



2020

國泰人壽專題——人臉辨識BMI

國泰人壽 沈維倫、陳宗霆
指導老師

台大 張智星 老師

政大 彭金隆 老師

東吳 蔡芸琇 老師

台大 資工系 徐子修

東吳 巨資系 簡大為

台大 森林系 朱毅豪

政大 風管系 周奕成

目錄



提案亮點介紹



技術說明



使用方式講解



應用場景



提案亮點介紹 | PART1

提案背景介紹



我們的產品是

F . A_S

Face analysis and application

客戶臉部分析及應用服務

Accuracy – 正確率的定義

定義：

預測BMI與實際BMI的誤差在正負10%之內

舉例：

若實際BMI為20，則預測數值落在18-22的區間內，皆算是正確

提案績效

經過經過國泰人壽的測試集驗證
用臉部辨識所計算出的BMI數值

平均準確率

0.919701434 91.97014

高達 **91.97 %** 的判斷準確率

以最高標準解決業務上的需求！





技術説明 | PART2

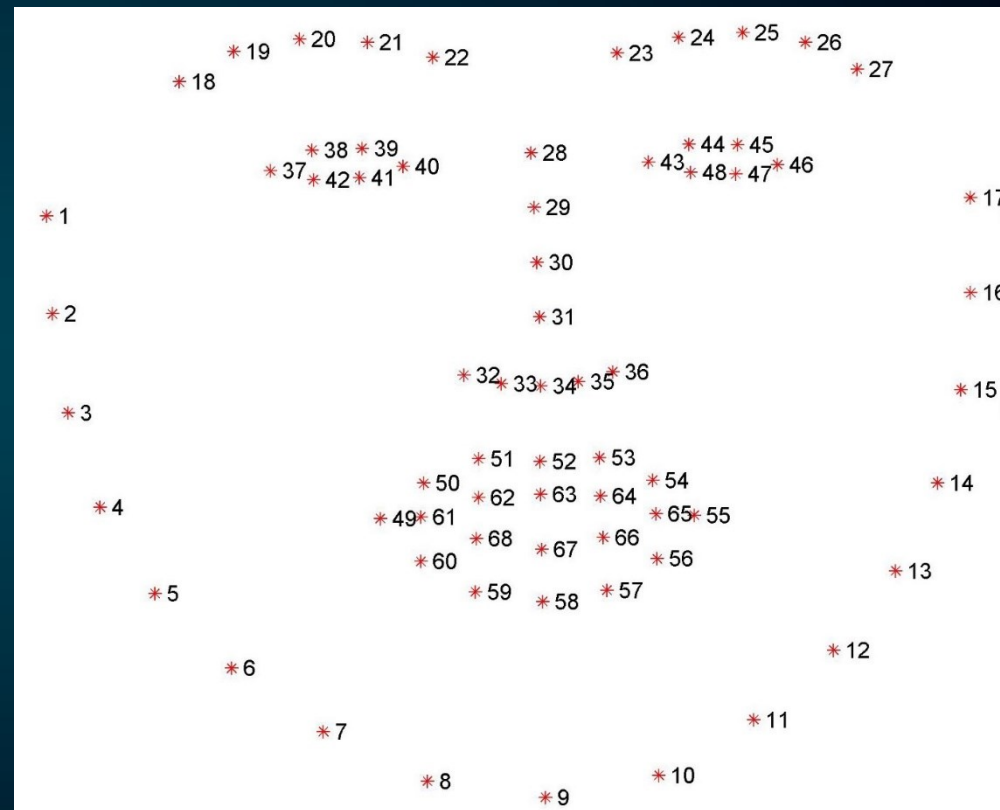


臉部比例特徵之SVR分析法

核心技術 — 動機



- 鑒於資料筆數只有5000多筆
- 透過增加臉部特徵，去放大BMI高與低的差異性
- 彌補資料樣本過少的問題
- 使用 Python - dlib 套件
- dlib 的 Landmarks 數相對多，多達 68 個點
- 同時 dlib 的使用率是非常高的





核心技術 — 特徵設計

1、參考論文：

- A computational approach to body mass index prediction from face images
- Deciphering Faces: Quantifiable Visual Cues to Weight
- 基於人臉圖像的 BMI 預測演算法研究

2、自行設計：

除了參考上述論文有使用到的臉部特徵外，我們也有再自行設計新的特徵，同時依照特徵之算法，將特徵分成三類：

- A. Ratio Type（長度比例類型） - 臉部座標間距離的比例數值
- B. Area Type（面積比例類型） - 臉部座標圍出的面積之間的比例
- C. Index Type（複合計算類型） - 有運用長度比例與面積比例做計算



