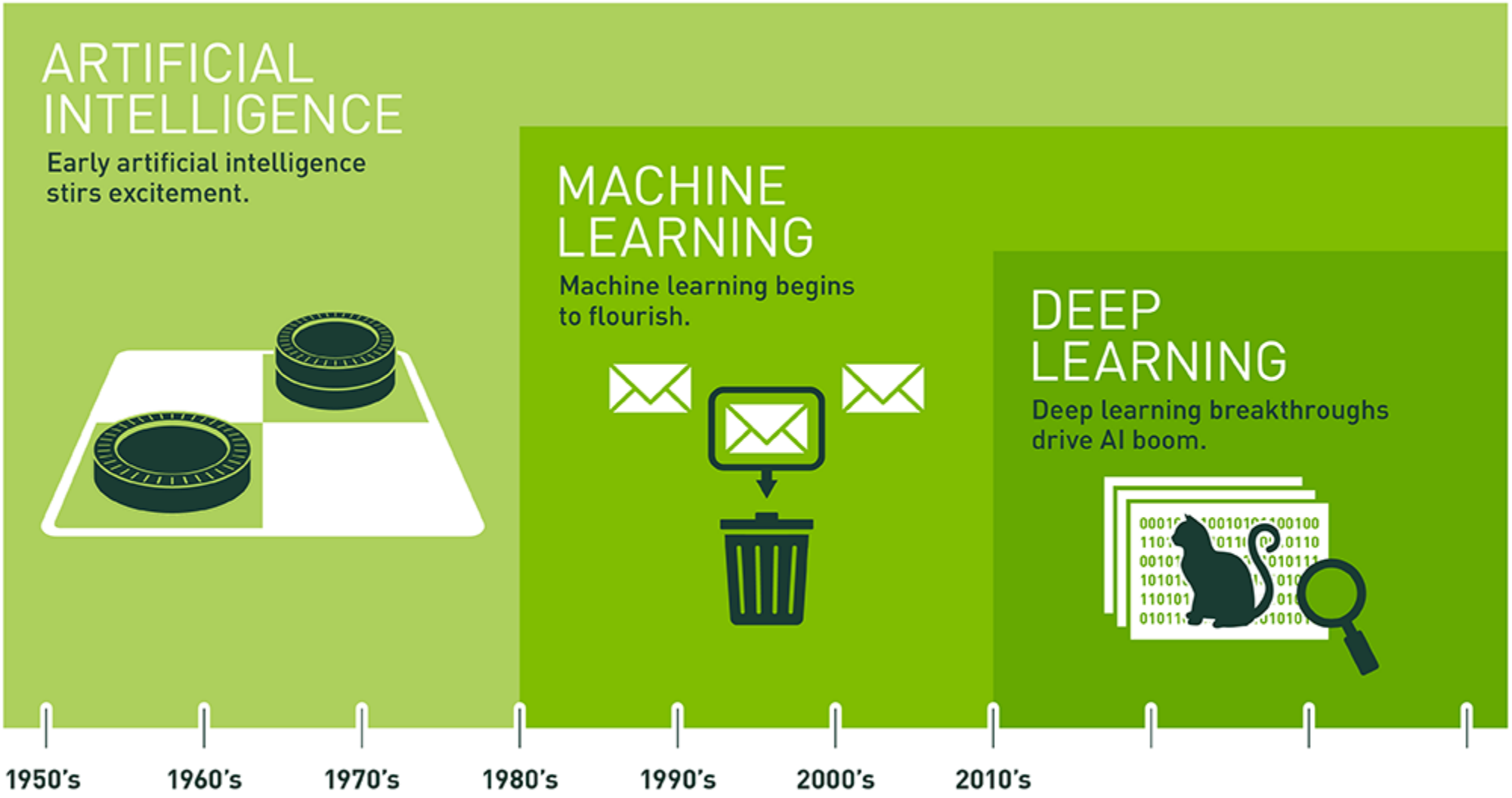
機器學習於材料資訊的應用
Machine Learning on Material Informatics

陳南佑(NAN-YOW CHEN)

楊安正(AN-CHENG YANG)



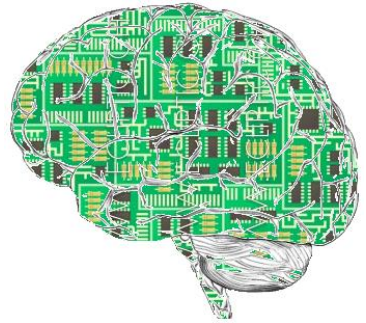
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

https://blogs.nvidia.com.tw/2016/07/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

What is AI, ML, DL?

- **AI** (Artificial Intelligence): *the ability of the computer to simulate/model human thinking processes to imitate human ability or behavior.*

  (consciousness, cognitive, mind, …)

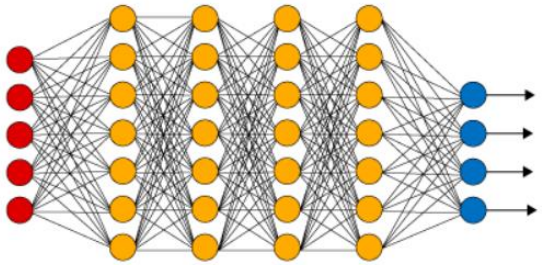


- **ML** (Machine Learning): *is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead.*

  (learning from the data, data-driven model)

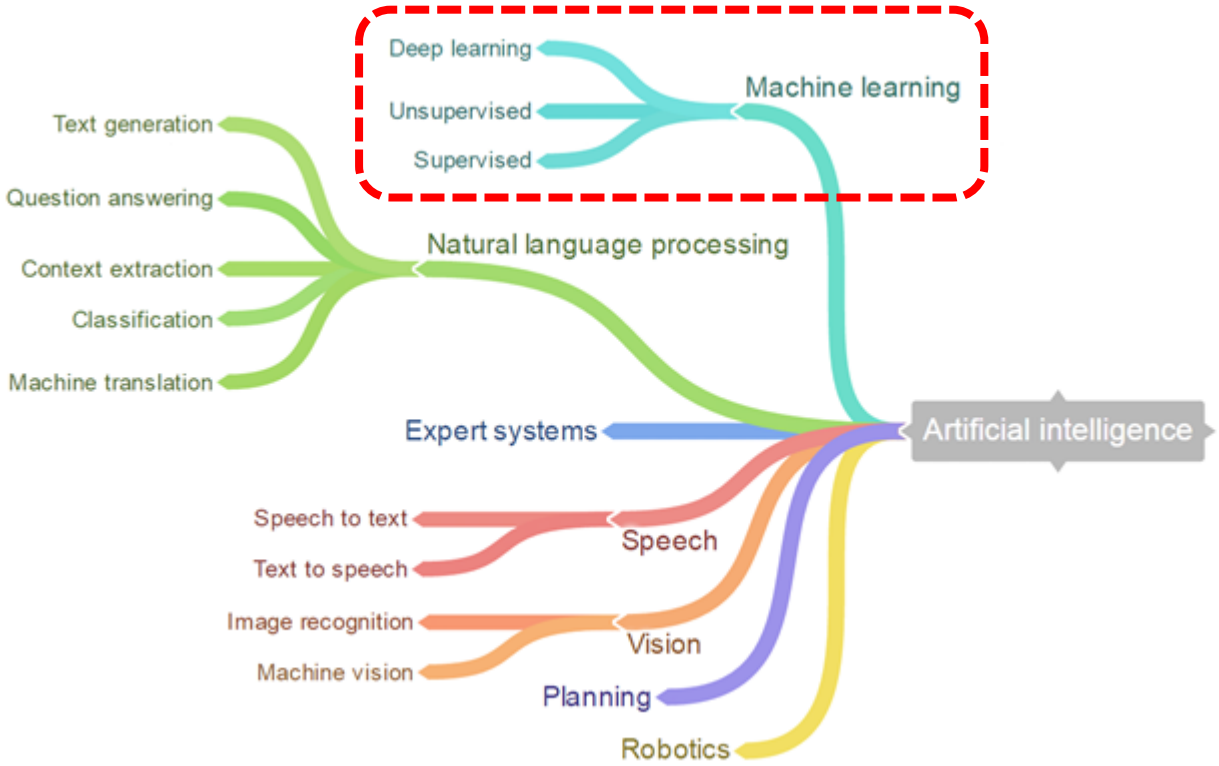
- **DL** (Deep Learning): *use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input.*

  (learning from the data, multi-layer neural network)



https://zh.wikipedia.org/wiki/

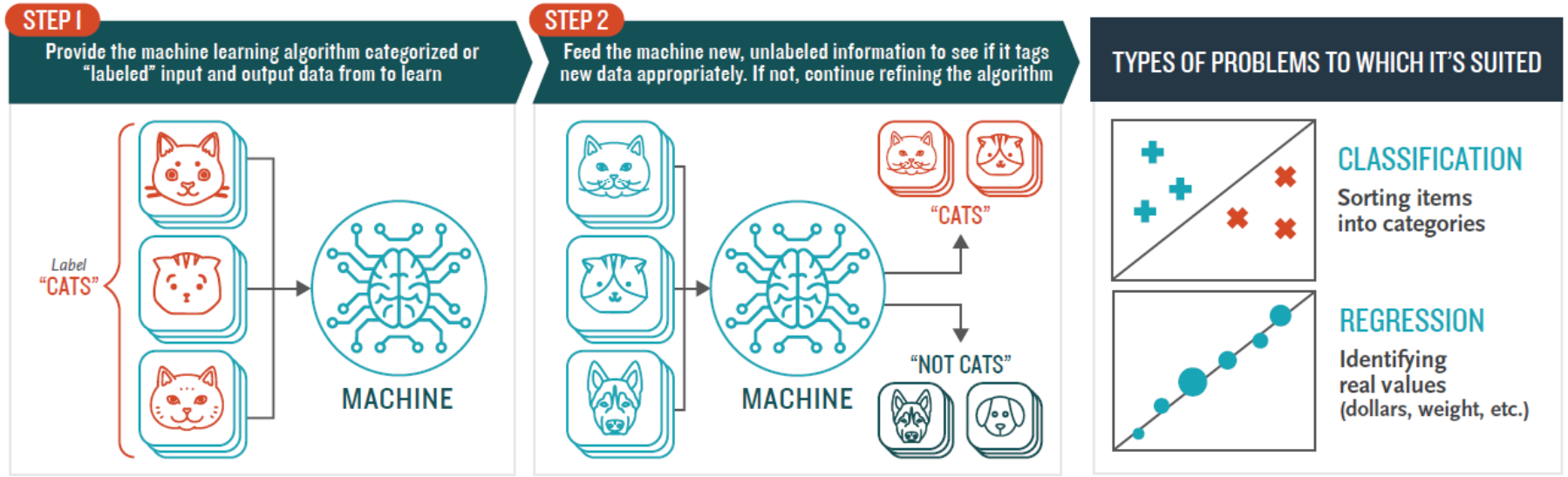
Applications of AI



https://hackernoon.com/jump-start-to-artificial-intelligence-f6eb30d624ec

Types of ML

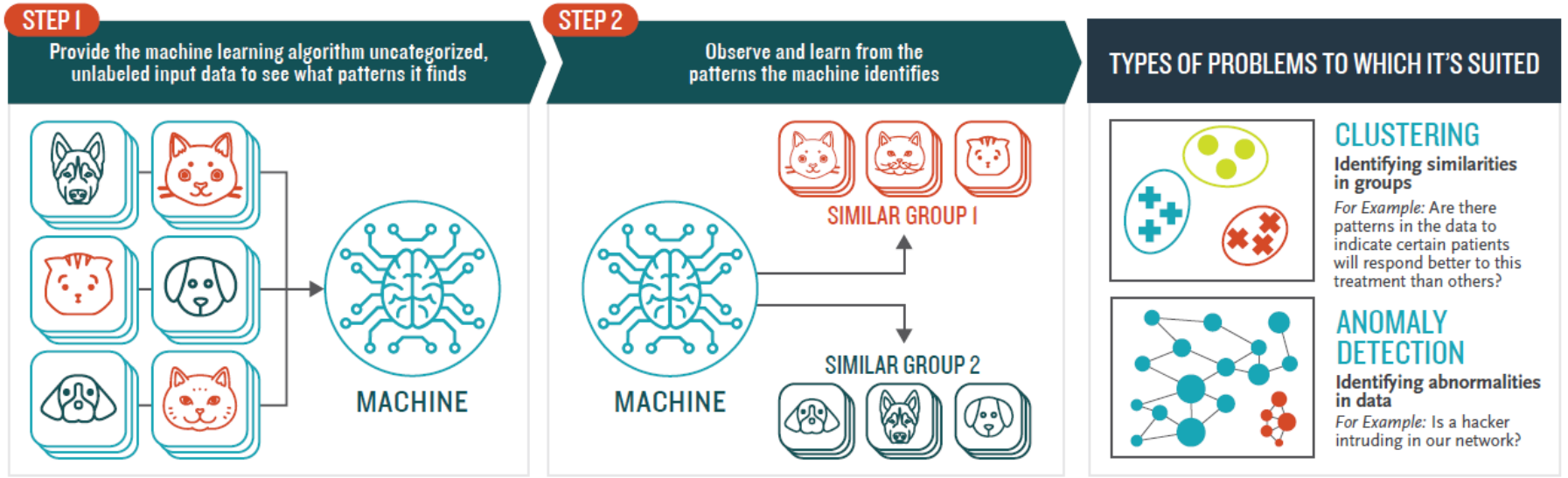
- Supervised Learning (監督式學習):
  - 可以由訓練資料中學到或建立一個模式(函數 / learning model)，並依此模式推測新的實例。訓練資料是由輸入物件(通常是向量)和預期輸出所組成。函數的輸出可以是一個連續的值(稱為迴歸分析)，或是預測一個分類標籤(稱作分類)。
  - 人工神經網路(Artificial Neural Network, ANN)、決策樹(Decision Tree)、支援向量機(Support Vector Machine, SVM)、...



https://zh.wikipedia.org/wiki/ & https://towardsdatascience.com/wtf-is-machine-learning-a-quick-guide-39457e49c65b

Types of ML

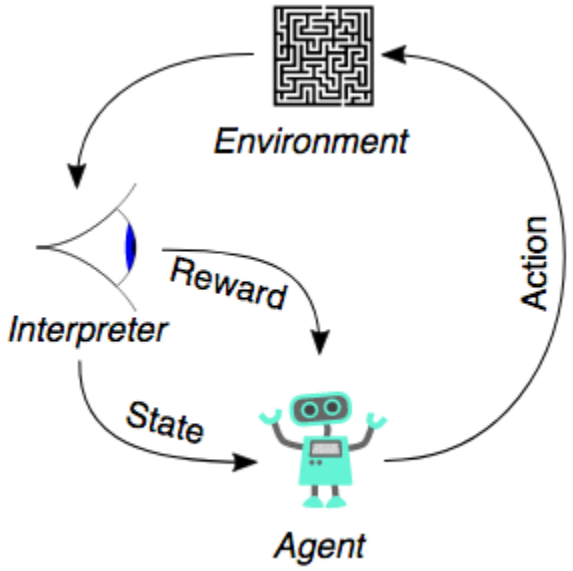
- Unsupervised Learning (非監督式學習):
  - 沒有給定事先標記過的訓練範例，自動對輸入的資料進行分類或分群。
  - 分群(Cluster Analysis)、關聯規則(Association Rule)、維度縮減(Dimensionality Reduce) 、…



https://zh.wikipedia.org/wiki/ & https://towardsdatascience.com/wtf-is-machine-learning-a-quick-guide-39457e49c65b

