Machine Learning Basics

By

Yasir Hussein Shakir

Yasir Hussein Shakir | Master | Kaggle

(12) Yasir Hussein Shakir – YouTube

yasserhessein (Yasir Hussein Shakir) · GitHub

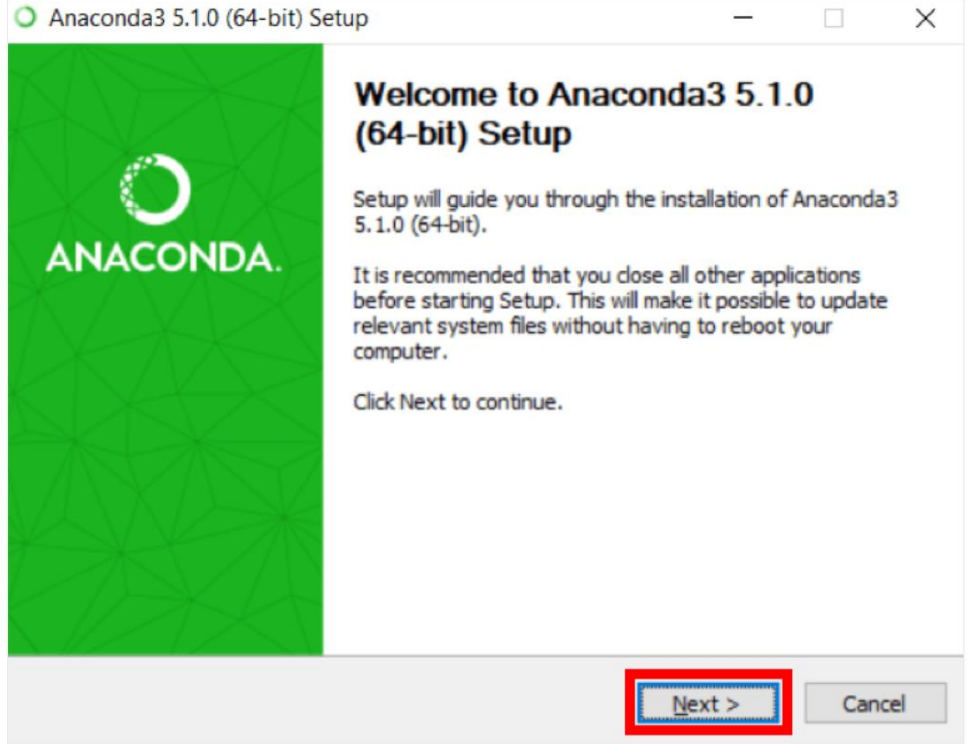
Yasir Hussein - A - yasir | LinkedIn

Contents

- How to Environment Setup Command Line
- Why Python?
- How to use Jupyter Notebook
- Numbers and Boolean Values
- Input & Output
- Pip package managers
- Variables
- What the Data
- Data types
- What is the relationship between DS, AI, ML, and DL.
- Machine Learning
- Stages of Machine Learning
- Machine Learning vs Traditional Programming
- Types of Machine Learning
- What is Deep Learning
- Deep Learning Applications

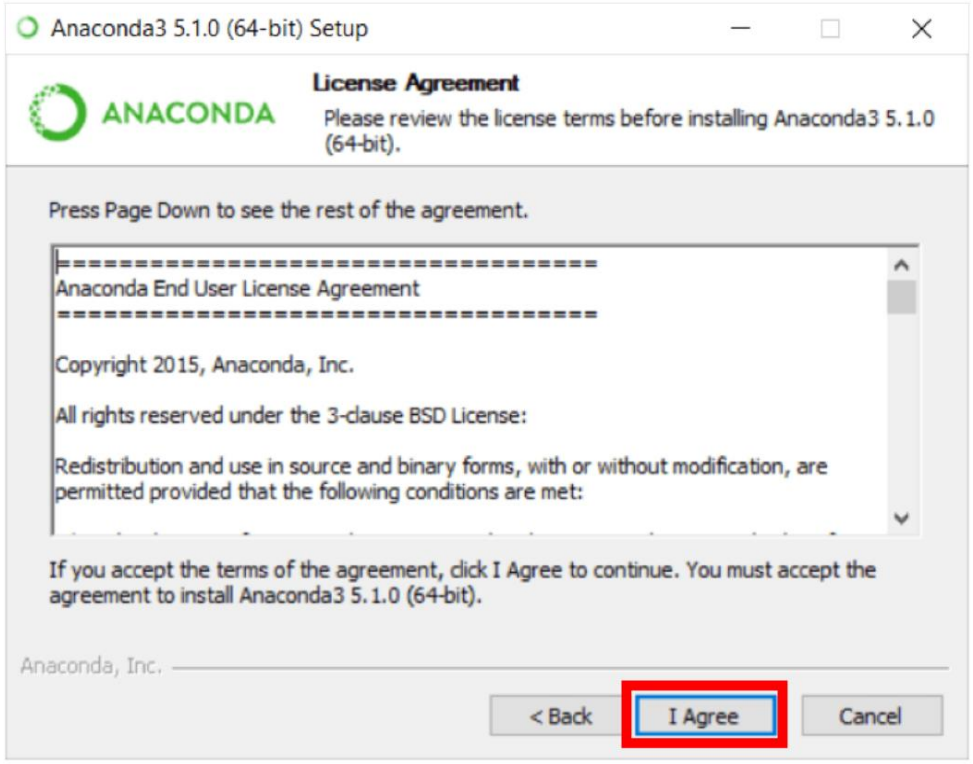
How to Environment Setup

- Download Anaconda
- (1) New Messages! (anaconda.com)
- **Download**
- When the screen below appears, click on Next.



