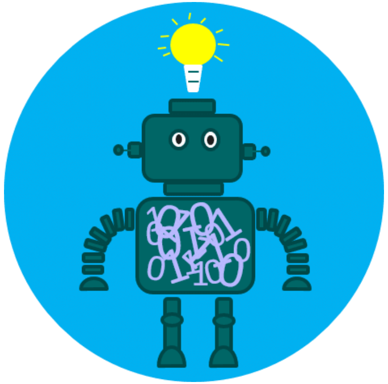
Applied Deep Learning

Introduction

September 7th, 2023 http://adl.miulab.tw

National Taiwan University
國立臺灣大學



2

What is Machine Learning?

什麼是機器學習？
白話文讓你了解！

AI & ML

Artificial intelligence (AI) is intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by animals and humans.

Machine learning (ML) is a field of inquiry devoted to understanding and building methods that "learn", that is, methods that leverage data to improve performance on some set of tasks.

It is seen as a part of artificial intelligence.

What Computers Can Do?

→ Programs can do the things you ask them to do

Program for Solving Tasks

- Task: predicting positive or negative given a product review

"I love this product!" → program.py → +

"It claims too much." → program.py → -

"It's a little expensive." → program.py → ?

if input contains "love", "like", etc.
output = positive

if input contains "too much", "bad", etc.
output = negative

"台灣第一波上市!" → program.py → 推

"規格好雞肋…" → program.py → 噓

"樓下買了我才考慮" → program.py → ?

Some tasks are complex, and we don't know how to write a program to solve them.

Learning ≈ Looking for a Function

- Task: predicting positive or negative given a product review

"I love this product!" ↓ *f* +

"It claims too much." ↓ *f* -

"It's a little expensive." ↓ *f* ?

if input contains "love", "like", etc.
output = positive

if input contains "too much", "bad", etc.
output = negative

"台灣第一波上市!" ↓ *f* 推

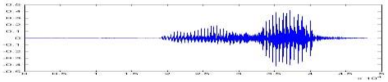
"規格好雞肋..." ↓ *f* 噓

"樓下買了我才考慮" ↓ *f* ?

Given a large amount of data, the machine learns what the function *f* should be.

Learning ≈ Looking for a Function

- Speech Recognition

$f($  $)=$ "你好"

- Handwritten Recognition

$f($  $)=$ "2"

- Weather forecast

$f($ Thursday $)=$ "  Saturday"

- Play video games

$f($  $)=$ "move left"

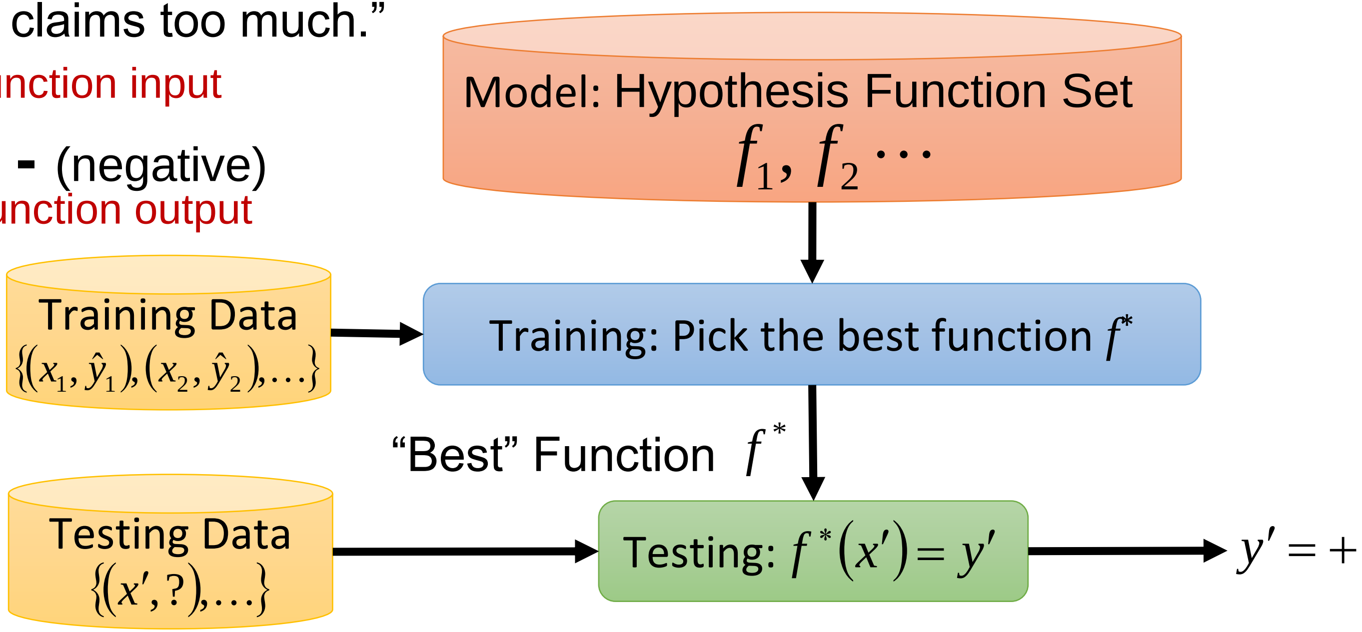
Machine Learning Framework

x : "It claims too much."

function input

$\hat{y}$: - (negative)

function output



Training is to pick the best function given the observed data
Testing is to predict the label using the learned function

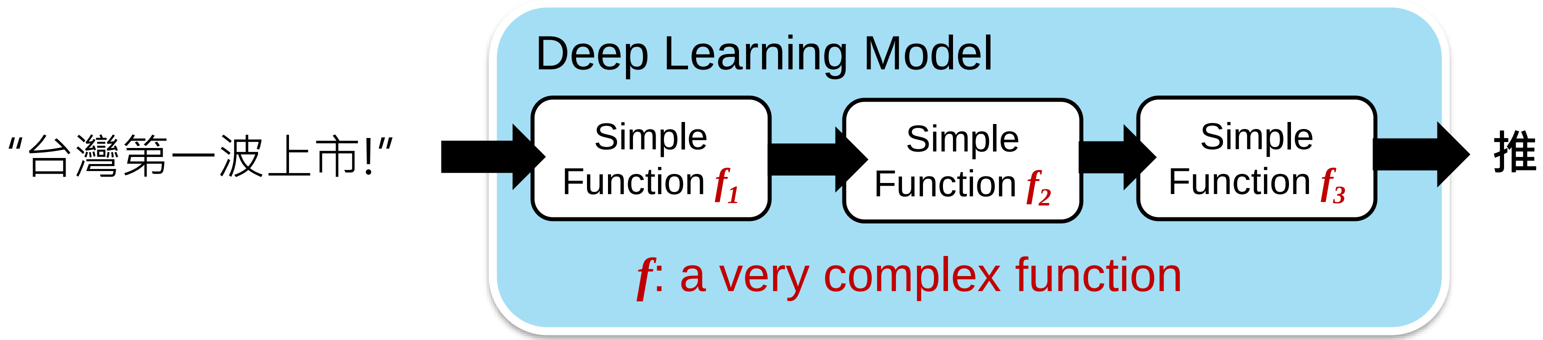
What is Deep Learning?