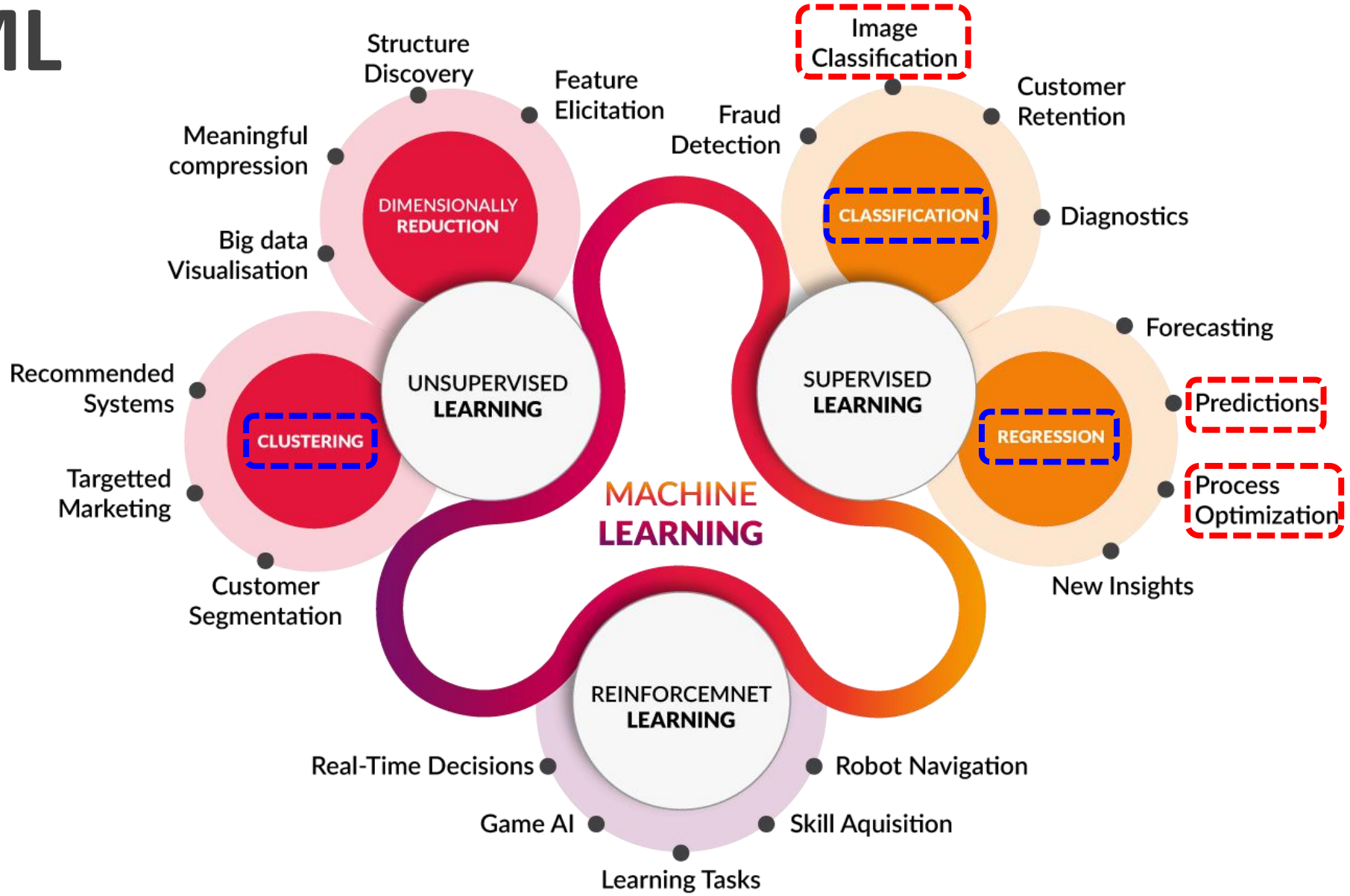
Types of ML

- Reinforcement Learning (**強化學習**):
  - 強調如何基於環境而行動，以取得最大化的預期利益。其靈感來源於心理學中的行為主義理論，即有機體如何在環境給予的獎勵或懲罰的刺激下，逐步形成對刺激的預期，產生能獲得最大利益的習慣性行為。



https://zh.wikipedia.org/wiki/ & https://towardsdatascience.com/wtf-is-machine-learning-a-quick-guide-39457e49c65b

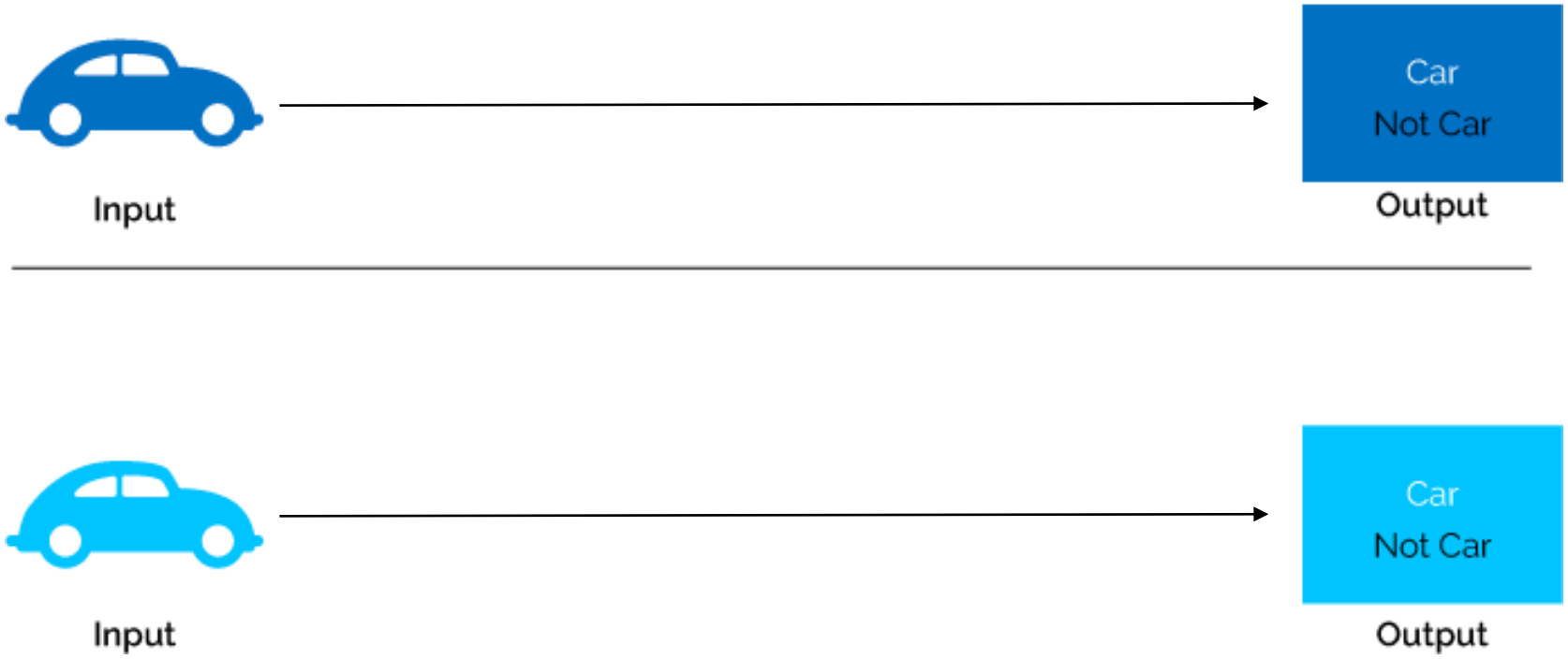
Types of ML



http://www.cognub.com/index.php/cognitive-platform/

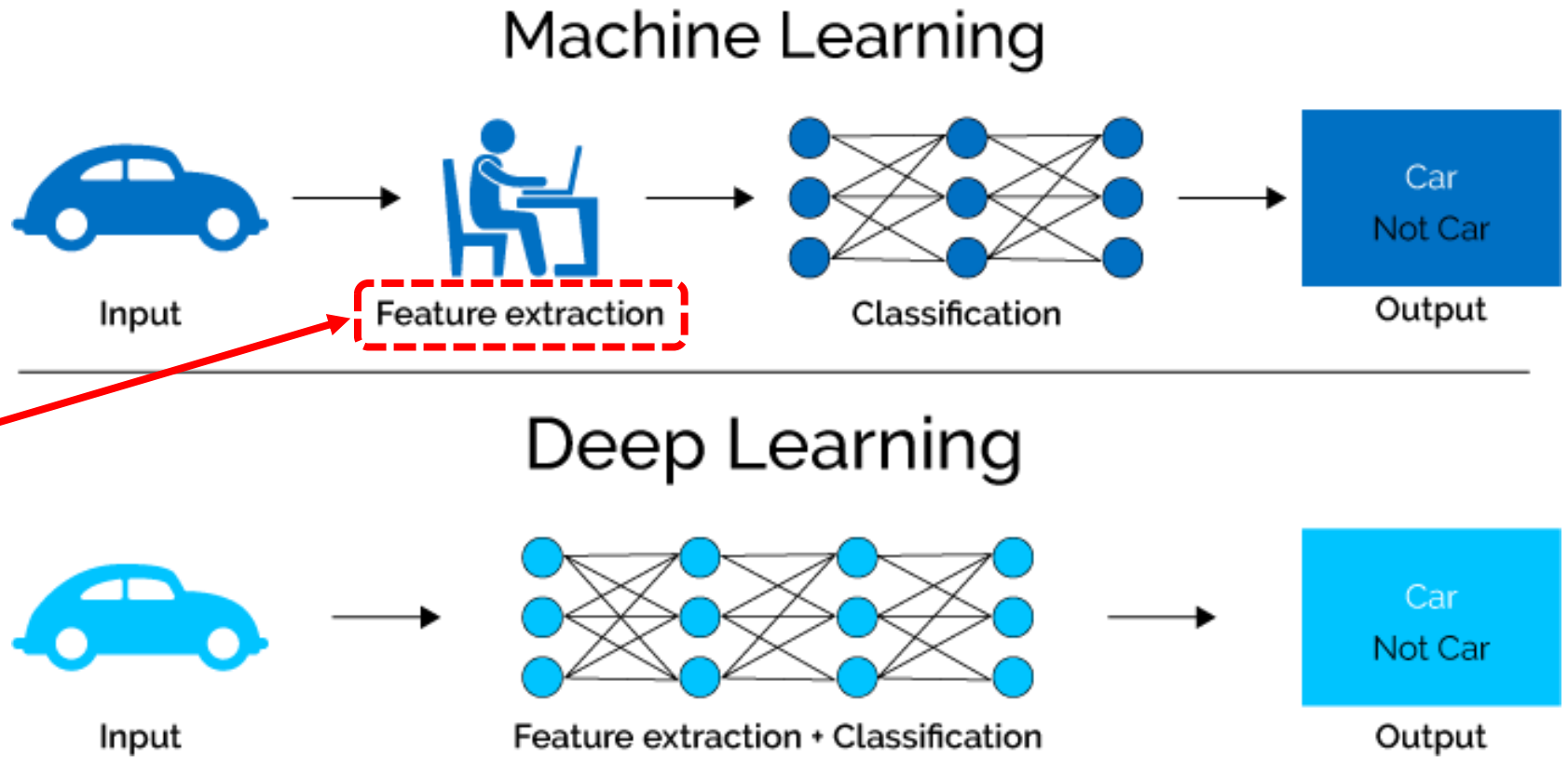
What the same between ML & DL?

Need labeling data for training



https://semiengineering.com/deep-learning-spreads/

What difference between ML & DL?



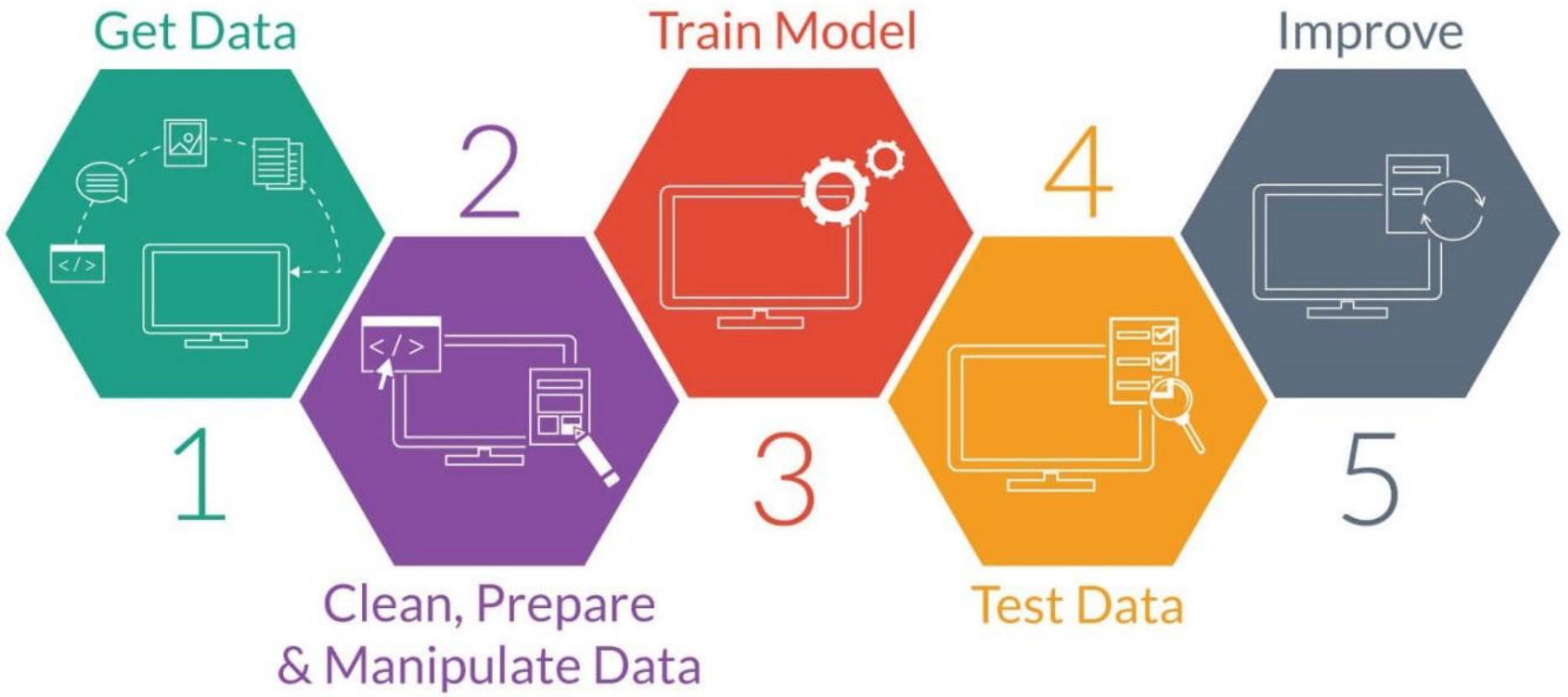
Feature engineering: the process of using **domain knowledge** of the data to create features that make machine learning algorithms work.

Labeling data are fewer

Labeling data are more

https://semiengineering.com/deep-learning-spreads/

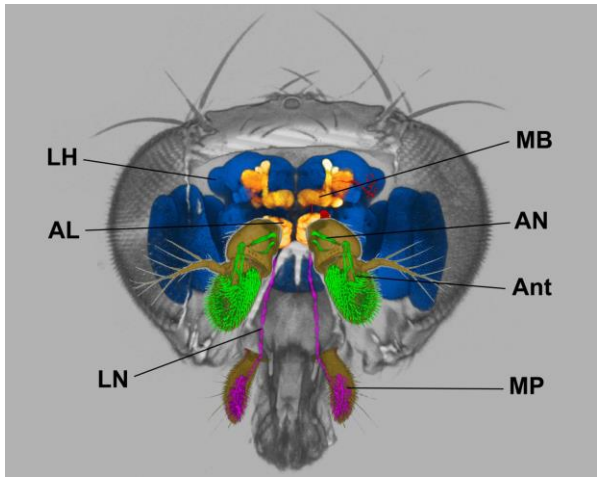
Work flow for ML



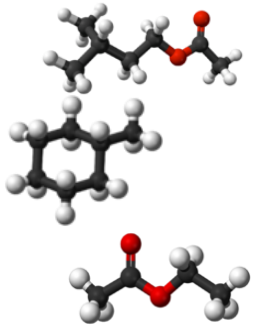
https://towardsdatascience.com/building-a-deployable-ml-classifier-in-python-46ba55e1d720

What is neuron?