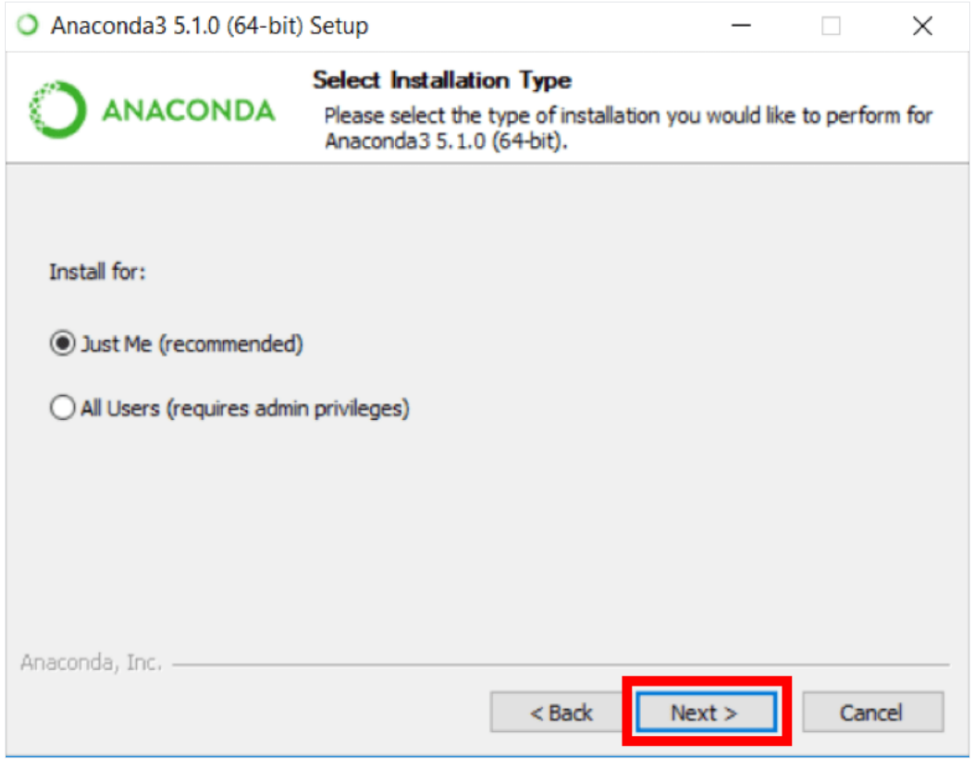
How to Environment Setup

Read the license agreement and click on I Agree.



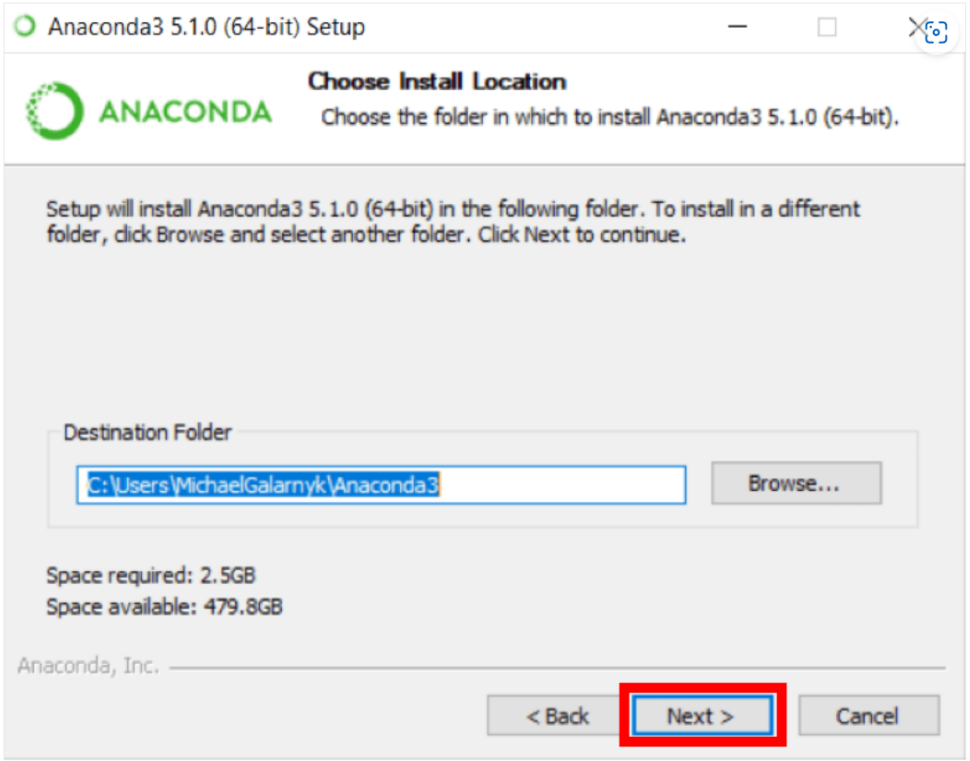
How to Environment Setup

Click on Next.



