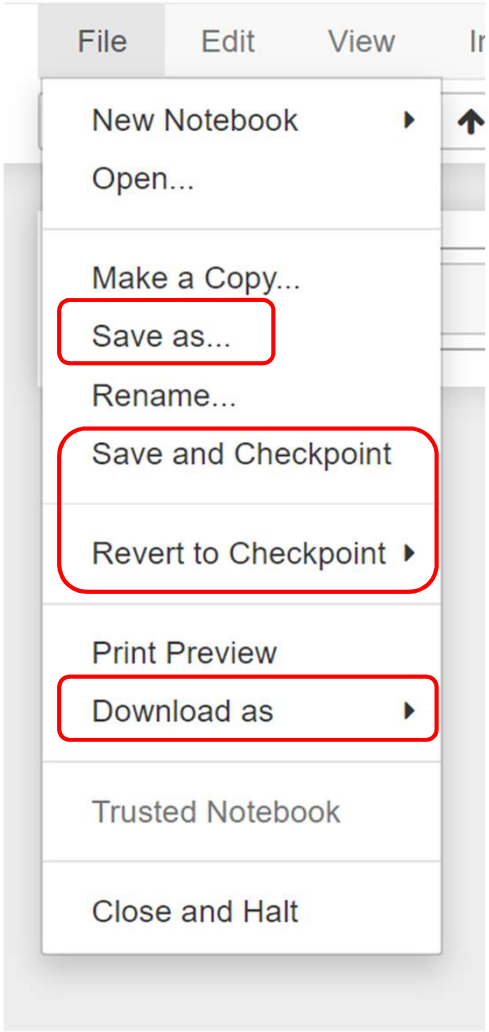
Jupyter Notebook 사용방법

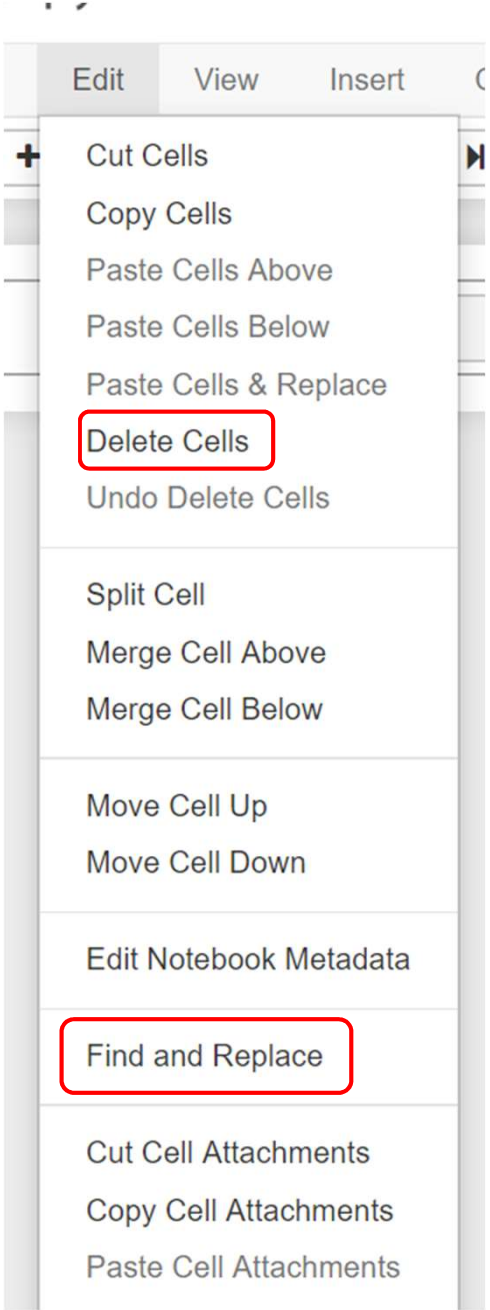


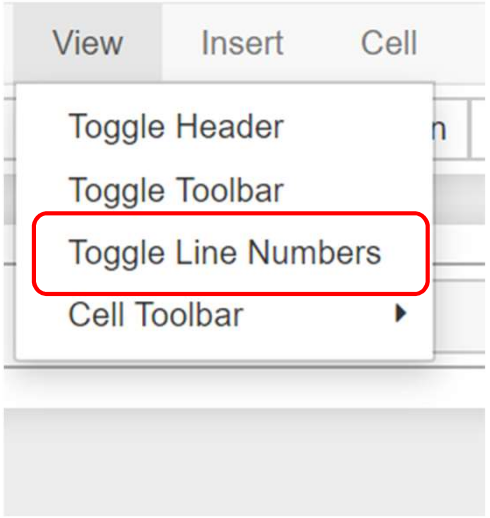




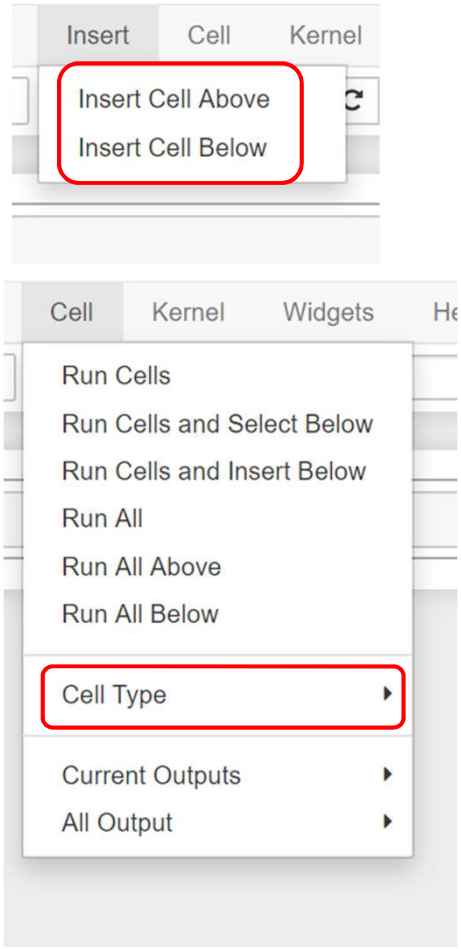
Jupyter Notebook 사용방법

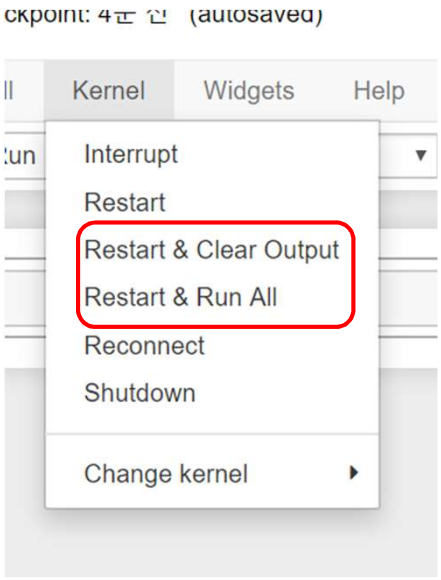


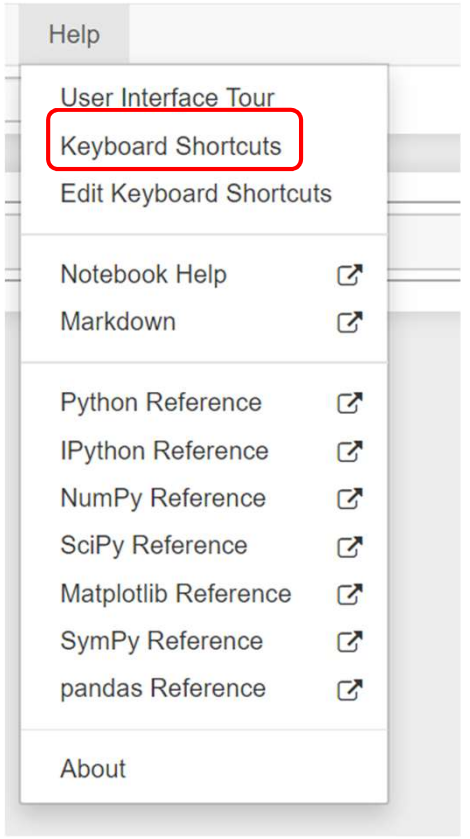




Jupyter Notebook 사용방법



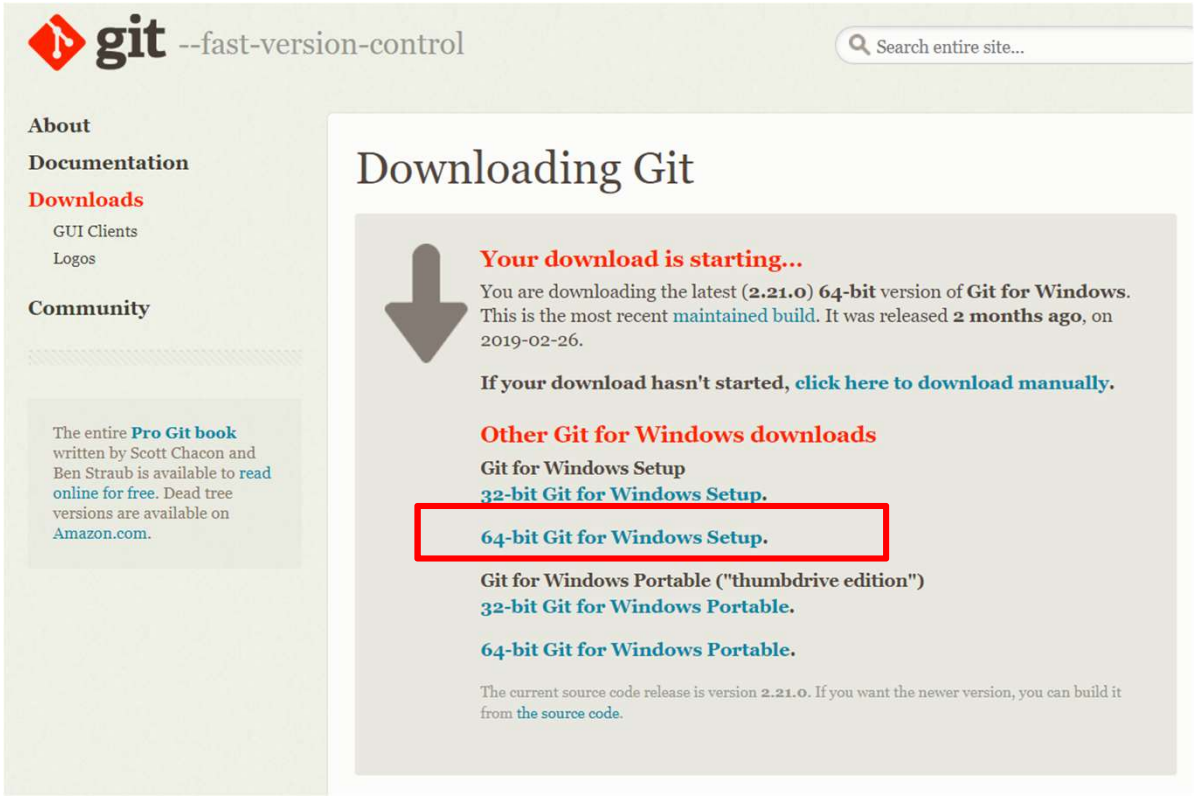




