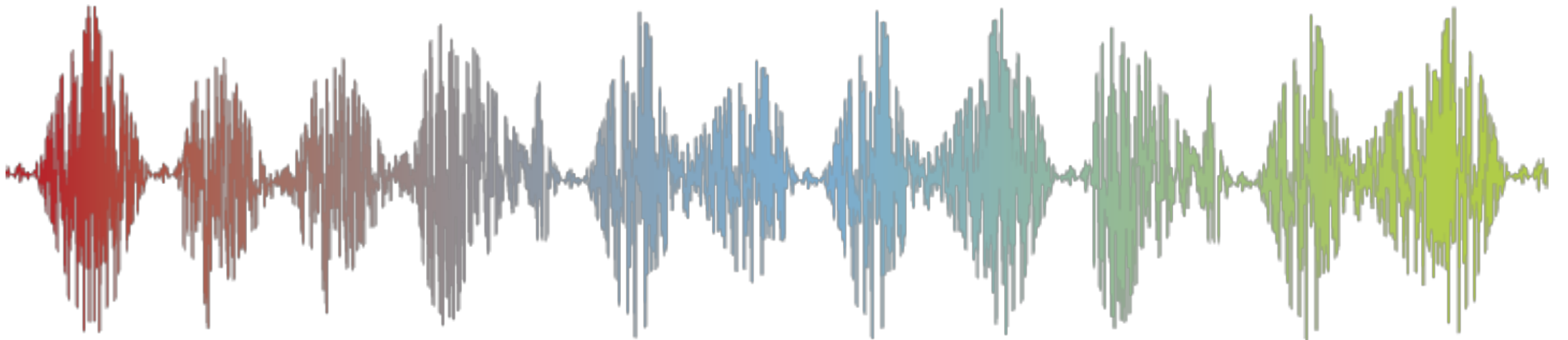
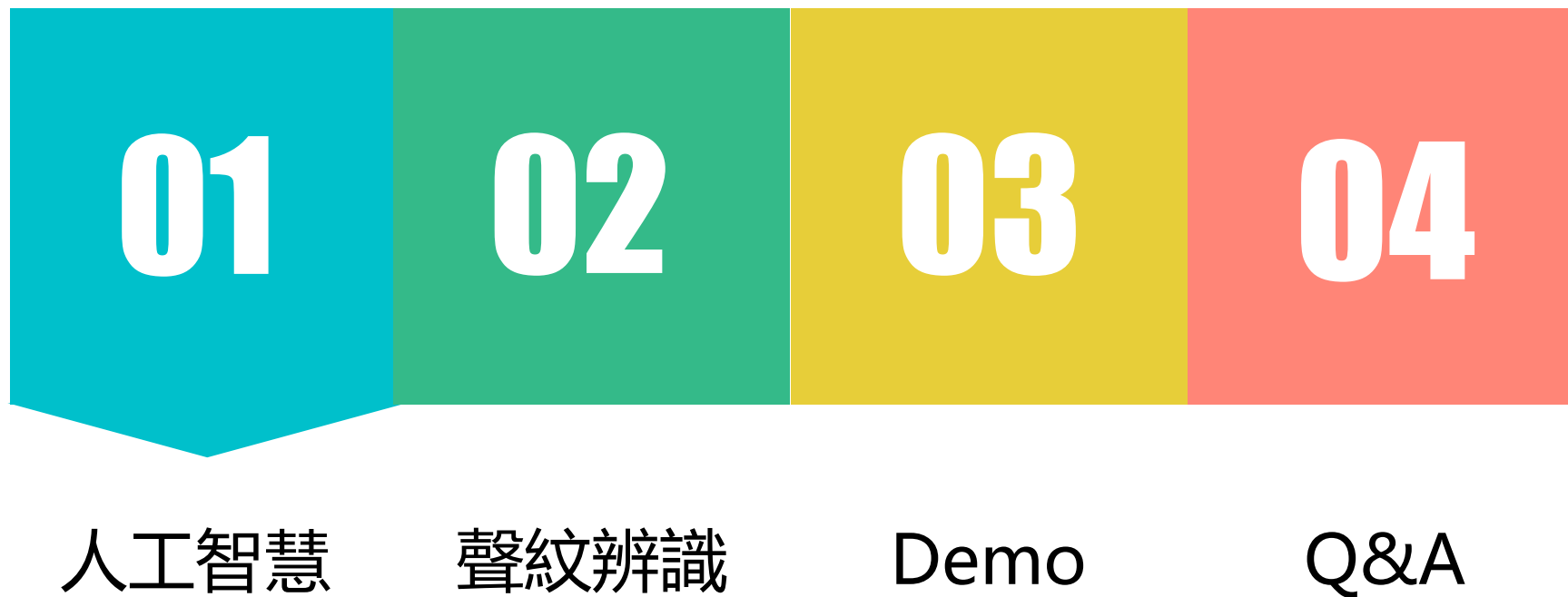