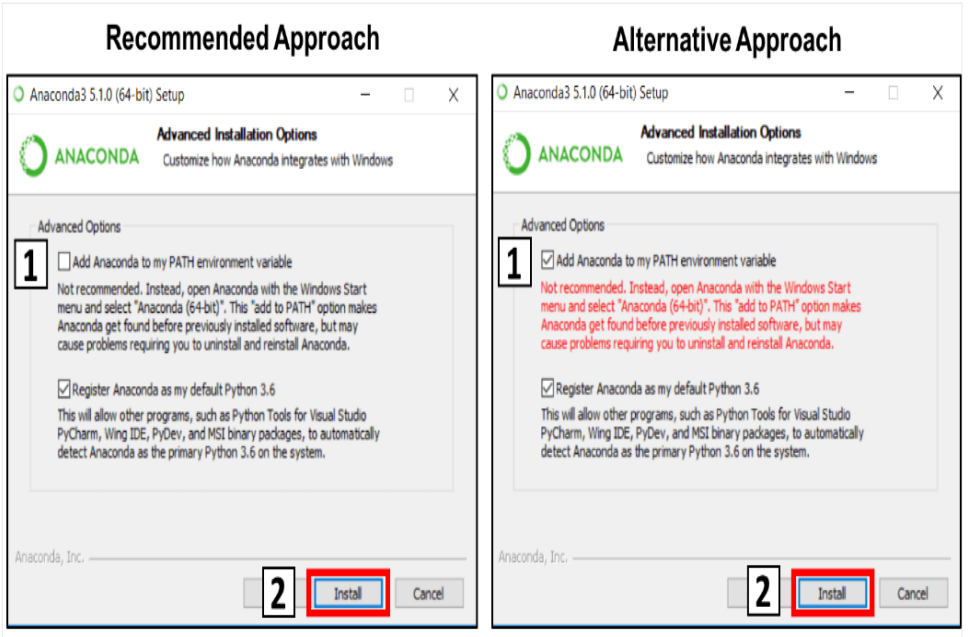
How to Environment Setup

- Note your installation location and then click Next.



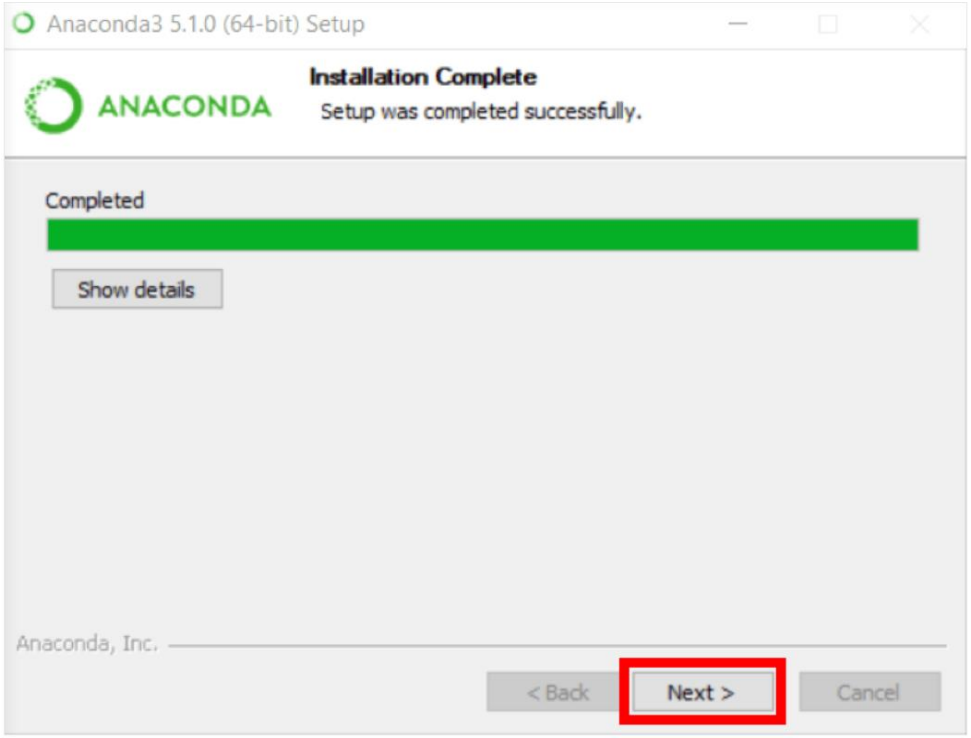
How to Environment Setup

- This means you will have to use Anaconda Navigator or the Anaconda Command Prompt (located in the Start Menu under "Anaconda") when you wish to use Anaconda (you can always add Anaconda to your PATH later if you don't check the box). If you want to be able to use Anaconda in your command prompt (or git bash, cmder, powershell etc), please use the alternative approach and check the box.



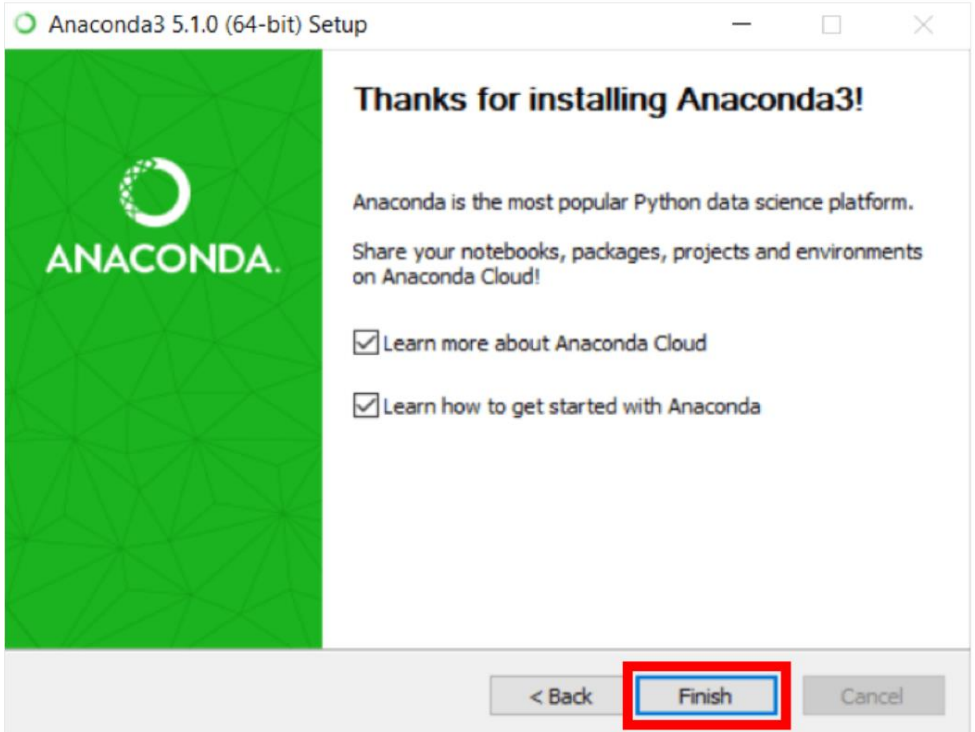
How to Environment Setup

- Click on Next.



