

## 實務上遇到的問題

# 大綱

---

- ▶ 企業如何導入人工智慧
- ▶ 導入過程中會遇到之問題
- ▶ 真實案例分享

# 企業如何導入AI

---

- ▶ 蒐集“質”夠好且“量”夠大之資料
  - ▶ 最有價值的一部分
- ▶ 與懂**AI**的人/機構合作
- ▶ 使用開源軟體降低進入門檻

# 蒐集”質”夠好且”量”夠大之資料

---



更多時間



更多成本



更多勞力



更多感應器

# 蒐集”質”夠好且”量”夠大之資料

---

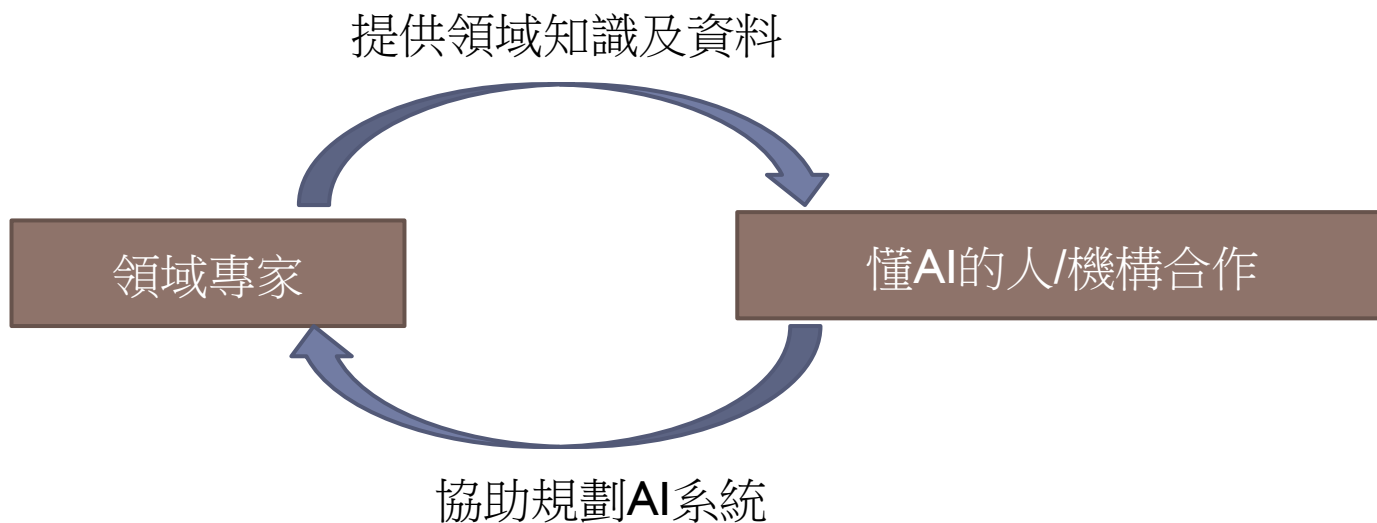
- ▶ 蒐集資料的難處
  - ▶ IT系統可能需要更動(資料庫)
  - ▶ 短期需要勞力去篩資料
  - ▶ 跨部門的溝通

•  
•  
•  
•  
•

# 與懂AI的人/機構合作

---

## ▶ 常見合作方式



# 使用開源軟體降低進入門檻

---

Caffe

Microsoft  
CNTK

  
TensorFlow

 PyTorch

 torch

dmlc  
**mxnet**

  
Chainer



# 使用開源軟體降低進入門檻

---

- ▶ 物件偵測

- ▶ YOLO專案

- ▶ <https://github.com/thtrieu/darkflow>

- ▶ 語音辨識

- ▶ DeepSpeech專案

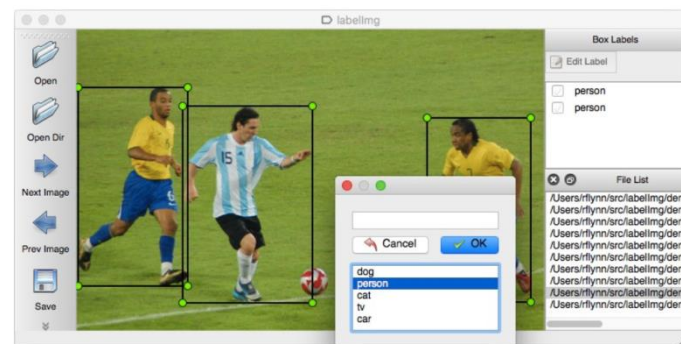
- ▶ <https://github.com/mozilla/DeepSpeech>