暑期實習-成果驗收報告

科技研發部 實習生
蘇玫如

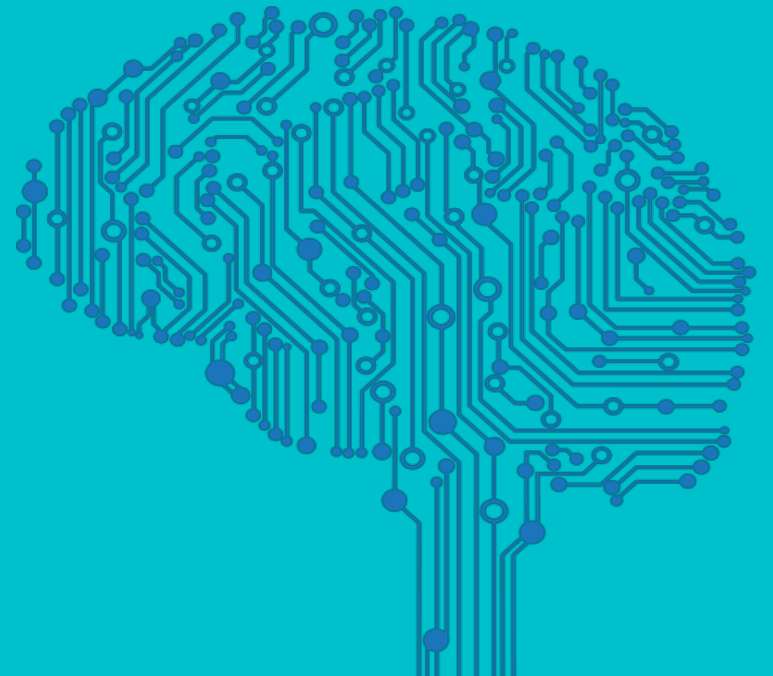


我在富邦人壽實習中學到什麼？

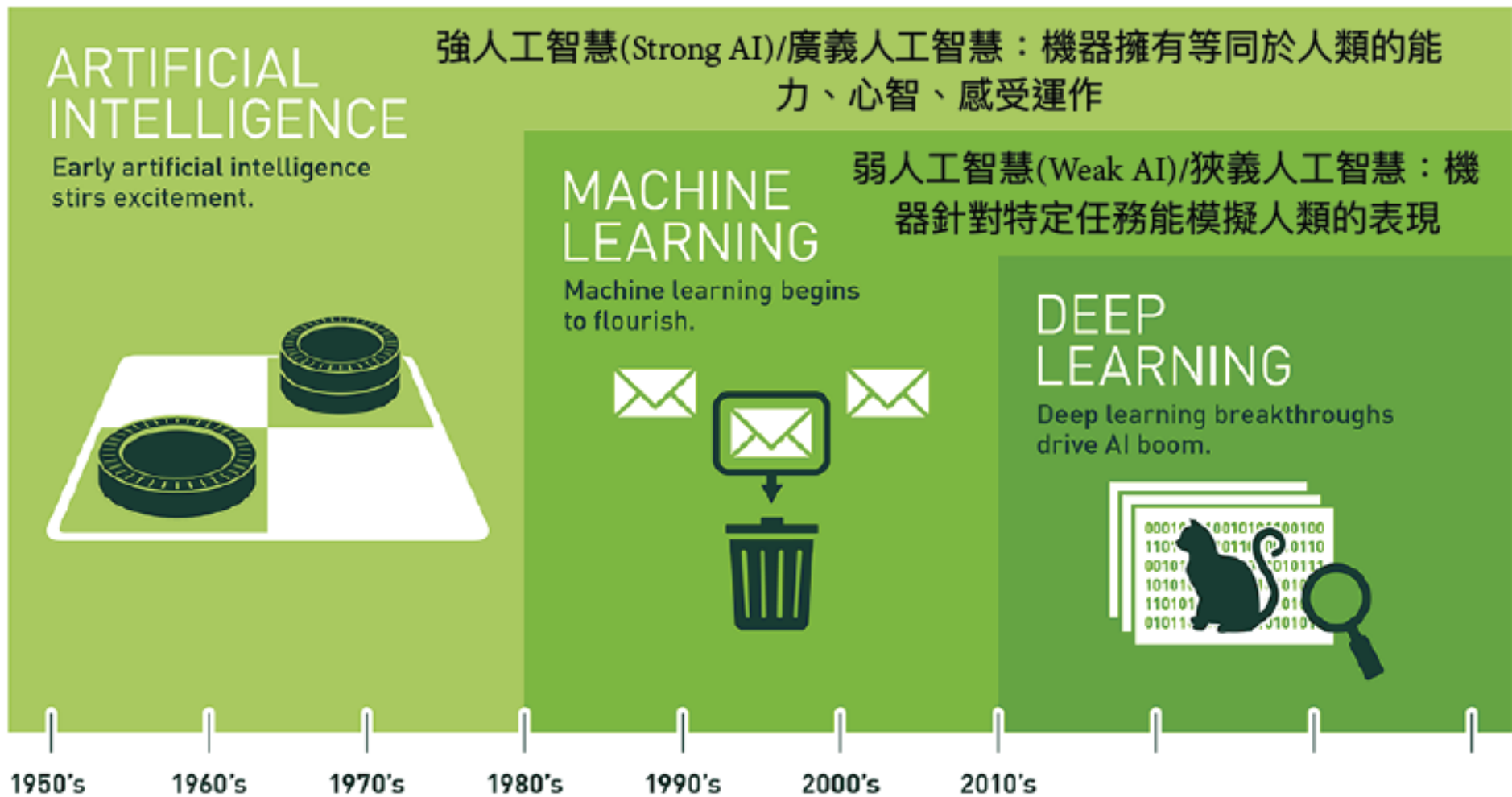


PART ONE

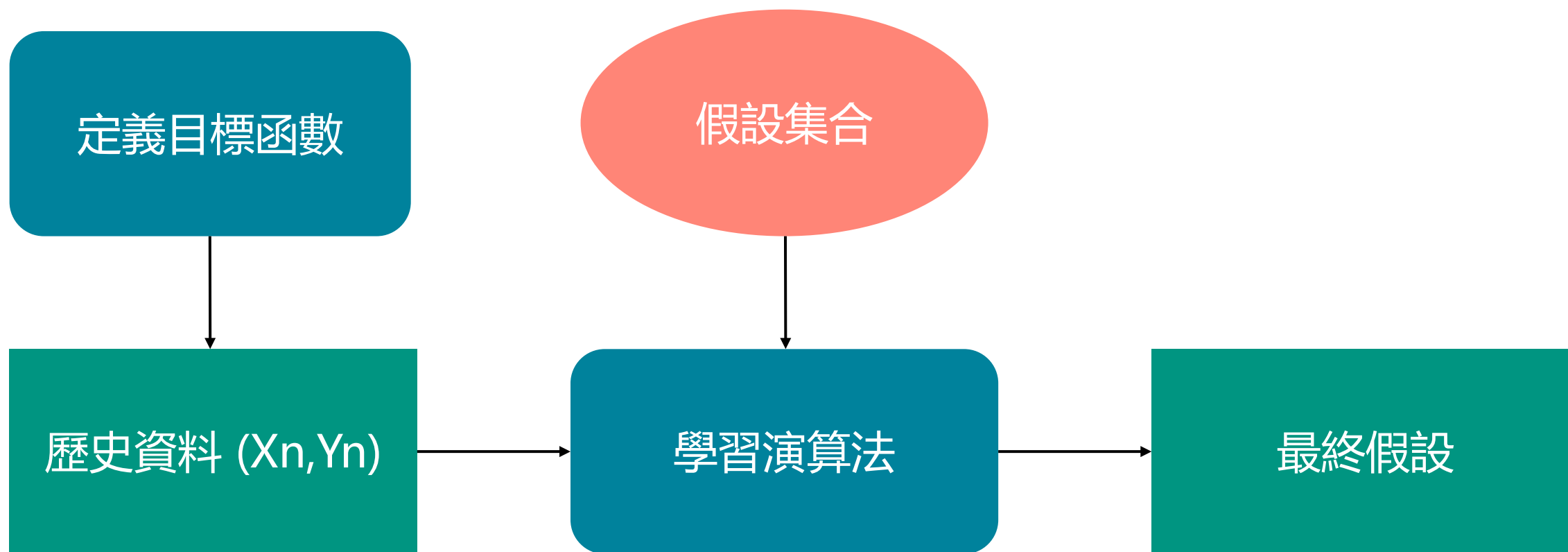
Artificial Intelligence



什麼是人工智慧？



機器學習



求得一個最好的假設，可以最逼近我們的目標函數

比較

機器學習 ML	資料挖礦 DM	人工智慧 AI	統計 Stat
<ul style="list-style-type: none">求得一個最好的假設，可以最逼近目標函數	<ul style="list-style-type: none">用大量資料去尋找特徵注重在計算效率強大的大型資料庫如果 特徵=假設，那 DM=ML。如果兩者相關，DM可協助ML	<ul style="list-style-type: none">表現出智慧的行為 如果假設是智慧的行為，ML 是實現 AI 的一種方式	<ul style="list-style-type: none">使用資料去做關於未知過程的推論Stat 對 ML 有助得到最佳假設

流程圖

Q: 如何選擇假設函式？

假設(Hypothesis)

資料擷取

Q: 如何最小化cost？

成本函數(Cost function)

資料預處理

Q: 如何處理不同種類資料？

Q: 如何調節超參數？

Q: 如何選擇學習演算法？

學習方法

資料切分

Q: 如何切分資料？

Q: 如何評估模型預測效果？

模型評估

Q: 模型是否有過適現象？

學習曲線/驗證曲線



Good Model

學習方法

	監督式學習	集體學習	非監督式學習
概念	有X跟Y	集思廣益 整合多種方法，來降低 因不同學習演算法特性 所產生的誤差	有X沒Y 透過資料的彼此作分群
方法	<ul style="list-style-type: none">• KNN• 羅吉斯迴歸• SVM• 決策樹	<ul style="list-style-type: none">• 投票分類器• 裝袋法• 隨機森林• XGBoost	<ul style="list-style-type: none">• KMeans• DBSCAN

