

Introduction of this course

李宏毅

Hung-yi Lee

Welcome our TAs

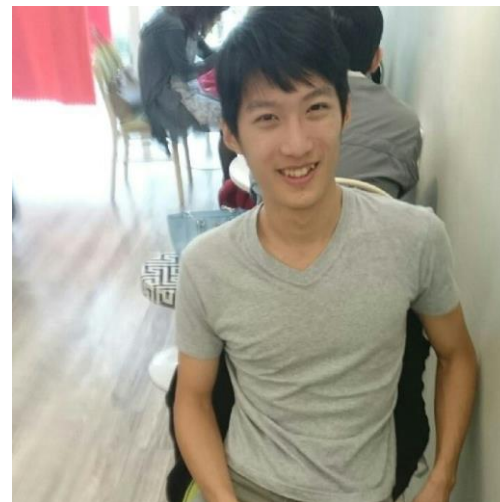
TA 信箱：mldsntu2017@gmail.com

Tensorflow 教學

鍾佩宏



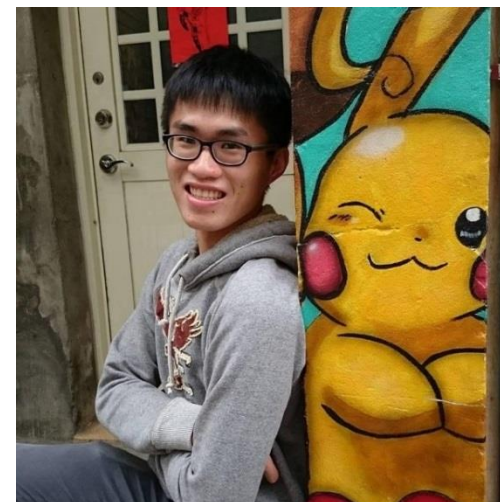
陳鴻佑



戴敬倫



陳尚甫



作業一



萬家宏



袁培傑



蘇上育

作業二

楊棋宇



莊舜博



陳鴻佑



袁培傑



戴敬倫



施順耀

作業三



樊恩宇



曾柏翔



段逸林



黃淞楓



陳尚甫

作業四

李致緯



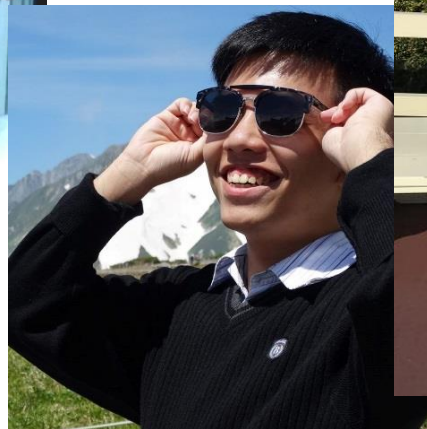
施順耀



段逸林



吳彥諶



蘇上育



協助處理運算資源



Chiung-Chih Chang 我的照片超級可愛不用換

Unlike · Reply · 2 · 5 hrs

作業零

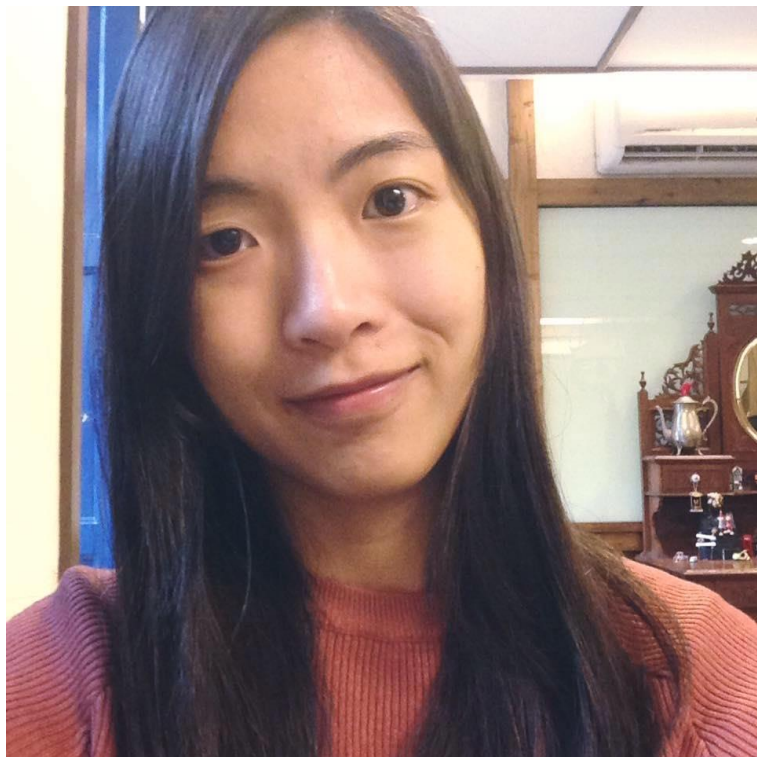


張瓊之



黃邦齊

期末專題



宋昀蓁

大助教



楊棋宇

非常感謝台大計中 支援運算資源

感謝張傑生先生、顏嗣鈞主任大力幫忙

What are we
going to learn?