How to Environment Setup

- Click on Finish.



File Help

ANACONDA.NAVIGATOR

Upgrade Now

Connect

Home

Environments

Learning

Community

Documentation

Anaconda Blog

All applications on base (root) Channels

DataSpell

DataSpell is an IDE for exploratory data analysis and prototyping machine learning models. It combines the interactivity of Jupyter notebooks with the intelligent Python and R coding assistance of PyCharm in one user-friendly environment.

Install

Anaconda Notebooks

Cloud-hosted notebook service from Anaconda. Launch a preconfigured environment with hundreds of packages and store project files with persistent cloud storage.

Launch

CMD.exe Prompt

0.1.1

Run a cmd.exe terminal with your current environment from Navigator activated

Launch

JupyterLab

3.6.3

An extensible environment for interactive and reproducible computing, based on the Jupyter Notebook and Architecture.

Notebook

6.5.4

Web-based, interactive computing notebook environment. Edit and run human-readable docs while describing the data analysis.

Powershell Prompt

0.0.1

Run a Powershell terminal with your current environment from Navigator activated

How to use Jupyter Notebook

- New

Why Python?

- 1- Open-source software (OSS) Open-source means it is free. Python has a large and active scientific community with access to the software's source code and contributes to its continuous development and upgrading, depending on users' needs.

- 2- General-purpose There is a broad set of fields where Python could be applied: web programming, analysis of financial data, analysis of big data, and more.

- 3- High-level High-level languages employ syntax a lot closer to human logic, which makes the language easier to learn and implement.

Numbers and Boolean Values

Term	Definition
Integer	Positive or negative whole numbers without a decimal point *Example: 5, 10, -3, -15*
Floating point (float)	Real numbers. Hence, they have a decimal point *Example: 4.75, -5.50, 11.0*
Boolean value	a True or False value, corresponding to the machine's logic of understanding 1s and 0s, on or off, right or wrong, true or false. *Example: True, False*

Input & Output

```
#Input & Output:

# Input
name = input("Enter your name: ")
print("Hello, " + name + "!")
```

```
Enter your name: guys
Hello, guys!
```

Pip package managers

- **Pip Package Managers**

pip is a package installer for Python.

```
pip install package_name
```

```
Defaulting to user installation because normal site-packages is not writeableNote: you may need to restart the kernel to use up
dated packages.

Collecting package_name
  Downloading package_name-0.1.tar.gz (782 bytes)
  Preparing metadata (setup.py): started
  Preparing metadata (setup.py): finished with status 'done'
Building wheels for collected packages: package_name
  Building wheel for package_name (setup.py): started
  Building wheel for package_name (setup.py): finished with status 'done'
  Created wheel for package_name: filename=package_name-0.1-py3-none-any.whl size=1235 sha256=7db71f6c2c95f2c3499972938979095d0
466f4d7c0fdf40525bc958cfc42b6bf
  Stored in directory: c:\users\lenovo\appdata\local\pip\cache\wheels\b3\c1\6f\538e951eb00f535f43151173b4c55e463a35c17b9e90ab3b
1a
Successfully built package_name
Installing collected packages: package_name
Successfully installed package_name-0.1
```

Variables

Variables Table

Variable Name	Role	Type	Demographic	Description	Units	Missing Values
Sample_code_number	ID	Categorical				no
Clump_thickness	Feature	Integer				no
Uniformity_of_cell_size	Feature	Integer				no
Uniformity_of_cell_shape	Feature	Integer				no
Marginal_adhesion	Feature	Integer				no
Single_epithelial_cell_size	Feature	Integer				no
Bare_nuclei	Feature	Integer				yes
Bland_chromatin	Feature	Integer				no
Normal_nucleoli	Feature	Integer				no
Mitoses	Feature	Integer				no

What is Data

- We are living in a data-driven economy. It’s a world were having tons of data, understanding it, and knowing what to do with data is power.
- Understanding your data is not one of the most difficult things in data science, but it is time-consuming.
- Interpretation of data is effective when we know about the source of data.

What is Data (Types of Data)

- Categorical

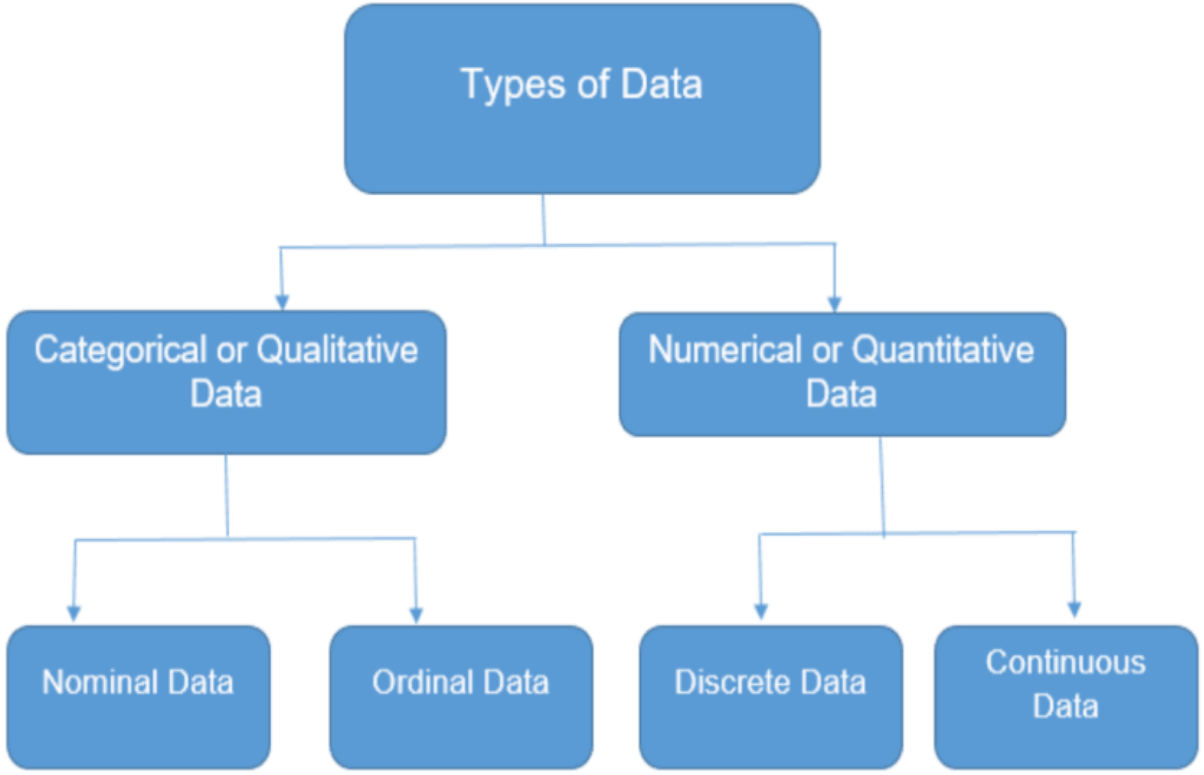
This represents qualitative data with no apparent inherent mathematical meaning. Example: Yes or No, Sex, Race, Marital Status etc.