Example 1: 成年果蠅的嗅覺系統

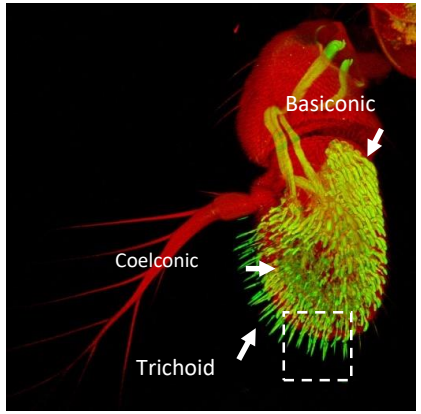


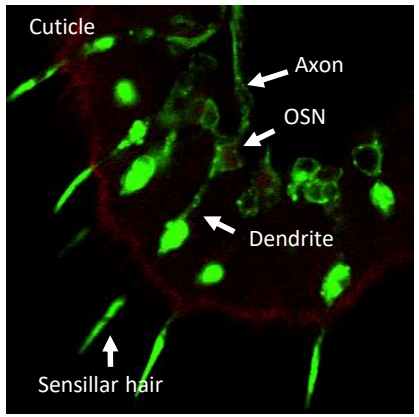
Odorant molecules



Input

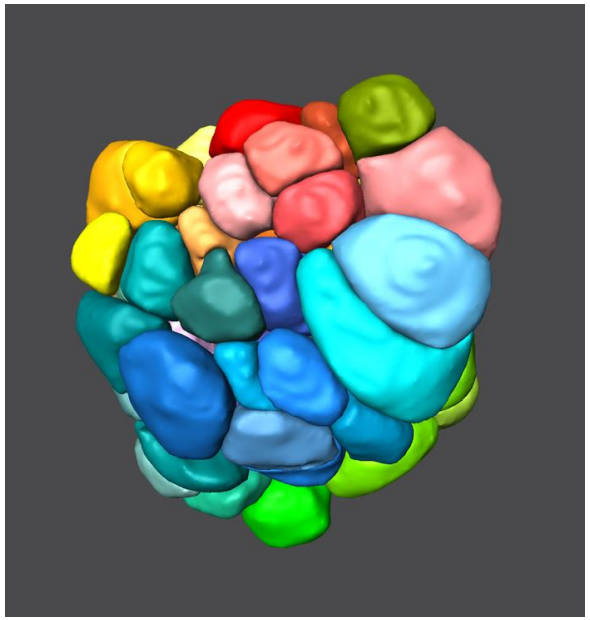
Olfactory Sensory Neurons





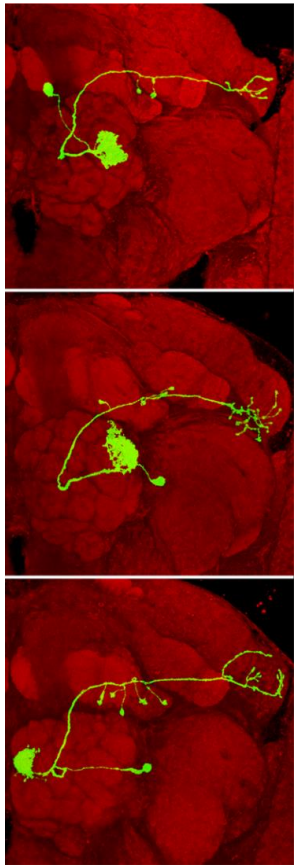
Antennal Lobe

Local processing unit for olfactory information



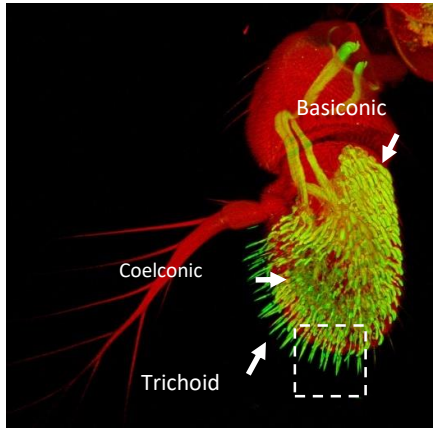
Output

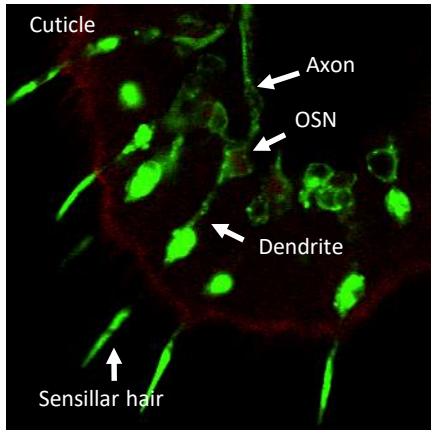
Projection Neurons



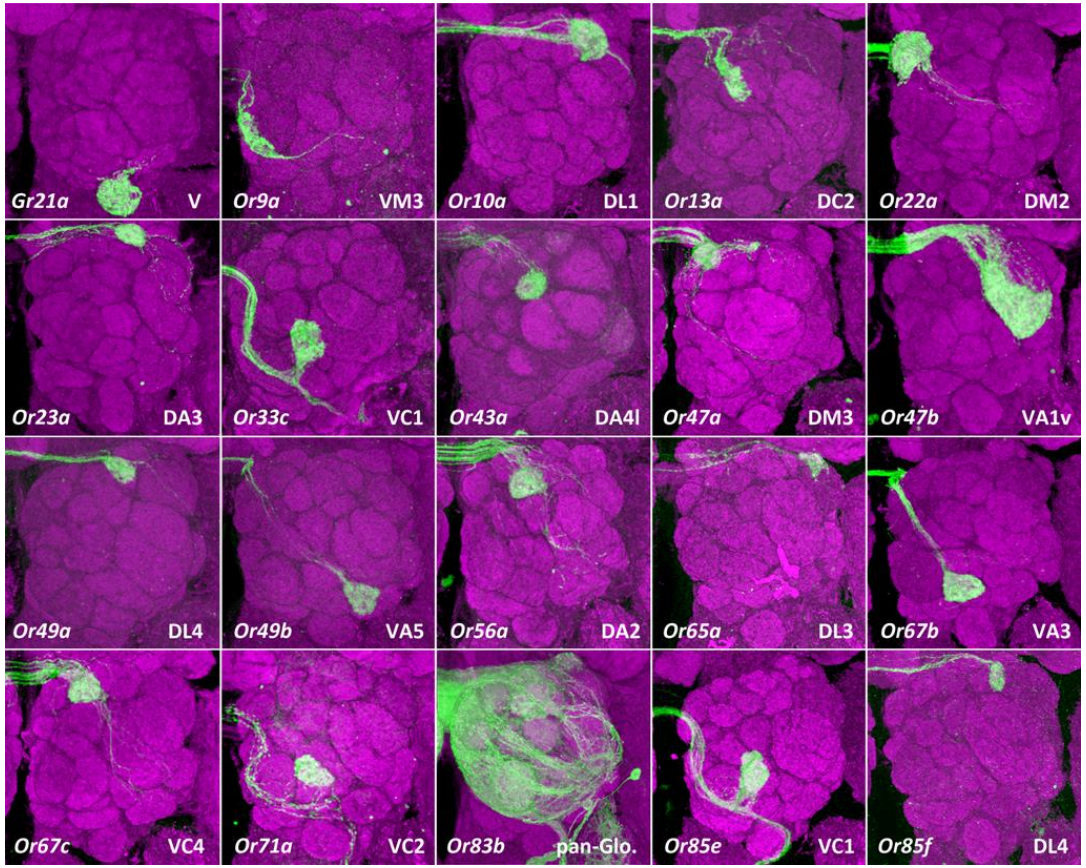
AL Input – Olfactory Sensory Neurons

~1300 OSNs expressing 62 odorant receptors convey olfactory information from antennae and maxillary pulps to ~50 glomeruli





“one neuron – one receptor”

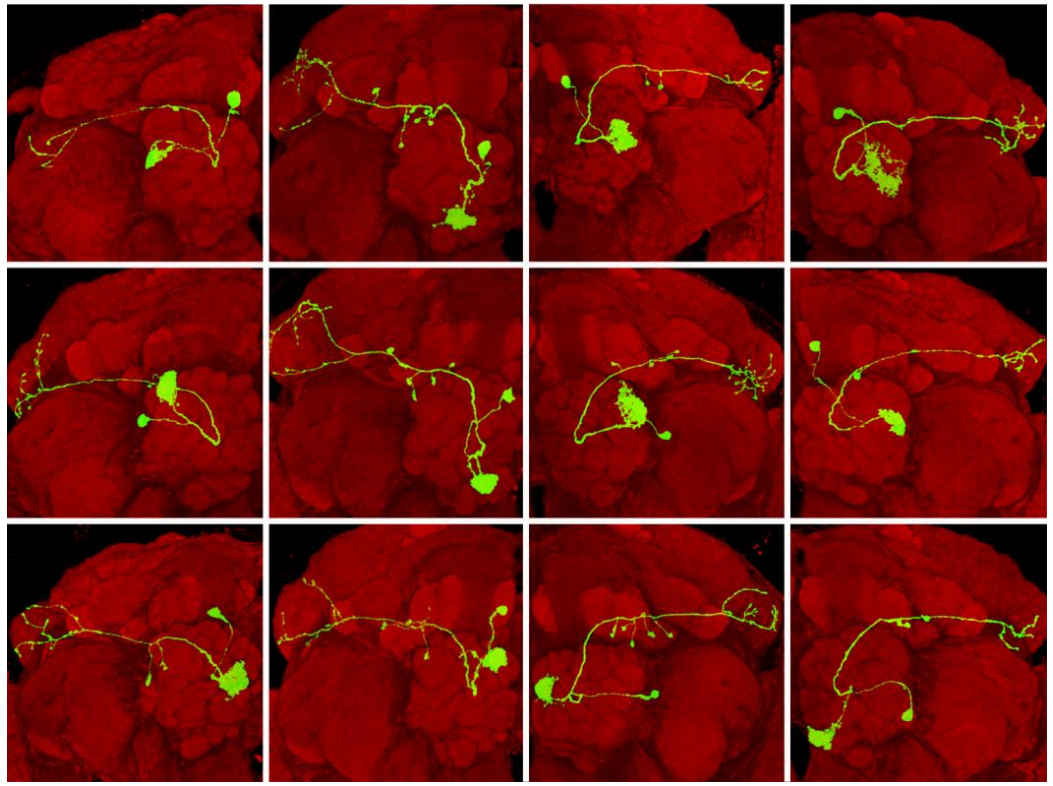


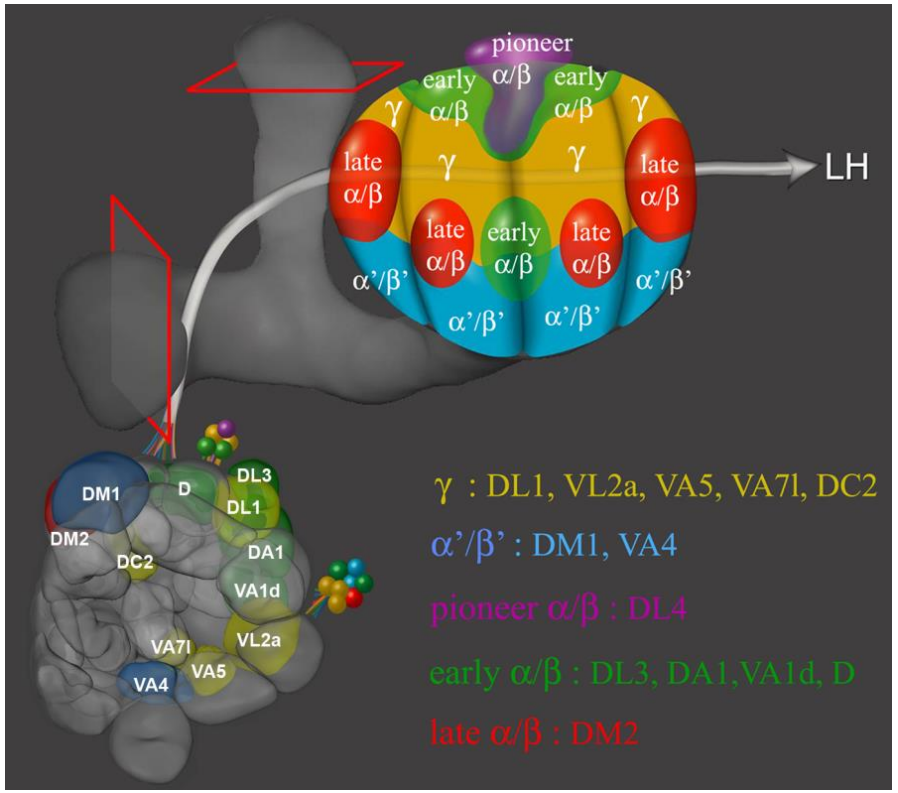
“one glomerulus – one receptor”

AL Output – Projection Neurons

~150 typical PNs convey olfactory information from antennal lobe to higher brain centers

Stereotyped PN terminals

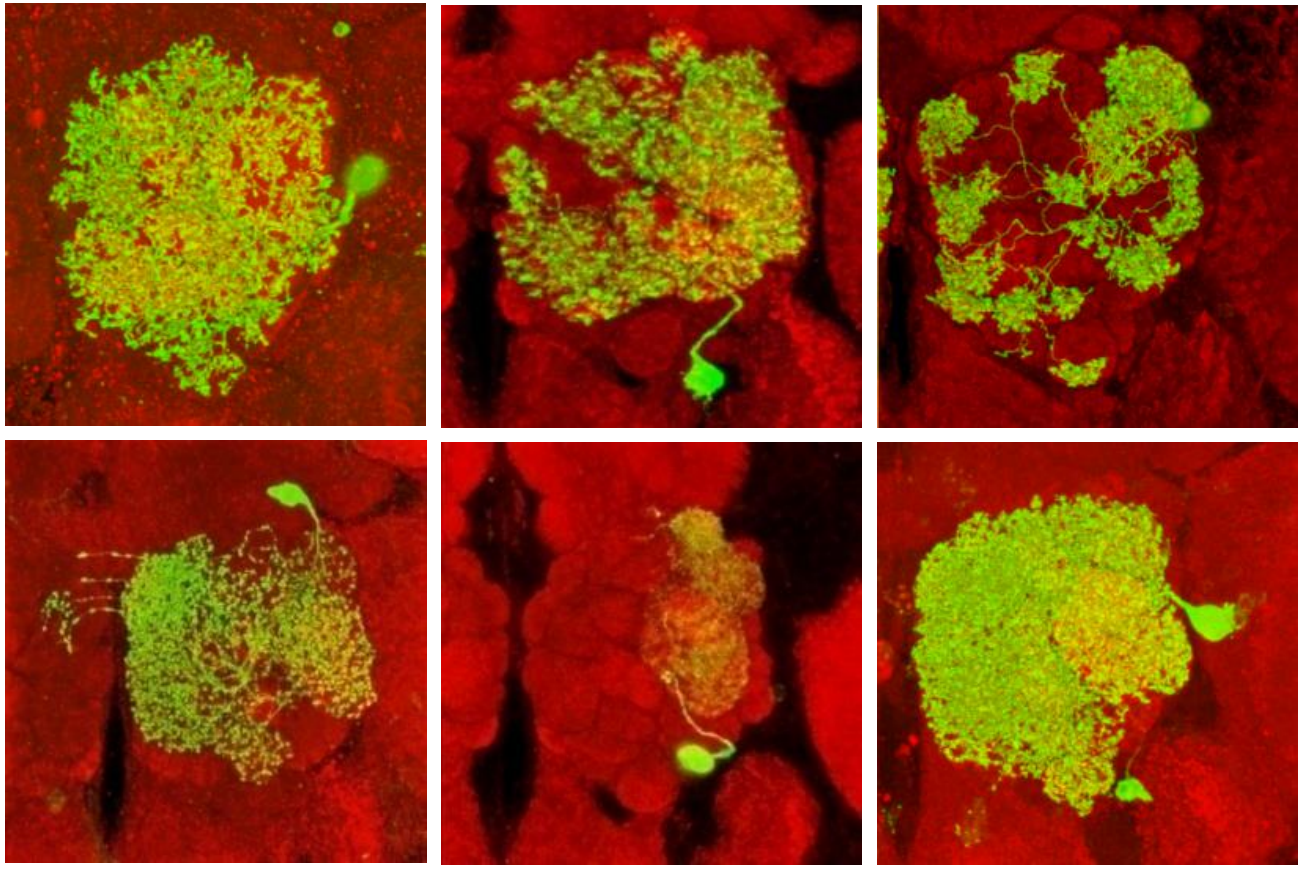




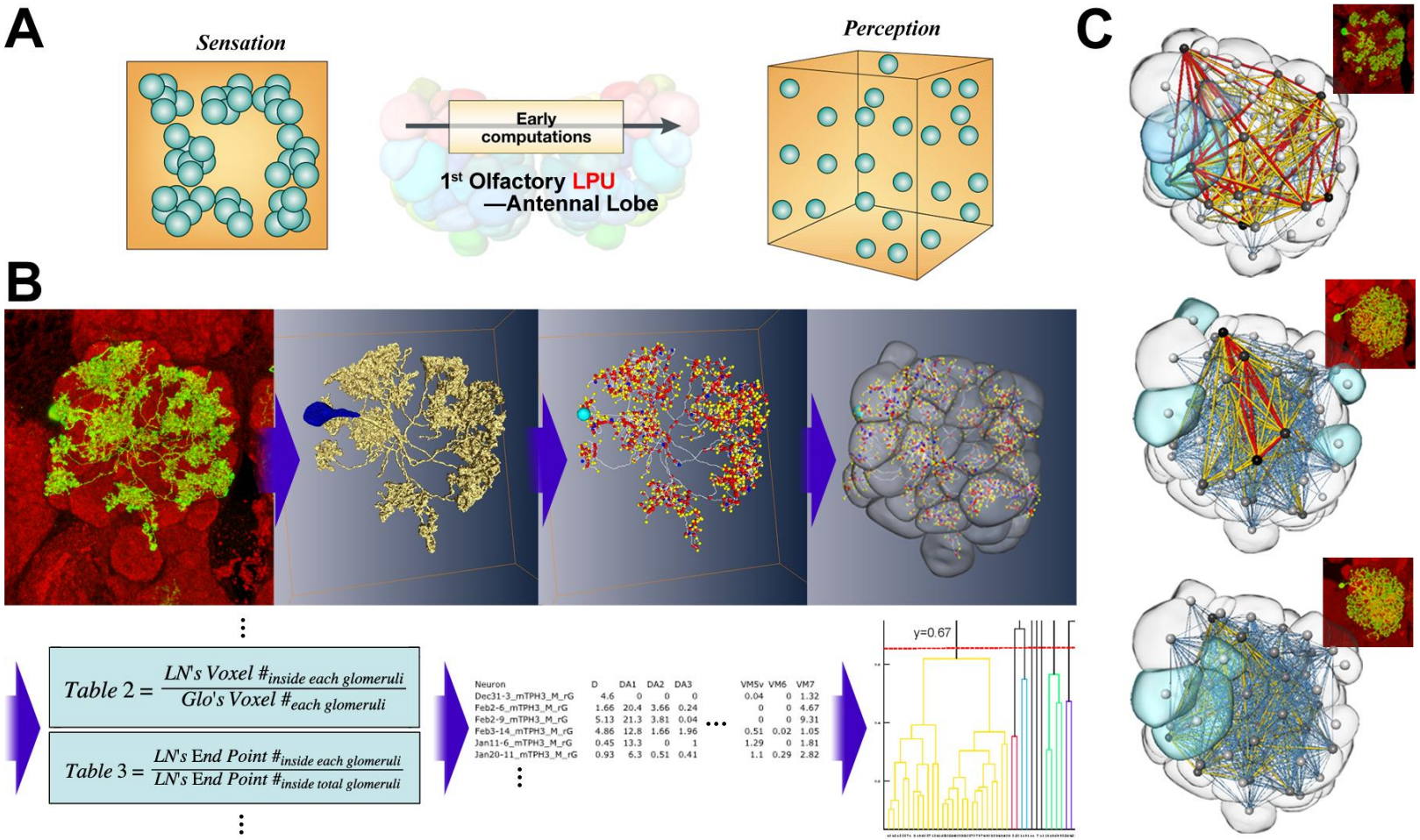
H.-H. Lin, J. S.-Y. Lai, A.-L. Chin, Y.-C. Chen and A.-S. Chiang, Cell **128**, 1205 (2007)