* Login 비밀번호 요구할 경우 사용자 계정의 ~/.jupyter/jupyter_notebook_config.py 삭제

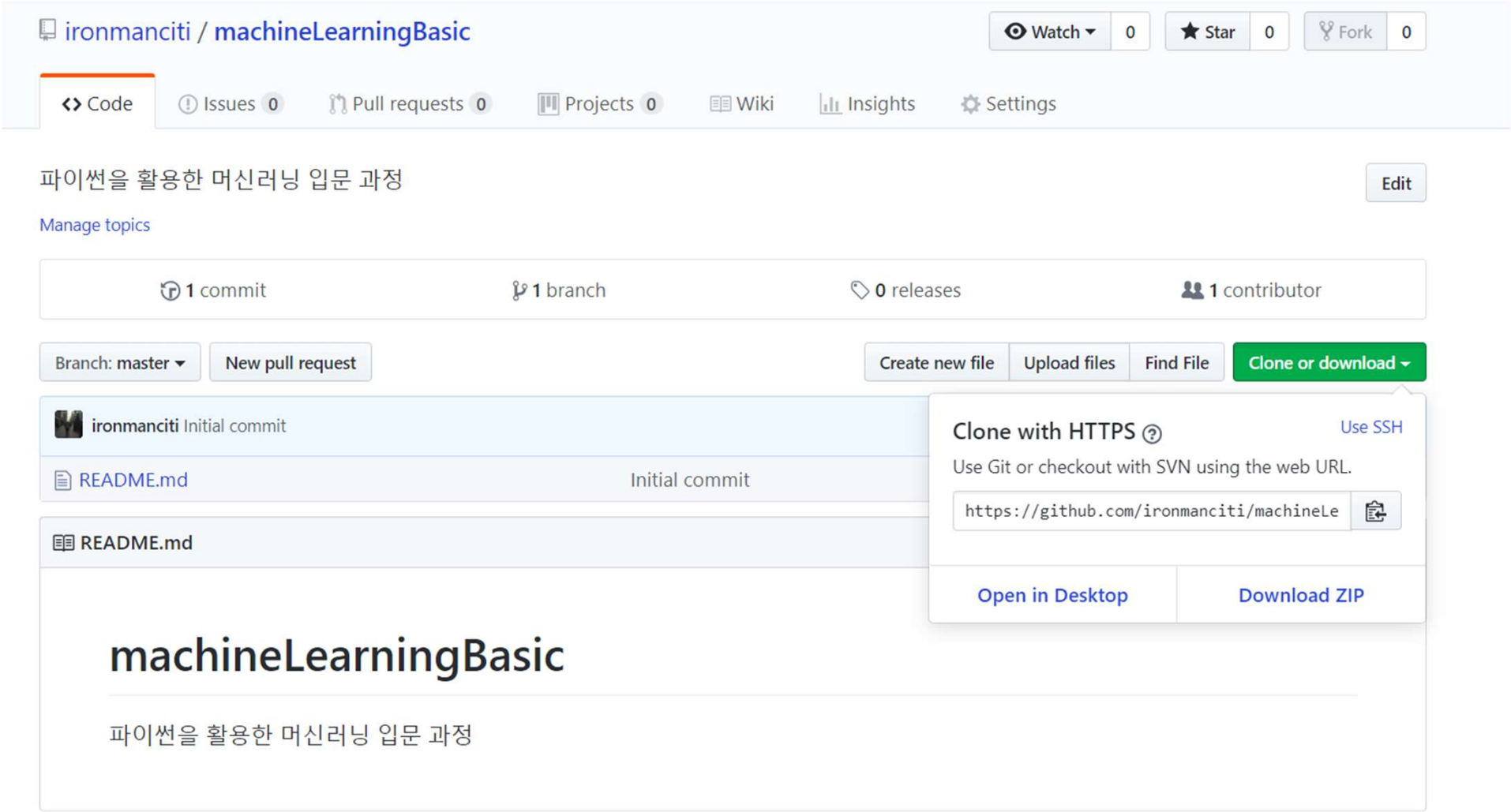
git Installation

https://git-scm.com/download/win



Github Repository

https://github.com/ironmanciti/machineLearningBasic



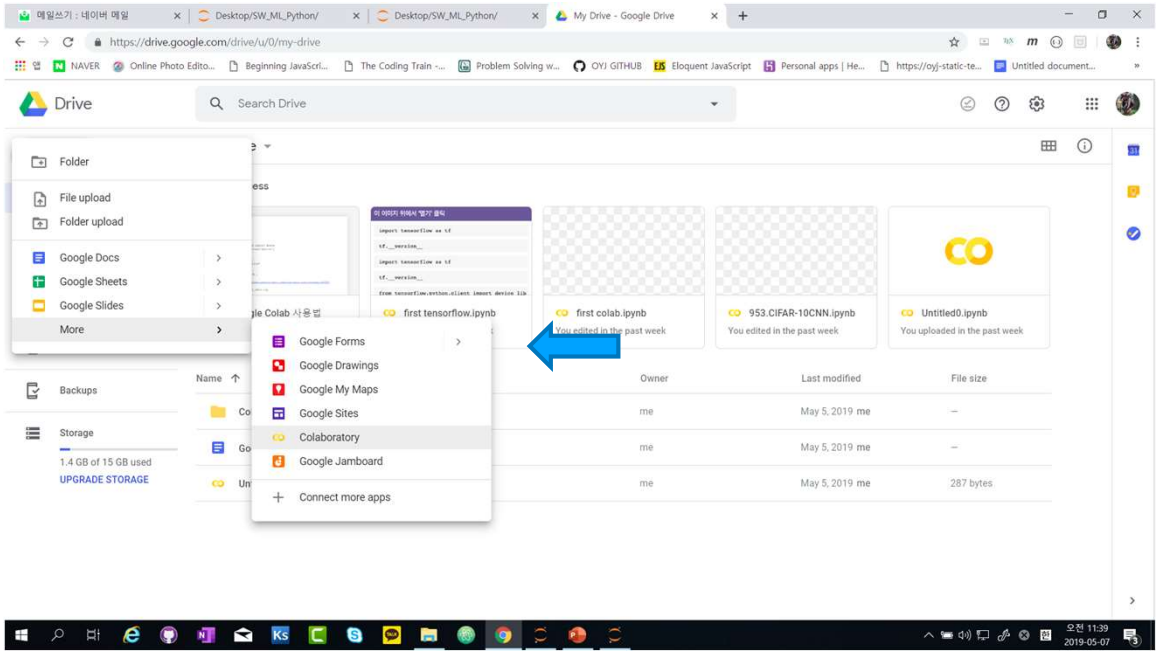
Google Colaboratory

- Free GPU 제공
- Google Drive 와 연동
- Jupyter Notebook 환경
- Deep Learning beginner 를 위한 최적의 환경
- 각종 snippet 제공