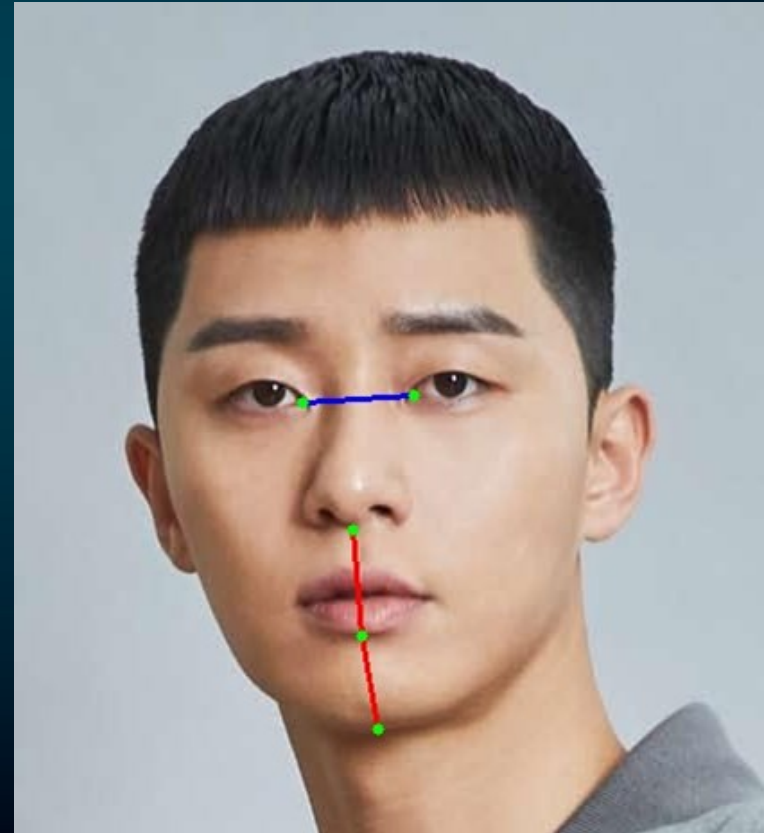
臉部比例特徵說明



由於照片中的每張臉當初與鏡頭的距離都不一樣，造成每張臉的像素大小差異很大，為了盡量消除，臉部像素大小差異帶了的影響，我們找出了一些比較不會隨著年紀增長而改變的特徵，作為臉部的衡尺。

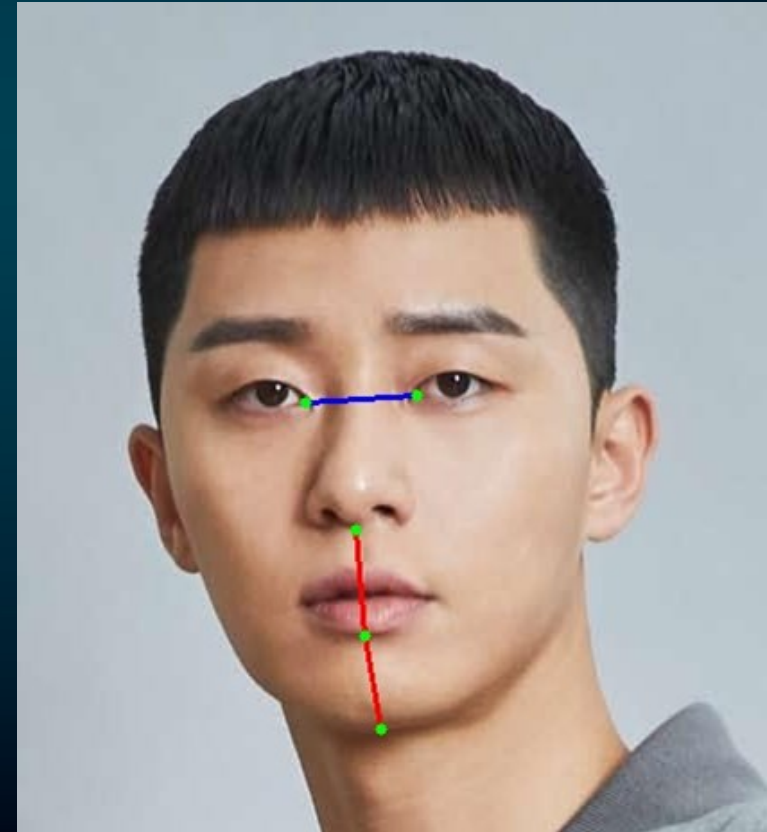
衡尺分別為：

- 水平向的眼距 (藍色線段)
- 垂直向的人中
到下巴距離 (紅色線段)