Antennal Lobe – Local Neurons

Local processing unit for olfactory information

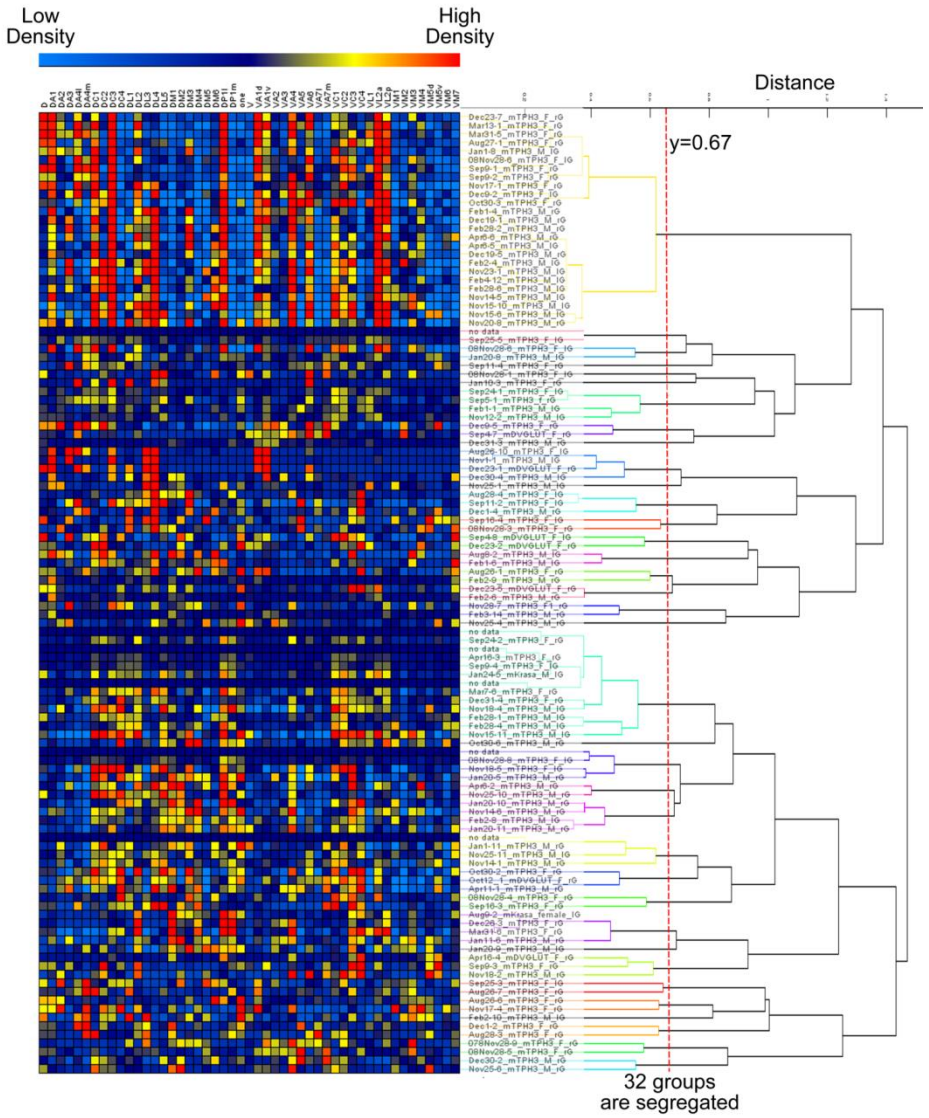


建構果蠅嗅覺神經元連結網路的程序

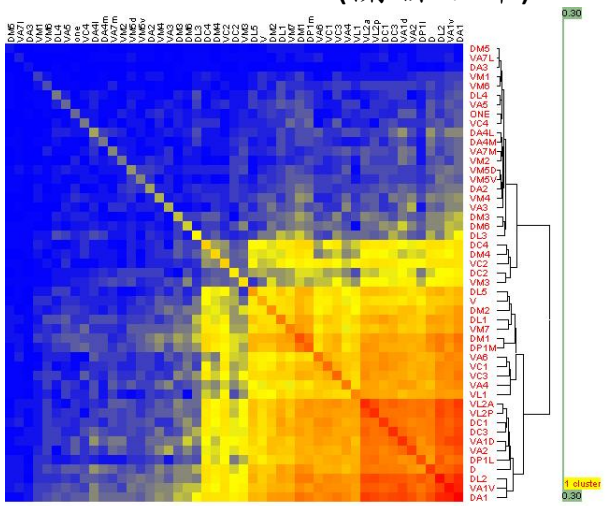


果蠅嗅覺神經元連結網路

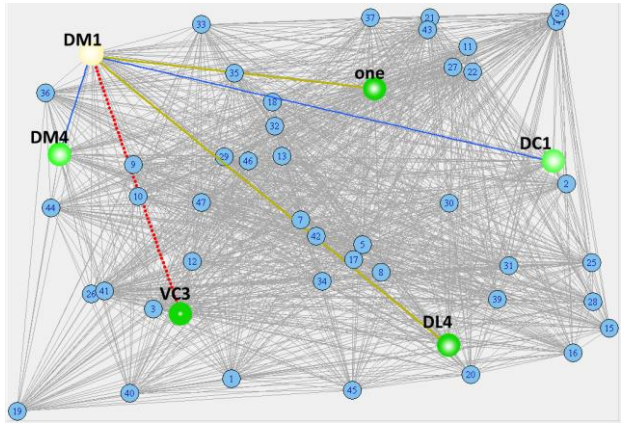
Innervation table (神經元分佈表)



Association matrix (關聯矩陣)



Network of local neurons
(局域神經元連結網路)



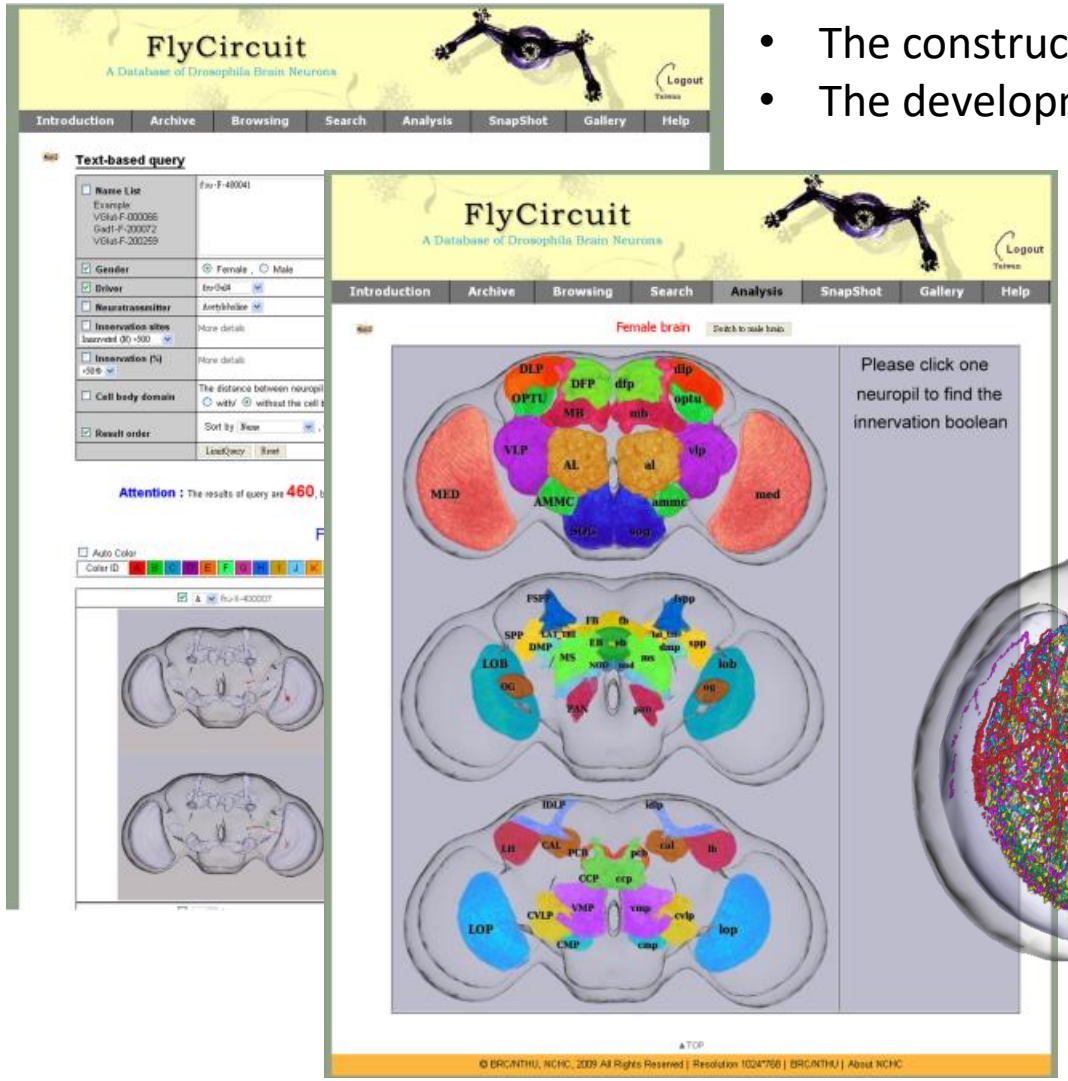
Association matrix A for each ALLN image

If A_{ij} < threshold, A_{ij} = 0 $\quad (A_{ij} = Table2_i \cdot Table2_j)$

Example 2：*FlyCircuit* v1.0

An Image Database of *Drosophila* Brain Neurons

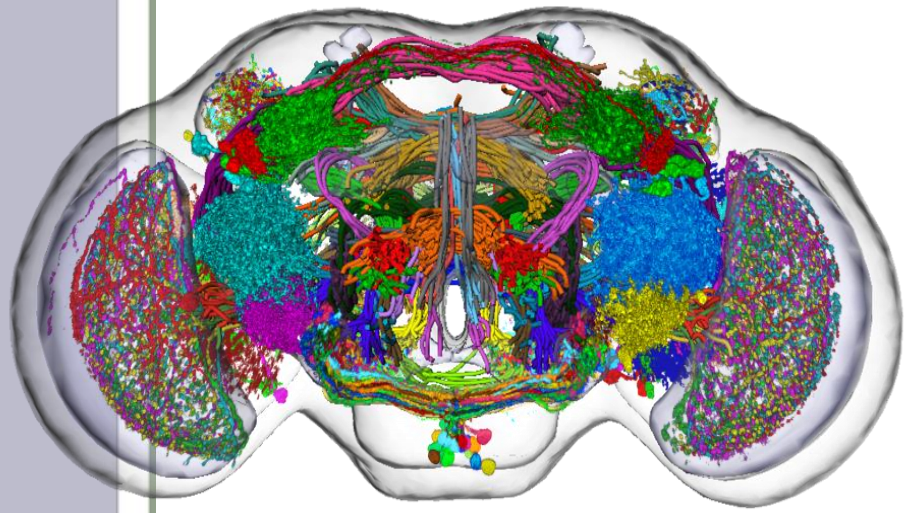
3-D map of fly brain is to neuroscience what genome is to genetics



- The construction of 3D neuron images database
- The development of searching and browsing tool

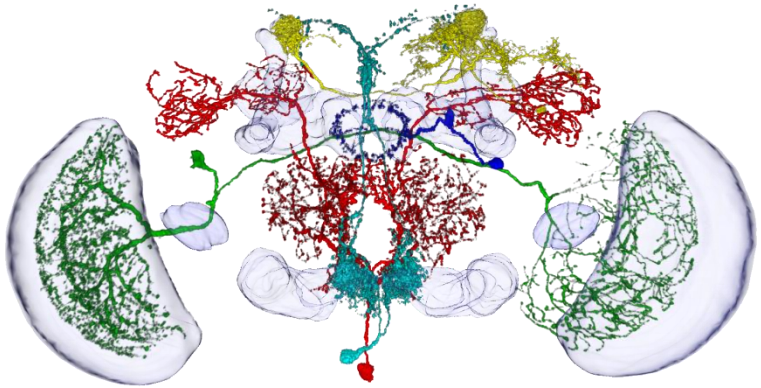
Current status:30,000 neurons

(23.5% of 130,000 total brain)

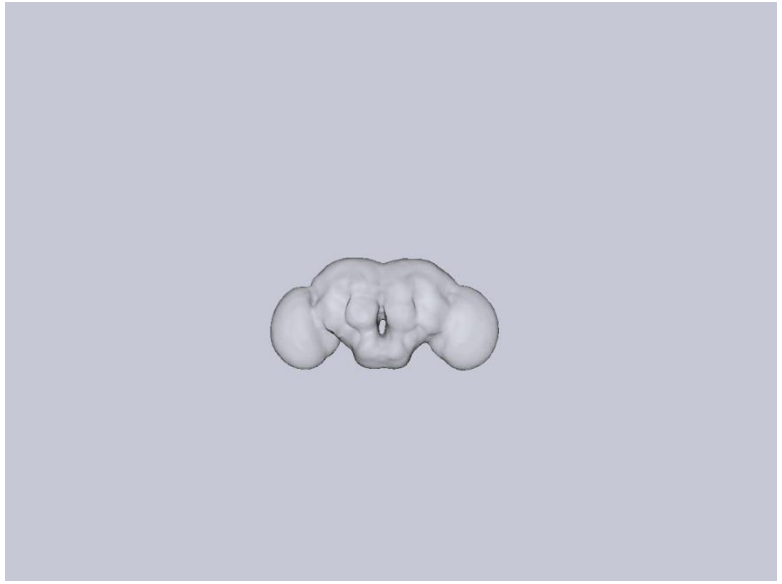


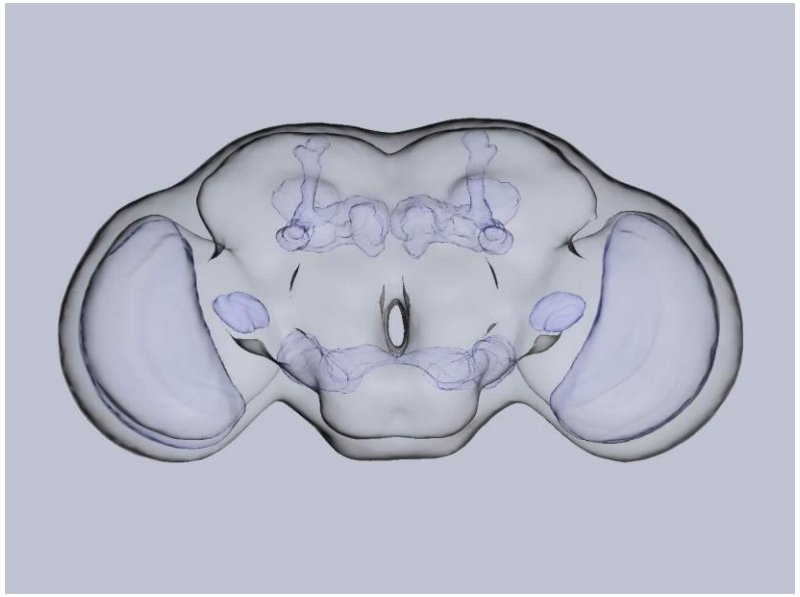
Current Biology **21**:1–11 (2011)