可以針對現有的資料及假設問題，選擇最佳的流程、分析、評估

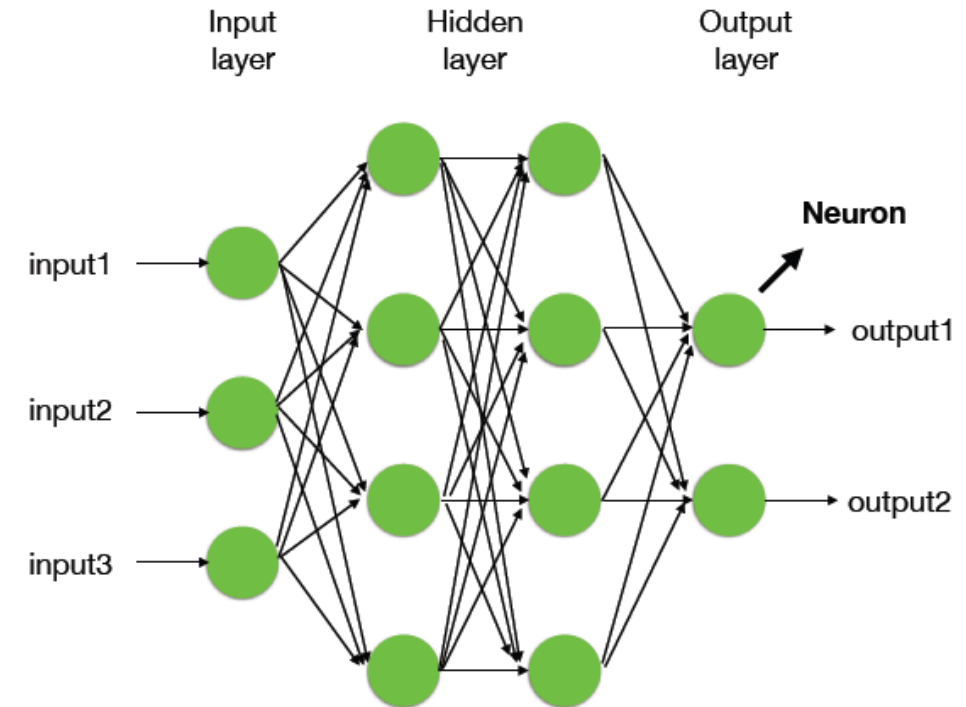
深度學習

- 常應用在語音辨識、圖像辨識
- 模擬大腦神經元運作方式

$f(\text{audio waveform}) = \text{'Hello World'}$

$f(\text{cat image}) = \text{'cat'}$

$f(\text{Go board}) = (K, 10)$



學習方法

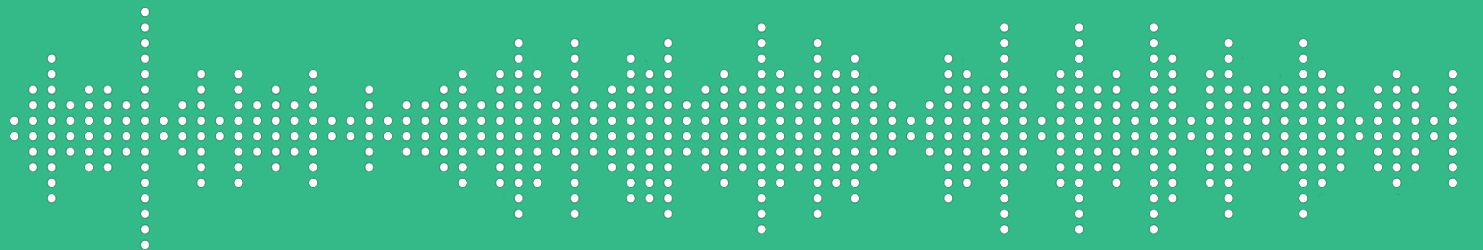
	卷積神經網路 CNN	遞歸神經網路 RNN	長短期記憶 LSTM
架構	<ul style="list-style-type: none">• 卷積層• 填補層• 池化層	<ul style="list-style-type: none">• X 本期• Y 下期	<ul style="list-style-type: none">• X 本期• Y 下期
特性	無時間性	<ul style="list-style-type: none">• 有時間性• 有遺忘現象	<ul style="list-style-type: none">• 有時間性• 可決定是否遺忘• 效果遠大於 RNN
主要應用	影像辨識	<ul style="list-style-type: none">• 語音辨識• 預測股價• 對話生成	<ul style="list-style-type: none">• 語音辨識• 預測股價• 對話生成

可以針對問題的需求，來選擇最適合應用的方法

2

PART TWO

Voiceprint Recognition



生物辨識

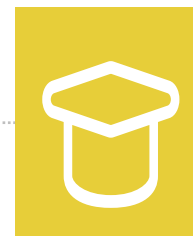


	指紋	掌紋	虹膜	人臉	靜脈	聲紋
使用簡易度	高	高	低	中	中	高
使用者接受度	中	中	中	中	中	高
準確度	高	超高	超高	高	高	高
成本	中	超高	超高	高	超高	低
遠端授權	可行	可行	可行	可行	可行	已實施
手機收集	部分已實施	可行	已實施	已實施	可行	已實施