課程名稱解釋

機器學習 及其深層與結構化

Machine Learning
and having it Deep and Structured



Method



Task

Deep Learning

- 上學期的「機器學習」錄影
 - DNN: <https://www.youtube.com/watch?v=Dr-WRIEFefw>
 - Tips for DNN: <https://www.youtube.com/watch?v=xki61j7z-30>
 - CNN: <https://www.youtube.com/watch?v=FrKWiRv254g>
 - RNN (Part 1): <https://www.youtube.com/watch?v=xCGidAeyS4M>
 - RNN (Part 2): <https://www.youtube.com/watch?v=xCGidAeyS4M>
 - Why Deep: <https://www.youtube.com/watch?v=XsC9byQkUH8>
 - Auto-encoder: <https://www.youtube.com/watch?v=Tk5B4seA-AU>
 - Deep generative model (Part 1):
<https://www.youtube.com/watch?v=YNUek8ioAJk>
 - Deep generative model (Part 2):
<https://www.youtube.com/watch?v=8zomhgKrsmQ>

Deep Learning

- In this course

深度学习
深度学习

Structured (Output) Learning

Machine learning is to find a function f

$$f : X \rightarrow Y$$

Regression: output a scalar

Classification: output a “class” (one-hot vector)



Class 1



Class 2



Class 3

Structured Learning: output a sequence, a matrix, a graph, a tree

Output is composed of components with dependency

Output Sequence

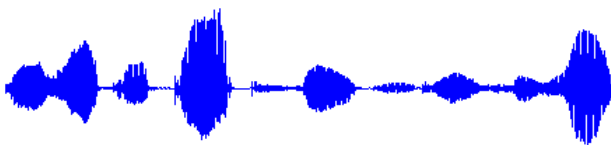
$$f : X \rightarrow Y$$

Machine Translation

X : “機器學習及其深層與結構化”
(sentence of language 1)

Y : “Machine learning and having it deep and structured”
(sentence of language 2)

Speech Recognition

X : 
(speech)

Y : “歡迎大家來修課”
(transcription)

Chat-bot

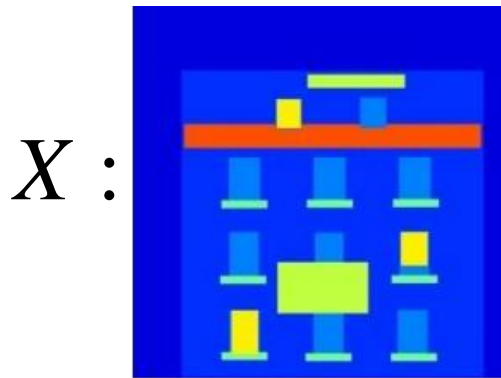
X : “How are you?”
(what a user says)

Y : “I’m fine.”
(response of machine)

Output Matrix

$$f : X \rightarrow Y$$

Image to Image



Colorization:



Ref: <https://arxiv.org/pdf/1611.07004v1.pdf>

Text to Image

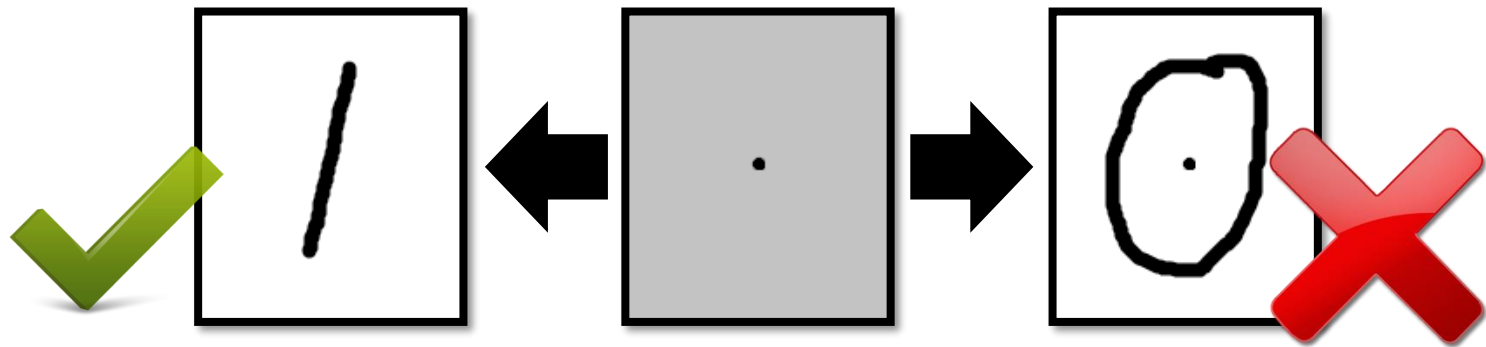
$X :$ “this white and yellow flower
have thin white petals and a
round yellow stamen”



ref: <https://arxiv.org/pdf/1605.05396.pdf>

Challenge of Structured Output

- The output space is very sparse:
 - In classification, each class has some examples.
 - In structured learning, most of the possible outputs never exist
- Because the output components have dependency, they should be considered globally.