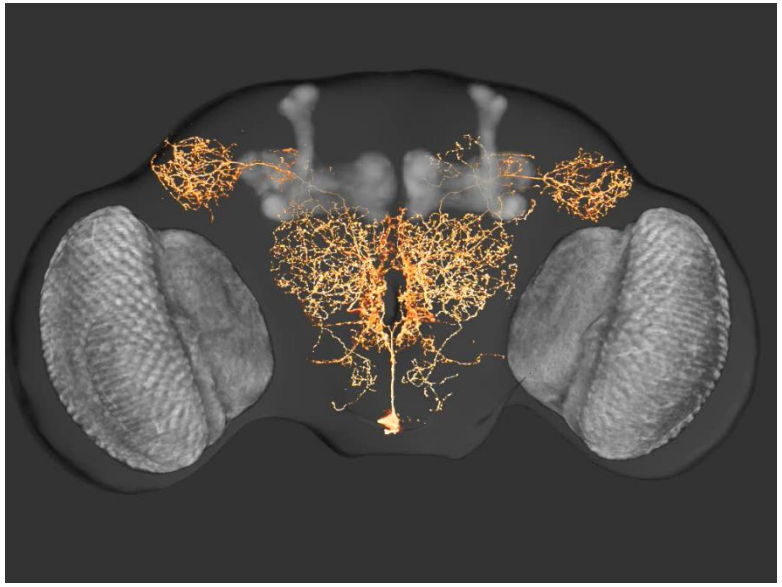
3D Compilation & Visualization



VGlut TpH TH GAD Tdc2



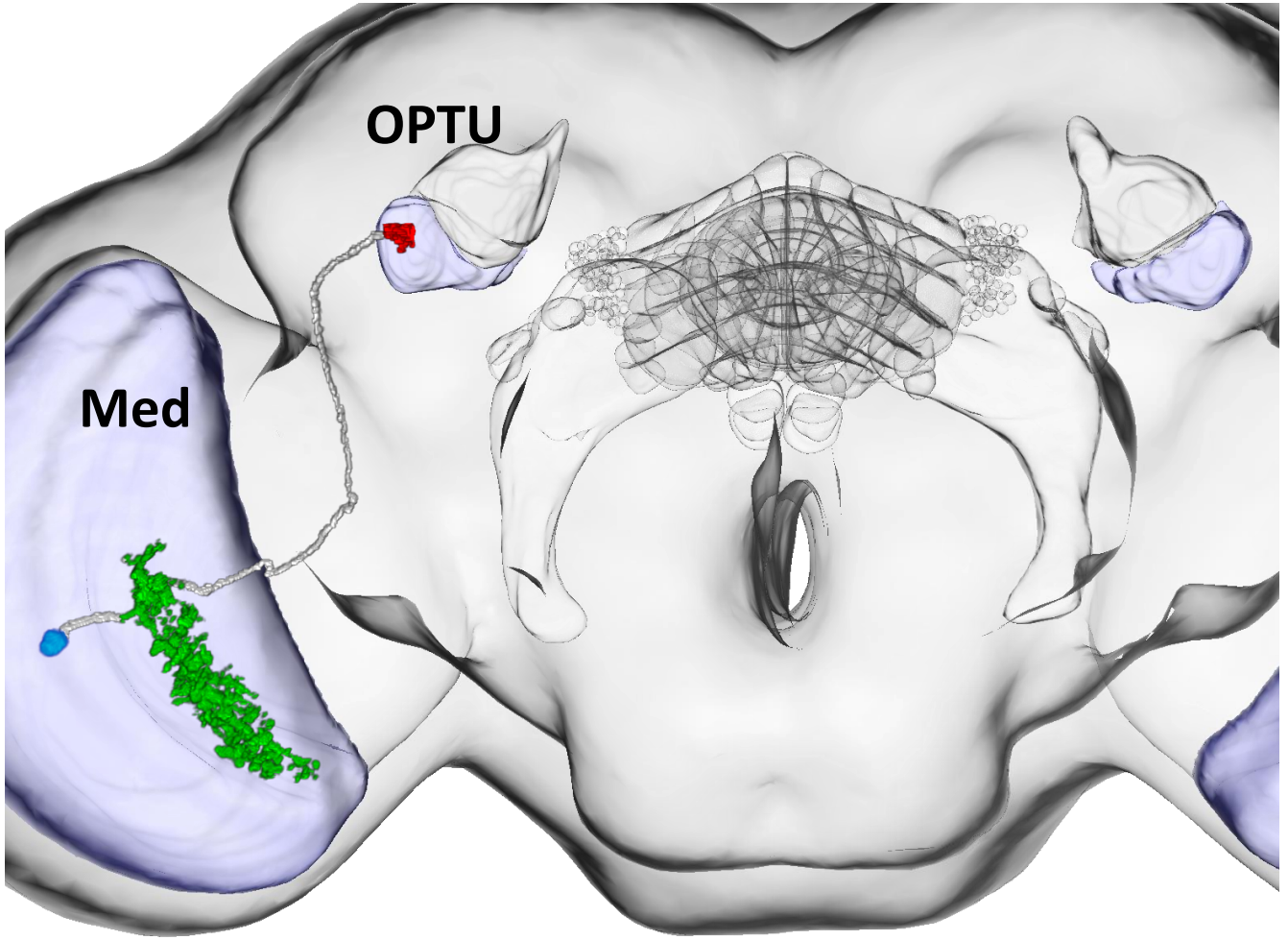




FlyCircuit v2.0

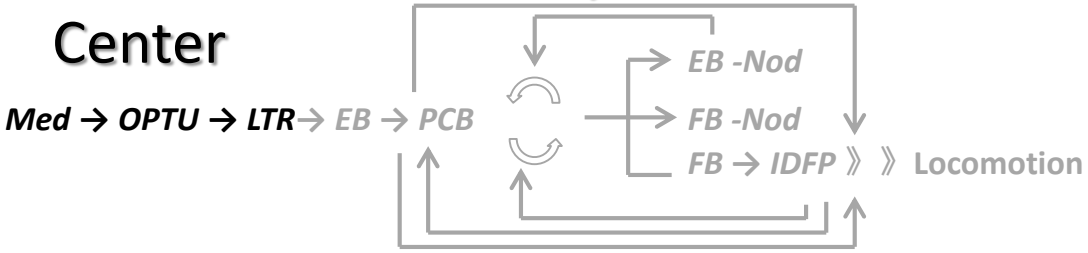
Brain-wide wiring networks in the *Drosophila* Brain