若臉部特徵的算法會需要用到水平或垂直的支線段，在計算上就會先將線段與臉部衡尺一起做計算，透過這樣的方法讓每張照片在大小差異上造成的誤差下降





透過 XGBoost 方法，可以利用 gain 值去評估每個特徵，對於電腦去區分 BMI 高低的貢獻，數值越高，代表該特徵對於整體預測越重要。



核心技術

CFI - Index Type

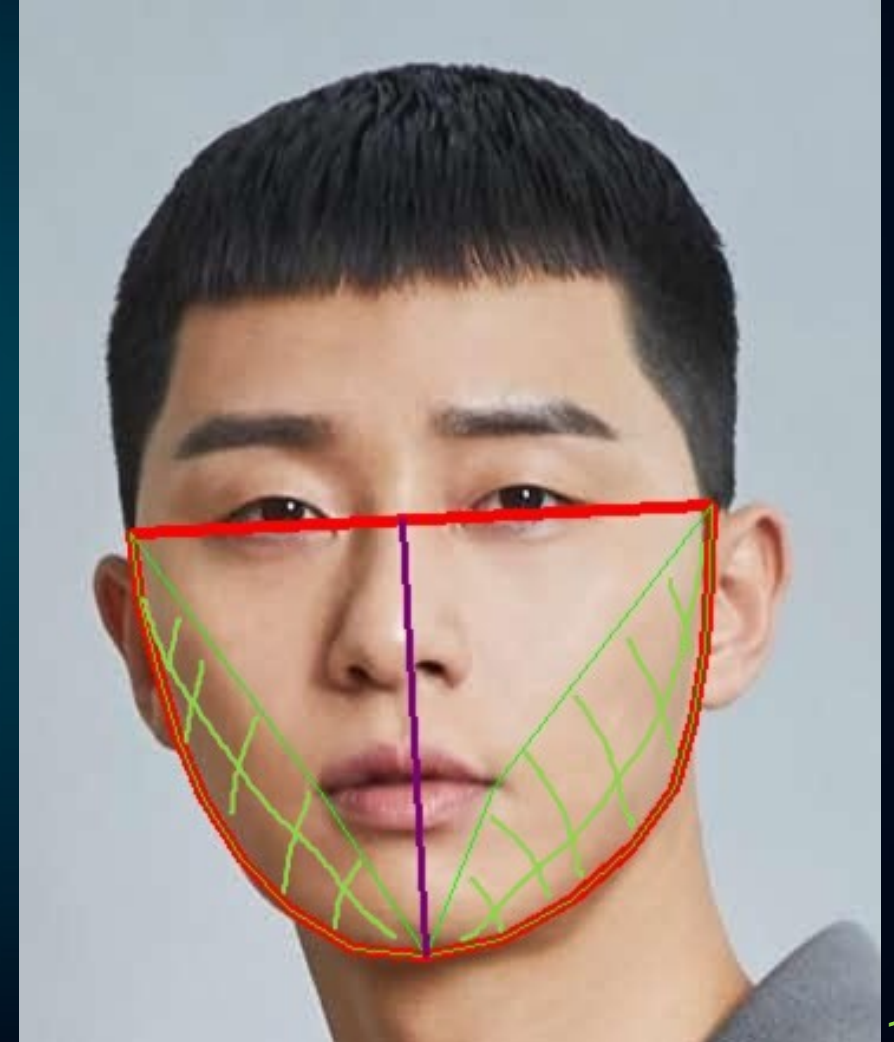


Cheek Fat Index

Ref: 原創

公式:

$$\div \frac{(\text{綠色區域面積} \div \text{紅色線段矩形面積})}{(\text{紫色線段長度})}$$



核心技術

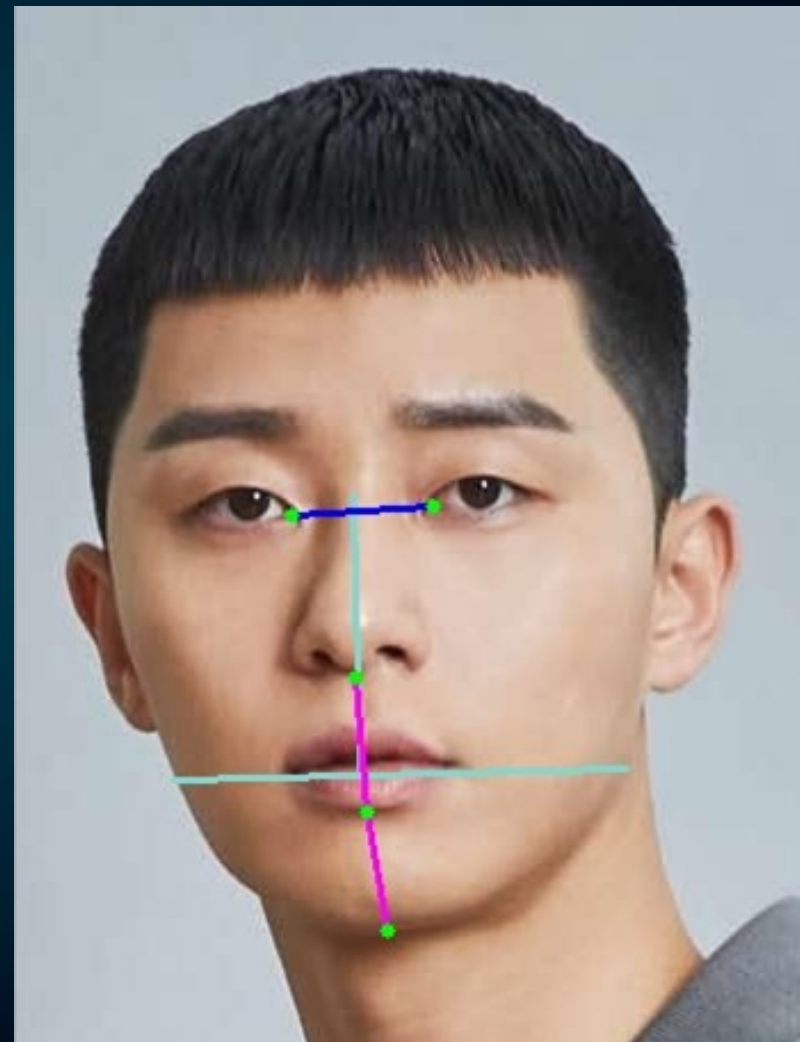
ITA - Area Type



Inverted T sign Area

Ref: 原創

公式：
$$\frac{(\text{水藍色線段相乘})}{(\text{藍色線段} \times \text{紫色線段})}$$



核心技術

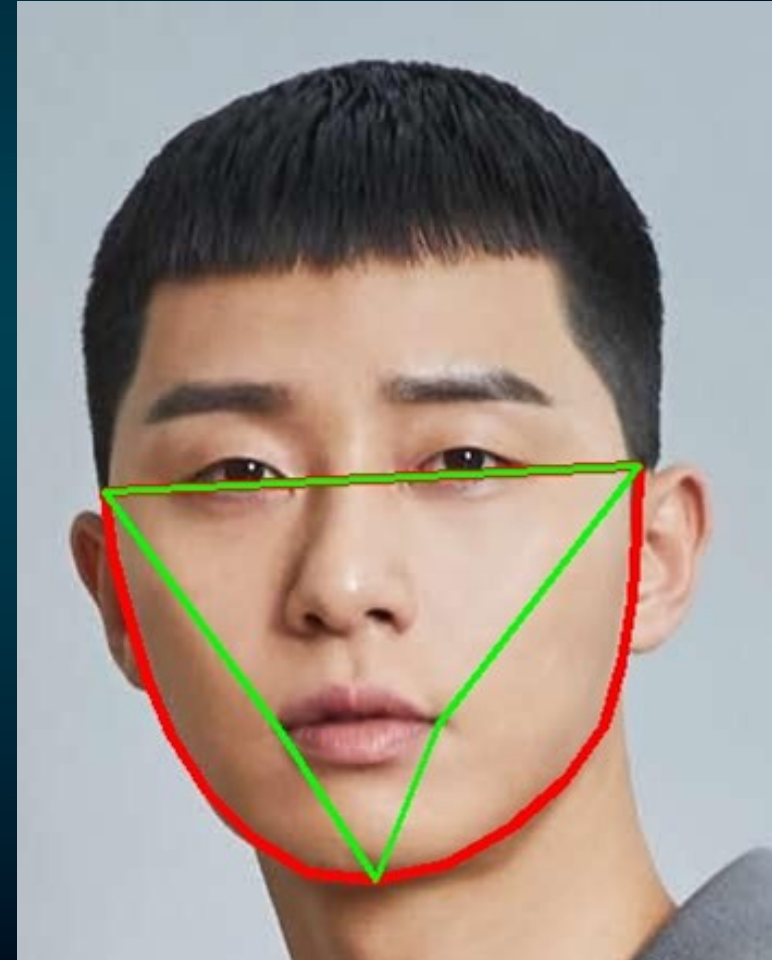
FPYA - Area Type