什麼是深度學習？

A subfield of machine learning

Stacked Functions Learned by Machine

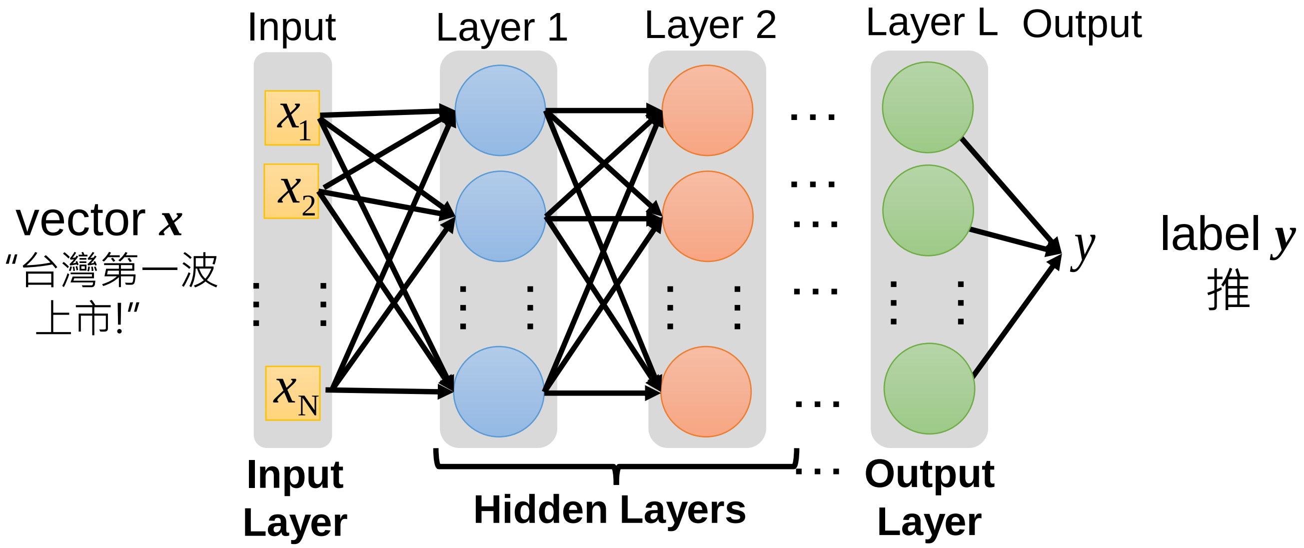
- Production line (生產線)



End-to-end training: what each function should do is learned automatically

Deep learning usually refers to *neural network* based model

Stacked Functions Learned by Machine



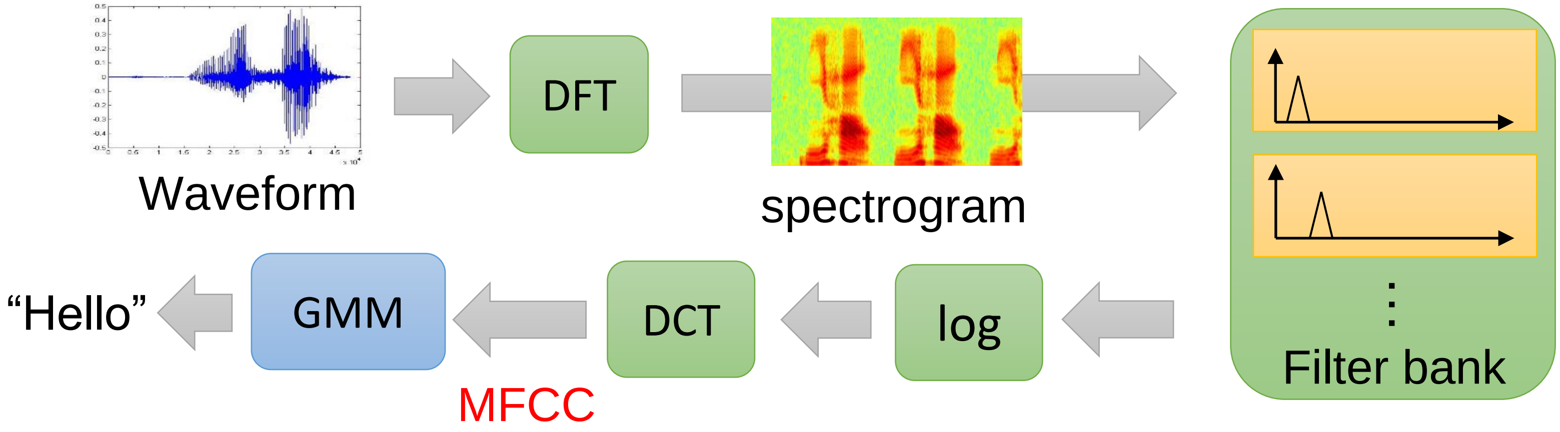
Features / Representations

Representation Learning attempts to learn good features/representations

Deep Learning attempts to learn (multiple levels of) representations and an output

Deep v.s. Shallow – Speech Recognition

- Shallow Model



Each box is a simple function in the production line:

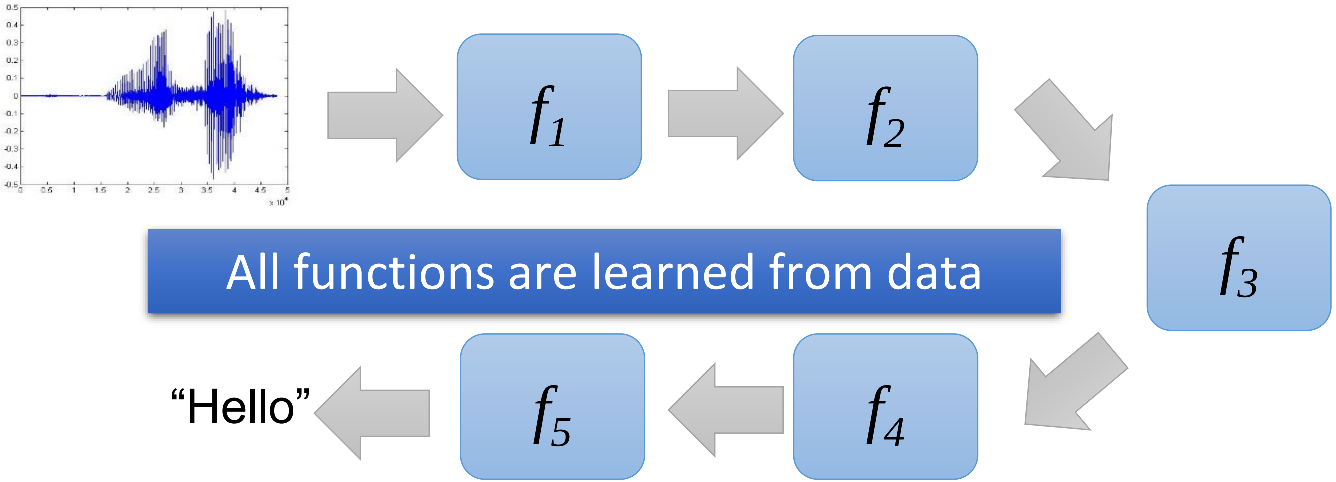
:hand-crafted

:learned from data

Deep v.s. Shallow – Speech Recognition

"Bye bye, MFCC" - Deng Li in Interspeech 2014

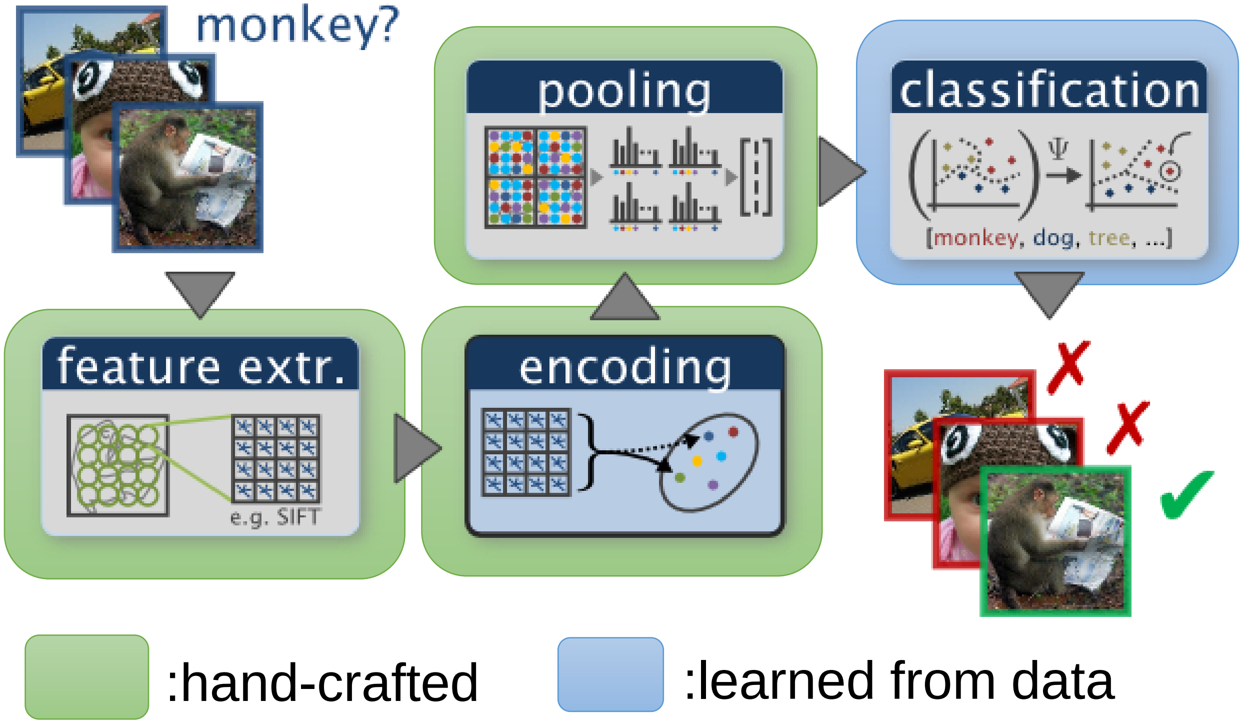
- Deep Model



Less engineering labor, but machine learns more

Deep v.s. Shallow – Image Recognition

- Shallow Model

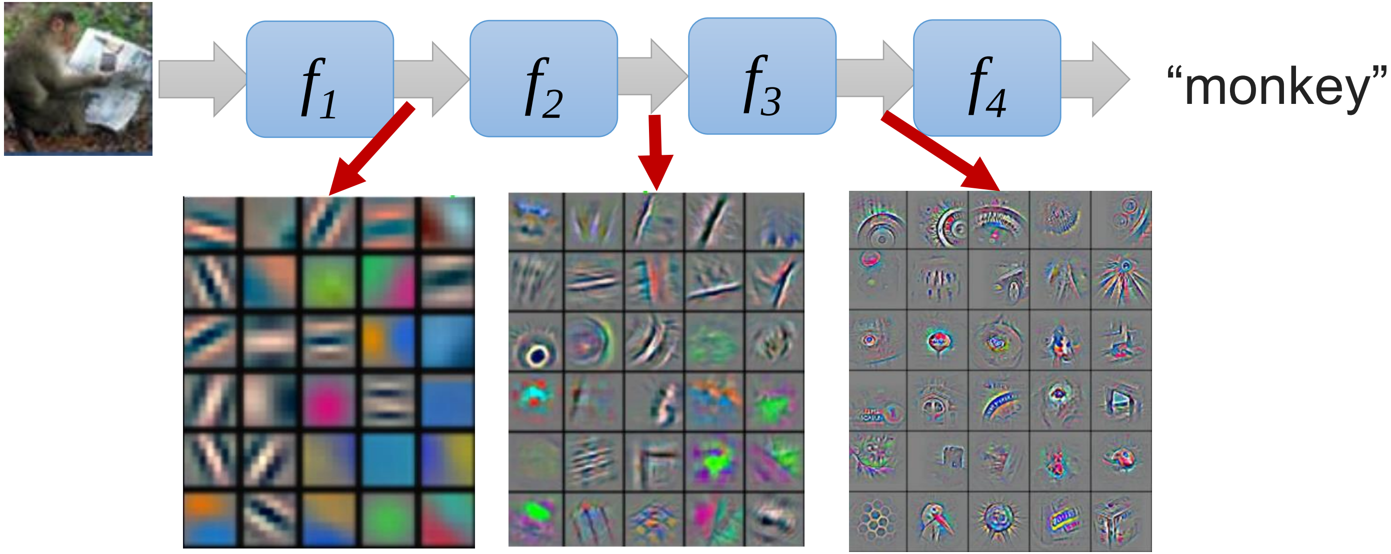


Deep v.s. Shallow – Image Recognition

Reference: Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *Computer Vision–ECCV 2014* (pp. 818-833)