These can take be assigned numbers like Yes(1) and No(0), but numbers have no mathematical meaning.

- Numerical

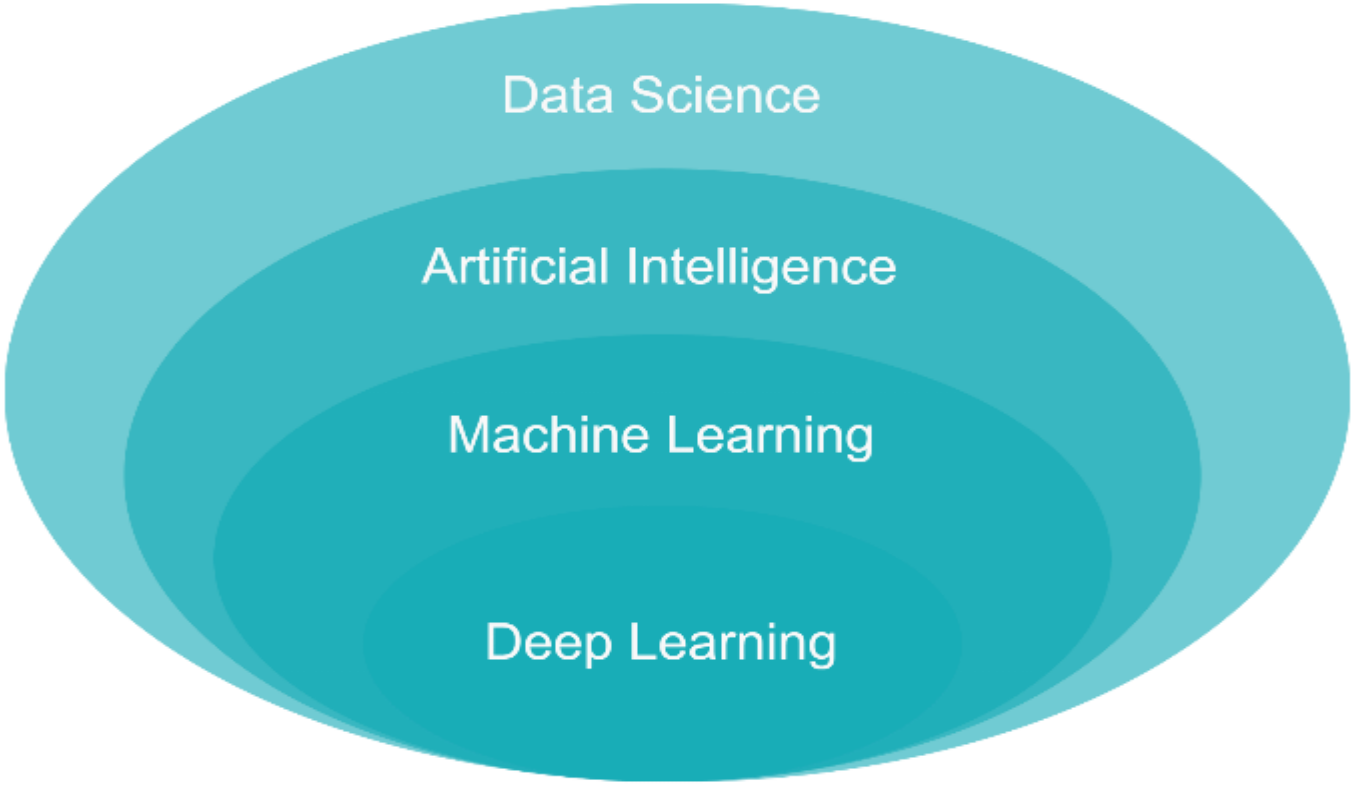
This represents some sort of quantitative measurement.