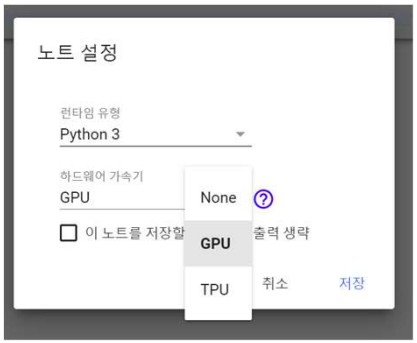
Google Colaboratory 사용하기

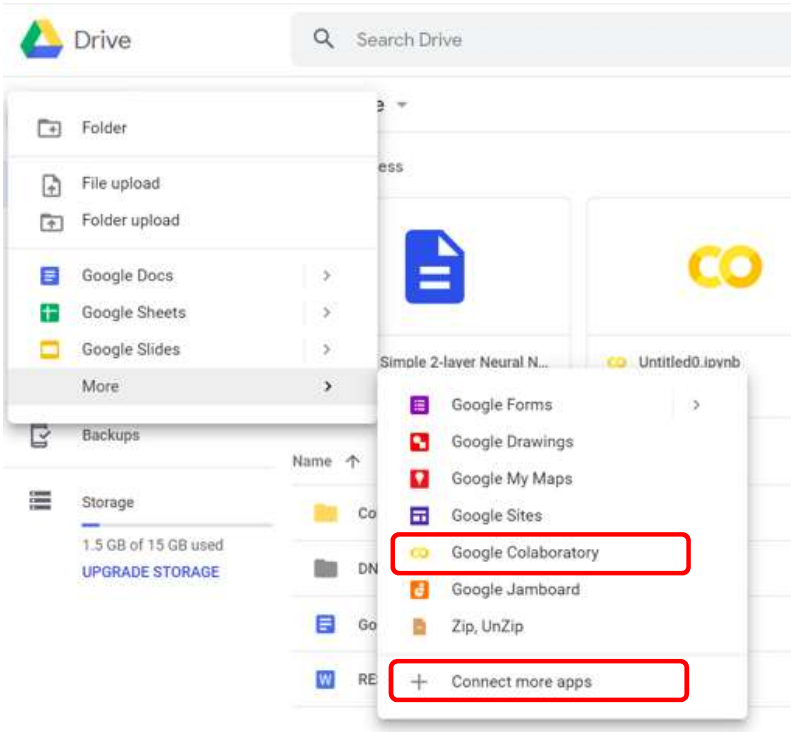
- **Google drive**

  ➔ Colaboratory 연결



- **런타임 → 런타임 유형변경 → GPU 선택**





보이지 않을 경우 Colab 추가

- gpu check

```
[1] from tensorflow.python.client import device_lib
    device_lib.list_local_devices()
```

```
The default version of TensorFlow in Colab will soon switch to TensorFlow 2.x.
We recommend you upgrade now or ensure your notebook will continue to use TensorFlow 1.x via the
%tensorflow_version 1.x magic: more info.
[name: "/device:CPU:0"
 device_type: "CPU"
 memory_limit: 268435456
 locality {
 }
 incarnation: 11925876377786261971, name: "/device:XLA_CPU:0"
 device_type: "XLA_CPU"
 memory_limit: 17179869184
 locality {
 }
 incarnation: 5168198067976361036
 physical_device_desc: "device: XLA_CPU device"]
```

CPU

```
incarnation: 1135439848827926988
physical_device_desc: "device: XLA_GPU device", name: "/device:GPU:0"
device_type: "GPU"
memory_limit: 11330115994
locality {
  bus_id: 1
  links {
  }
}
incarnation: 18206703667281236574
physical_device_desc: "device: 0, name: Tesla K80, pci bus id: 0000:00:04.0, compute capability: 3.7"]
```

GPU

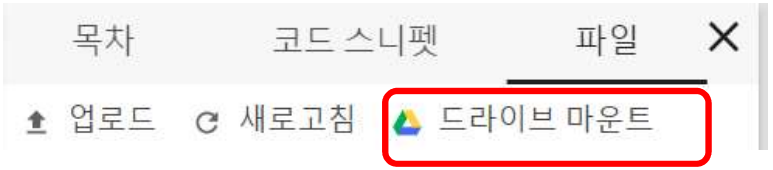
- Google Drive & Colab 연결

1) from google.colab import auth

```
from google.colab import drive
drive.mount('/gdrive')
%cd /gdrive
```

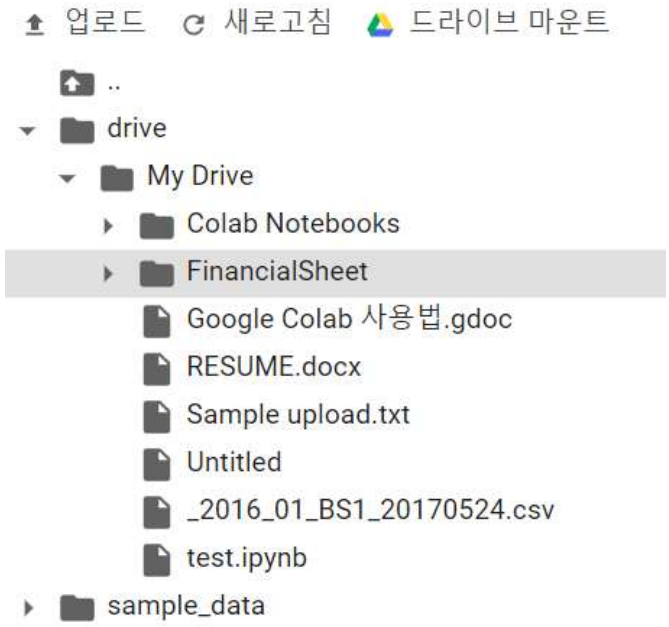
Go to this URL in a browser: https://accounts.google.com/o/oauth2/auth?client_id=947318989803-6bn6qk8qdgf4n4g3pfee6491hc