Face Pyramid Area

Ref: 原創

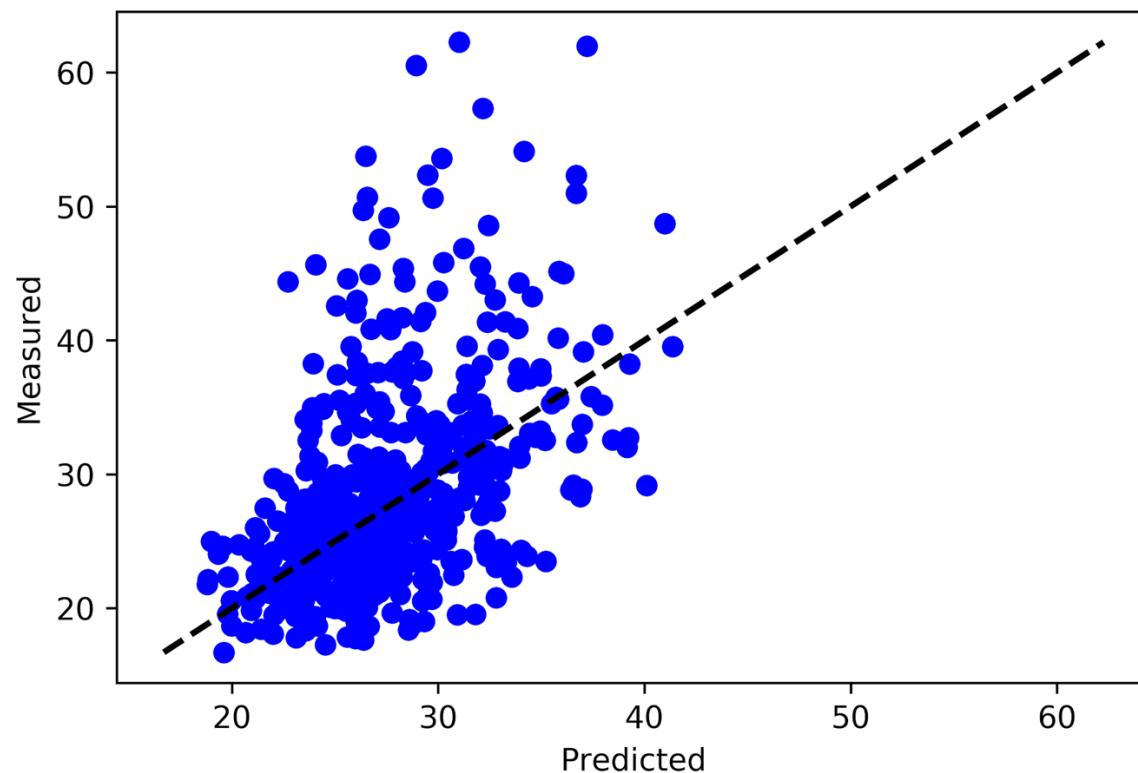
公式: $\frac{(\text{綠色線段矩形面積})}{(\text{紅色線段矩形面積})}$



核心技術 — Evaluation



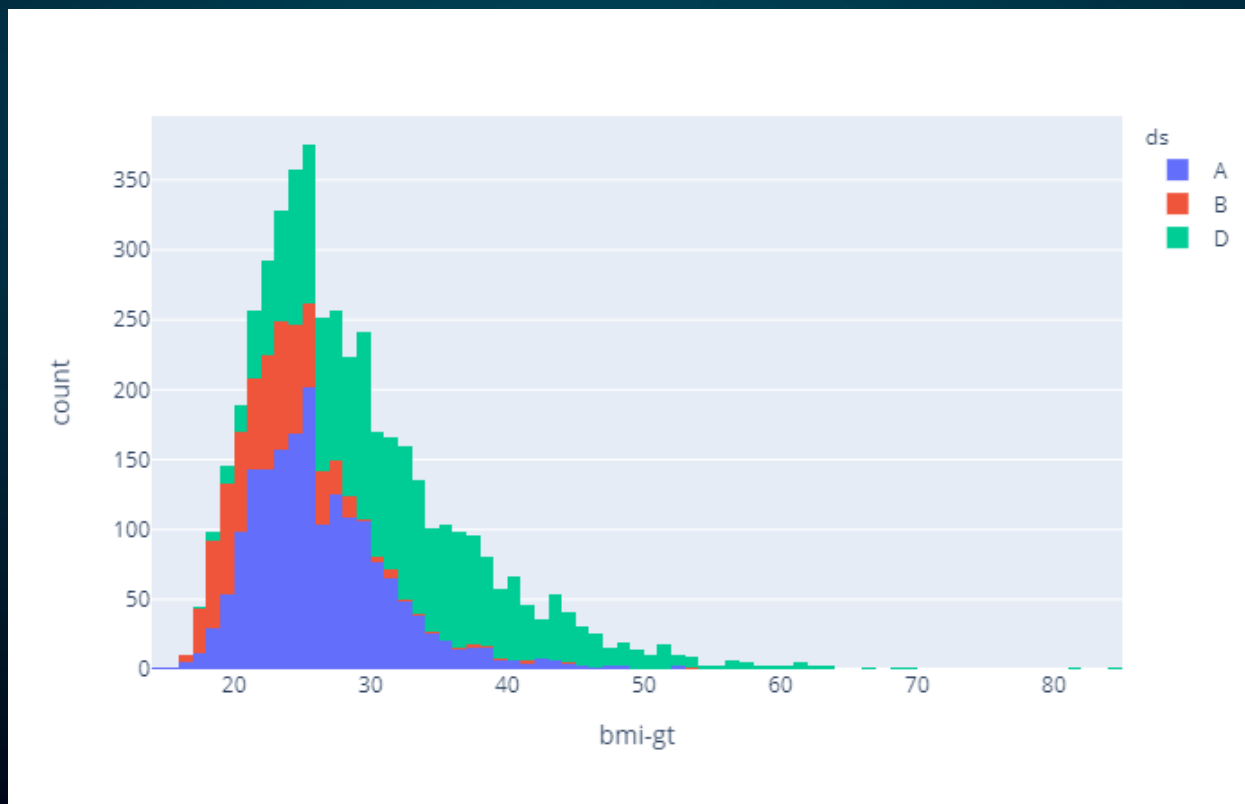
透過上述 18 種類型的特徵，利用 SVR (支援向量機) 方法建出的預測模型。若預測之 BMI 值與實際之 BMI 值在正負 10% 之內算是正確的話，這個方法的 Accuracy 平均落在 46% - 52% 左右。