•  
•  
•  
•  
•



# 清理資料及標註資料

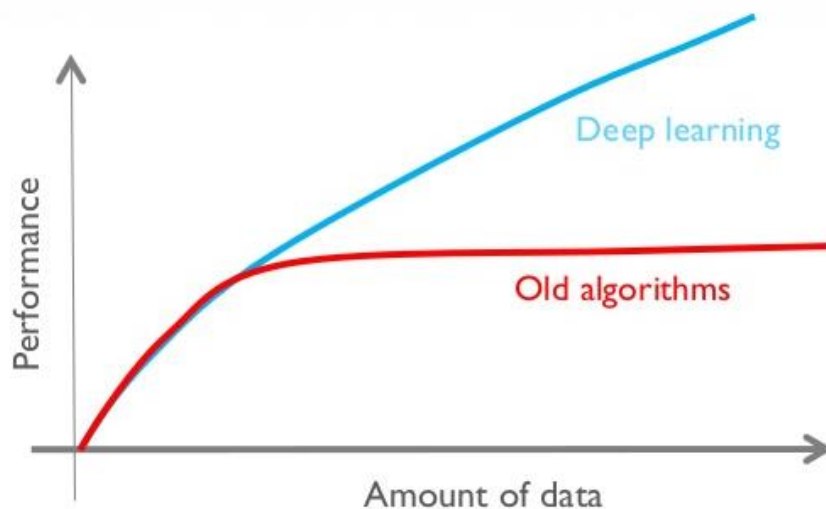
- ▶ 清理資料
  - ▶ 把不乾淨、不好之資料處理掉
- ▶ 標註資料
  - ▶ 使用標註工具
  - ▶ <https://github.com/tzutalin/labelImg>