Example: height of people, stock price, page load time, etc



What is the relationship between DS, AI, ML, and DL

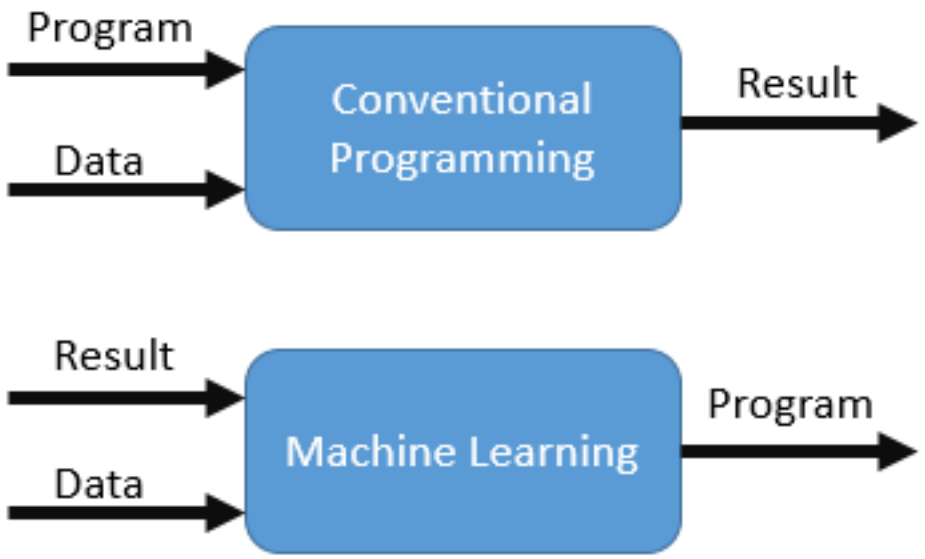
- The relationship between data science, artificial intelligence, machine learning, and deep learning



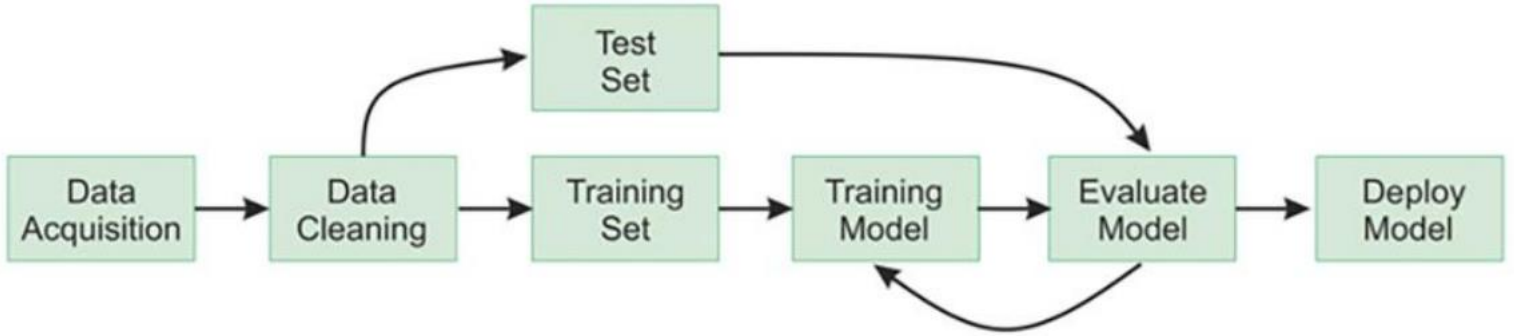
Machine Learning

- Machine learning involves computers discovering how they can perform tasks without being explicitly programmed to do so. It involves computers learning from data provided so that they carry out certain tasks

Machine Learning vs Traditional Programming



Stages of Machine Learning



Types of Machine Learning

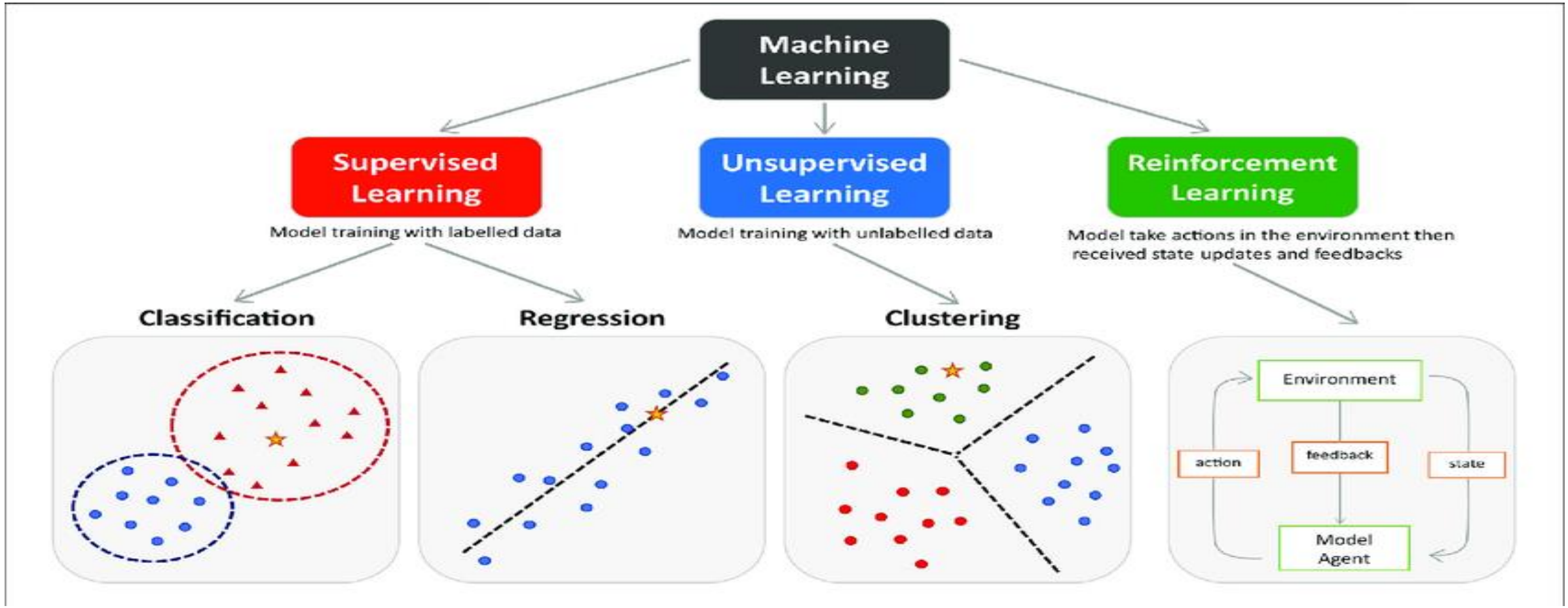
- Supervised learning : Task Driven (Classification, Regression)
- Unsupervised learning : Data Driven (Clustering)
- Reinforcement learning :

· Close to human learning.