概論

聲紋模式

主要以人體解剖學及行為差異
為評斷標準



研究機構

中國科學研究院、部分美國大學、
台灣調查局

方法

訊號處理、GMM、HMM、
DNN、監督式分類



應用

目前: Siri、監聽、銀行生物辨識
未來: 台灣鑑識科學應用、語音遙控、
防止中國盜領社會保險金

使用疑慮

模仿、合成、變聲器、
受情緒、老化、生病等影響



總結

使用人工智慧
多增加一道生物辨識驗證的保障

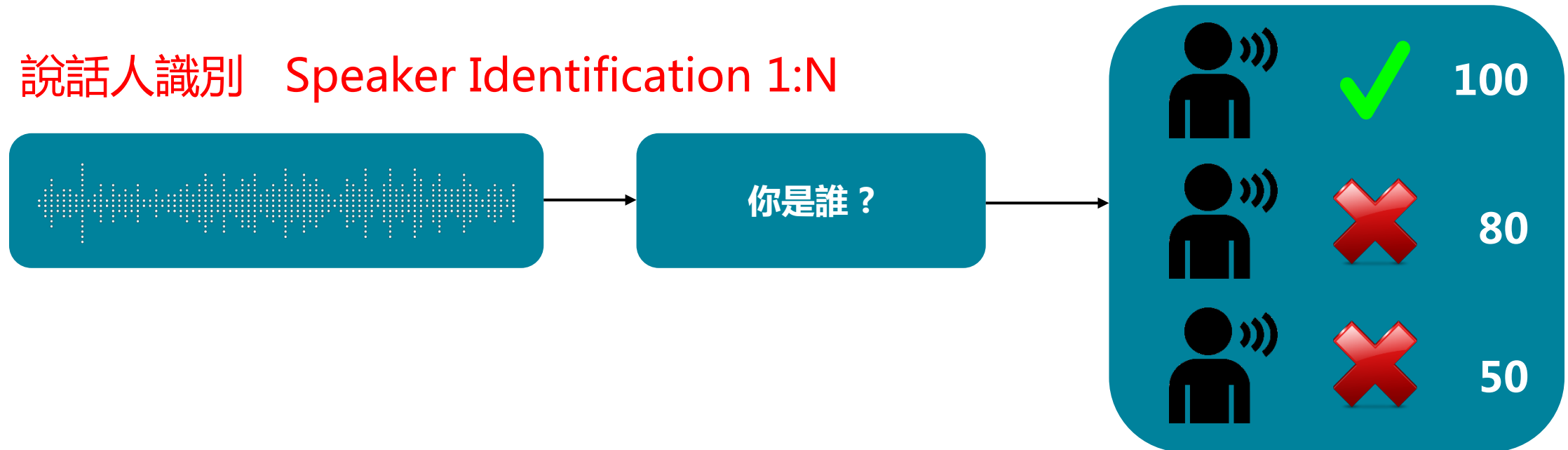
以上疑慮雖能改變音色導致人類觀察不出來，
但實際上很難改變生理構造所產生的聲紋特徵

功能

說話人驗證 Speaker Verification 1:1



說話人識別 Speaker Identification 1:N



Voice Gender Detection

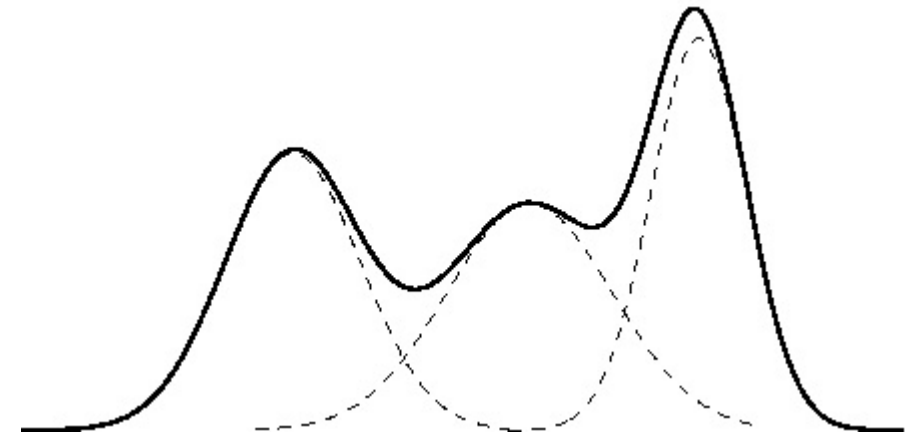
- Input: Google Audio Dataset (Size: 1116)
- Feature: MFCC_delta2