核心技術 — Evaluation



在預測 BMI 較低與 BMI 超過 40、50 的人的時候，其誤差非常大，我們認為這和資料本身之傾斜程度有關，訓練資料的 BMI 值大都是落在 21 - 30 之間，我們這次總用了三個資料集，而三個資料集都有非常相似的分布情形

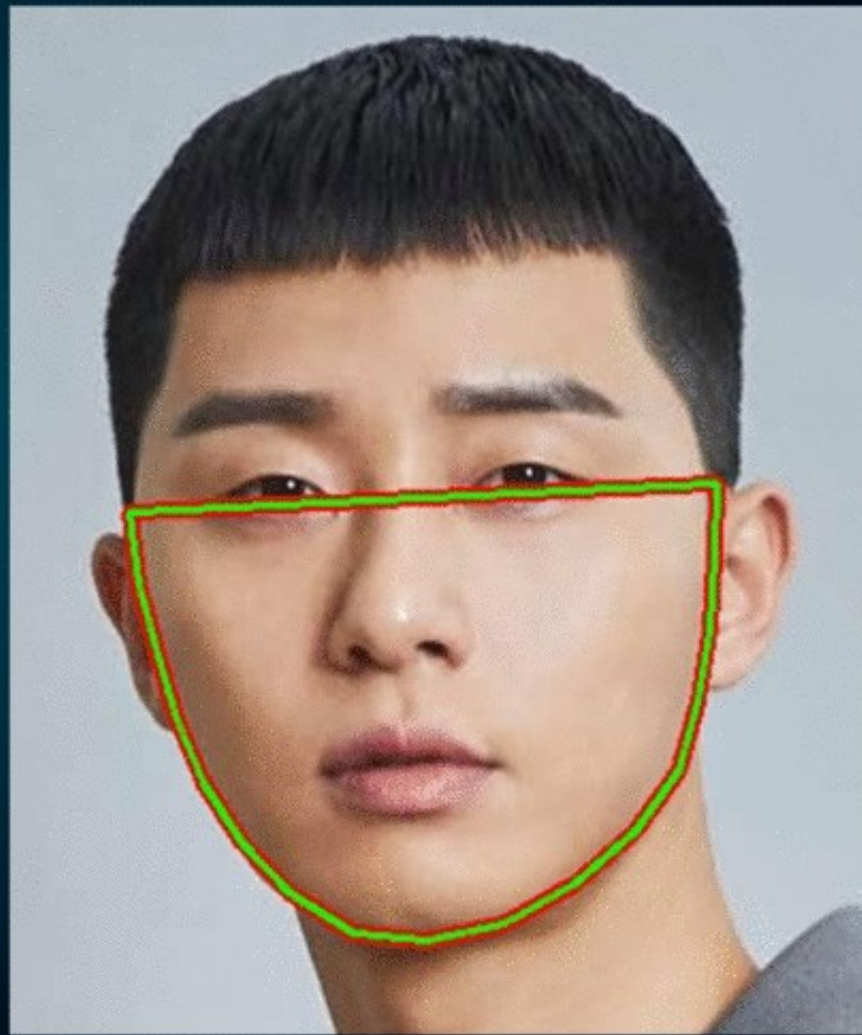




Lower Face Index

Ref: A computational approach to body from face images

公式:
$$\frac{\text{(綠色線段長度)}}{\text{(紅色線段矩形面積)}}$$



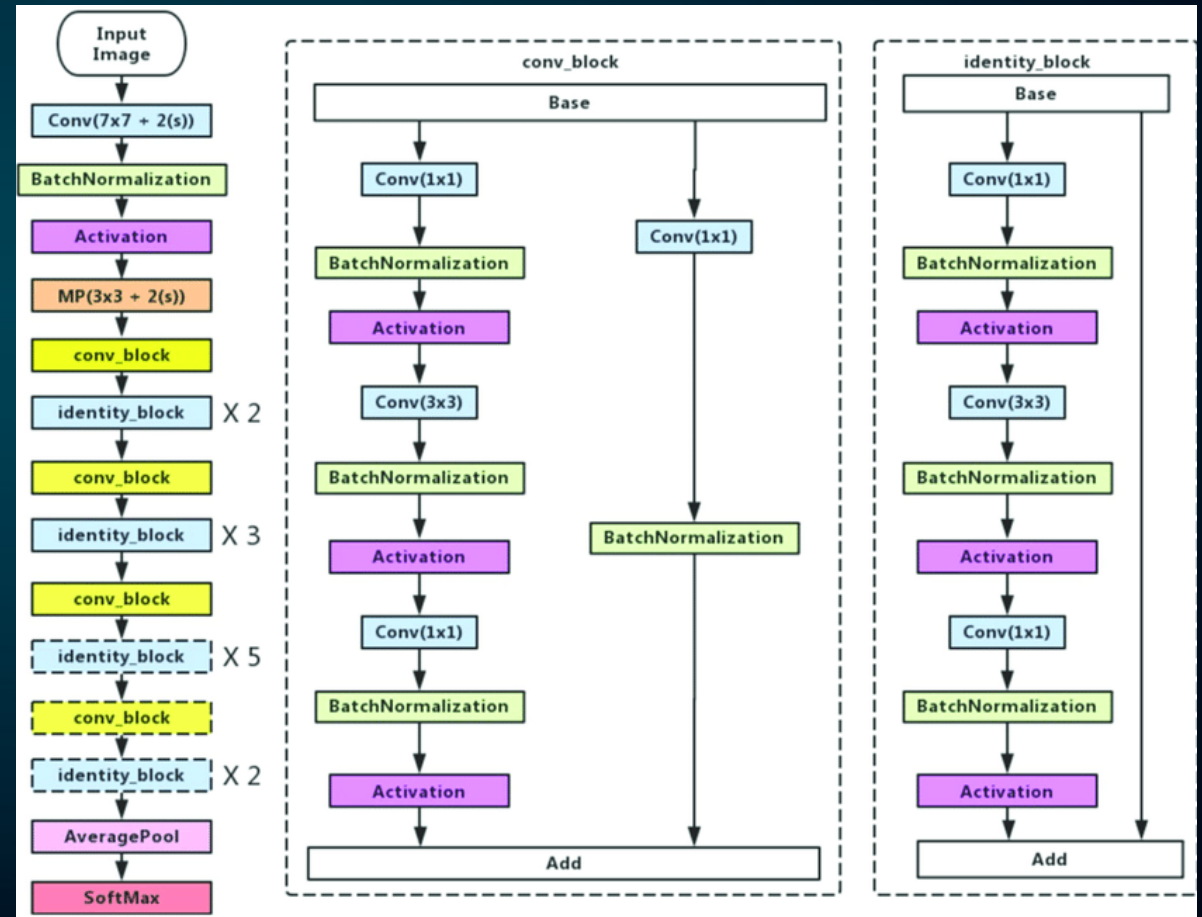
核心技術 — 遷移式學習



Why train when you can fine tune ?

模型架構選擇 ResNet50。

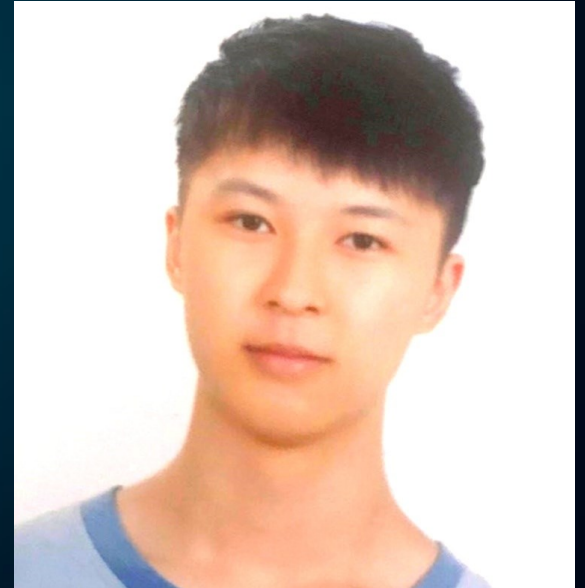
並選擇了一個基於人臉圖像預測年齡的優秀專案作為載入的 Pre-trained model 權重。





核心技術 — 照片前處理

透過 dlib 與 haar 找出人臉範圍後，裁切並做資料增強 (Data Augmentation)。在訓練的每一個 Epoch 對圖片做旋轉、翻轉、亮度、對比、顏色的調整。