- Deep Model

All functions are learned from data



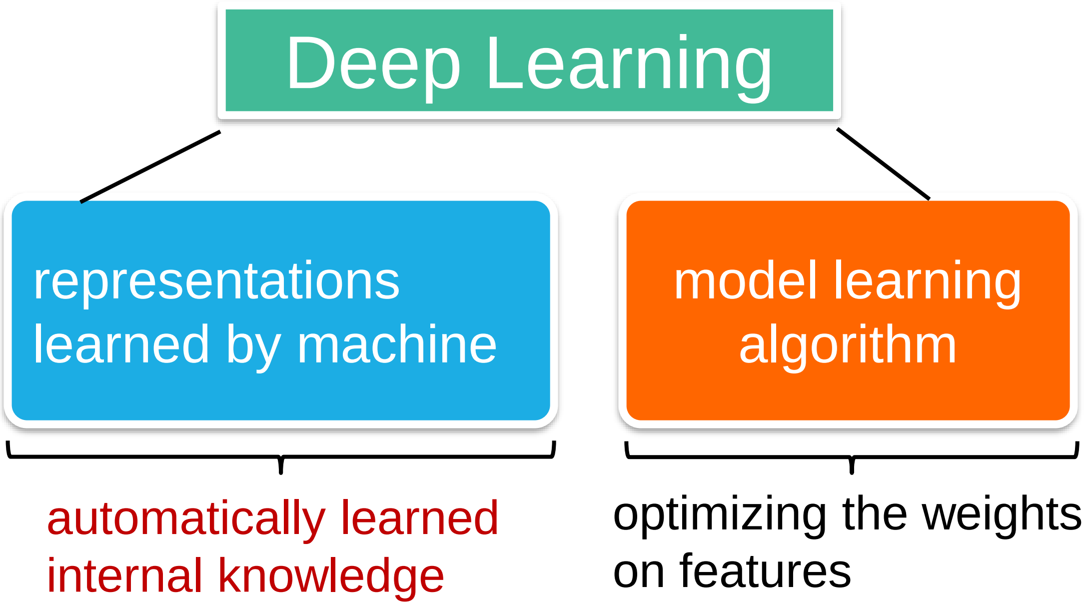
Features / Representations

Machine Learning v.s. Deep Learning

Machine Learning

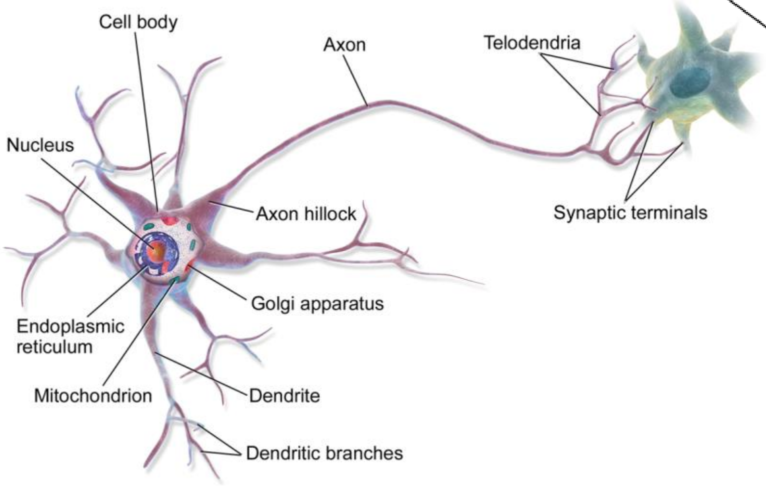
- describing your data with features a computer can understand
  - hand-crafted domain-specific knowledge
- model learning algorithm
  - optimizing the weights on features

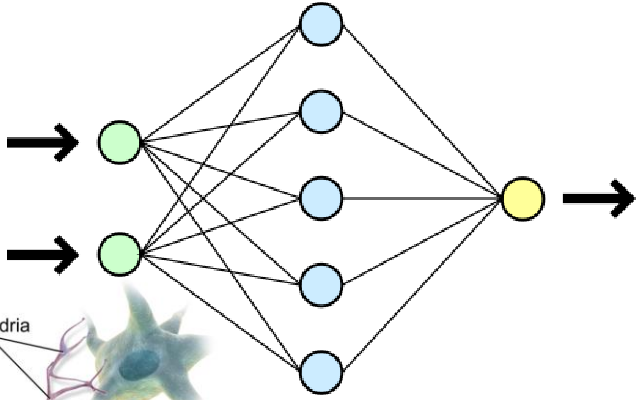
Machine Learning v.s. Deep Learning



Deep learning usually refers to *neural network* based model

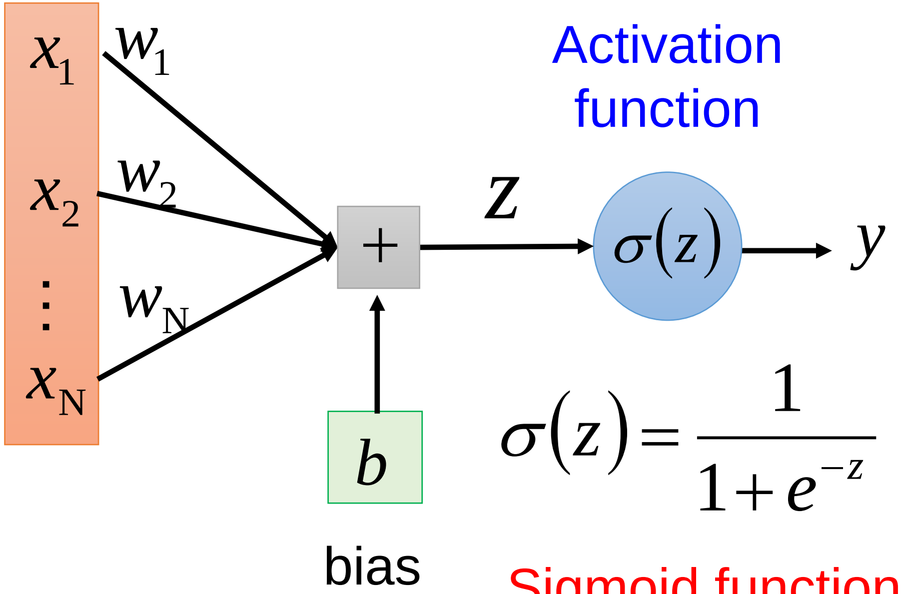
Inspired by Human Brain





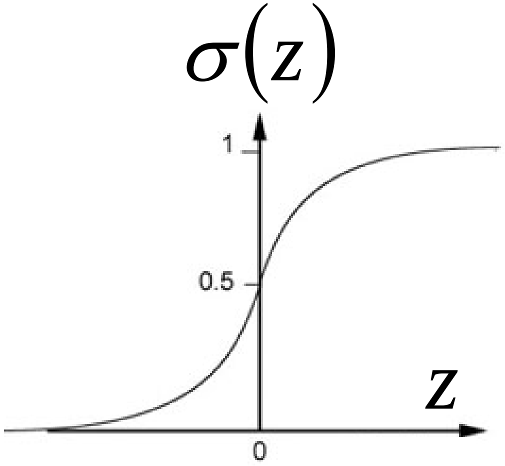
Pictures from web

A Single Neuron



$$\sigma(z) = \frac{1}{1+e^{-z}}$$

Sigmoid function

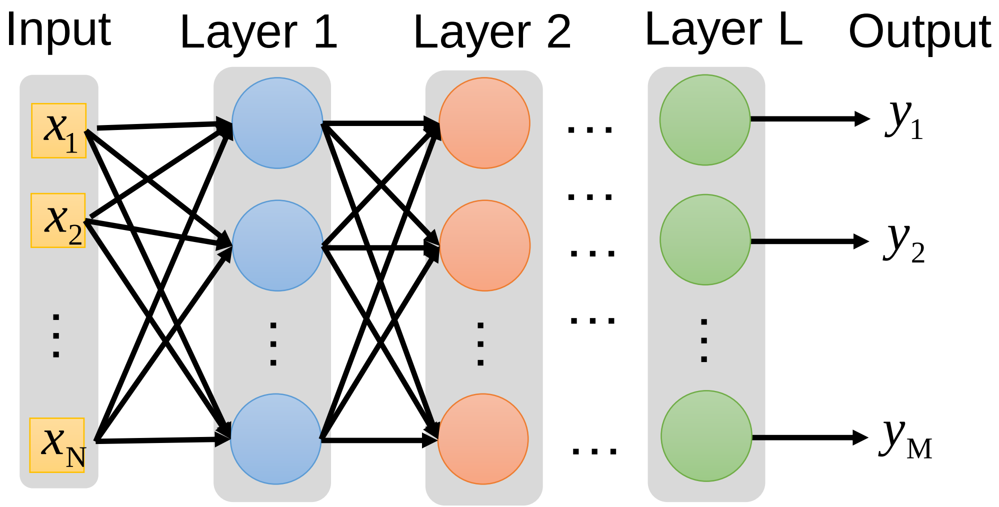


Each neuron is a very simple function

Deep Neural Network

A neural network is a complex function: $f : R^N \rightarrow R^M$

- Cascading the neurons to form a neural network

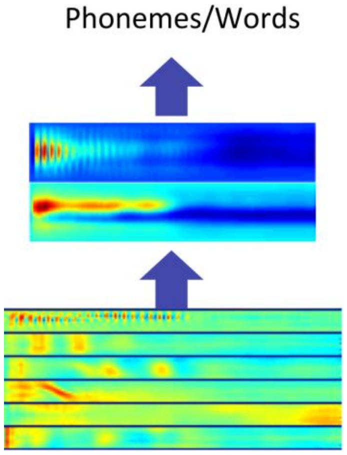


Each layer is a simple function in the production line

History of Deep Learning

- 1960s: Perceptron (single layer neural network)
- 1969: Perceptron has limitation
- 1980s: Multi-layer perceptron
- 1986: Backpropagation
- 1989: 1 hidden layer is “good enough”, why deep?
- 2006: RBM initialization (breakthrough)
- 2009: GPU
- 2010: breakthrough in Speech Recognition (Dahl et al., 2010)
- 2012: breakthrough in ImageNet (Krizhevsky et al. 2012)
- 2015: “superhuman” results in Image and Speech Recognition
- 2016: AlphaGo “superhuman” results in Go playing
- 2022: ChatGPT “human-level” results in diverse domains