- Typical approach: structured SVM, CRF ... they are not deep

Next Wave

Deep and Structured

(e.g. Generative Adversarial Network, GAN)

Policy

FAQ

- Q: 這門課和週四的 “Machine Learning” (ML) 有何不同？
- A: 這門課的內容和 ML 完全不同
 - 這門課會著重於 deep learning 和 structured learning
 - 和 ML 內容不重複
 - 另外，相較於過去的同門課，會增加不少新的內容
 - 本課程較適合有機器學習背景同學

組隊

- 每組 2 ~ 4 人，其中一人為組長
- 以組為單位進行所有作業和專題
- 隊伍登記方法於下週說明作業一時順便公布
- 組內互評：學期結束前會有組內互評，會影響成績

評量方式

- 不點名、不考試
- 作業一 (20%) : 3/03 (三週)
- 作業二 (30%) : 3/24 (五週)
- 作業三 (20%) : 4/28 (三週)
- 作業四 (20%) : 5/19 (三週)
- 期末專題 (30%)

成績是相對的
成績是相對的
成績是相對的

FB 社團

- 社團: “Machine learning and having it deep and structured (2017 spring)”
 - <https://www.facebook.com/groups/1042636162508032/>
- 有問題可以直接在 **FB**社團上發問
 - 如果有同學知道答案請幫忙回答
- 有想法也可以在 **FB**社團上發言
- 會紀錄好的問題、答案、留言，期末會加分

上課

- 上課投影片和錄音會放到李宏毅的個人網頁上
 - 李宏毅的個人網頁：
http://speech.ee.ntu.edu.tw/~tlkagk/courses_MLSD17.html
- 3/10 Tensorflow 教學
- 4/14 期中考前一週放假
- 6/16 期末考週不上課
- 6/23 期末成果發表 (暫定)

需要的基礎能力和知識

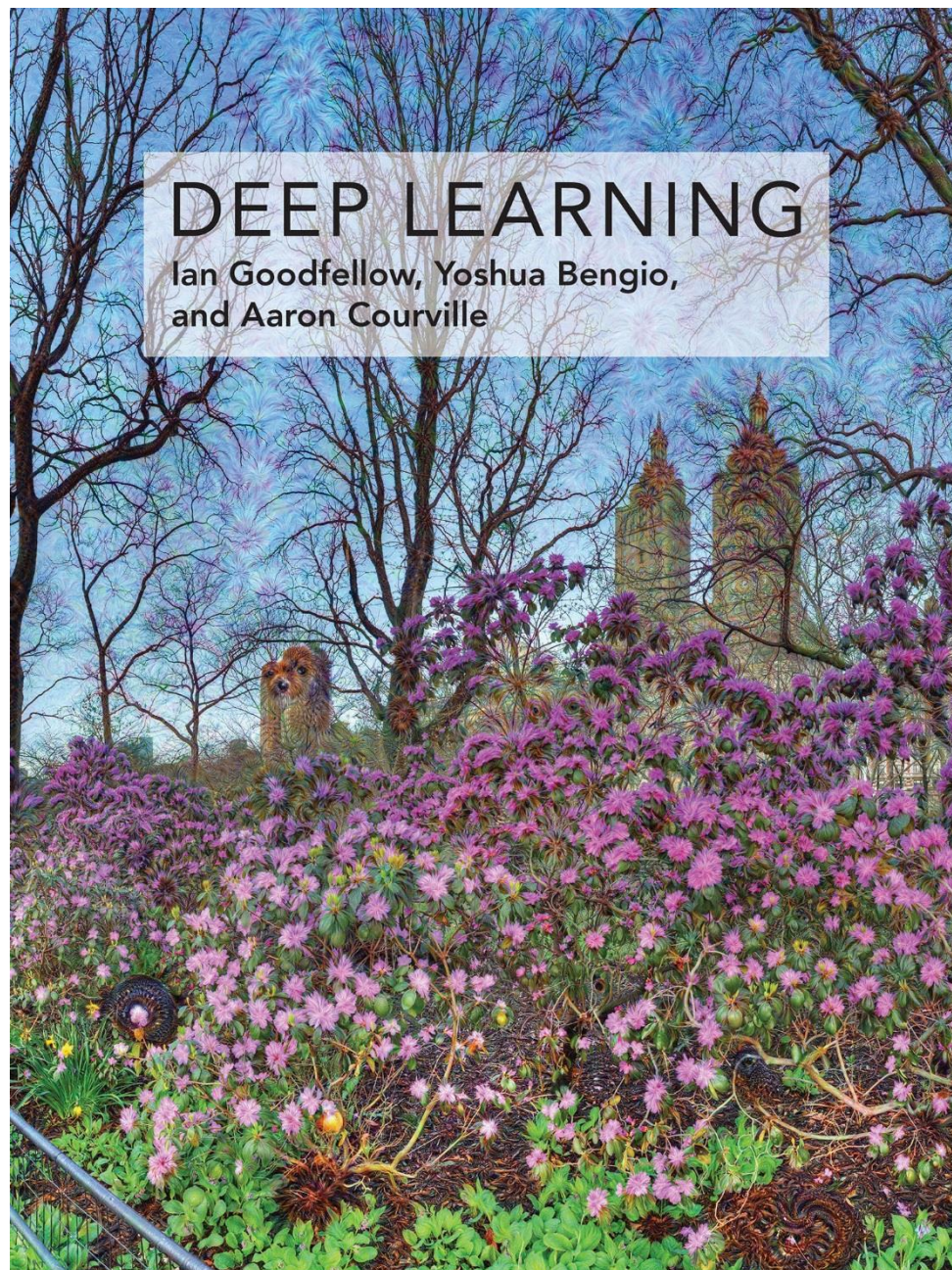
- 程式能力：python 會很有幫助
- 基礎知識：希望聽課同學具備機器學習的基礎知識
 - 沒有基礎知識的同學希望可以在上課前用上學期「機器學習」這門課的錄影預習
 - 3/03 上課前先預習
 - DNN: <https://www.youtube.com/watch?v=Dr-WRIEFefw>
 - CNN: <https://www.youtube.com/watch?v=FrKWiRv254g>
 - RNN (Part 1): <https://www.youtube.com/watch?v=xCGidAeyS4M>
 - RNN (Part 2): <https://www.youtube.com/watch?v=xCGidAeyS4M>