Enter your authorization code:




```
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

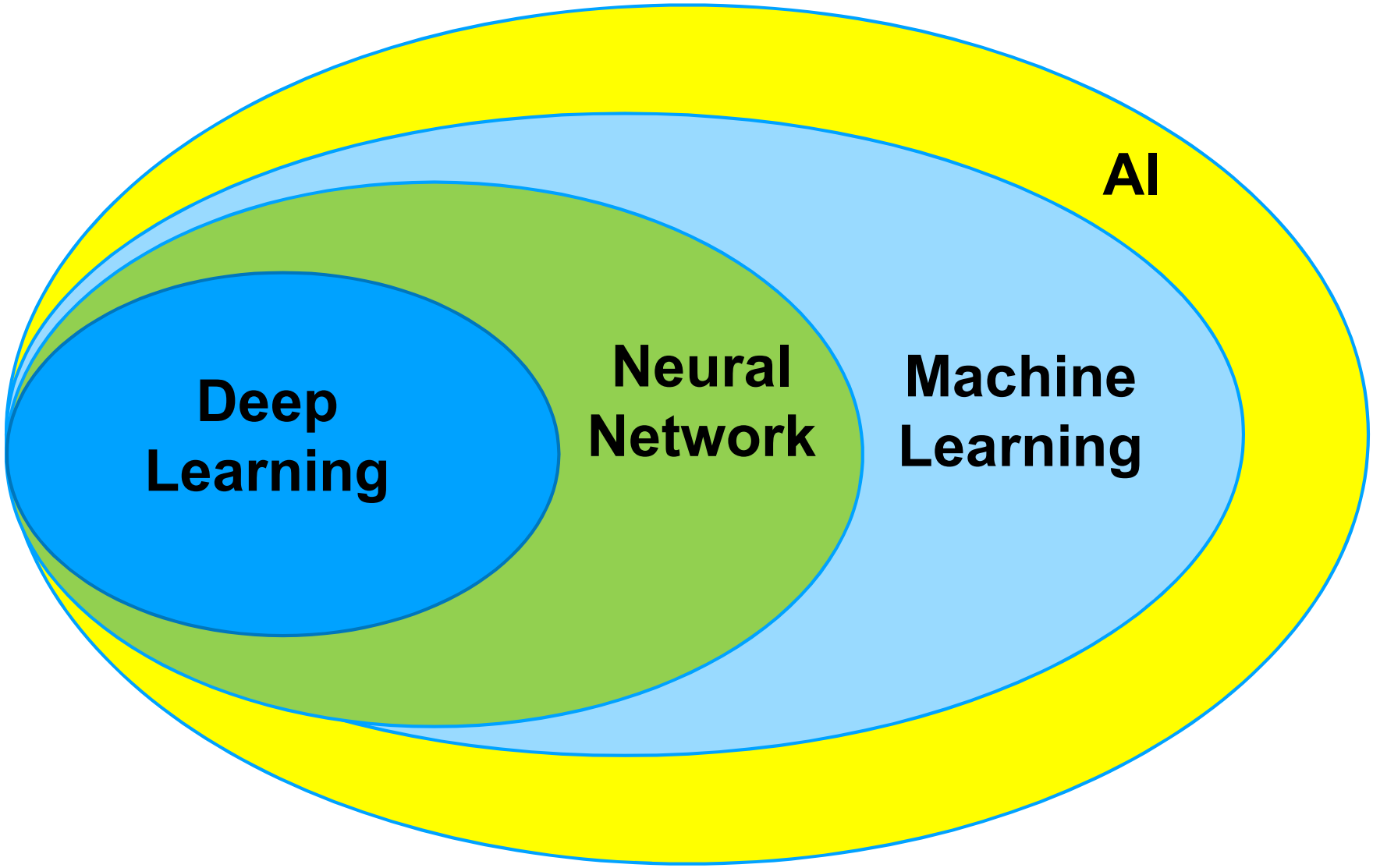


2) Read csv file