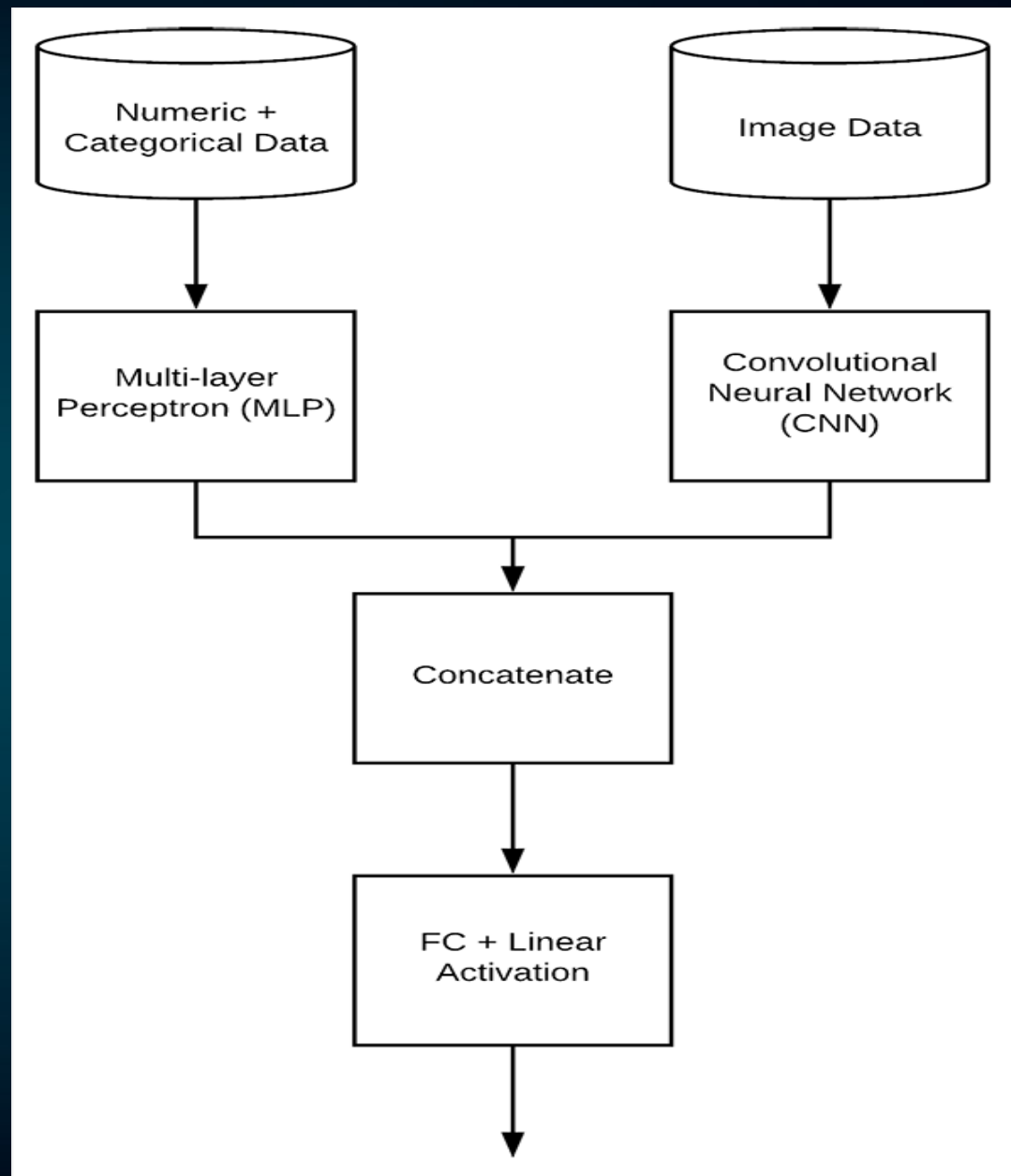


核心技術

架構設計

我們基於 Keras 設計了一個的模型架構，模型接受多種輸入，包括數字、分類和圖像數據。

並使用在 SVR 分析法中整理出的 18 種特徵作為模型的輸入，在新的 Layer 與圖像的輸出混合並一起訓練。





使用方式講解 | PART3

使用場景介紹



1

光線：在亮度充足但光線不會直射臉部的咖啡廳或室內

2

完整臉部：臉部至少占照片60 %以上，且不裁切到臉部

3

乾淨背景：背景避免過於鮮豔，且盡可能不超過2個人入鏡

4

正面照片：臉部水平擺動與垂直擺動都要在30度以內

使用流程 — 影片介紹



影片網址

<https://www.youtube.com/watch?v=IA4xpFDWqO8>



對話機器人網址

https://t.me/cathy_cathaylife_bot





FAS

Face Analysis Applications



應用場景 | PART4

配套措施



業務員層面

客戶

公司

● 推力

1. 親晤保戶的證明

以公司規定面來試辦，為確保業務員親見親簽，需要業務員拍照存證



配套措施



業務員層面



● 拉力

1. 加速核保效率

醫療險相關保單若有使用拍照上傳之業務員可以享有優先核保的福利

2. 競賽業績加成

讓有使用拍照上傳之業務員有在出國競賽或獎金排名時有業績上的加成 (5%-15%)

客戶

公司

配套措施



業務

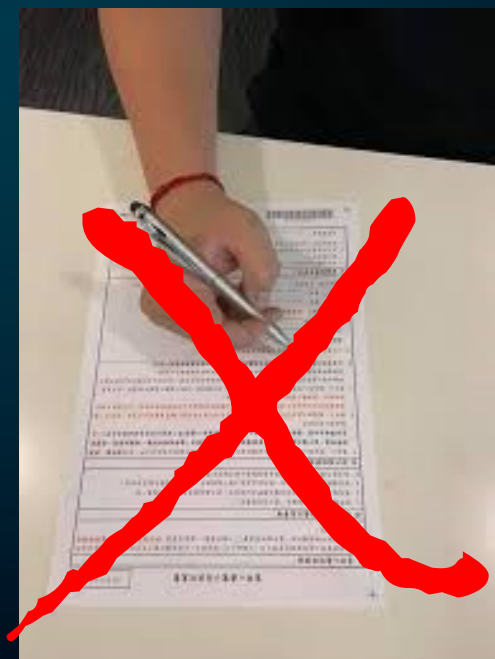
● 拉力

1. 可免簽電子投保同意書

使用拍照存證的方式即可享有免再簽電子投保同意書(克服法律問題)