Mapping the neural network

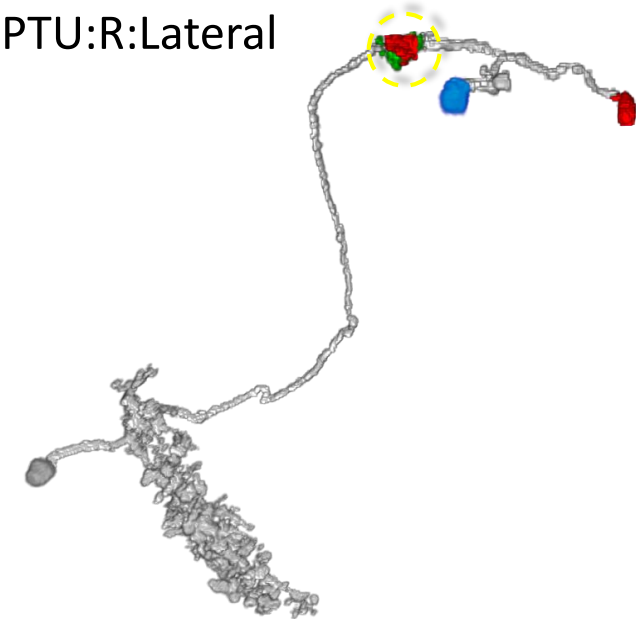


Visual input ➡ Med:R→OPTU:R:LaterI

An UV Circuit from Eye to Pre-motor Center



OPTU:R:Lateral

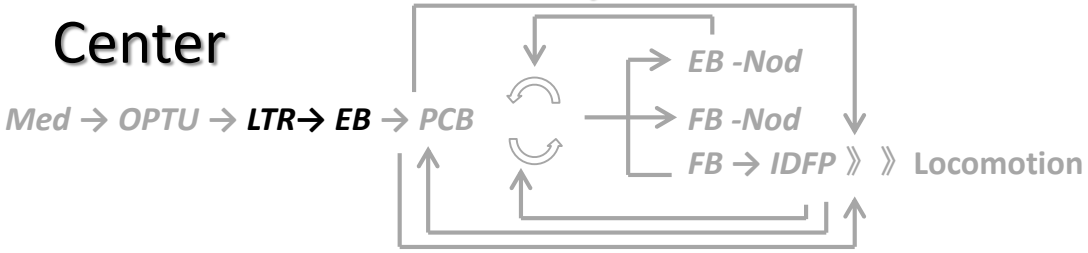


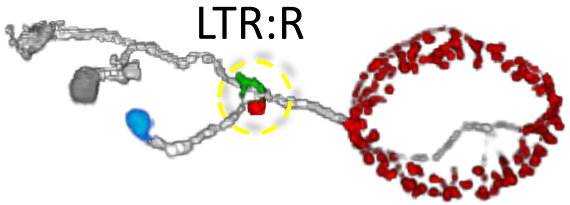
Med:R→OPTU:R:Laterl

↓

OPTU:R:Laterl→LTR:R

An UV Circuit from Eye to Pre-motor Center



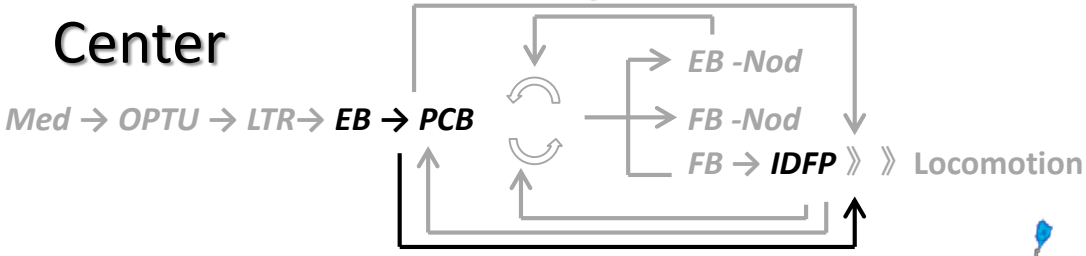


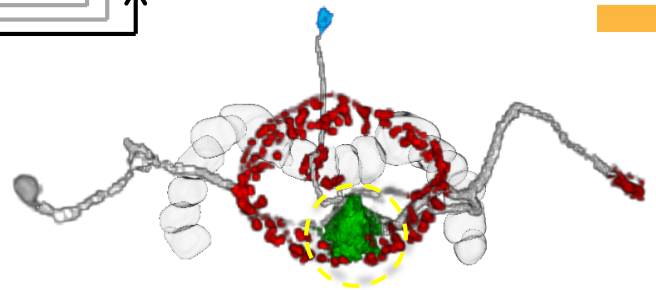
OPTU:R:Laterl→**LTR:R**

↓

LTR:R→EB:R:O1~8/L:O1~8

An UV Circuit from Eye to Pre-motor Center





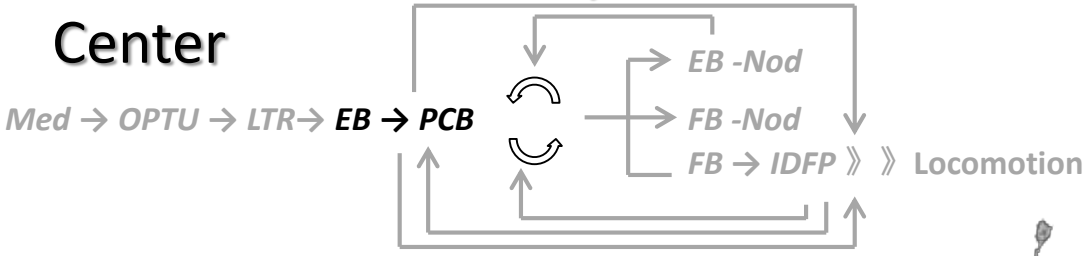
EB:L:O8

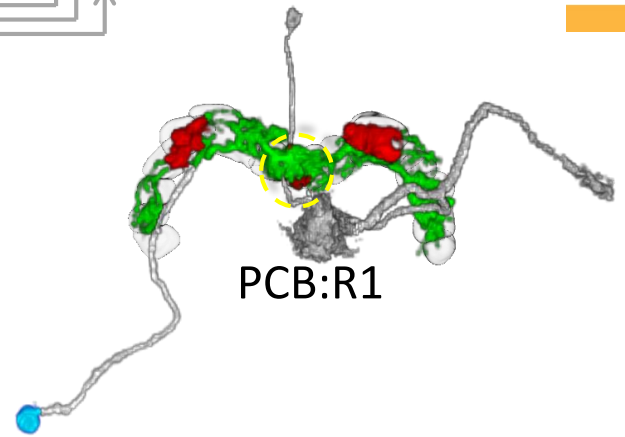
LTR:R→EB:R:O1~8/L:O1~8

↓

Class II PCB:R1←/EB:L:C8,P8,A8,O8→IDFP:L:DSB

An UV Circuit from Eye to Pre-motor Center

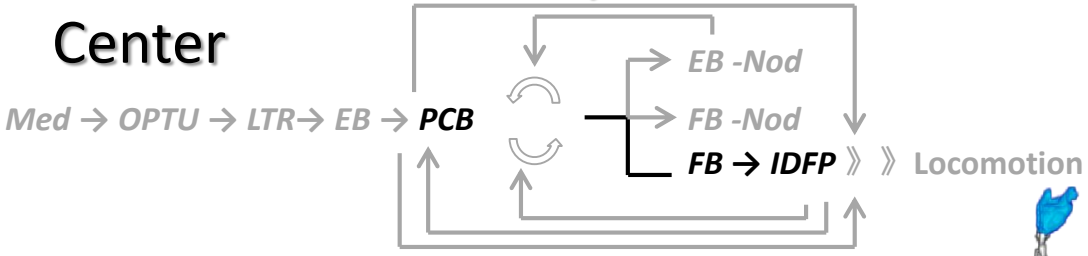


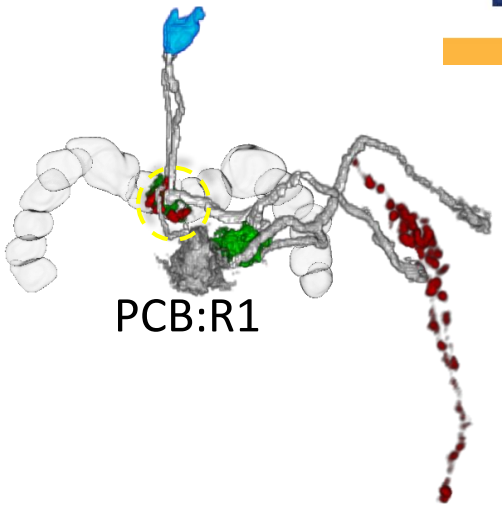


Class II **PCB:R1**←/EB:L:C8,P8,A8,O8→IDFP:L:DSB

Class I **PCB:R1~3,5~8/L1,2,4~7→PCB:R4/L3**

An UV Circuit from Eye to Pre-motor Center

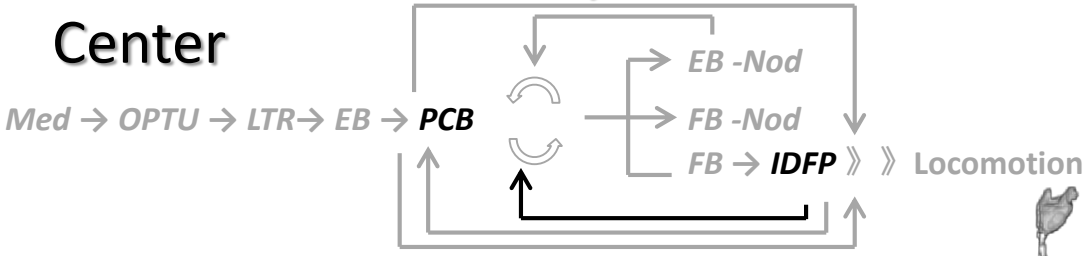


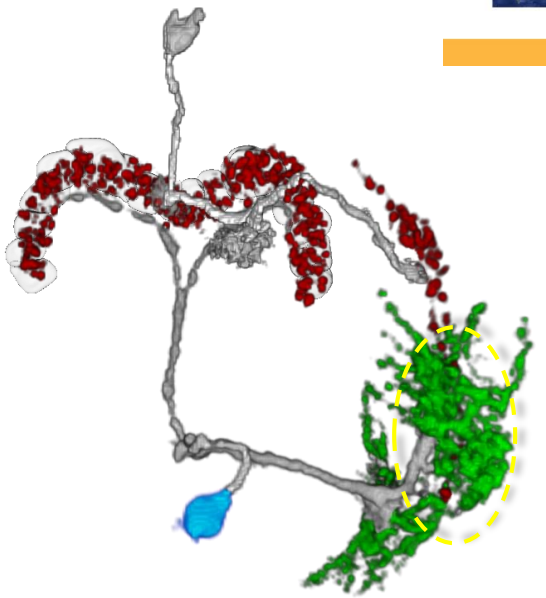


Class II **PCB:R1←/EB:L:C8,P8,A8,O8→IDFP:L:DSB**

Class VI **PCB:R1-/FB:A5→IDFP:L:LLAL**

An UV Circuit from Eye to Pre-motor Center



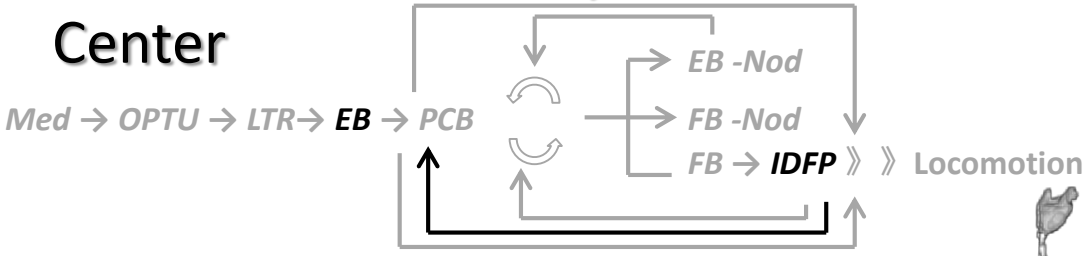


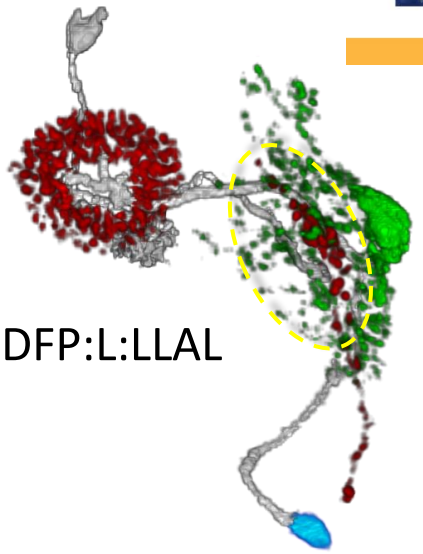
IDFP:L:LLAL

Class VI PCB:R1-/FB:A5→IDFP:L:LLAL

Class III CVLP:L-VMP:L-IDFP:L:LLAL→PCB:R1~8/L1~8

An UV Circuit from Eye to Pre-motor Center



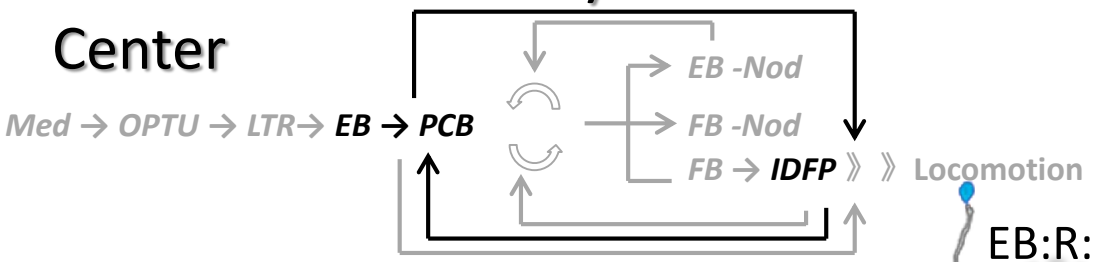


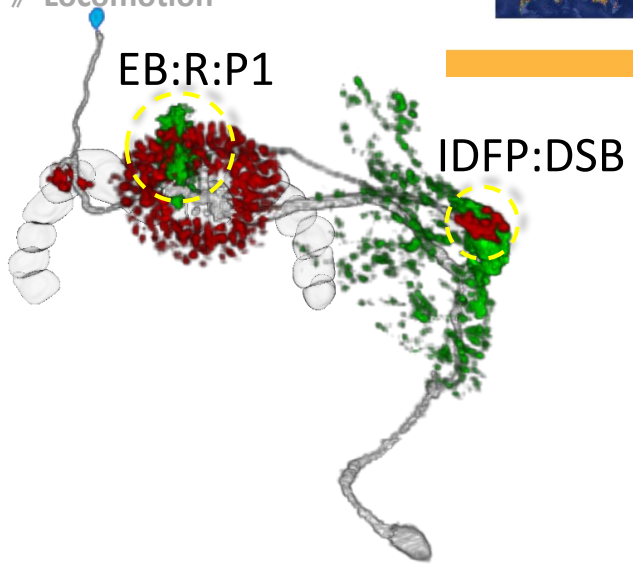
IDFP:L:LLAL

Class VI PCB:R1-/FB:A5→IDFP:L:LLAL

IDFP:L:LLAL,MLAL,DSB,VSB→ EB:R:P1~8/L:P1~8

An UV Circuit from Eye to Pre-motor Center

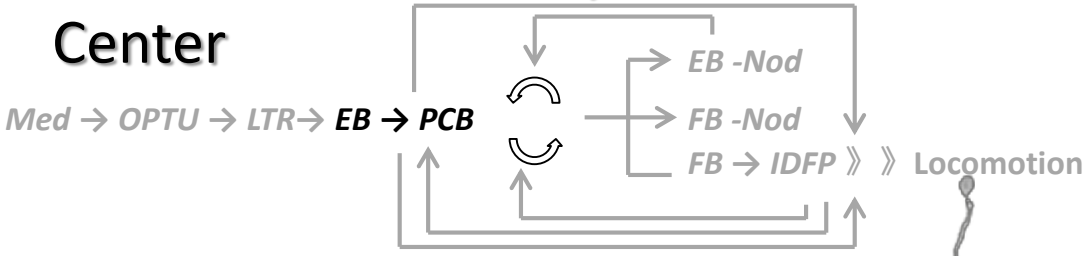




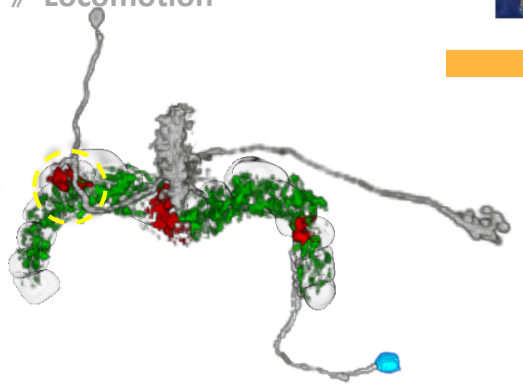
IDFP:L:LLAL,MLAL,DSB,VSB→ EB:R:P1~8/L:P1~8

Class II **PCB:R4←EB:R:C1,P1,A1,O1→/IDFP:L:DSB**

An UV Circuit from Eye to Pre-motor Center



PCB:R4

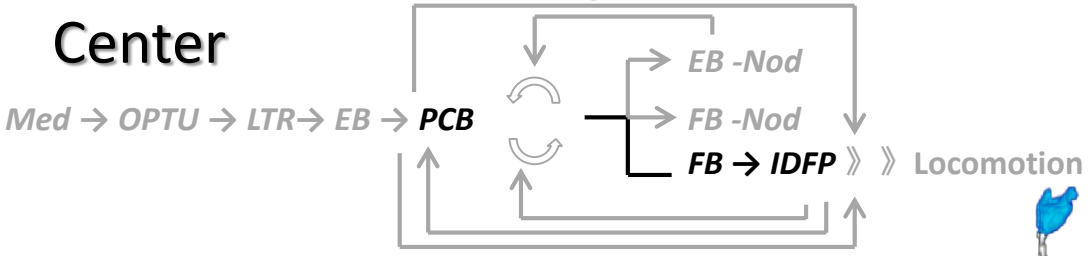


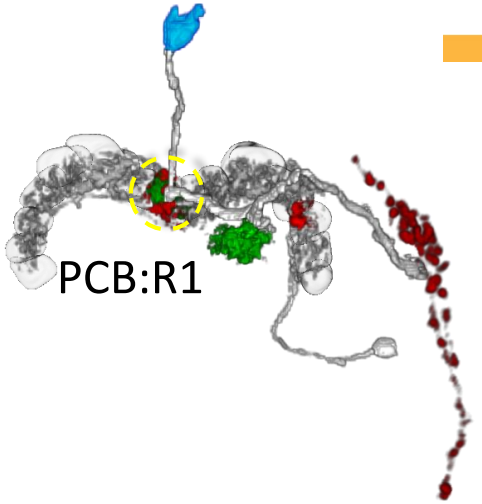
Class II **PCB:R4**←EB:R:C1,P1,A1,O1→/IDFP:L:DSB

↓

Class III **PCB:R2~7/L1~5,7,8→PCB:R1/L6**

An UV Circuit from Eye to Pre-motor Center

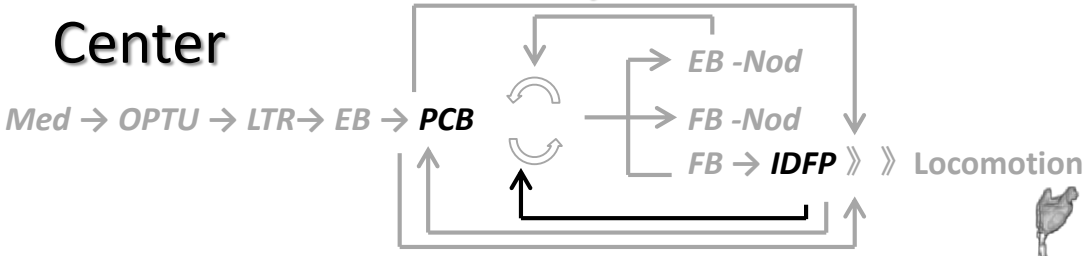


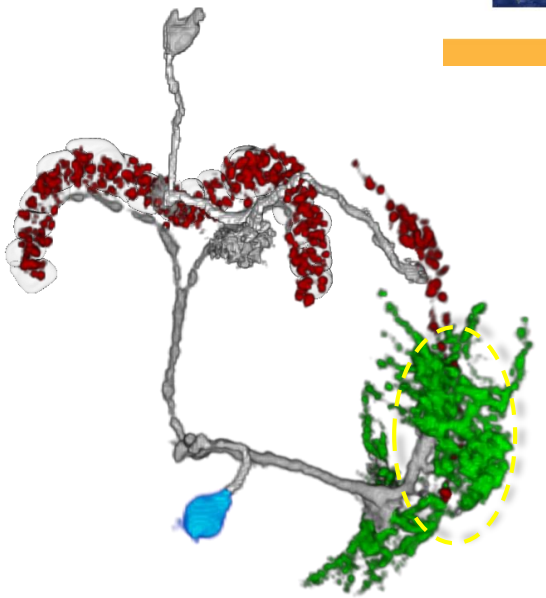


Class II **PCB:R2~7/L1~5,7,8→PCB:R1/L6**

Class III **PCB:R1-/FB:A5→IDFP:L:LLAL**

An UV Circuit from Eye to Pre-motor Center



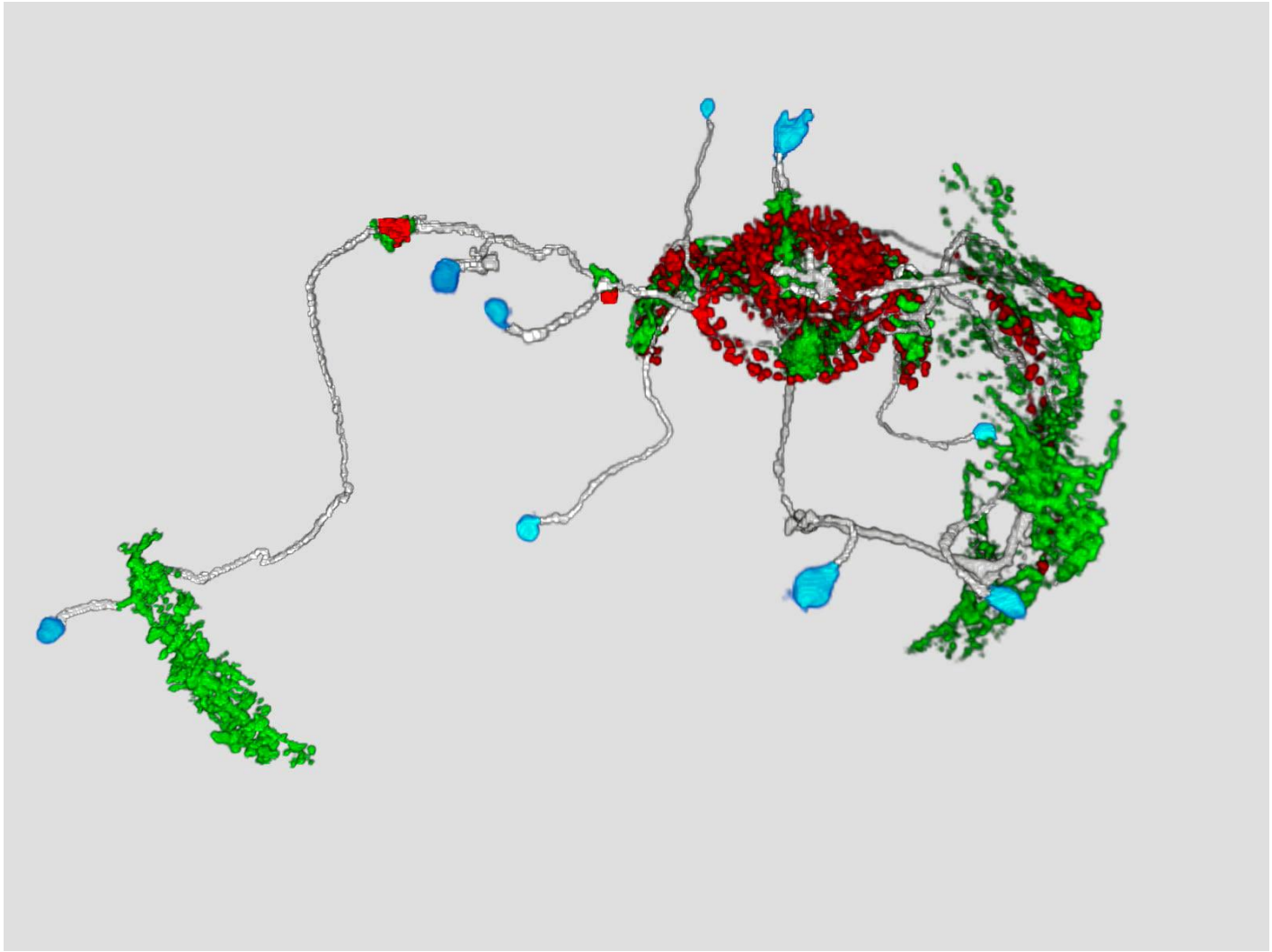


IDFP:L:LLAL

Class VI PCB:R1-/FB:A5→IDFP:L:LLAL

Class III CVLP:L-VMP:L-IDFP:L:LLAL→PCB:R1~8/L1~8

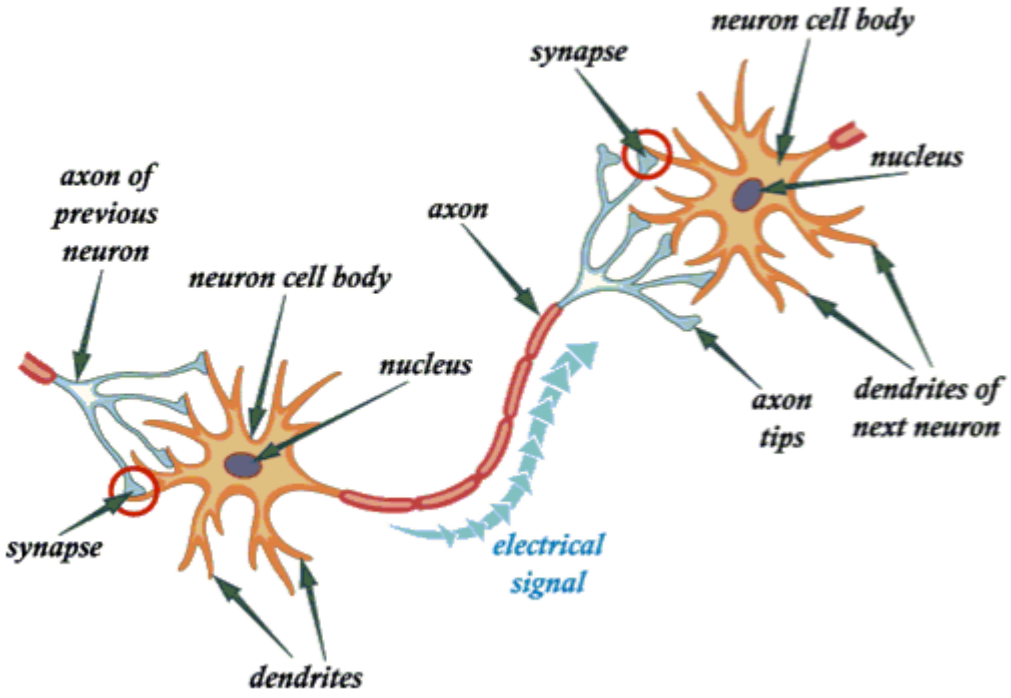
An UV Circuit from Eye to Pre-motor Center