# 資料的重要性

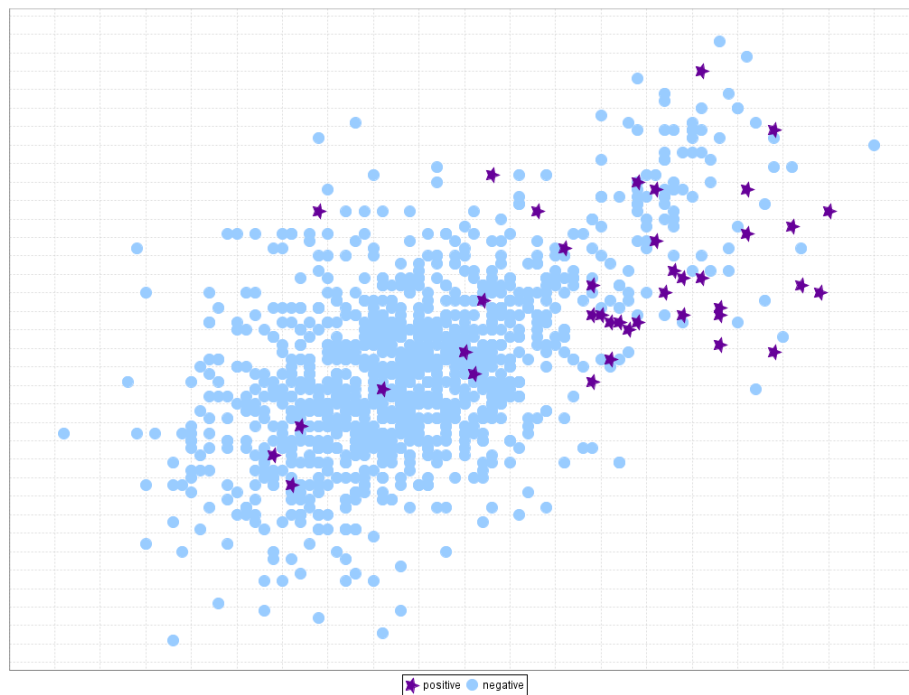
---

- ▶ 越多的資料代表越有競爭力
- ▶ 越罕見的資料代表越有價值



# 資料不平衡

---



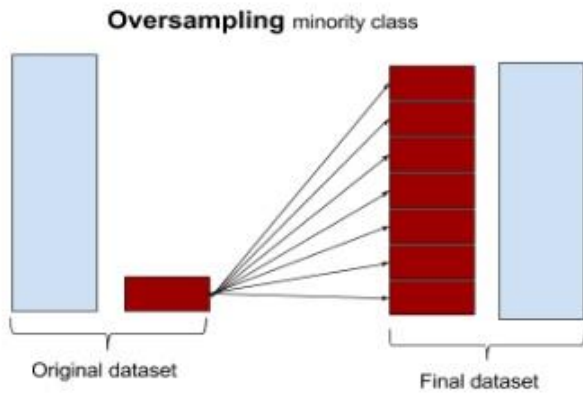
# 資料不平衡解決方法

---

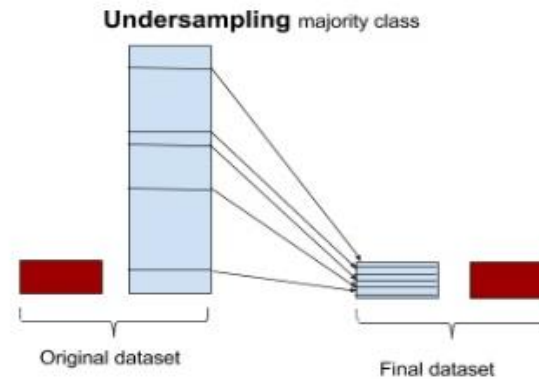
- ▶ 蒐集更多資料
- ▶ **Resample**
  - ▶ 上取樣
  - ▶ 下取樣
- ▶ **Generate Synthetic Samples**
- ▶ ....
- ▶ ....

# Resample

---



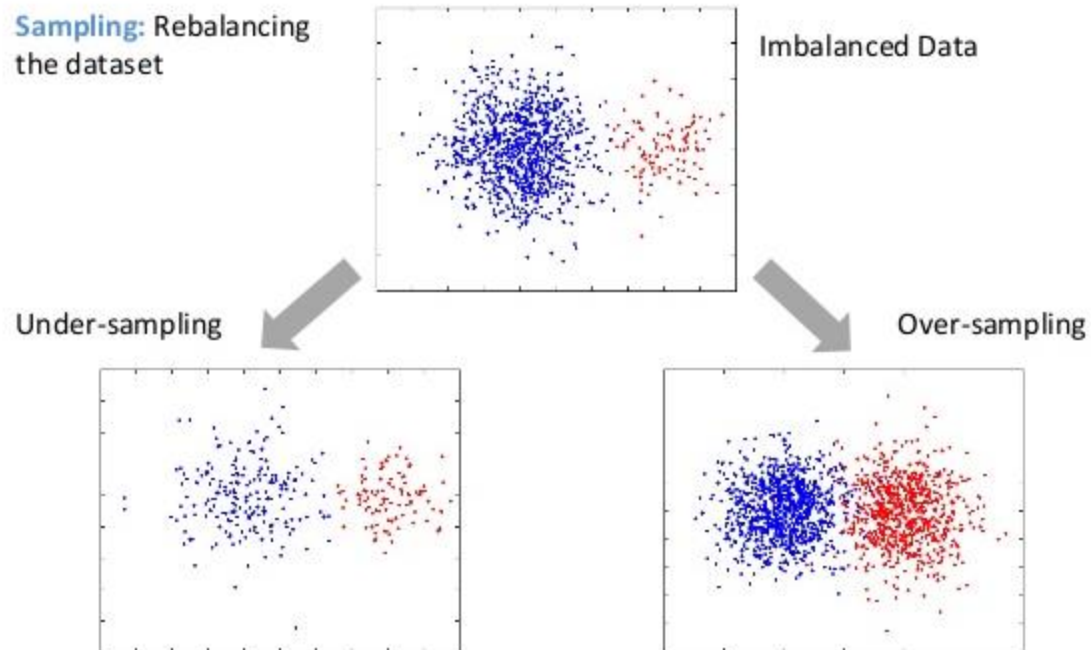
上取樣



下取樣

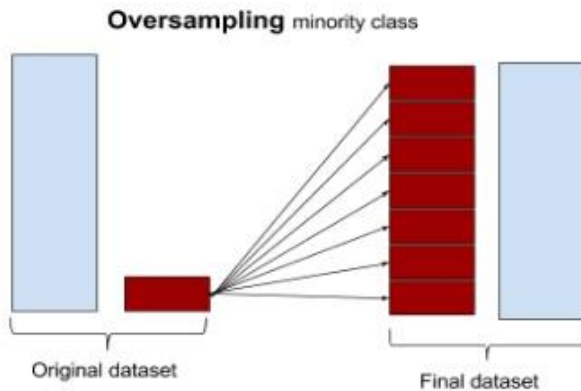
# Resample

---



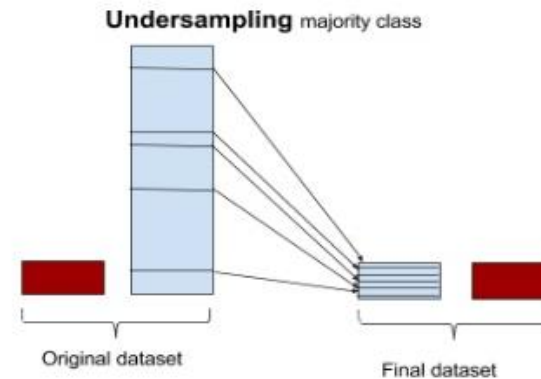
# Drawback on Resample

---



上取樣

容易過擬合  
(太多重複資料)