```
[8]  import pandas as pd
     pd.read_csv("/content/gdrive/My Drive/FinancialSheet/2015_4Q_FS/_2015_04_CF2_20160601.csv")
```

Machine Learning
개요

AI vs Machine Learning vs Deep Learning



History of Machine Learning

- **탄생 [1950 년대]**

  1958 년 코넬대 심리학자 프랭크 로센블래트가 인간의 뇌신경을 본떠 Perceptron 고안. 신경망 기반 인공지능 연구의 부흥기 시작

- **AI 의 첫번째 암흑기 [1970 년대]**

  Marvin Minsky 가 Perceptron 은 XOR 문제를 해결할 수 없음을 수학적으로 증명. 인공지능에 대한 대규모 연구 지원 중단

- **중흥기 [1980 년대]**

  산업계에 전문가 시스템(Expert System)이 도입되며 본격적으로 확산

- **AI 의 두번째 암흑기 [1987 – 1993]**

  투자대비 효용성의 한계가 노출. 슈퍼컴퓨터와 시뮬레이션 분야로 연구방향을 전환
  * Jeffery Hinton (Toronto 대학) back-propagation algorithm 개발
    → 민스키가 틀렸다는 것을 증명

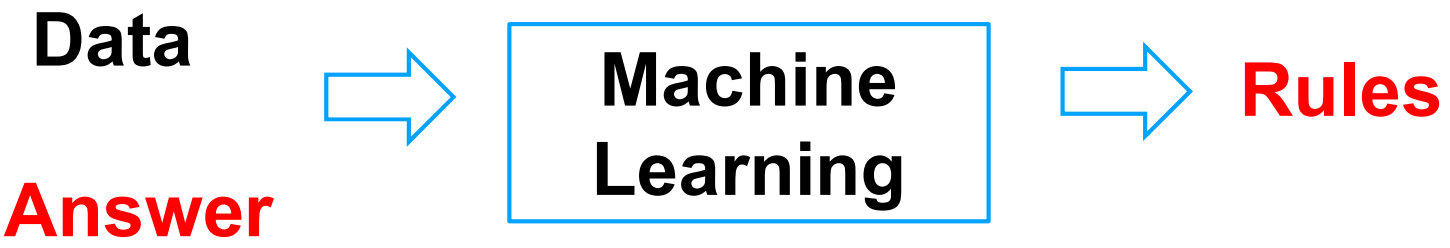
- **IBM Deep Blue 가 Garry Kasparov 에 승리 - 1996**