參考書籍

Original image:

<http://www.danielambrosi.com/Grand-Format-Collection/i-jbhqVhS/A>

<http://www.deeplearningbook.org/>



參考書籍

Part I: Applied Math and Machine Learning Basics

2. Linear Algebra

3. Probability and
Information Theory

4. Numerical
Computation

5. Machine Learning
Basics

Part II: Deep Networks: Modern Practices

6. Deep Feedforward
Networks

7. Regularization

8. Optimization

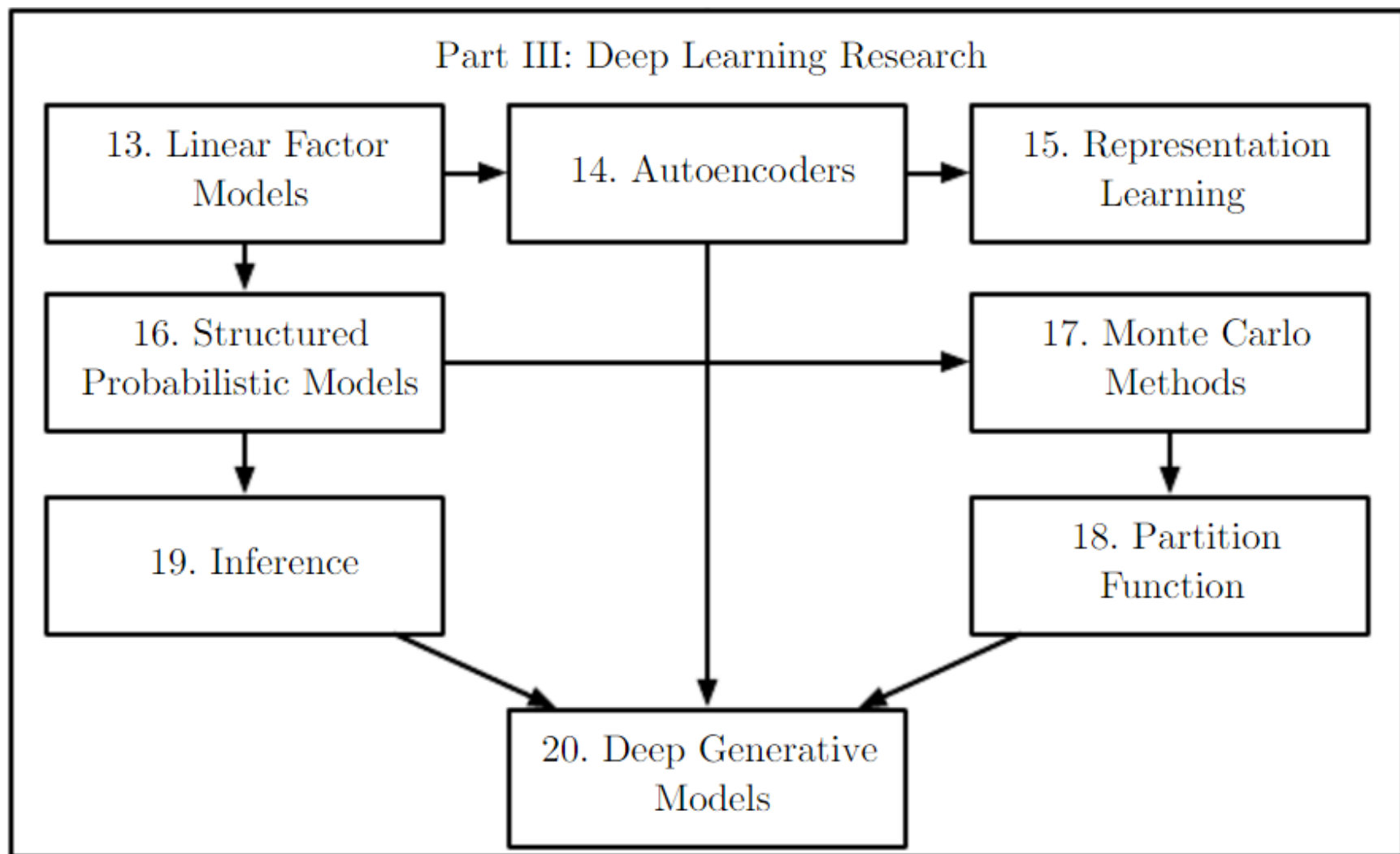
9. CNNs

10. RNNs

11. Practical
Methodology

12. Applications

參考書籍



作業內容

- 作業一：機器克漏字 (language modeling by RNN)
- 作業二：影片敘述自動生成 (sequence-to-sequence learning plus attention)
- 作業三：機器畫圖 (deep generative model)
- 作業四：聊天機器人 (sequence-to-sequence learning plus reinforcement learning)