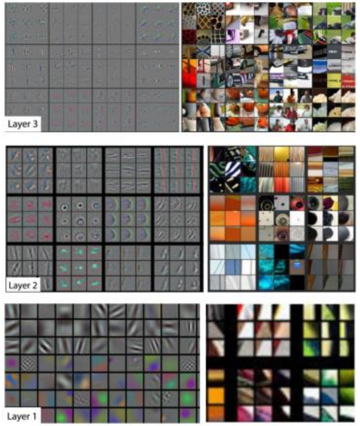
Deep Learning Breakthrough



- First: Speech Recognition

Acoustic Model	WER on RT03S FSH	WER on Hub5 SWB
Traditional Features	27.4%	23.6%
Deep Learning	18.5% (-33%)	16.1% (-32%)

- Second: Computer Vision



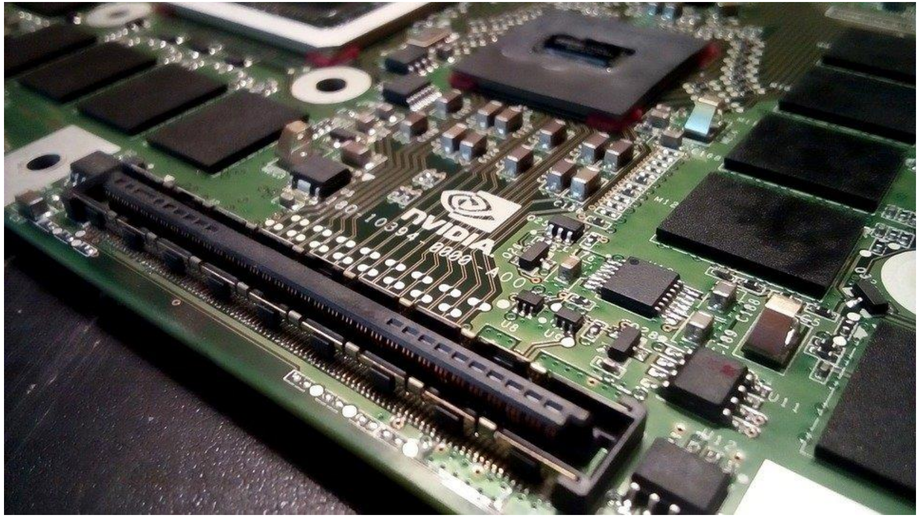
History of Deep Learning

- 1960s: Perceptron (single layer neural network)
- 1969: Perceptron has limitation
- 1980s: Multi-layer perceptron
- 1986: Backpropagation
- 1989: 1 hidden layer is "good enough", why deep?
- 2006: RBM initialization (breakthrough)
- 2009: GPU
- 2010: breakthrough in Speech Recognition (Dahl et al., 2010)
- 2012: breakthrough in ImageNet (Krizhevsky et al. 2012)
- 2015: "superhuman" results in Image and Speech Recognition
- 2016: AlphaGo "superhuman" results in Go playing
- 2022: ChatGPT "human-level" results in diverse domains

Why does deep learning show breakthrough in applications after 2010?

Why Deep Learning Works

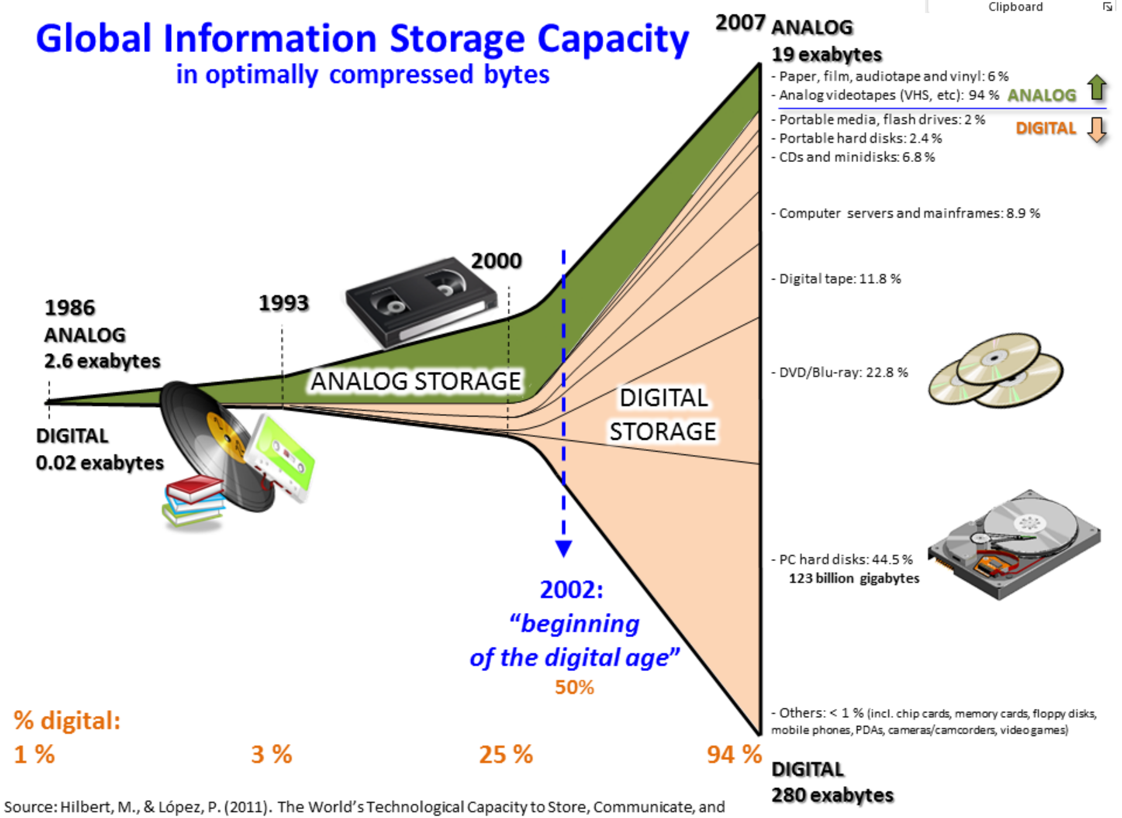
Big Data



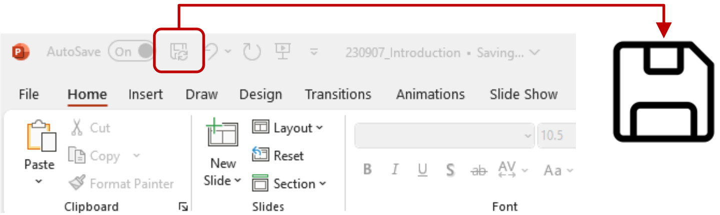
GPU

Pictures from web

Data Storage



Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60 –65. http://www.martinhilbert.net/WorldInfoCapacity.html