Pathway Movie

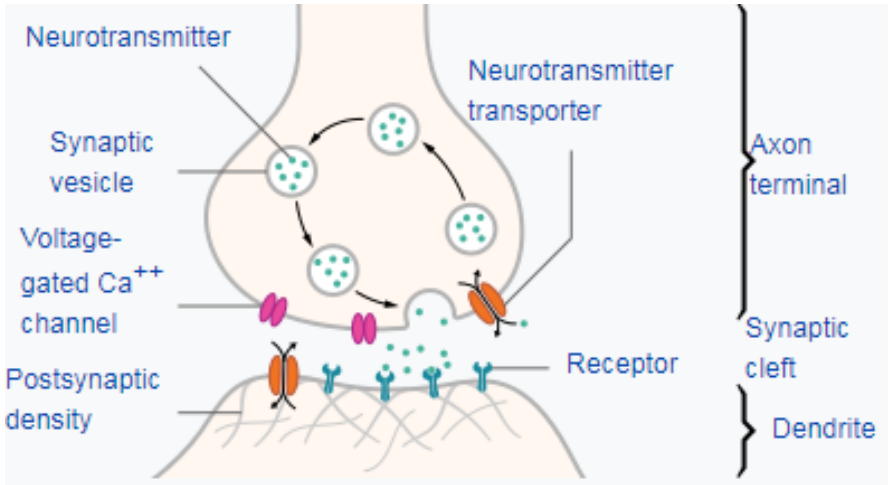
Cell Report **3**,1739 (2013)

A single neuron



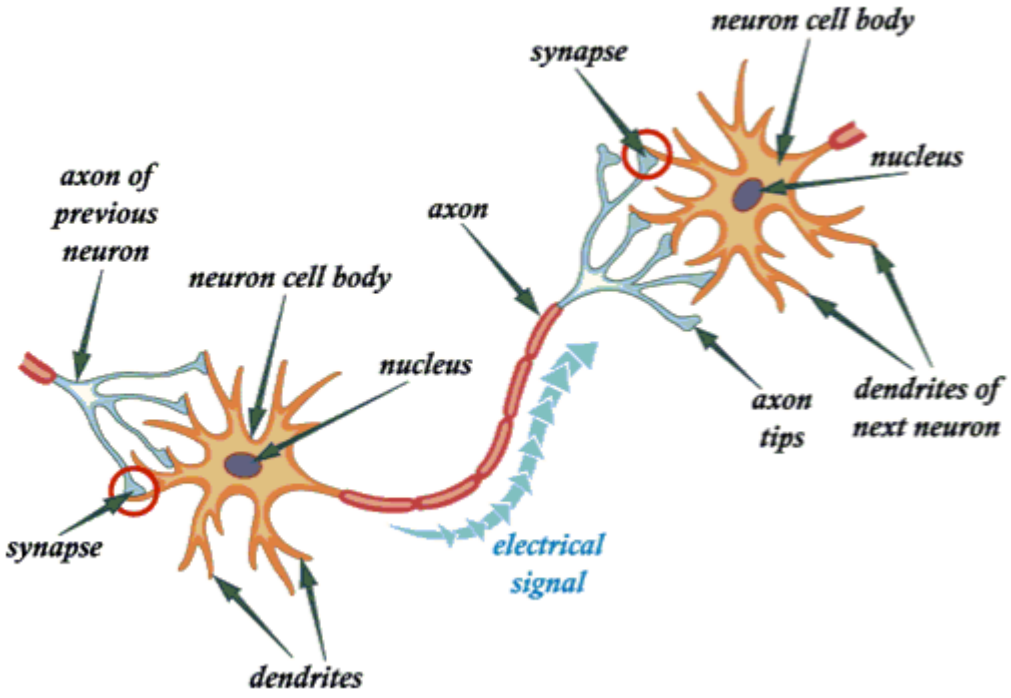
生物神經細胞基本構造：

- 細胞本體(soma)
- 樹突(dendrites) → input
- 軸突(axon) → output
- 突觸(synapse)

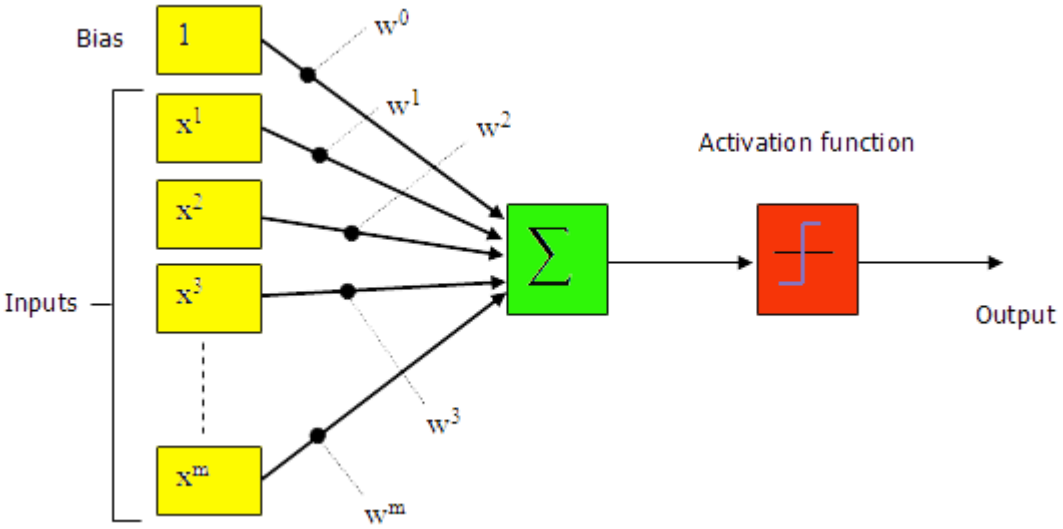


https://zh.wikipedia.org/wiki/ & https://hackernoon.com/jump-start-to-artificial-intelligence-f6eb30d624ec

Neuron → Perceptron

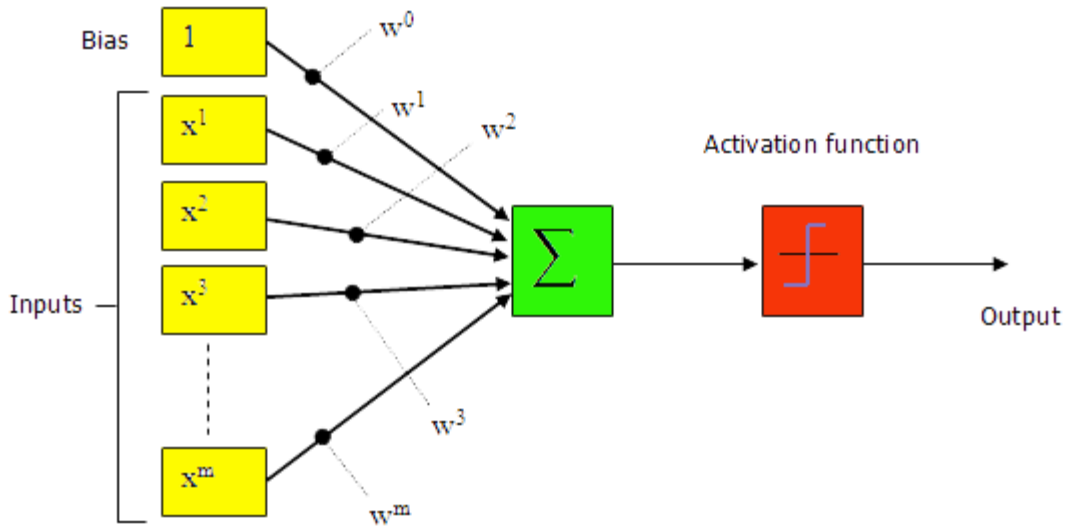


Perceptron(感知器)



A single perceptron

- Weights and Bias (權重與偏置)

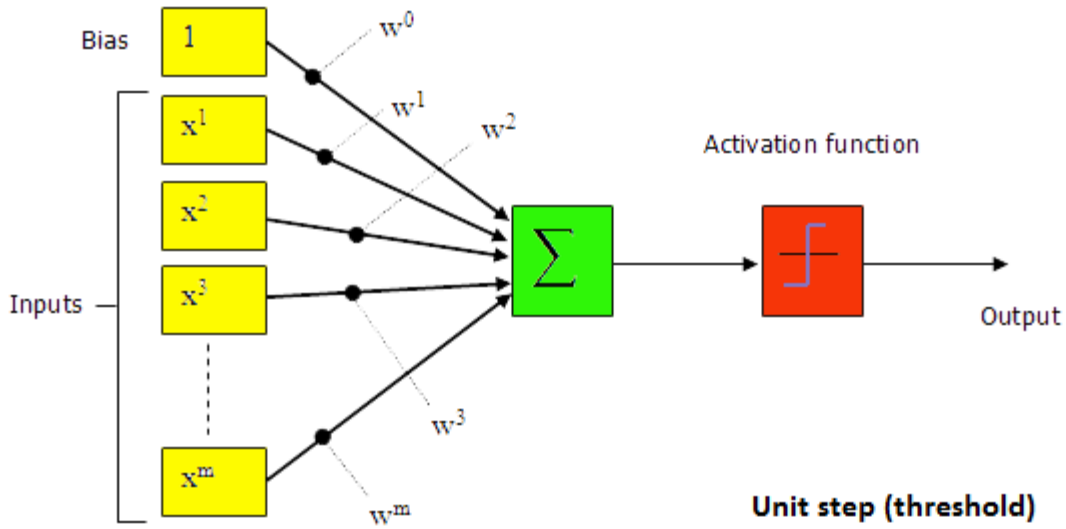


$$Output = f\left(\sum_{i=1}^{m} w^i x^i + w^0\right) = f(w^1x^1 + w^2x^2 + w^3x^3 + \cdots + w^m x^m + w^0)$$

Learning parameters

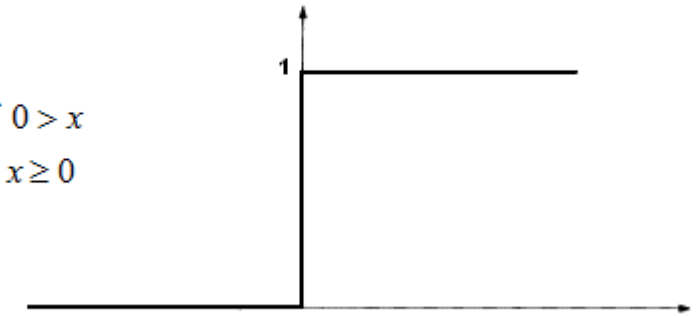
A single perceptron

- Activation Function (激活函數)



$$Output = f\left(\sum_{i=1}^{m} w^i x^i + w^0\right)$$

Unit step (threshold)



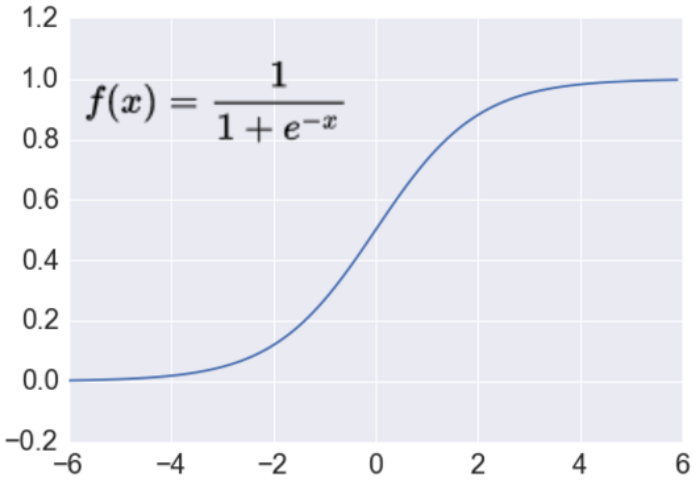
$$f(x) = \begin{cases} 0 \text{ if } 0 > x \\ 1 \text{ if } x \geq 0 \end{cases}$$

A single perceptron

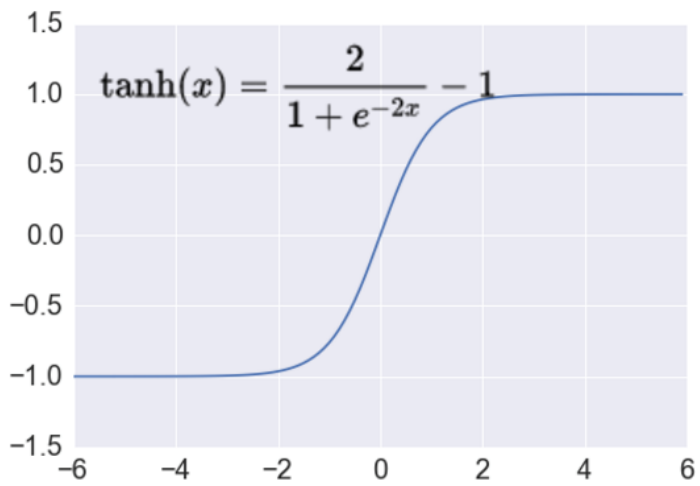
- Activation Function(激活函數)