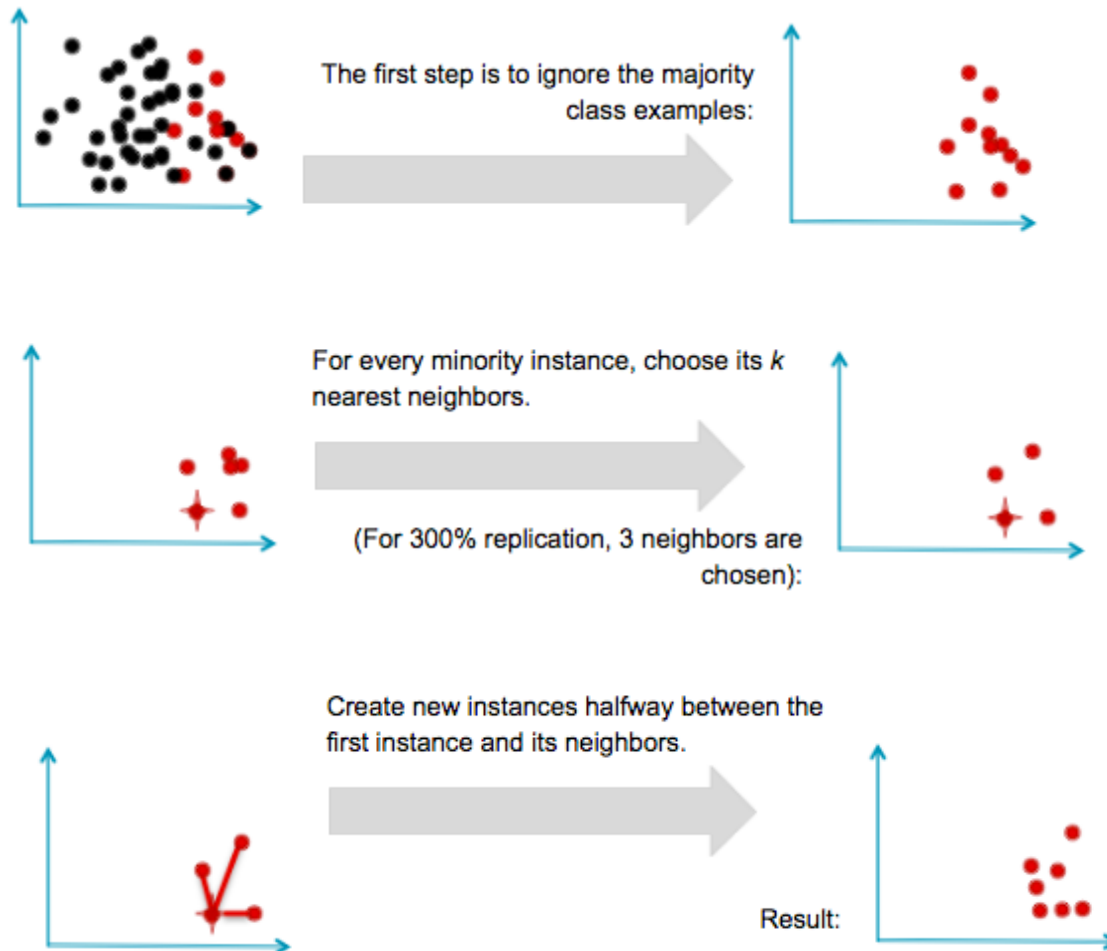


下取樣

浪費太多資料

# Generate Synthetic Samples

---





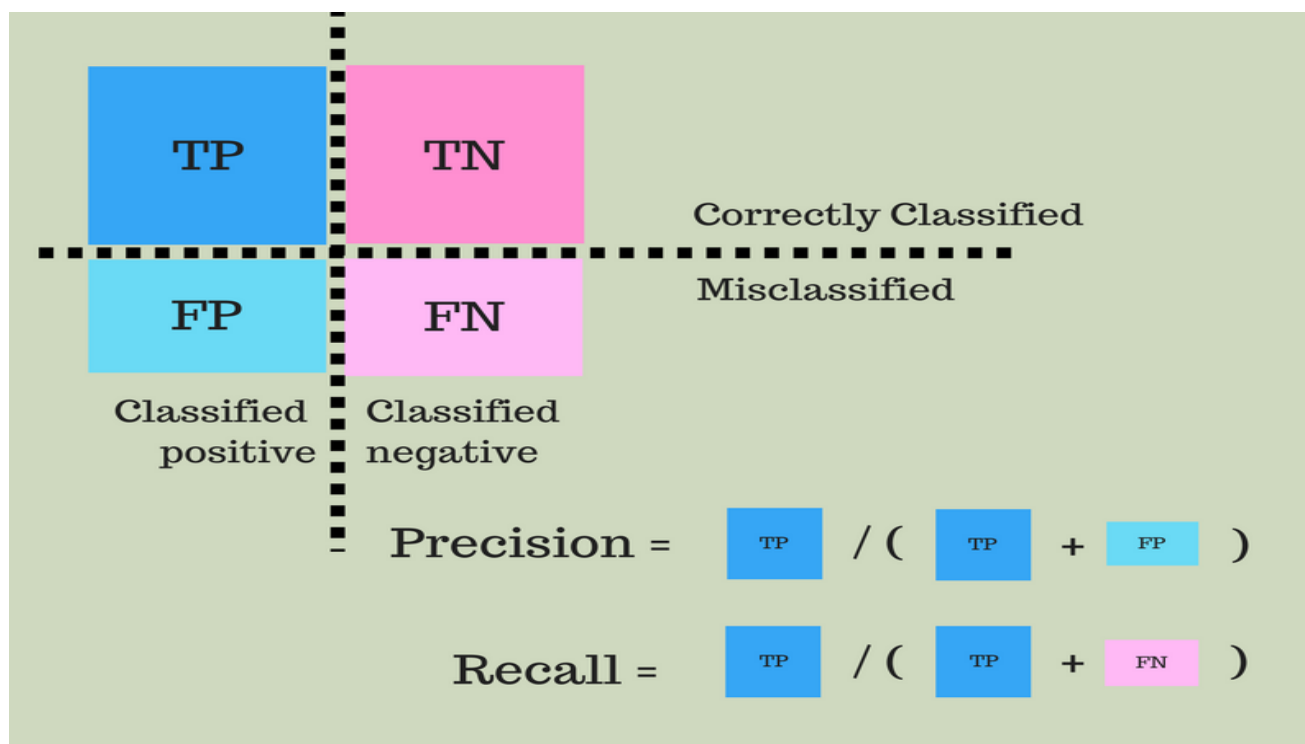
# 資料不平衡下的衡量指標

---

		actual value		
		<i>p</i>	<i>n</i>	total
prediction outcome	<i>p'</i>	True Positive	False Positive	P'
	<i>n'</i>	False Negative	True Negative	N'
total		P	N	