- **Google Brain 이 최초로 인간 얼굴 인식 - 2012**

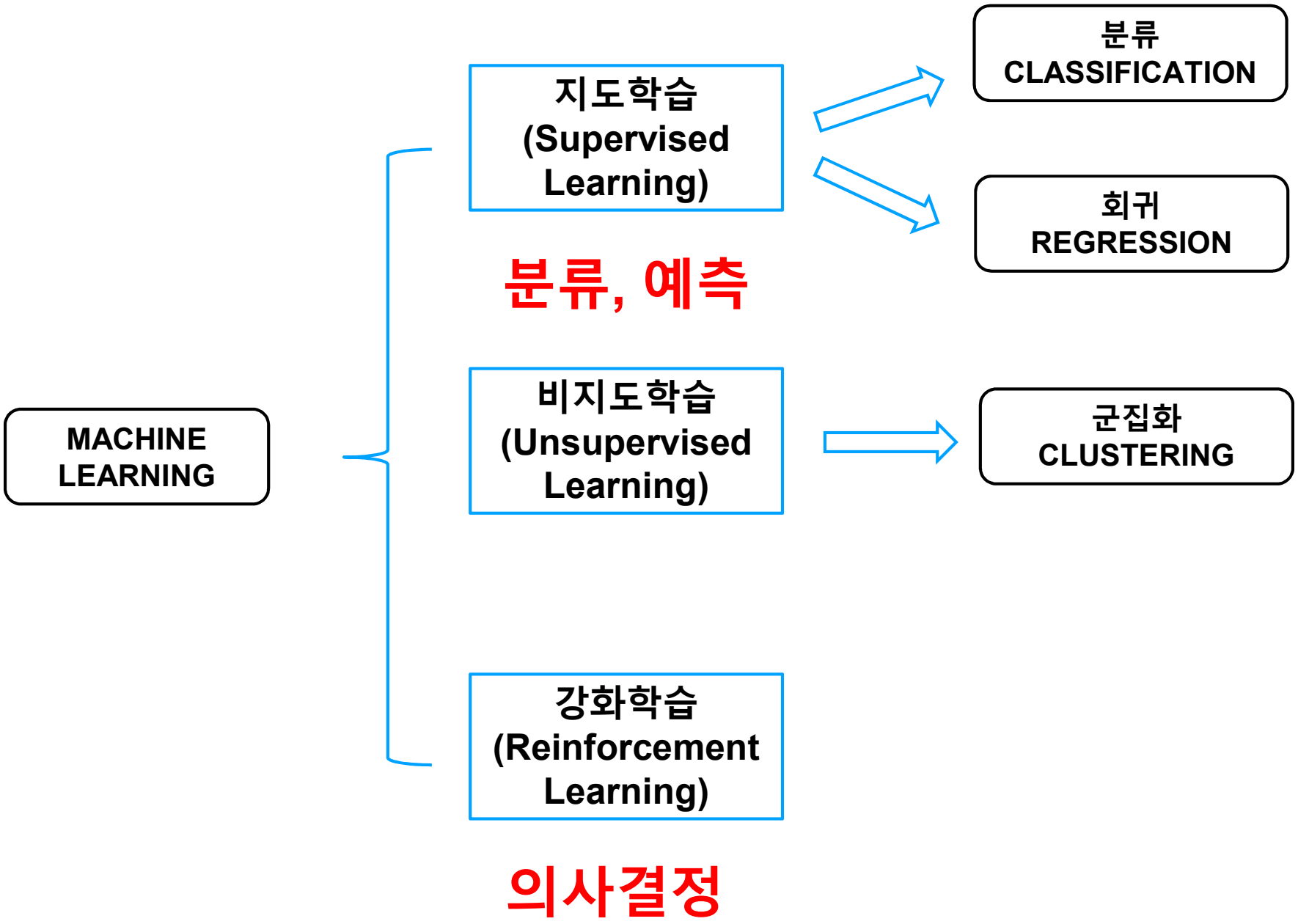
- **~ 상업적 대 폭발기**

- **Alphago 이세돌에 승리 – 2016**

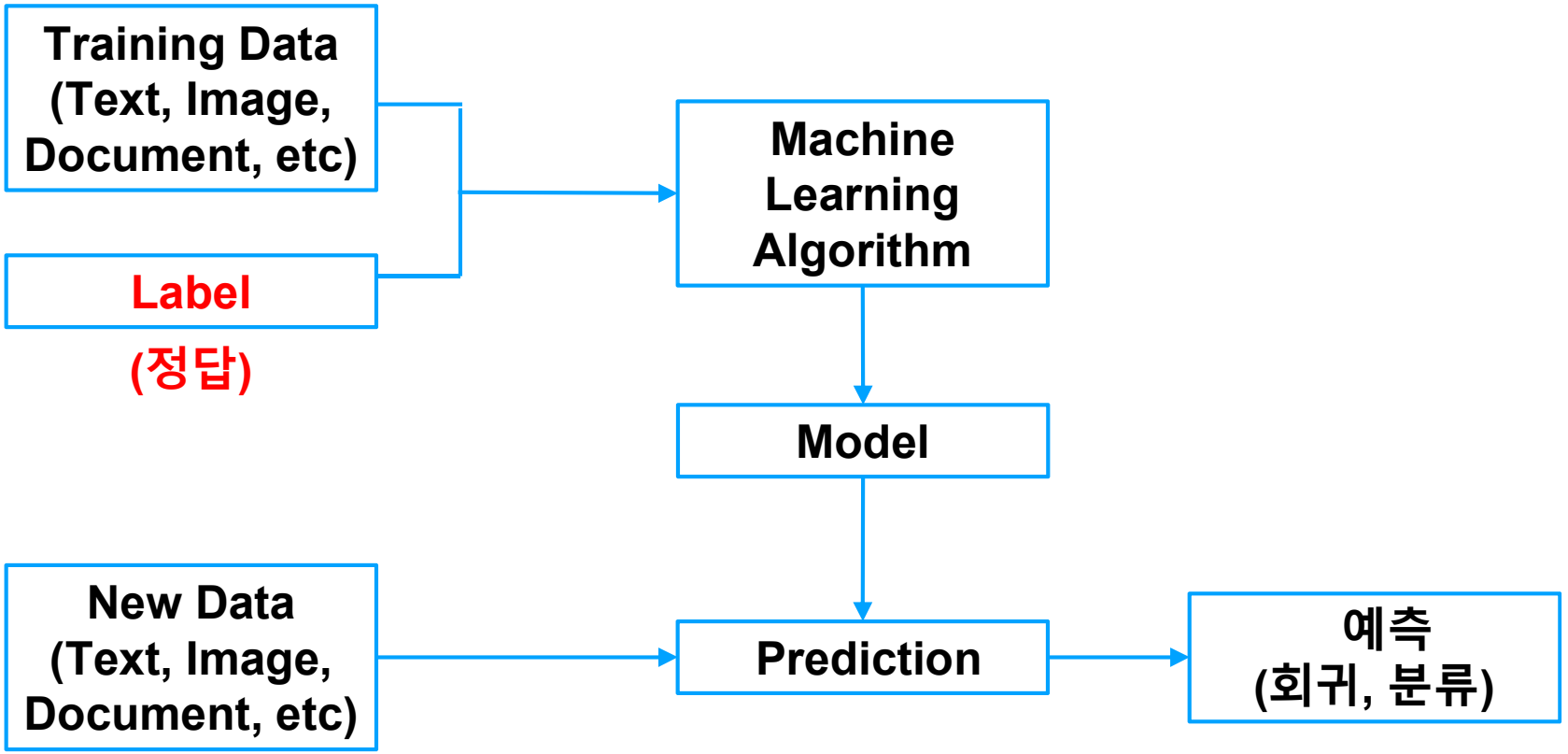
전통적 Programming
vs
Machine Learning



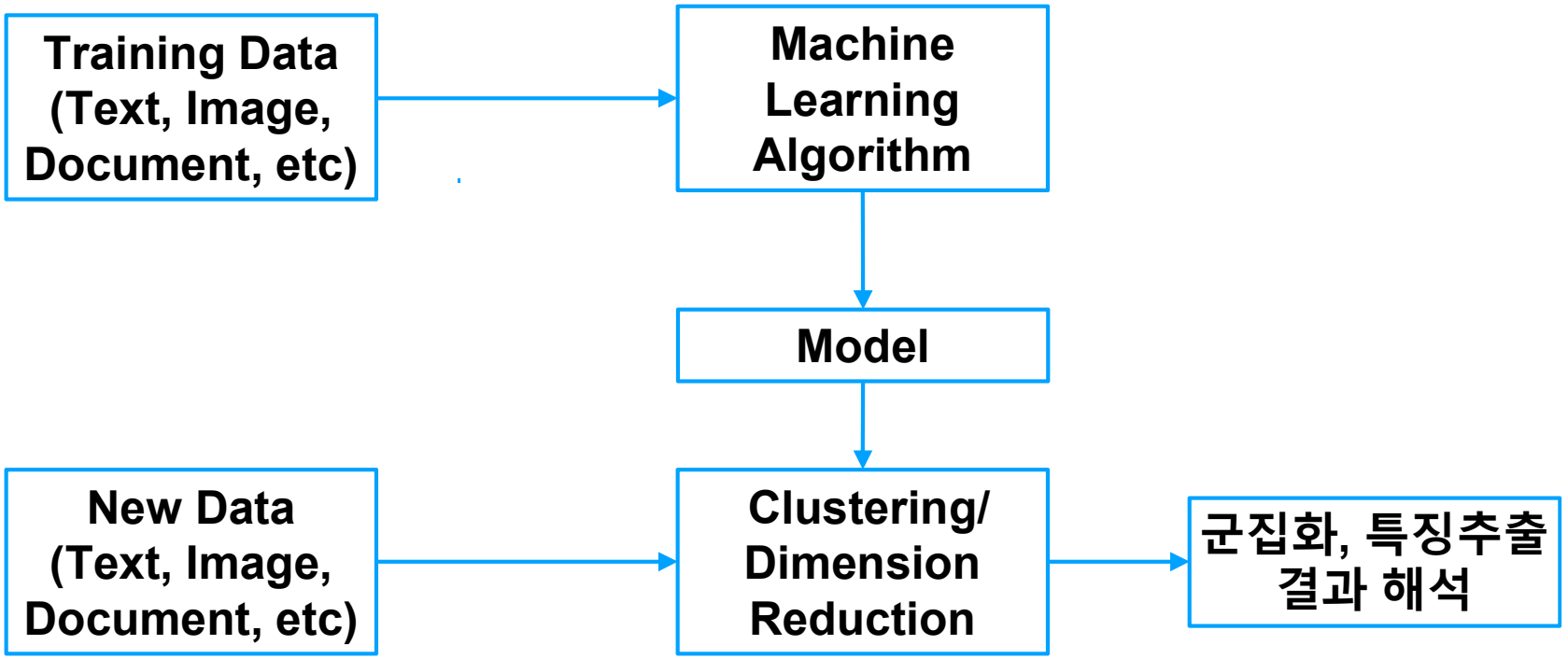
Machine Learning 의 종류



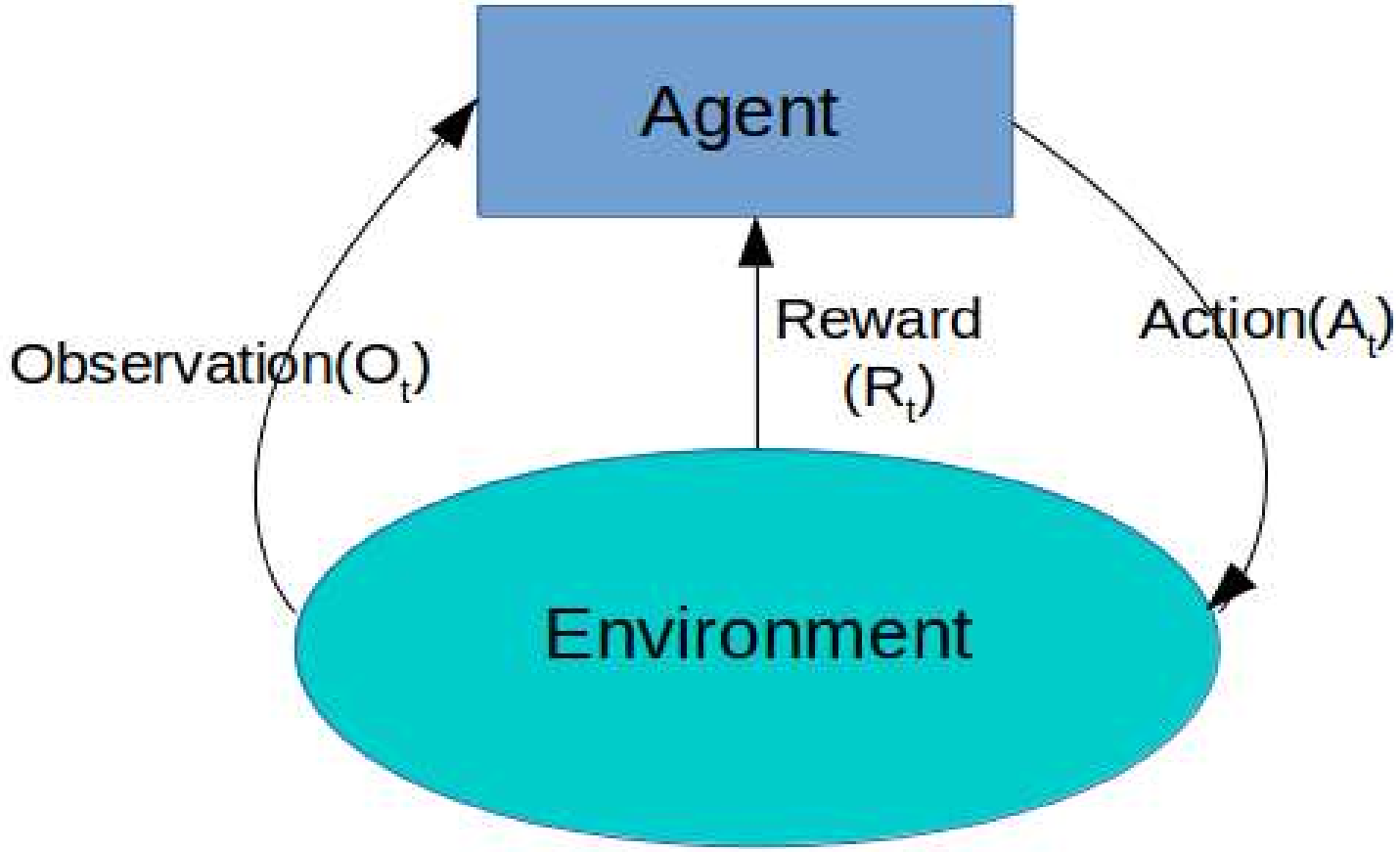
Supervised Learning (지도학습)



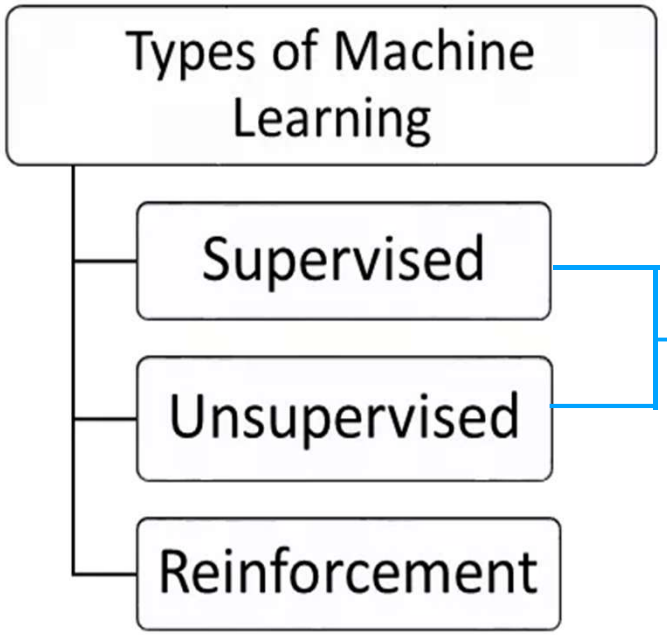
Unsupervised Learning (비지도학습)



Reinforcement Learning (강화학습)



Machine Learning 기법의 종류



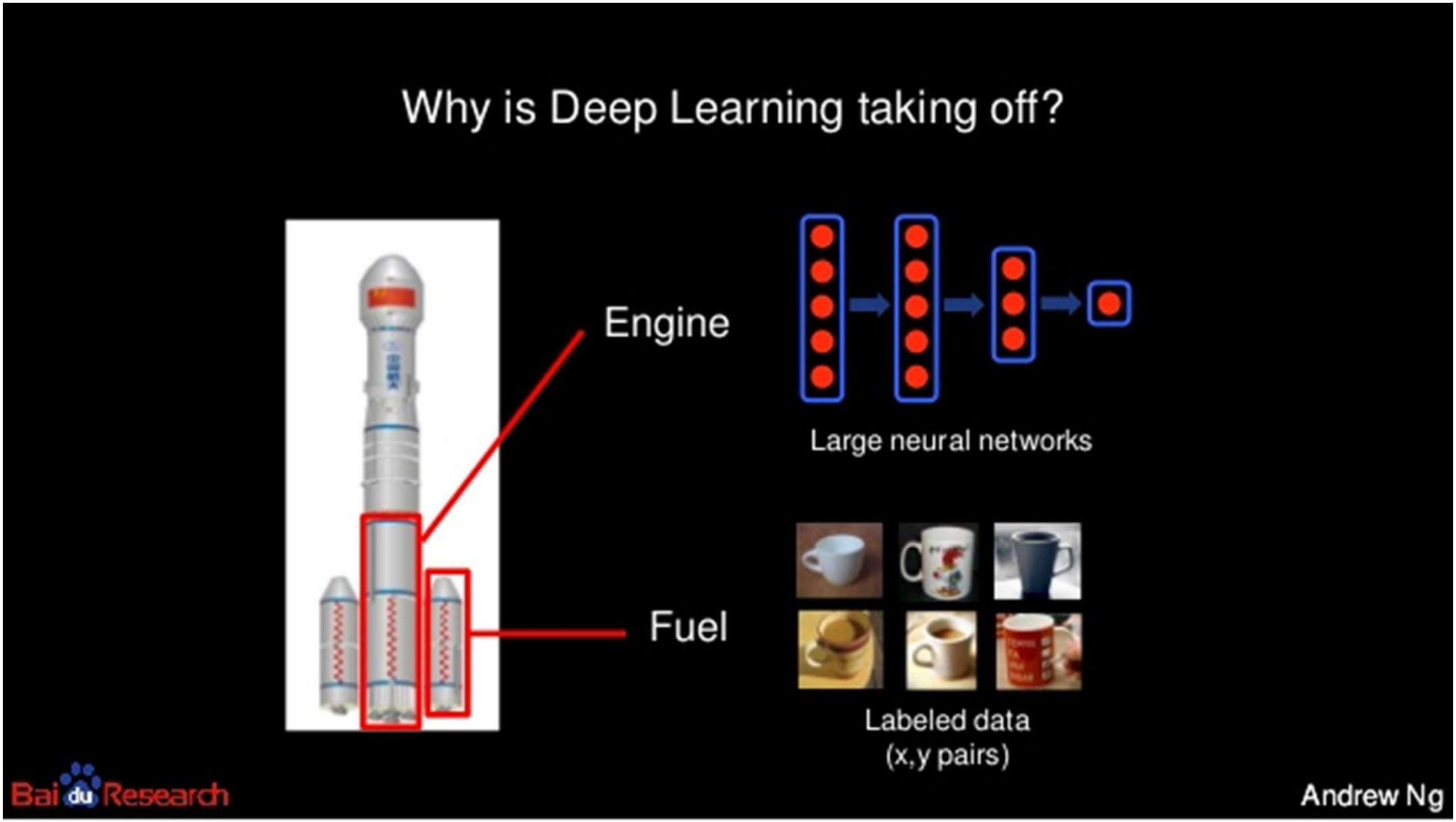
전통적 ML	• Linear & Logistic Regression • KNN • Decision Tree • SVM • Ensembles • K-Means Clustering • PCA
Deep Learning (Neural Network)	FC CNN RNN Auto Encoder GAN

전통적 ML의 기법

종류	용도
Linear / Logistic Regression	선형회귀/분류
KNN (K-Nearest Neighbor)	분류/회귀
Decision Tree (결정나무)	분류