作業進行方式

- 程式碼：程式碼符合指定格式可以順利執行，經助教要求修改後才能執行會被扣分
 - 本學期限制使用 **Tensorflow**
 - 請留意版本
- 繳交報告
- 機器學習成果評比詳見各作業說明

心理建設

- 機器學東西就是需要時間，等待的過程是虐心的
 - 作業早點開始
 - 死線前爆氣沒有什麼幫助
 - 健全的心靈
 - 試著調適等待過程的焦慮

運算資源

- 感謝計中本學期提供運算資源
- 其他免費運算資源
 - **Google Cloud Free Trial**
<https://cloud.google.com/free-trial/>
\$300 USD使用額度, 60天內有效, GPU: Nvidia Tesla K80 (0.8 USD per hour -> 約可連續用兩週)
 - **Rescale**
<https://www.rescale.com/deep-learning/>
提供 \$50 額度的 deep learning free trial, 含Tesla等級 GPU (Tensorflow, torch, pylearn已經裝好了)

Final Project

遇到問題，~~用c4~~就對了！

用 deep learning “硬train一發”

萬事皆可 train 的代表故事 – Fizz Buzz in Tensorflow

<http://joelgrus.com/2016/05/23/fizz-buzz-in-tensorflow/>

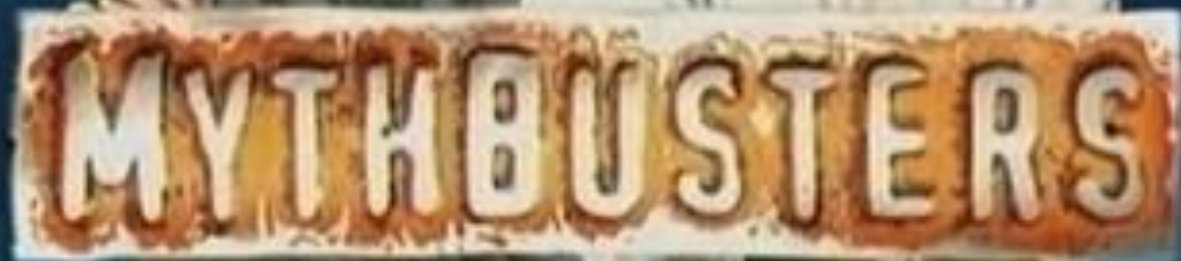
Motivation

- 人們嘗試用各種方法硬 train
 - “神農嘗百草”
- 在這過程中累積了大量尚待驗證的傳說



http://orchid.shu.edu.tw/upload/article/20110927181605_1_pic.png