Pictures from web

GPU for Efficient AI Training

為什麼訓練AI需要很多GPU的資源

Searching for: **為什麼訓練AI需要很多GPU的資源**

Generating answers for you…

訓練AI需要很多GPU的資源，主要是因為AI的一種方法叫做**深度學習**，它使用了大量的**神經網絡**來處理數據[1]。神經網絡是一種高度並行的計算任務，也就是說，它可以把一個大的任務分成很多小的任務，同時在不同的處理器上執行[1]。GPU是一種專門用於並行計算的處理器，它有很多個核心（core），每個核心都可以執行一個小任務[1][2]。相比之下，CPU是一種用於串行計算的處理器，它有較少的核心，但每個核心都可以執行較複雜的任務[2]。因此，GPU比CPU更適合人工智能計算，尤其是當數據量和計算量都很大時[3][2]。

Learn more: 1. zhuanlan.zhihu.com 2. zhihu.com 3. zhuanlan.zhihu.com

GPU Companies

有提供GPU的公司有甚麼?

✓ Searching for: **有提供GPU的公司有甚麼**

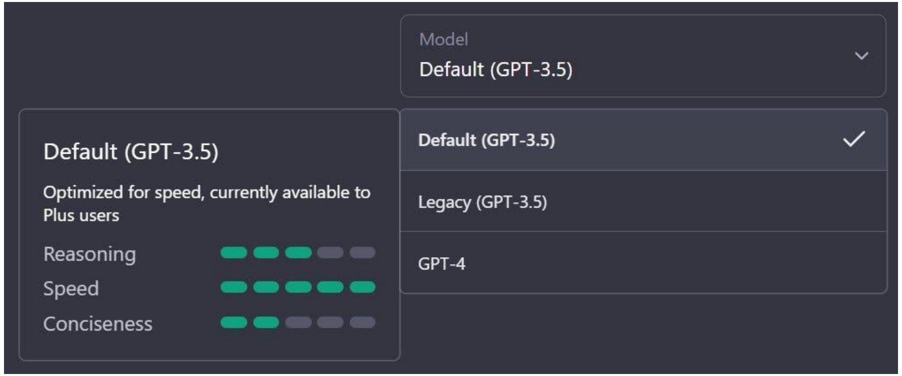
✓ Generating answers for you…

有提供GPU的公司有很多，其中最知名的是**英偉達**（NVIDIA）和**AMD**，它們分别生產了**GeForce**和**Radeon**系列的GPU [1]。這兩家公司的GPU主要用於遊戲和專業市場，也適合訓練AI [2] [1]。除此之外，還有一些國內的公司也在開發自己的GPU，例如長沙景嘉微電子股份有限公司，它擁有完全自主知識產權的GPU產品 [2]。另外，一些大型科技公司也在研發自己的GPU或類似產品，例如谷歌的TPU（Tensor Processing Unit），蘋果的M1芯片等 [1]。

Learn more: 1. sohu.com 2. thepaper.cn 3. gigabyte.com

Why Speed Matters?

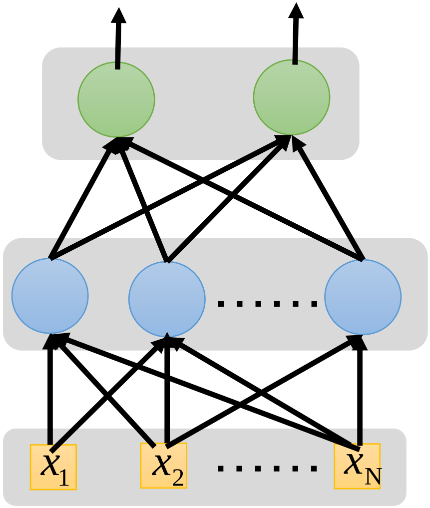
- Training time
  - Big data increases the training time
  - Too long training time is not practical
- Inference time
  - Users are not patient to wait for the responses



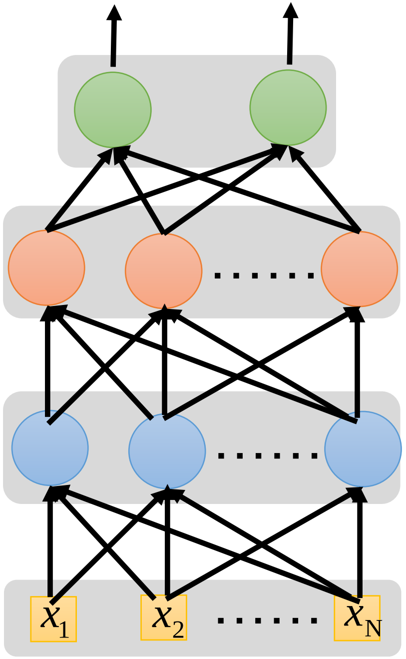
GPU enables the real-world applications using the computational power

Why Deeper is Better?

- Deeper → More parameters



Shallow



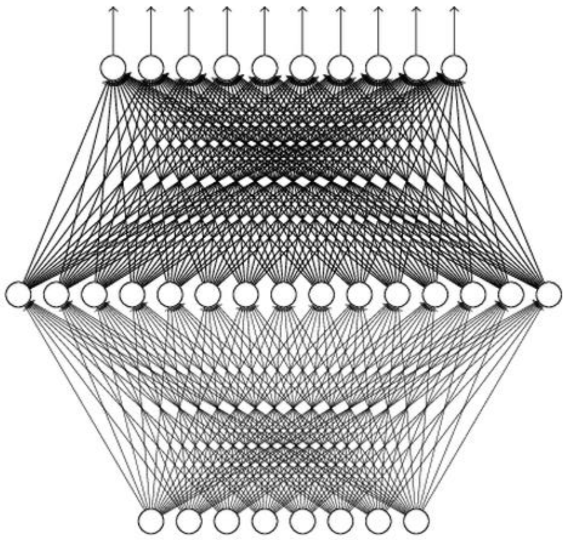
Deep

Universality Theorem

- Any continuous function f

$$f : R^N \rightarrow R^M$$

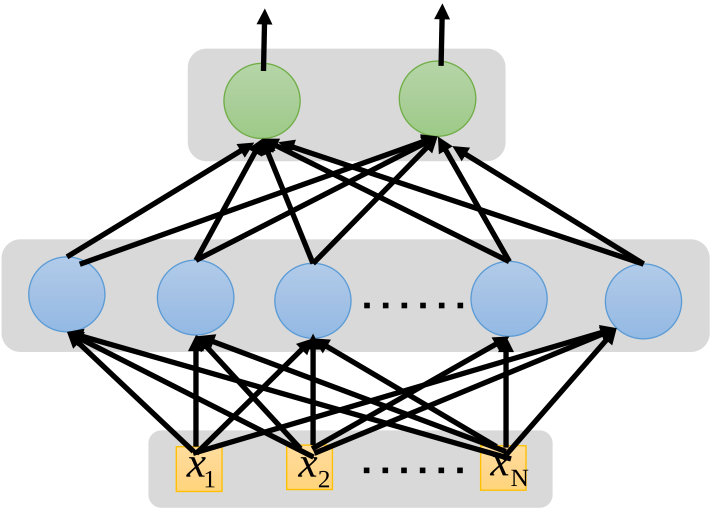
- can be realized by a network with only hidden layer