	GMM	XGBoost	Random Forest	MLP	CNN
Accuracy	89%	54%	55%	50%	60%

- 樣本達千筆時可選擇GMM
- 樣本高達萬筆時可以嘗試使用CNN，CNN效果或許比GMM更好
- 取特徵值時選擇MFCC_delta2

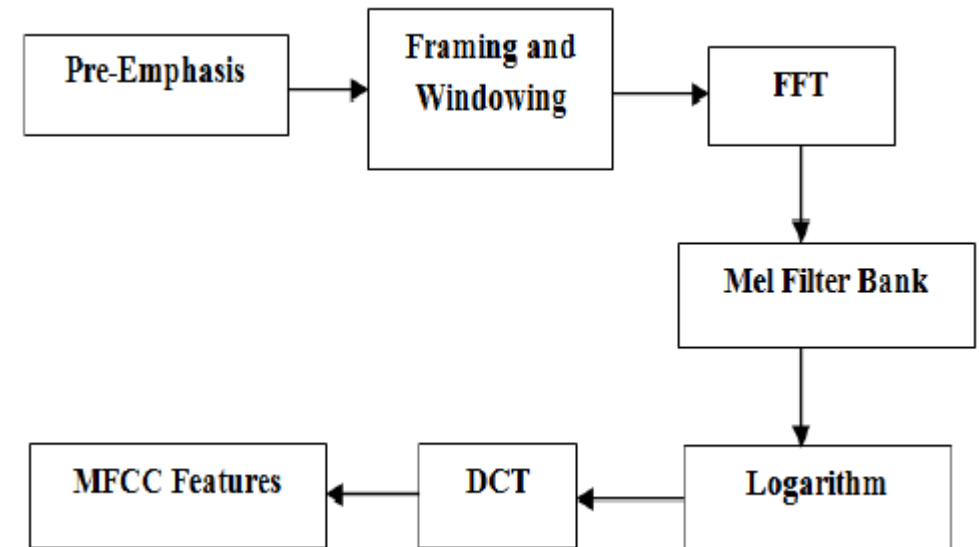
混合高斯模型 GMM

- 多個常態分佈混合的統計模型
- EM法 [Estimation(猜)、Maximum Likelihood(反思)]
- 重複迭代，直到得出最接近母體的參數
- 在應用於語音辨識上，就是做半監督式分群



梅爾頻率倒譜 MFCC

- 基於聲音頻率非線性梅爾刻度中，對數能量頻譜的線性變換
- 特徵提取包含兩個關鍵步驟：轉化到梅爾頻率，然後進行倒譜分析
- 二階微分後，變成 MFCC_delta2



CNN 結果

Layer (type)	Output Shape	Param #
conv1d_5 (Conv1D)	(None, 1205, 32)	1696
max_pooling1d_5 (MaxPooling1D)	(None, 301, 32)	0
conv1d_6 (Conv1D)	(None, 298, 64)	8256
max_pooling1d_6 (MaxPooling1D)	(None, 74, 64)	0
flatten_3 (Flatten)	(None, 4736)	0
dense_121 (Dense)	(None, 32)	151584
dense_122 (Dense)	(None, 2)	66
Total params: 161,602		
Trainable params: 161,602		
Non-trainable params: 0		

```

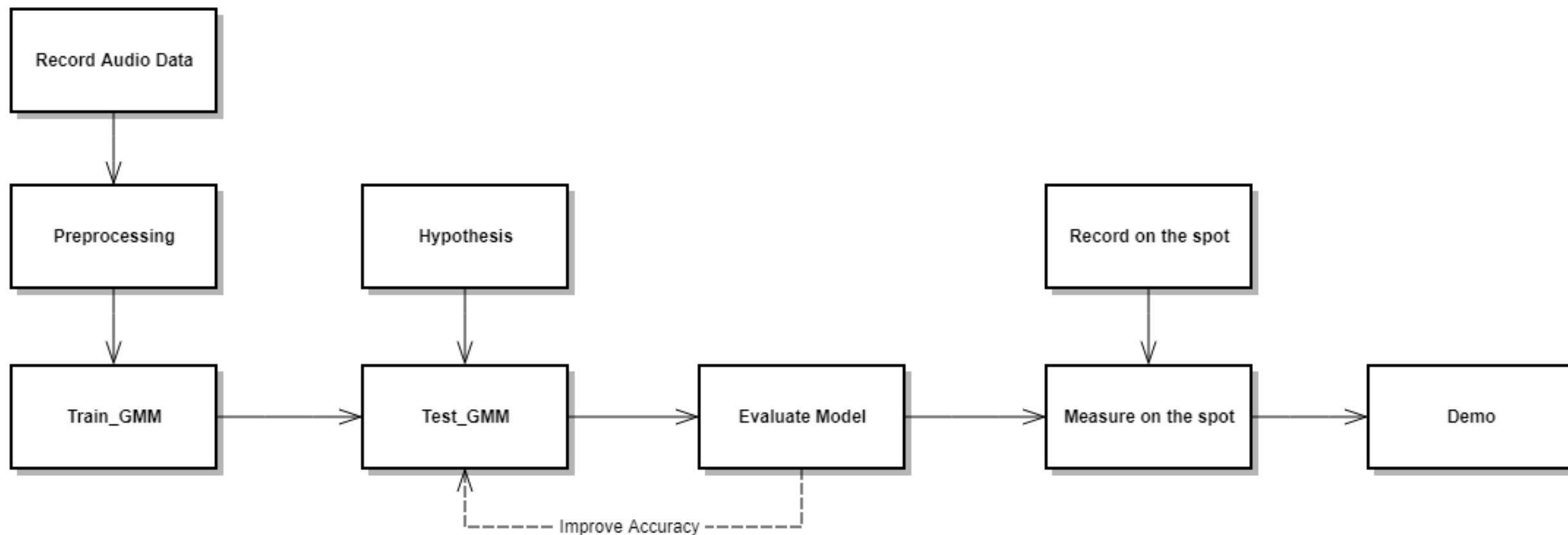
Epoch 1/10
781/781 [=====] - 9s 12ms/step - loss: 1.1153 - acc: 0.5147
Epoch 2/10
781/781 [=====] - 5s 6ms/step - loss: 0.6926 - acc: 0.4930
Epoch 3/10
781/781 [=====] - 5s 6ms/step - loss: 0.6889 - acc: 0.5352
Epoch 4/10
781/781 [=====] - 5s 6ms/step - loss: 0.6768 - acc: 0.5736
Epoch 5/10
781/781 [=====] - 5s 7ms/step - loss: 0.6409 - acc: 0.6095
Epoch 6/10
781/781 [=====] - 5s 6ms/step - loss: 0.5497 - acc: 0.7657
Epoch 7/10
781/781 [=====] - 5s 6ms/step - loss: 0.4317 - acc: 0.8592
Epoch 8/10
781/781 [=====] - 5s 6ms/step - loss: 0.3184 - acc: 0.9155
Epoch 9/10
781/781 [=====] - 5s 6ms/step - loss: 0.2116 - acc: 0.9706
Epoch 10/10
781/781 [=====] - 5s 6ms/step - loss: 0.1386 - acc: 0.9898

Train_loss, Train_accuracy = [0.10158014477787494, 0.9948783611518626]
Test_loss, Test_accuracy = [0.7486329093797883, 0.597014925728983]