Deep Learning



深度學習
流言終結者

Description

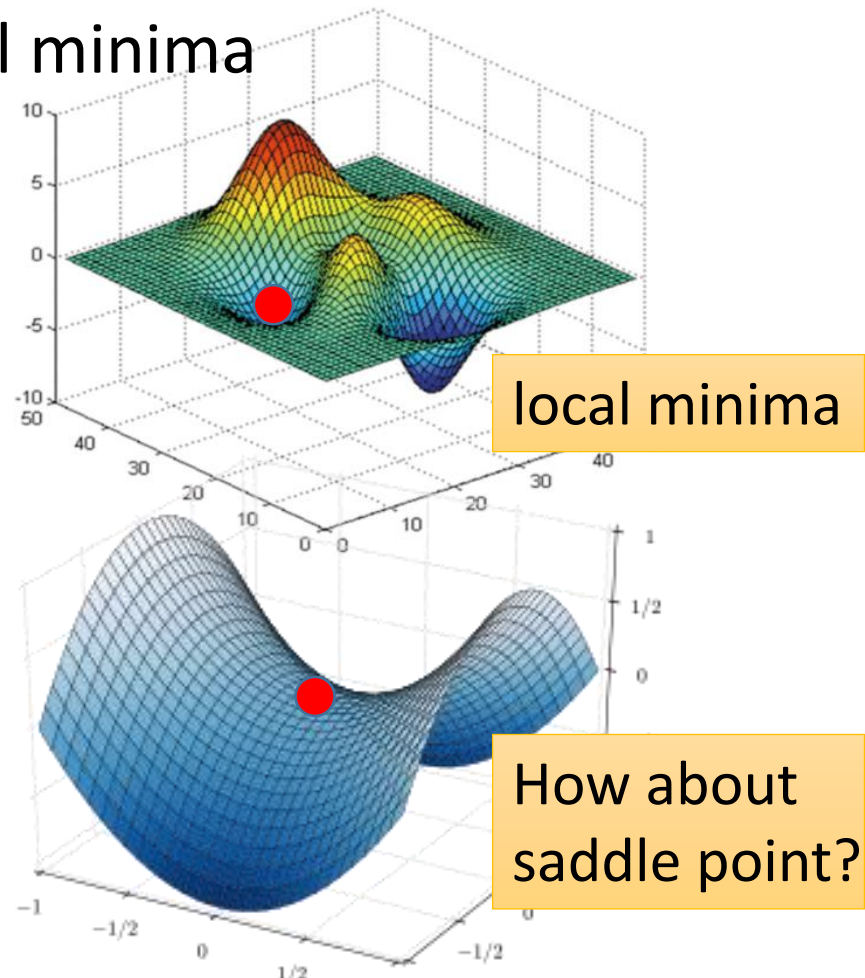
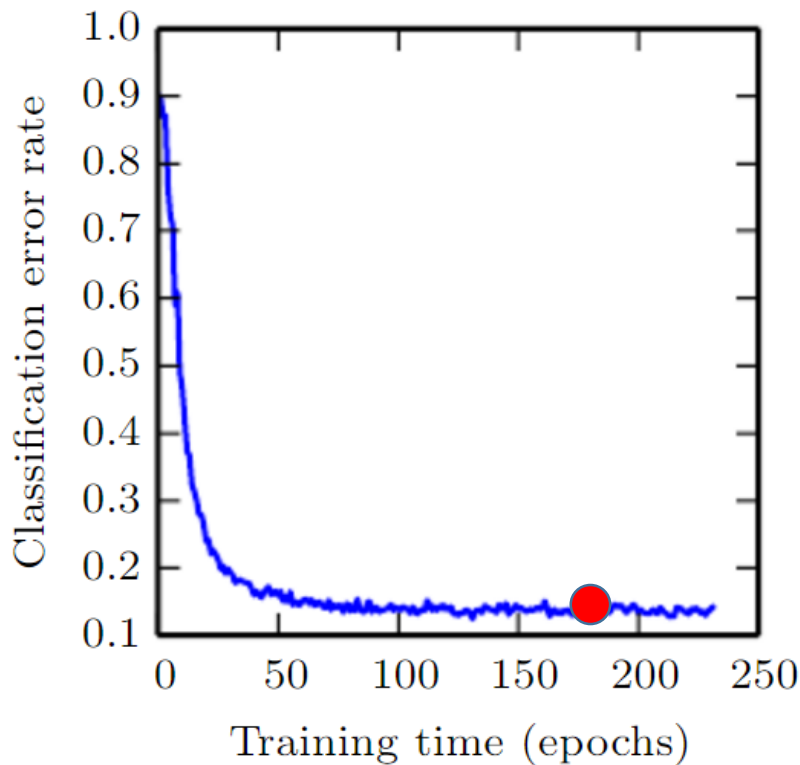
- 細節於 3/24 (和作業二) 一起公告
- Step 1. 尋找流言 (想要驗證的主題)
 - 尋找一個和深度學習有關的大眾信仰、問題、都市傳說 (e.g. 來自論文或論壇的訓練秘訣)
 - 就算已經有論文驗證也沒關係
 - 請於 4/28 前登記你的流言 (之後可修改)
- Step 2. 作實驗驗證自己的流言
 - 6/09 繳交報告初稿
 - 6/16 決定期末發表的組別
 - 6/23 期末發表

Example: RBM 是不是過譽了

- In the past, RBM initialization = Deep Learning
- Today, RBM is seldom used.
- Why it is not very helpful today?
 - We have more data today?
 - We are better at training today?
 - It is not very helpful at the beginning?

Example: Training stuck because ? (1)

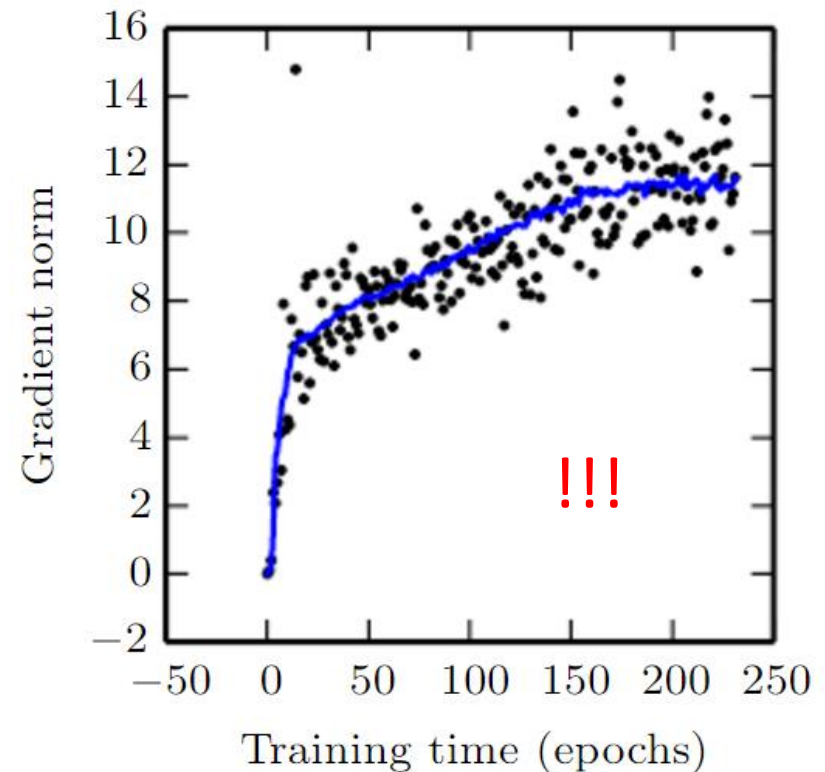
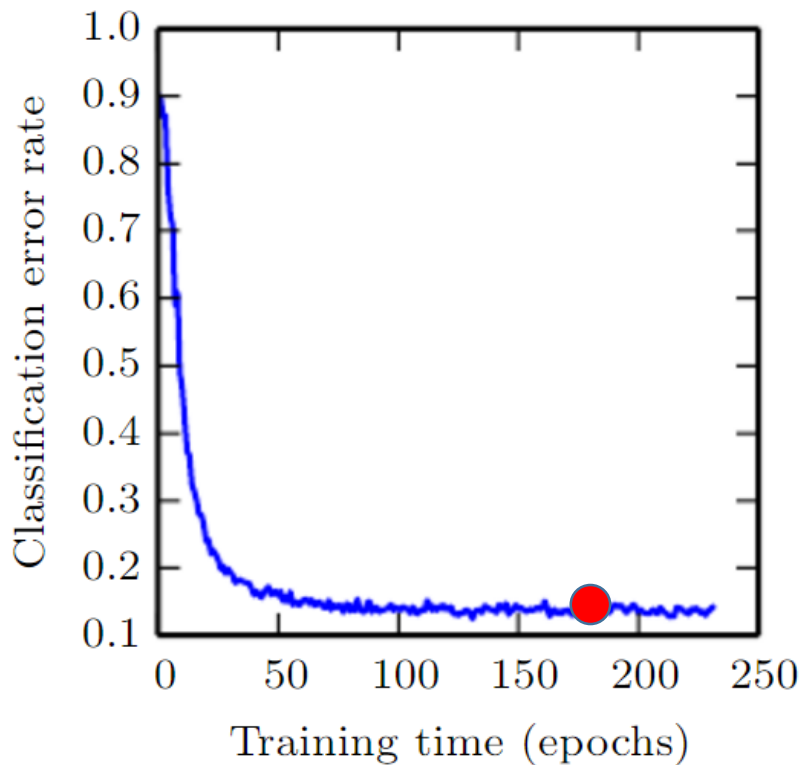
- People believe training stuck because the parameters are near a local minima