SVM (Support Vector Machine)	분류, 회귀
Ensemble (Random Forest, XGBoost, etc)	분류, 회귀
K-Means Clustering (K-평균 군집화)	군집화
PCA (Principal Component Analysis)	차원 축소

Neural Network 의 종류

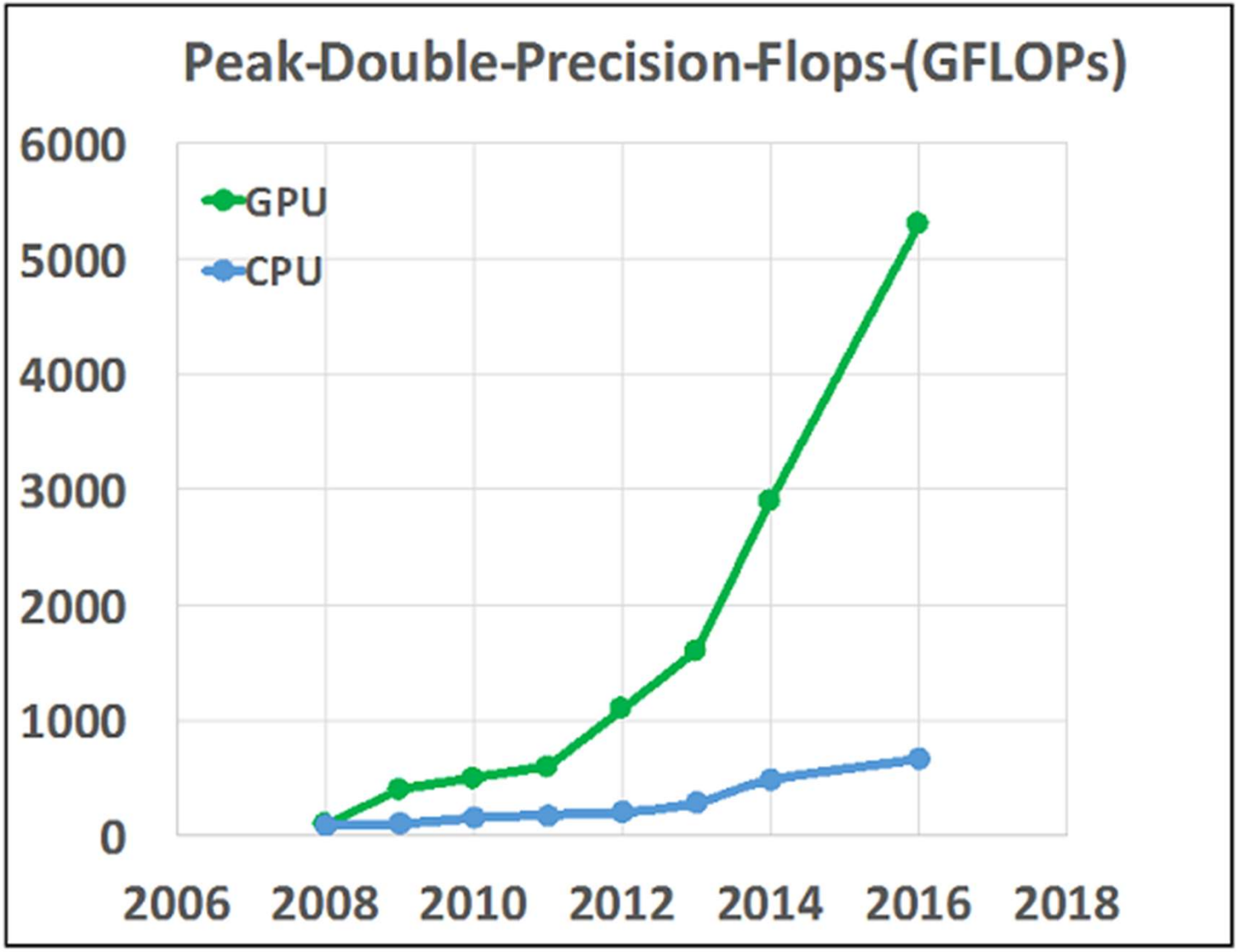
종류	용도
FC (Fully Connected Neural Network)	분류, 회귀
CNN (Convolutional Neural Network)	Image 인식
RNN (Recurrent Neural Network)	시계열 인식, 자동번역, 감성분석 등
AE (Auto Encoder)	비지도 학습, 차원축소
GAN (Generative Adversarial Nets)	적대적 생성모델, 이미지 위조



Rocket Engine : NVIDIA + Deep Learning Algorithm

Fuel : Data (25,000 pictures for cat)

CPU vs GPU Performance



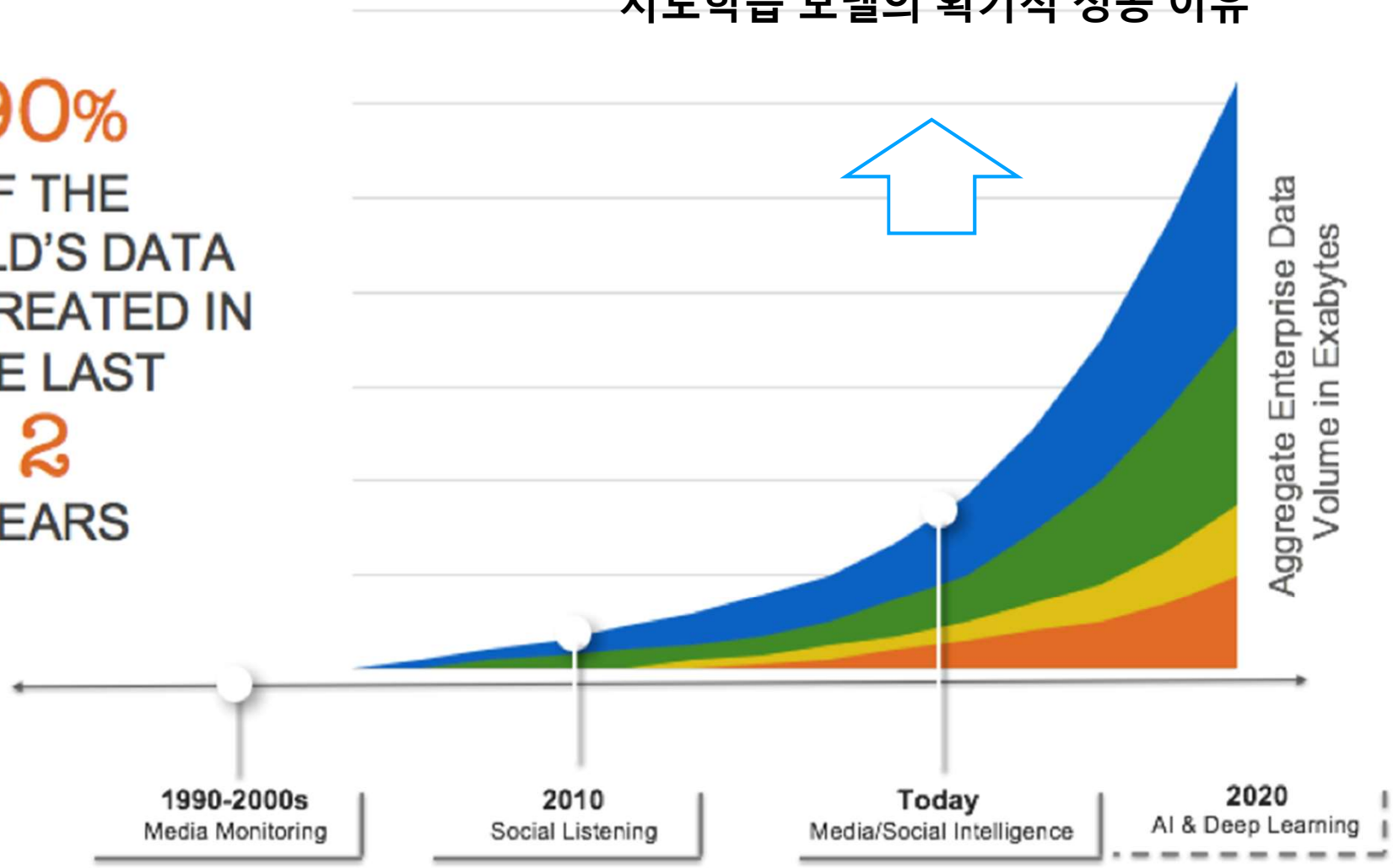
지도학습 모델의 획기적 성공 이유

90%

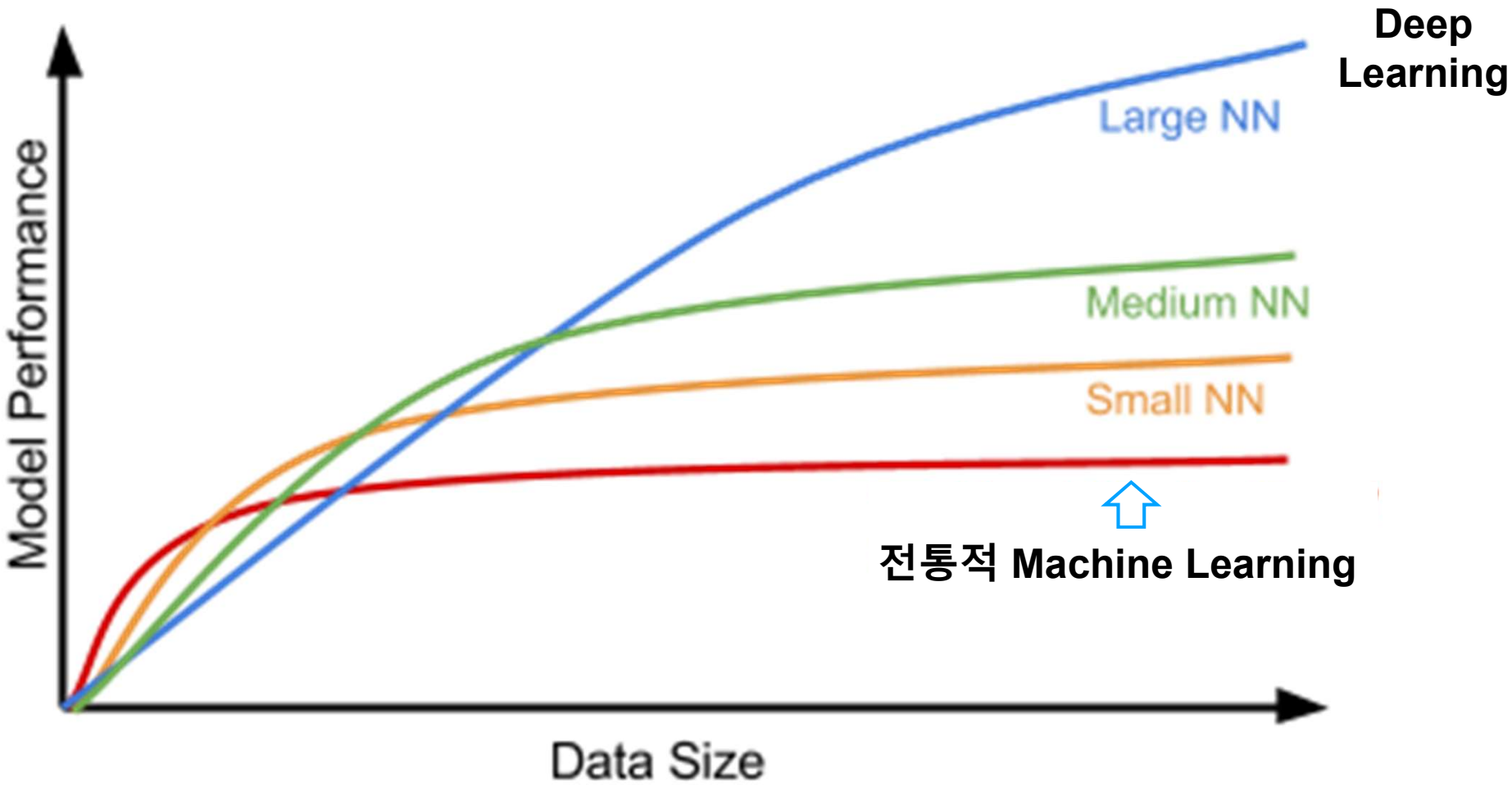
OF THE WORLD'S DATA WAS CREATED IN THE LAST

2

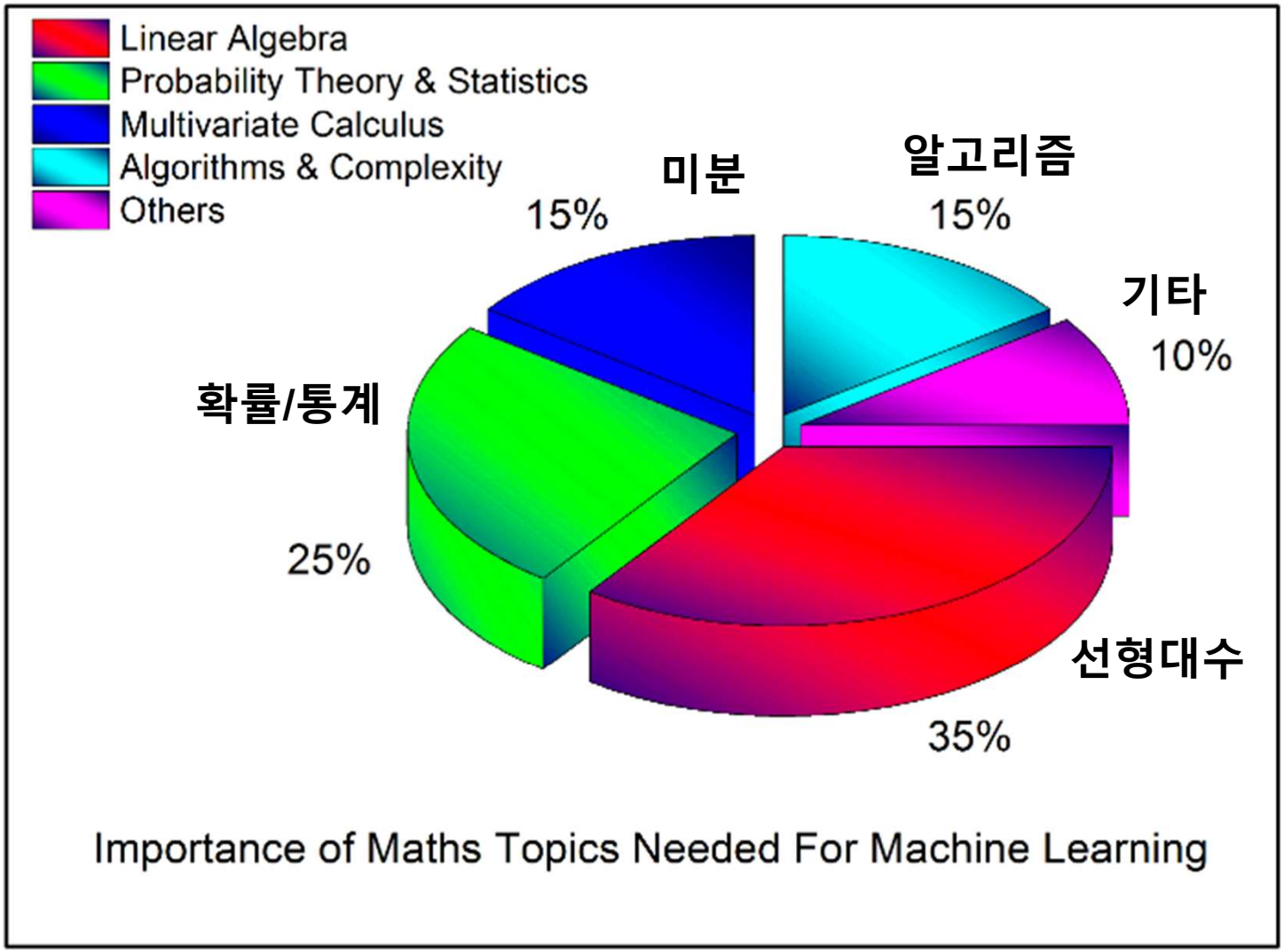
YEARS