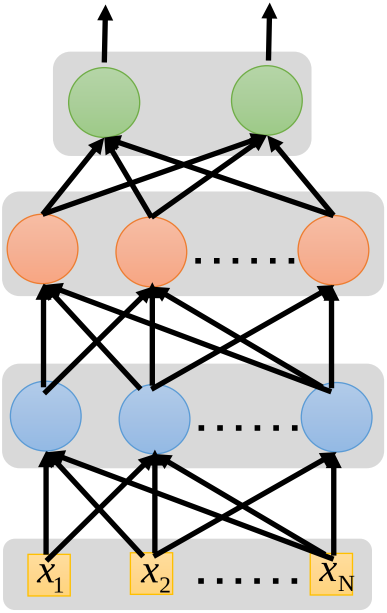


Why “deep” not “fat”?

Fat + Shallow vs. Thin + Deep

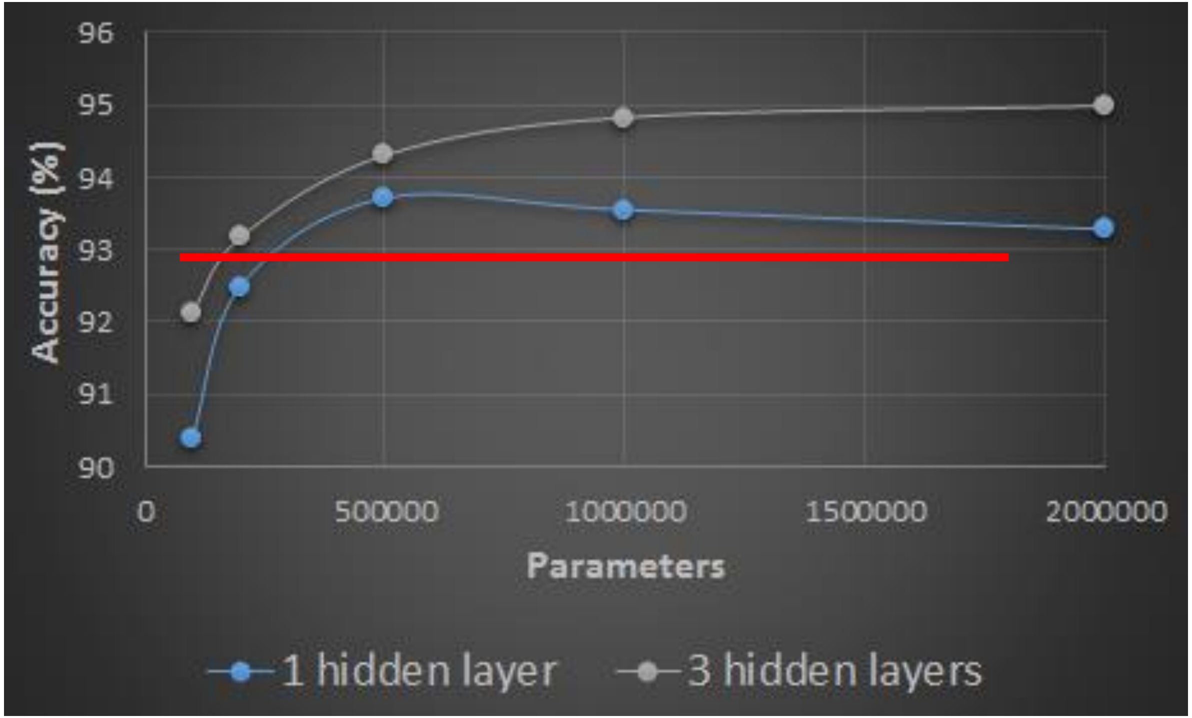
- Two networks with the same number of parameters





Fat + Shallow vs. Thin + Deep

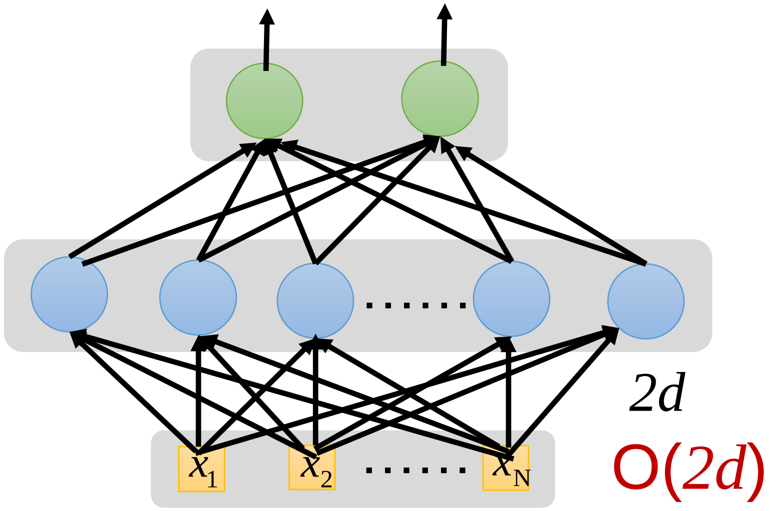
Hand-Written Digit Classification

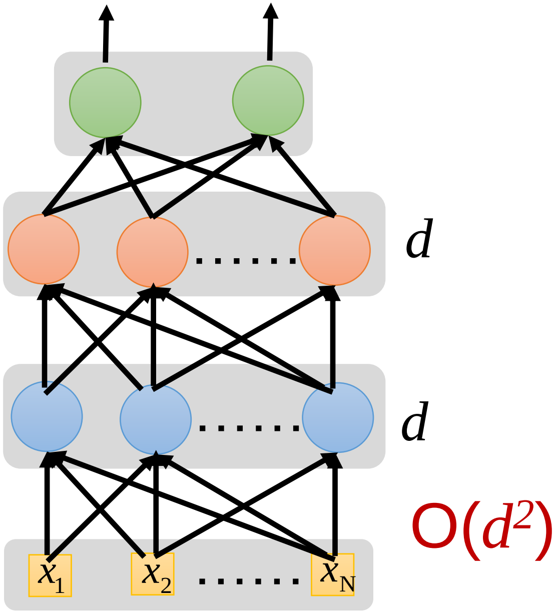


The deeper model uses less parameters to achieve the same performance

Fat + Shallow vs. Thin + Deep

- Two networks with the same number of parameters





34

How to Apply?

如何應用深度學習？

How to Frame the Learning Problem?

- The learning algorithm f is to map the input domain X into the output domain Y

$$f : X \rightarrow Y$$

- Input domain: word, word sequence, audio signal, click logs
- Output domain: single label, sequence tags, tree structure, probability distribution

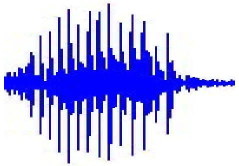
Output Domain – Classification

- Sentiment Analysis

"這規格有誠意!" → +

"太爛了吧~" → -

- Speech Phoneme Recognition



→ /h/

- Handwritten Recognition

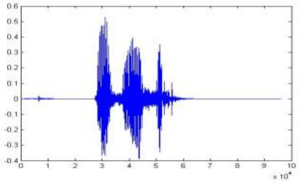
→ 2

Output Domain – Sequence Prediction

- POS Tagging

“推薦我台大後門的餐廳” ⟶ 推薦/VV 我/PN 台大/NR 後門/NN 的/DEG 餐廳/NN

- Speech Recognition



⟶ “大家好”

- Machine Translation

“How are you doing today?” ⟶ “你好嗎?”