Example:

Training stuck because ? (2)

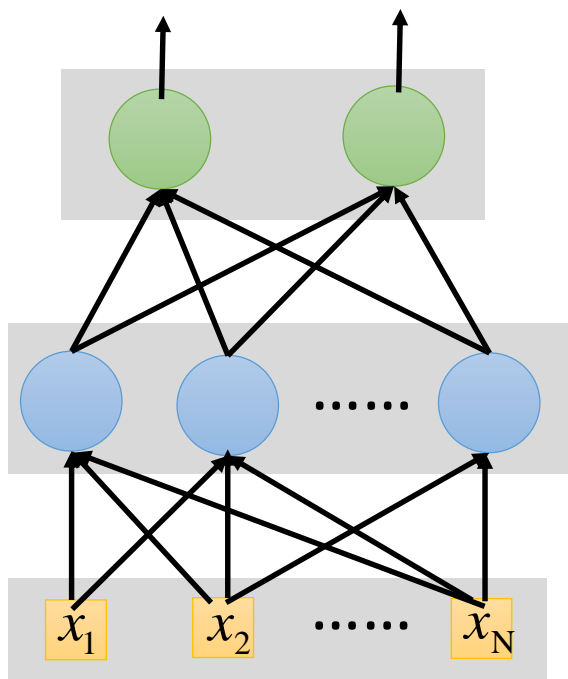
- People believe training stuck because the parameters are around a critical point



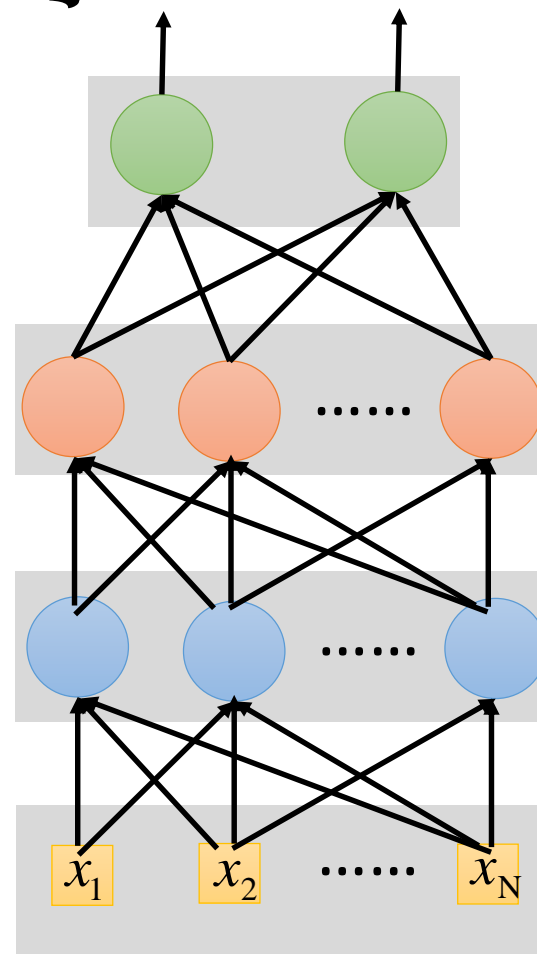
Example:

深度學習是不是過譽了？

Deep works better simply because it uses more parameters.