전통적 Machine Learning vs. Deep Neural Network



Machine Learning 학습에 필요한 수학 지식

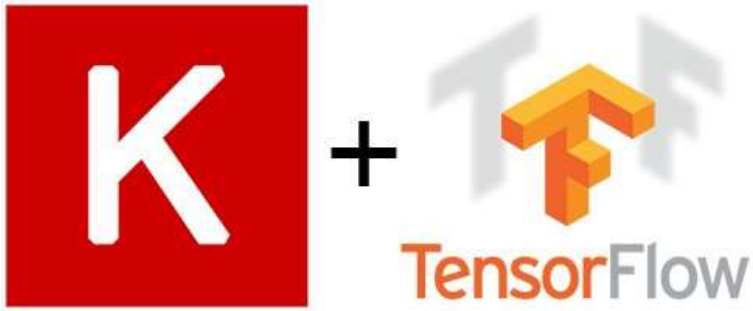


Machine Learning

전통적 Machine Learning Tool:

- 벤치마크용 데이터셋 예제
- 데이터 전처리(preprocessing)
- 지도 학습(Supervised learning)
- 비지도 학습(Unsupervised learning)
- 모형 평가 및 선택 (evaluation and selection)

Deep Learning



Deep Learning Tool:

- Pre-train model, Sample Dataset 제공
- Deep Learning 에 필요한 각종 함수 제공
- GPU support
- 각종 language 지원 API 제공

과정 SCOPE

- Crash Course – Numpy, Pandas, Matplotlib
- Linear Regression (선형회귀)
- non-Linear Regression (비선형회)
- KNN (K-Nearest Neighbor)
- Decision Tree
- Logistic Regression
- SVM (Support Vector Machine)
- Random Forest

- K-Means Clustering
- Tensorflow/Keras
- Simple Neural Network
- CNN
- RNN
- GAN(Optional)
- 실습문제