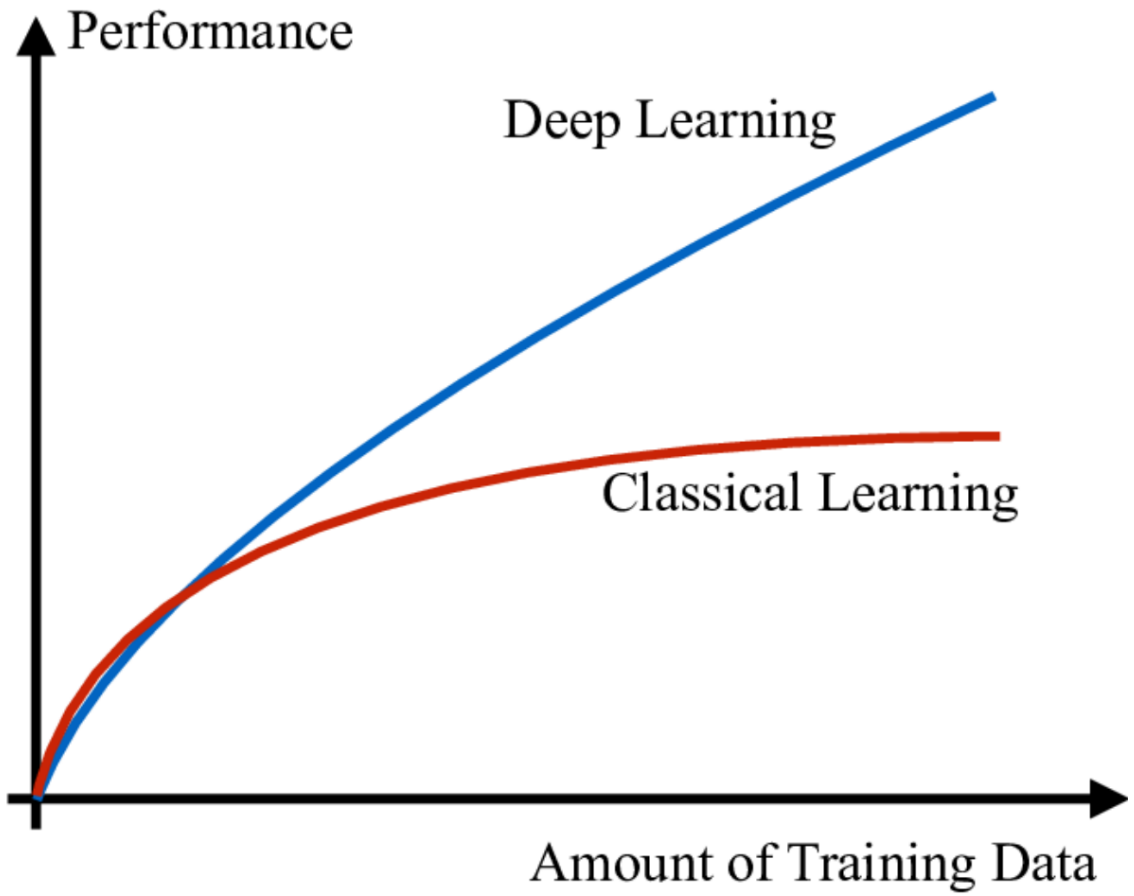
· Algorithm learns a policy of how to act in a given environment.

· Every action has some impact in the environment, and the environment provides rewards that guides the learning algorithm.