客戶層面

公司



配套措施

業務

客戶層面

公司



2. 加速理賠效率的獎勵

如果選擇使用臉部辨識BMI的話，可以在之後發生理賠時獲得更好的理賠效率

3. 保費折扣

跟高保額折扣或信用卡繳費一樣，使用臉部辨識的客戶也可享有0.5%-1%的保費折扣



配套措施

業務

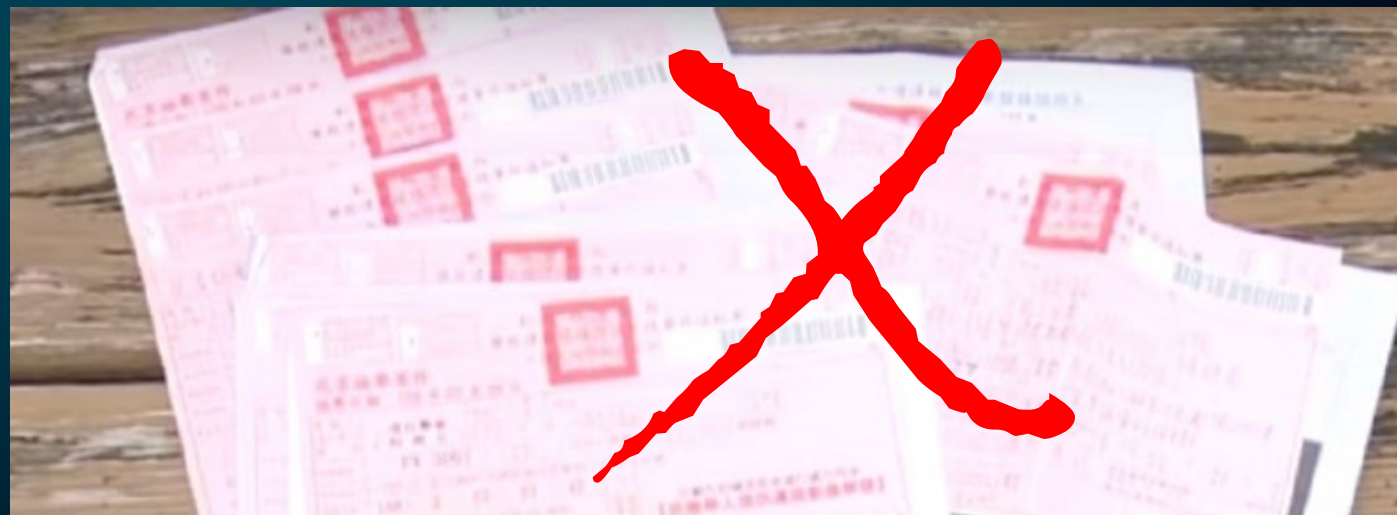
客戶

公司層面

● 推力

1. 降低金融爭議案件罰金

藉由拍照存證的方式減少業務員偽造簽名或保單未親見親簽的金融爭議裁罰頻率





業務

客戶

公司層面



● 拉力

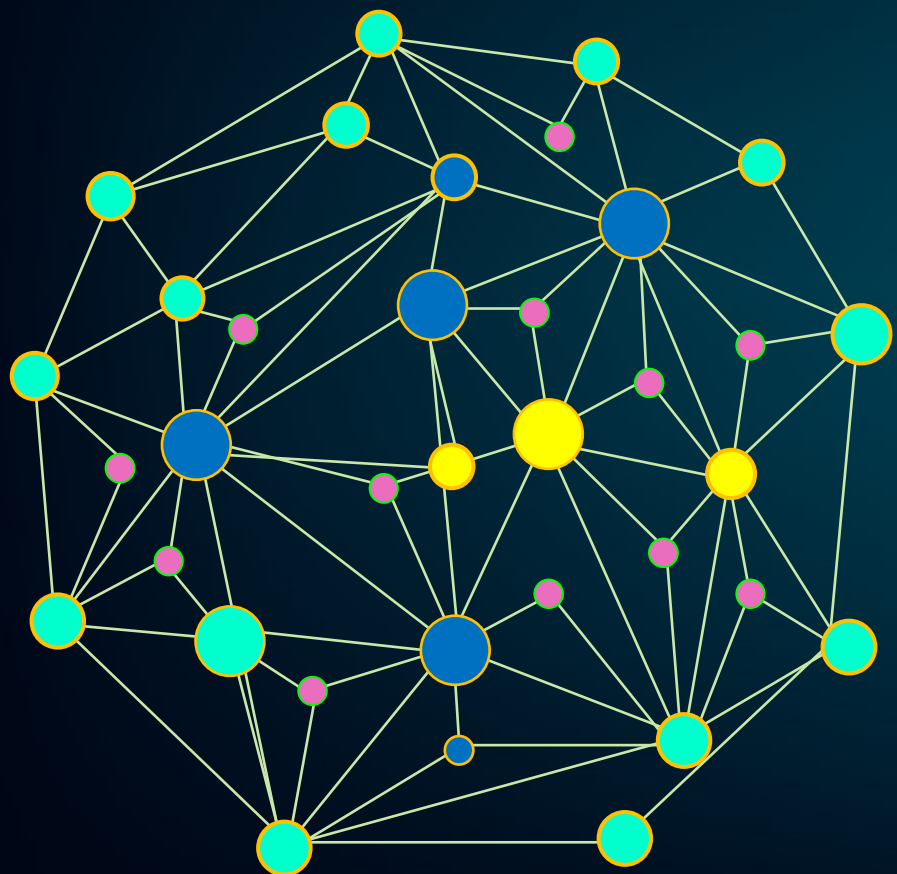
使公司可以用保費上的優惠吸引出險率相對較低的健康體，不只合乎公平費率，也促使業務員篩選客戶

未來可以延伸至客戶經營、進階辨識、區塊鏈 etc

配套措施



● 拉力 → 大量的後續延伸應用



如同臉書一樣有週年回顧，
加強客戶經營



透過大量資料，訓練出癌症、
失智、眼疾等核保需求



未來區塊鏈的運用，將臉部
資料與病例、理賠紀錄連結

績效評估



核保

- ✓ 可減少核保人力約10-20%，將人力著重於高風險案件
- ✓ 資料互相對照，增加資料準確度25%以上
- ✓ 減少確認被保險人本人的流程，加速核保

財務

收入

- ✓ 增加年輕族群投保率
- ✓ 獲得與醫院之間至少10%的資料回饋

支出

- ✓ 將節省至少50%的體檢成本
- ✓ 降低金管會裁罰約25%
- ✓ 可降低理賠率15%左右



潛力

- ✓ 藉由資料對照來進行業務員的評分
- ✓ 訓練模型進行更進階的其他警示性判斷
- ✓ 利用BMI及臉部驗證加強網路投保
- ✓ 將蒐集的資料利用大數據找出投保族群
- ✓ 資料上區塊鏈後增加跨領域使用的可能



謝謝大家 | THANK YOU