$$Output = f\left(\sum_{i=1}^{m} w^i x^i + w^0\right)$$

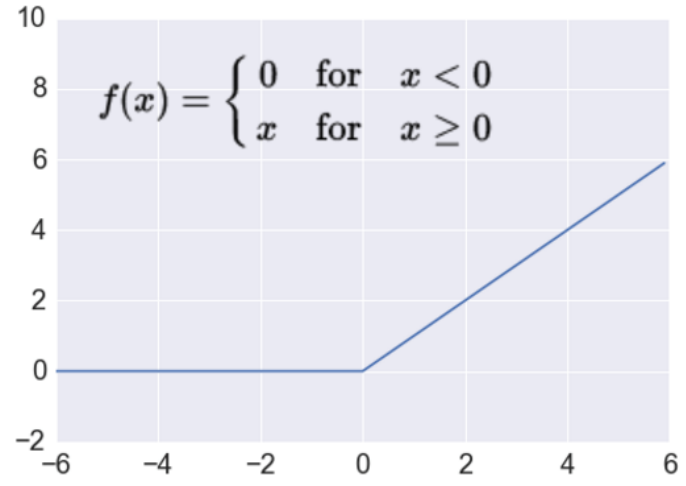
Sigmoid



TanH

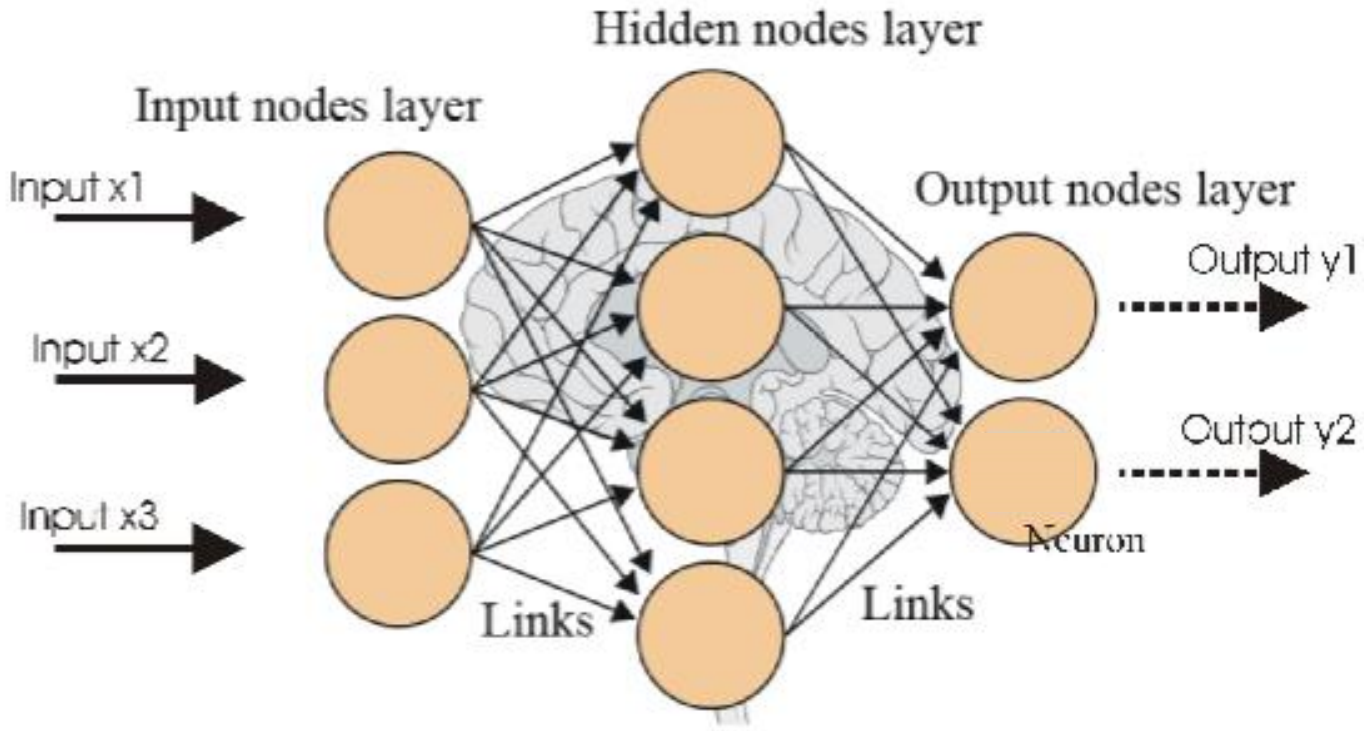


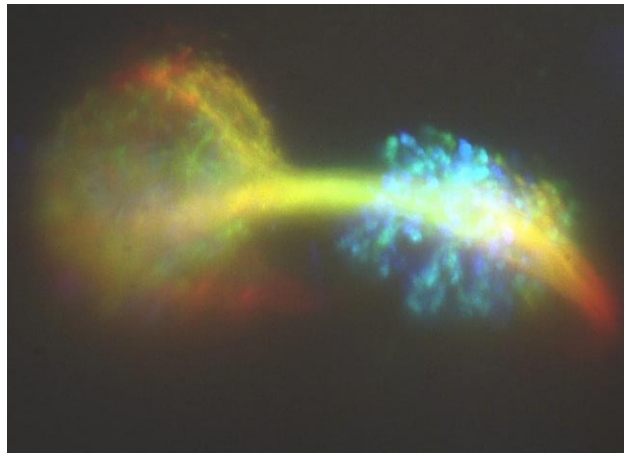
ReLU



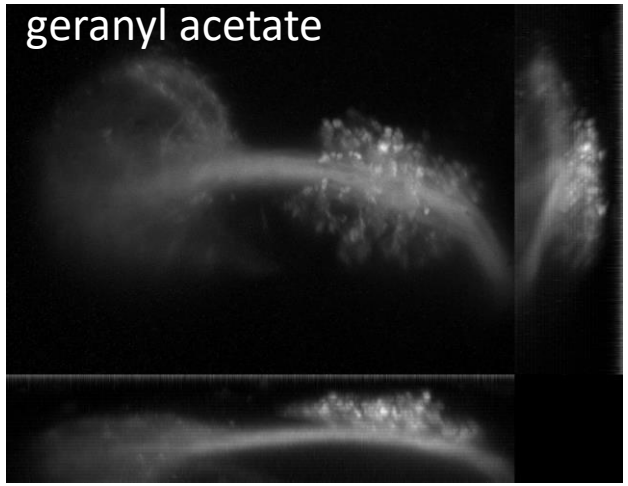
https://mropengate.blogspot.com/2017/02/deep-learning-role-of-activation.html

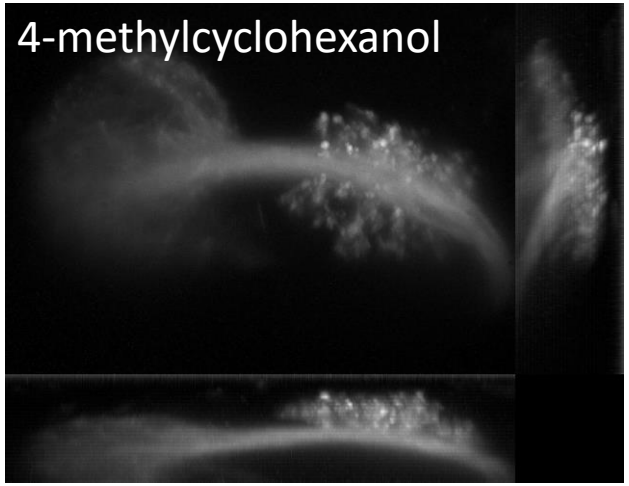
ANN = Multilayer Perceptron

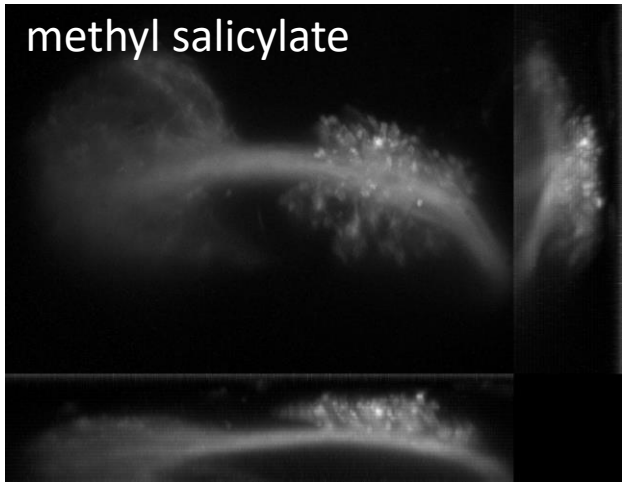


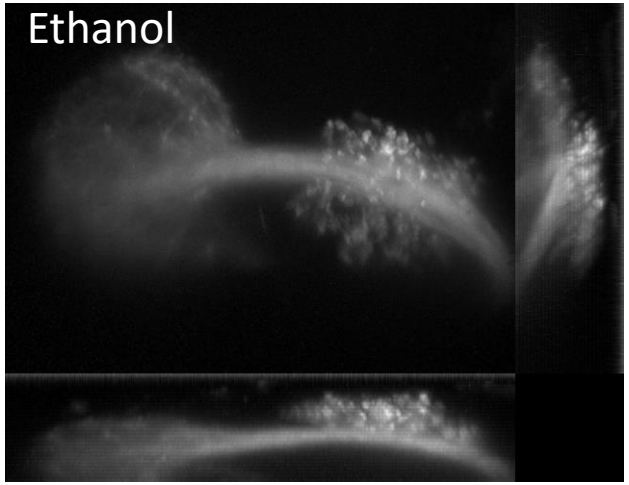


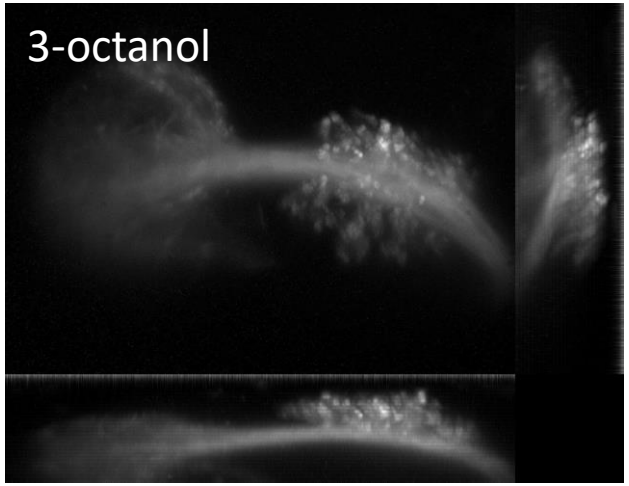
Real neurons’ firing

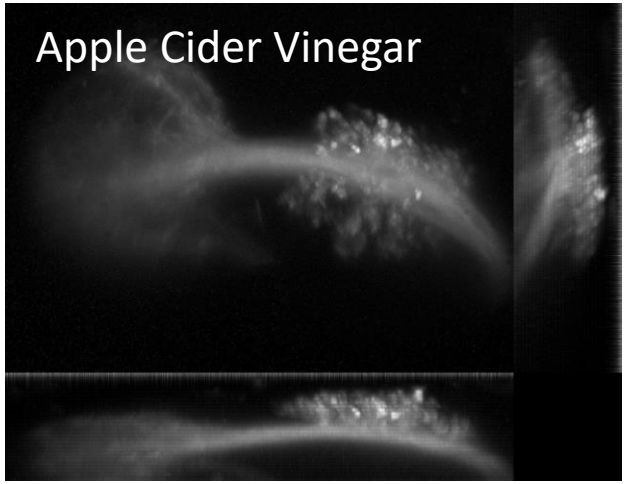




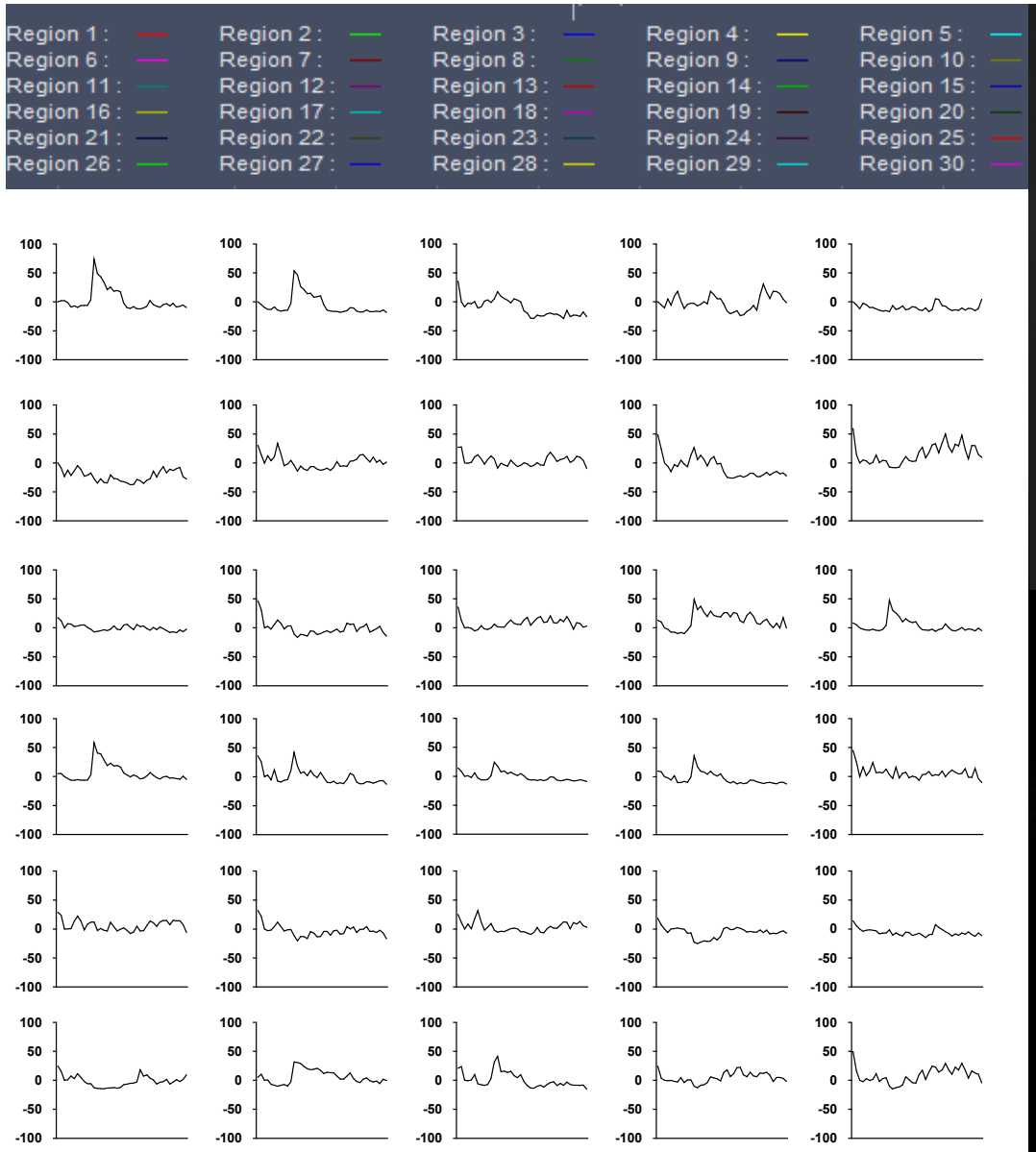


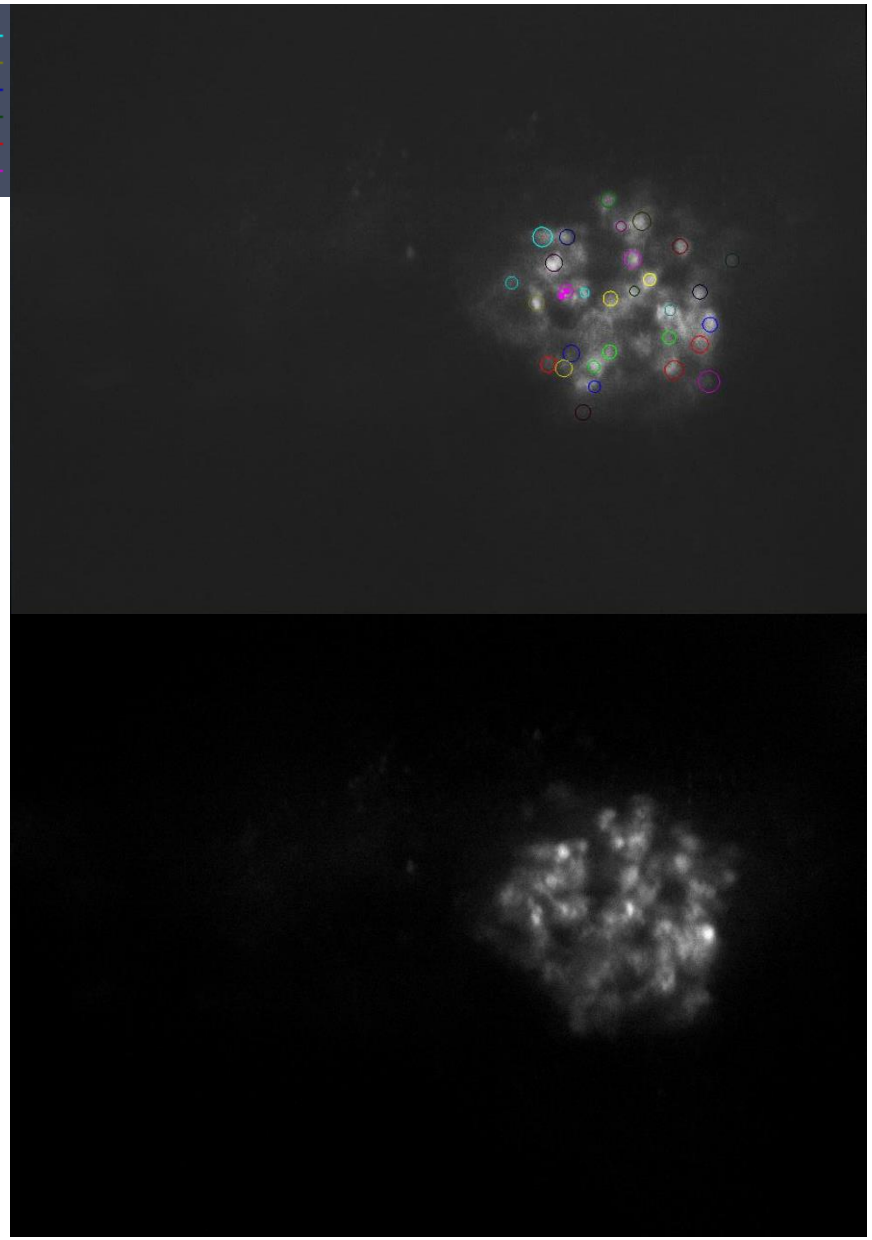






Real activation functions





How ANN learn?

Labeling data
(training set)

$$Input = (x^1, x^2, x^3, \dots, x^m)_j$$
$$Target = (y^1, y^2, y^3, \dots, y^n)_j$$

$$Output = f\left((\sum_{i=1}^{m} w^i x^i + w^0)_j\right) = (\hat{y}^1, \hat{y}^2, \hat{y}^3, \dots, \hat{y}^n)_j$$

$$Loss\ function = L(\sum_{j} (Target_j - Output_j))$$

The main goal is finding appropriate weights and biases to minimize the loss function!!!

Next class

- Regression
- Loss function
- Logistic regression
- Forward propagation
- Backward propagation
- Gradient descent
- Optimizers
- ……