Learning tasks are decided by the output domains

Input Domain – How to Aggregate Information

- Input: word sequence, image pixels, audio signal, click logs
- Property: continuity, temporal, importance distribution
- Example
  - CNN (convolutional neural network): local connections, shared weights, pooling
    - AlexNet, VGGNet, etc.
  - RNN (recurrent neural network): temporal information
  - Transformer: multiple inputs with interaction

Network architectures should consider the input domain properties

How to Frame the Learning Problem?

- The learning algorithm f is to map the input domain X into the output domain Y

$$f : X \to Y$$

- Input domain: word, word sequence, audio signal, click logs
- Output domain: single label, sequence tags, tree structure, probability distribution

Network design should leverage input and output domain properties

“Applied” Deep Learning

Deep Learning

representations learned by machine

automatically learned internal knowledge

model learning algorithm

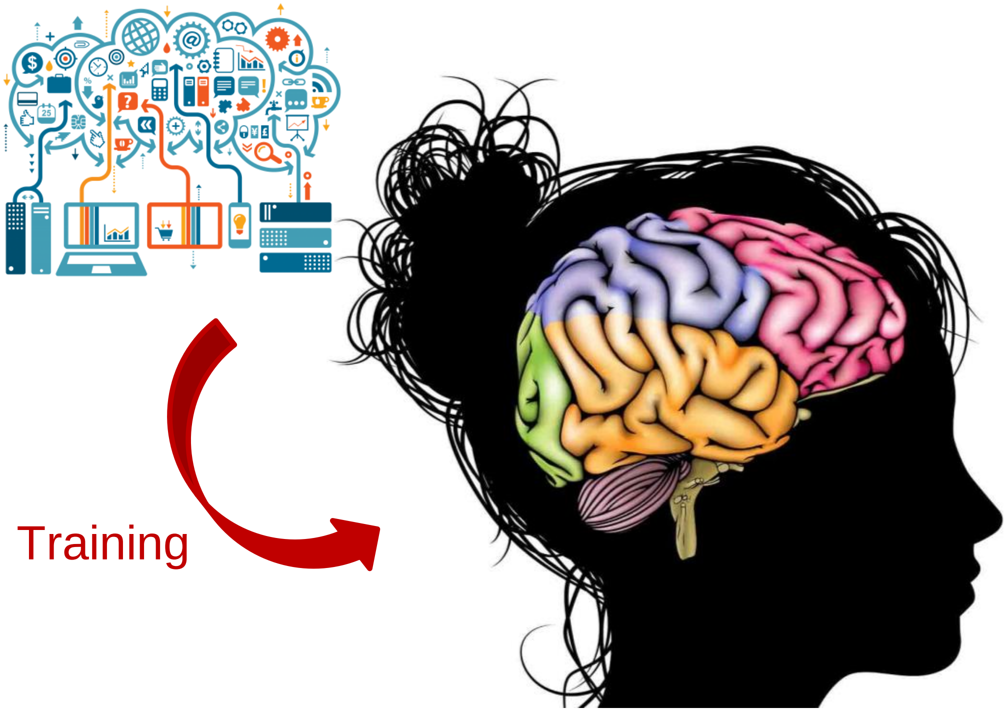
optimizing the weights on features

How to frame a task into a learning problem and design the corresponding model

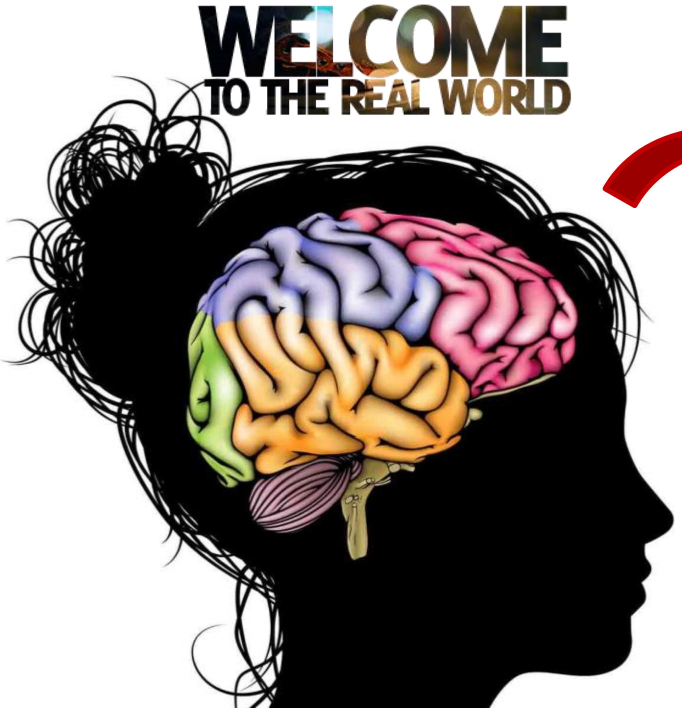
Core Factors for Applied Deep Learning

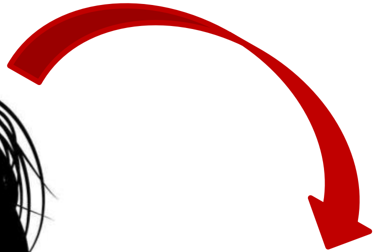
1. Data: big data
2. Hardware: GPU computing
3. Talent: design algorithms to allow networks to work for the specific problems

Concluding Remarks

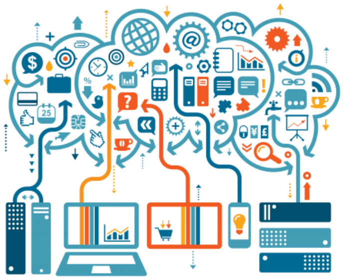


Concluding Remarks

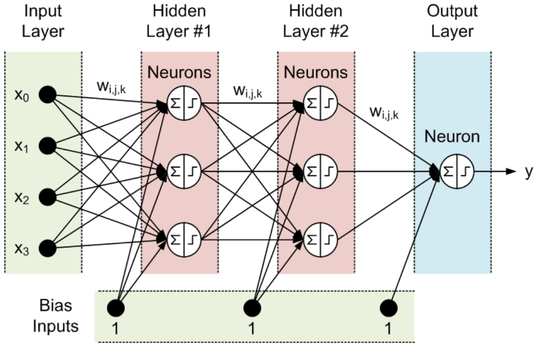




Concluding Remarks



Inference



Training

Main focus: how to apply deep learning to the real-world problems

Reference

- Reading Materials
  - Referenced academic papers can be found in the slides
- Deep Learning
  - Goodfellow, Bengio, and Courville, “Deep Learning,” 2016. http://www.deeplearningbook.org
  - Michael Nielsen, “Neural Networks and Deep Learning” http://neuralnetworksanddeeplearning.com

Thanks!

Any questions ?

You can find the course information at

- http://adl.miulab.tw
- adl-ta@csie.ntu.edu.tw
- slido: #ADL2023
- YouTube: Vivian NTU MiuLab