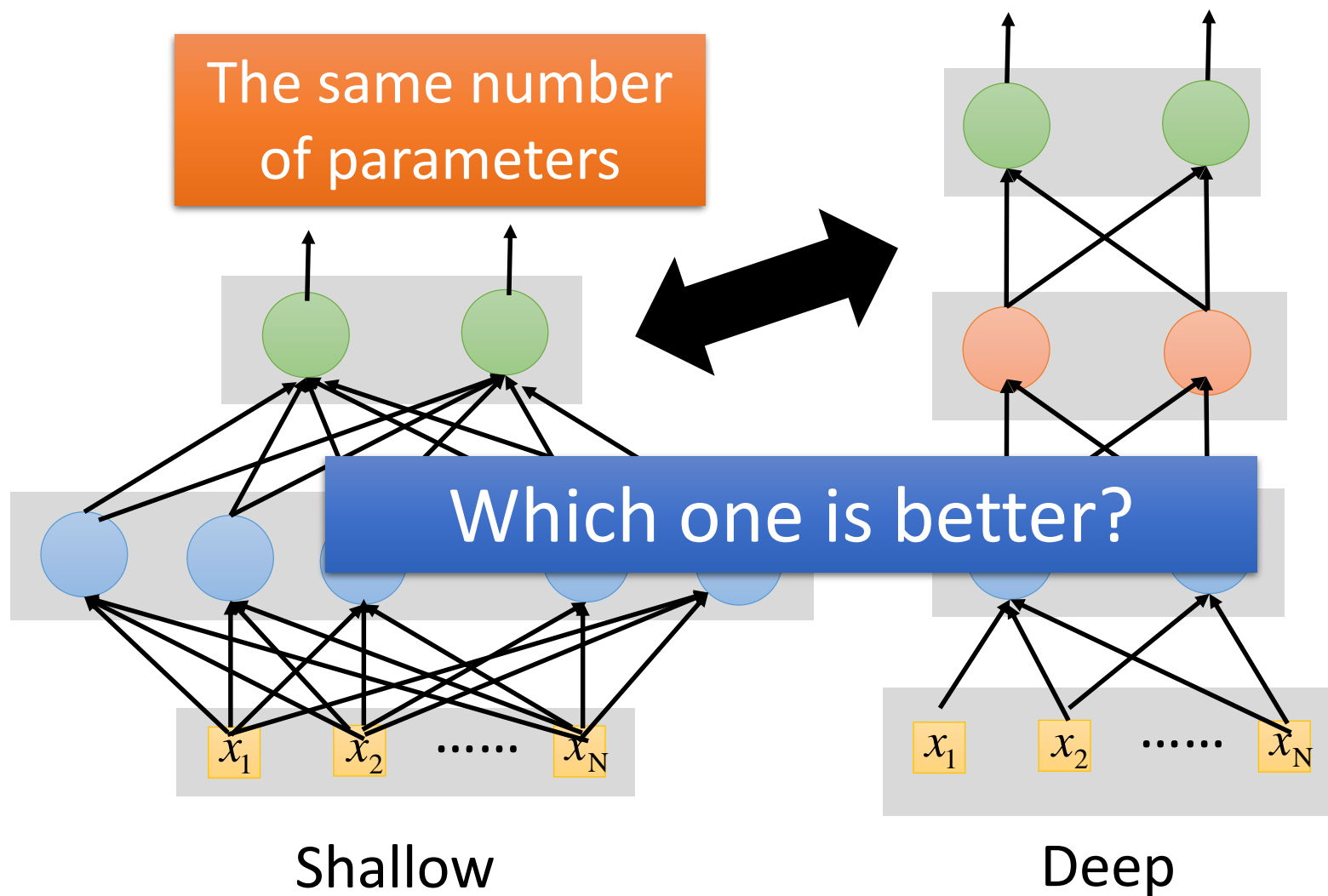
Shallow



Deep

Example:

深度學習是不是過譽了？



Example:

深度學習是不是過譽了？

- Discussion in the previous lecture:
 - [http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2015_2/Lecture/Brief%20ML%20\(v2\).ecm.mp4/index.html](http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2015_2/Lecture/Brief%20ML%20(v2).ecm.mp4/index.html)
 - <https://www.youtube.com/watch?v=XsC9byQkUH8>
- For some kinds of functions, deep structure can represent the them with less parameters
 - Shallow network is more likely to memorize the training data (overfitting)
 - We use deep learning because we don't have sufficient training data.

Example: Adam 是不是過譽了

- Usually Adam is the default optimization strategy.
- Adam harms the performance when training GAN.
<https://arxiv.org/pdf/1701.07875.pdf>



加簽

加簽

- 等一下助教會公告作業 0 ，週日(02/26)中午 12:00 前完成
 - 作業零是一個簡單的機器學習題目
 - 可以用任何機器學習方法完成，只要達到要求的正確率就行
 - 助教不會改作業零的程式，所以**不限套件**
 - 完成作業 0 後，助教會公告授權碼取得方式
 - **以個人為單位完成**

加簽

- 本課程預計加簽大約 **100** 人
- 如果完成作業零人數超過預計加簽人數，依照以下條件排序
 - 修越多和機器學習相關的課程且成績越好的同學就越優先
 - 相關程度由授課教師認定
- 如已經組好隊，隊伍中有兩人加簽，則加簽剩餘的同學
- 這門課下一個學期還會再開，如果有必須這學期修到這門課的理由，請寄信給老師