```

Overfitting !!

科技研發部_聲紋辨識



聲紋收集

• 彭瑋玉
林淑芳
陳冠霖

林上淳
吳偉誌
陳威良

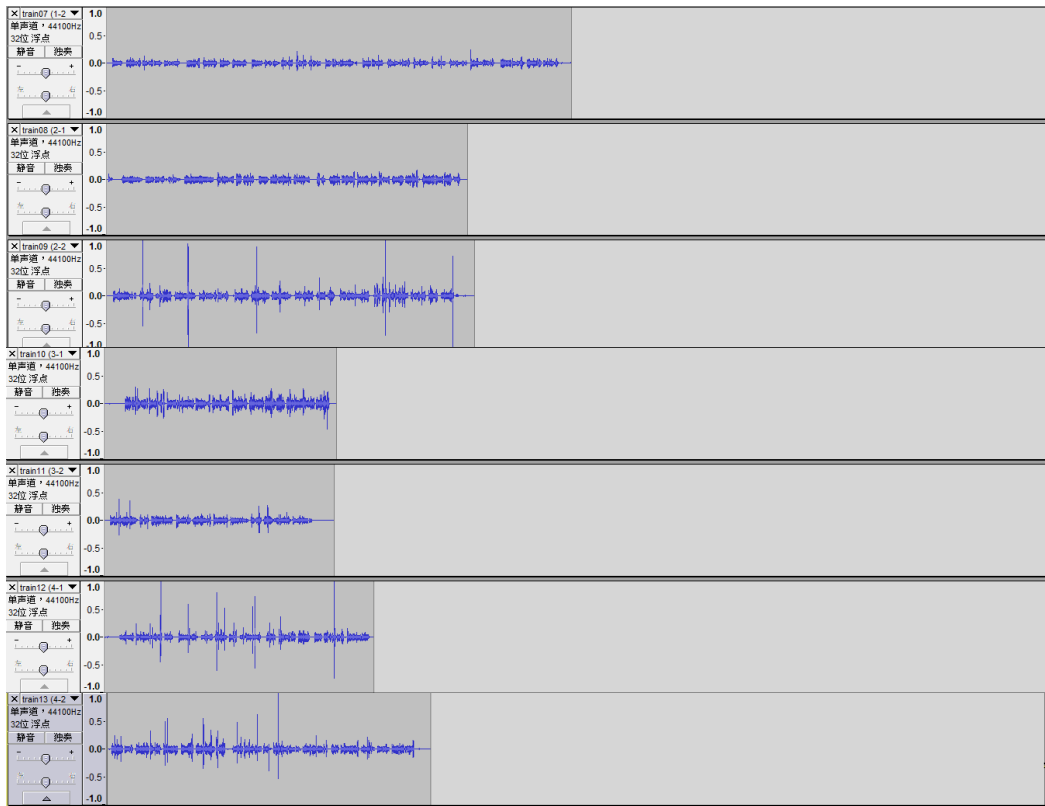
任卓英
賴昇頤
蘇玫如

徐天鴻
邱建豪

柯青宏
陳祐嘉

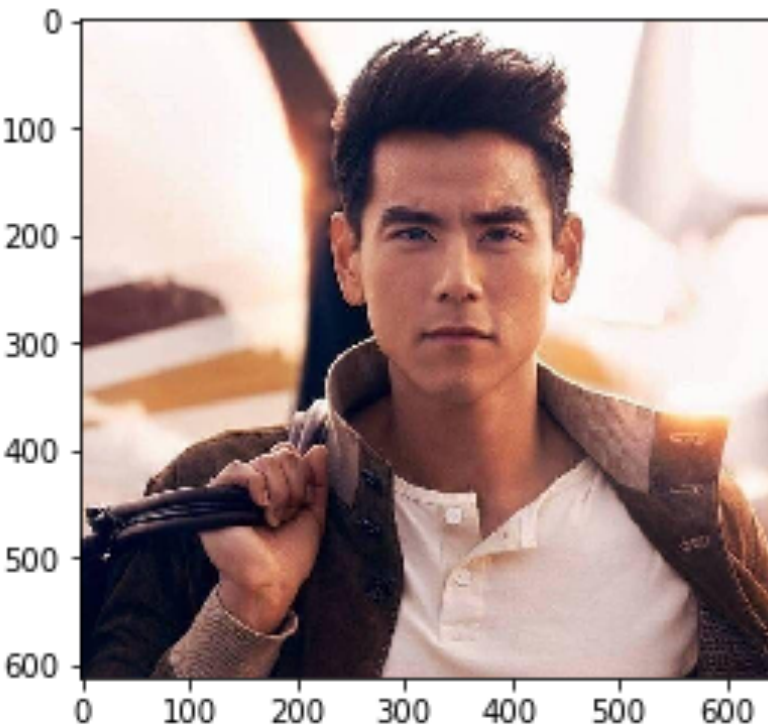
陳慶安
賴嘉平

洪立全
陳亭嘉



功能實作

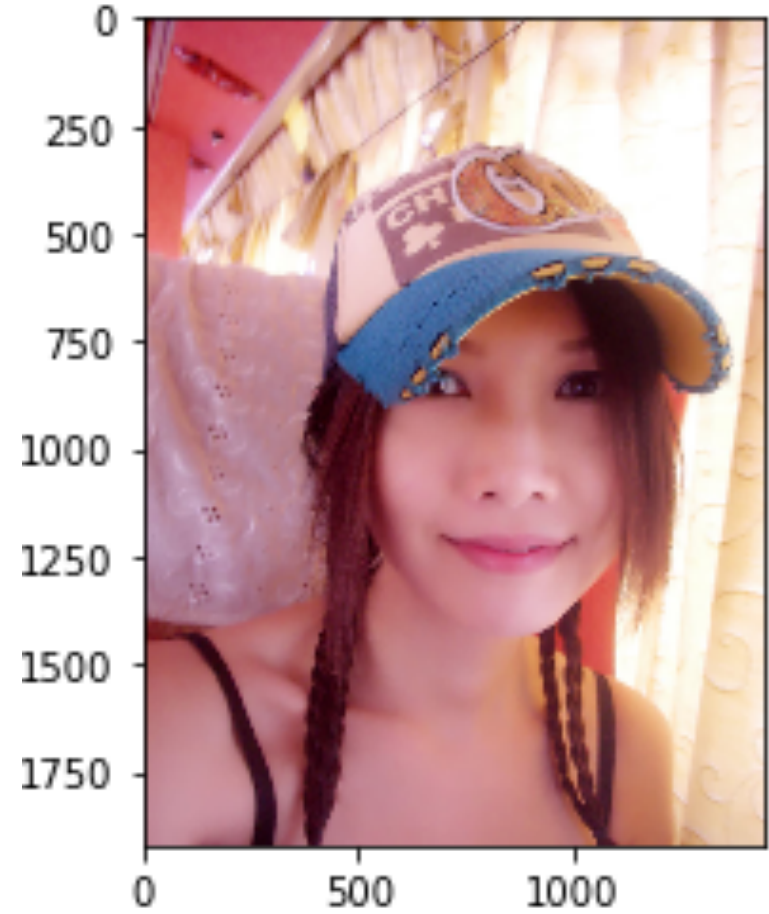
Detected as- 陳慶安
Similarity: 100 %



Correct~ May I help you, 陳慶安?

說話人驗證

detected as - 彭瑋玉



說話人辨識

困難

- 收音跟錄音時聲紋振幅不同
- 除錯前，某些模型效果不彰：

青宏、亭嘉、冠霖、立全、建豪、威良、祐嘉、玫如、昇頤

- 測試資料空白音過多會影響模型判斷
- 除錯前準確率: 76%
- 除錯後準確率: 96%





PART THREE

Demo

錄音時讓聲音盡量錄多點，
避免空白音過長喔～

DEMO

4

PART FOUR

Q & A

Have a
Question?

THANKS FOR YOUR PATIENCE



Thank You!