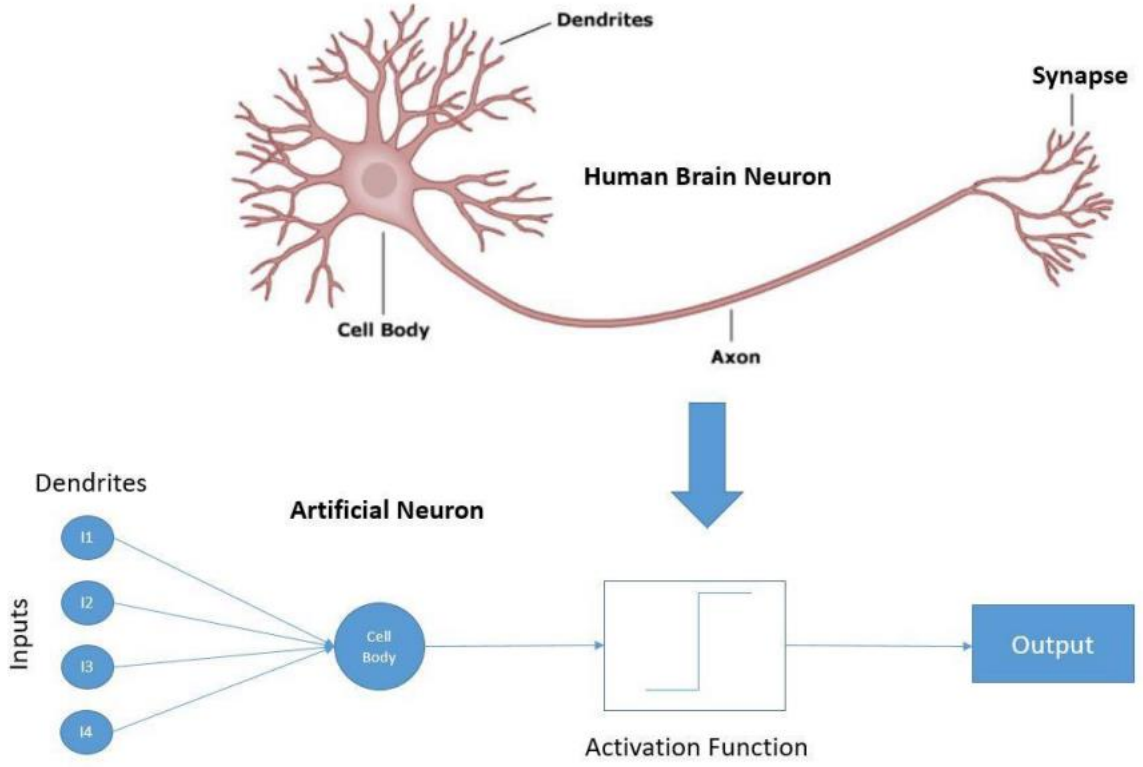
Types Machine Learning



What is Deep Learning

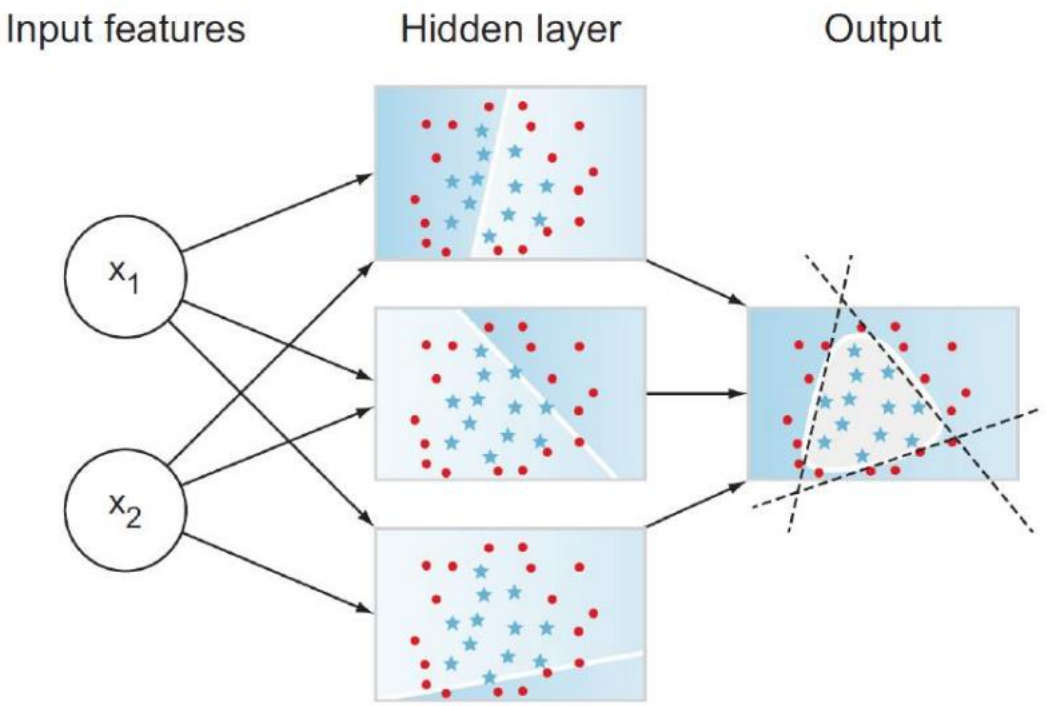


What is Deep Learning



What is Deep Learning

- A Neural Network Playground (tensorflow.org)



Deep Learning Applications

DL Today: Vision

[Krizhevsky 2012]

[Ciresan et al. 2013]

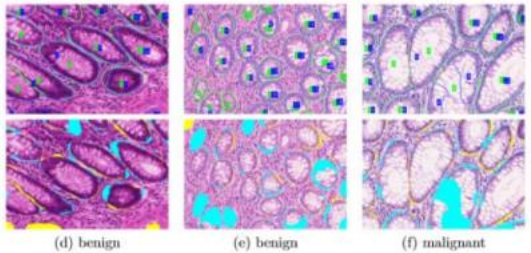
[Faster R-CNN - Ren 2015]

[NVIDIA dev blog]

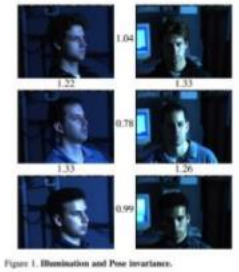
Deep Learning Applications

DL Today: Vision

[Stanford 2017]



[Nvidia Dev Blog 2017]

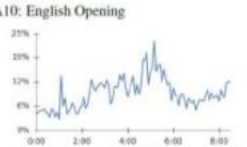


[FaceNet - Google 2015]

[Facial landmark detection CUHK 2014]

Deep Learning Applications

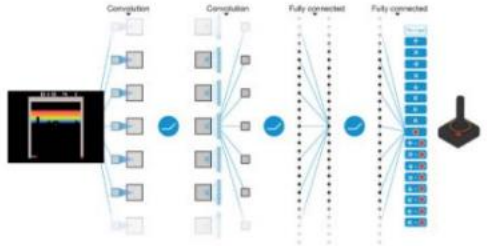
DL for AI in games



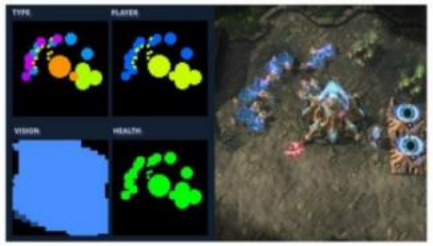
w 20/30/0, b 8/40/2

1...e5 g3 d5 cxd5 ♘f6 ♗g2 ♘xd5 ♘f3

[Deepmind AlphaGo / Zero 2017]



[Atari Games - DeepMind 2016]



[Starcraft 2 for AI research]

- Teachable Machine


ANY
QUESTIONS