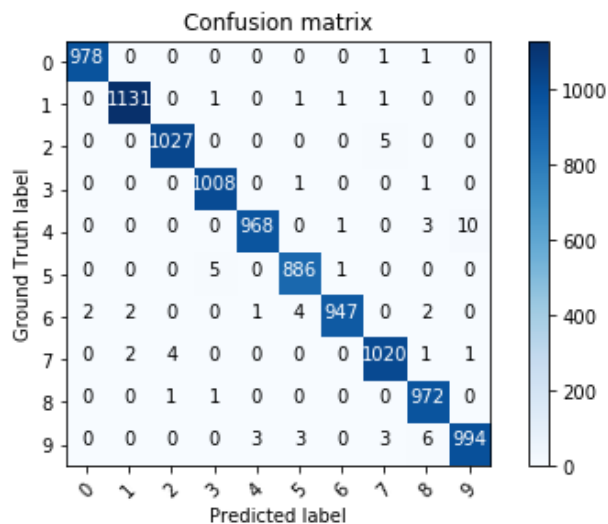
$$accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

# 資料不平衡下的衡量指標



# 資料不平衡下的衡量指標

- ▶ 混淆矩陣可以協助我們更清楚知道每種狀況誤判的情況
  - ▶ 非常推薦的衡量指標



# Kaggle平台介紹

---

- ▶ **Kaggle**是一個資料科學之平台
  - ▶ 平台上常會有與資料科學相關之比賽
  - ▶ 上面有各領域之數據集、專家解題之經驗
  - ▶ <https://www.kaggle.com/>
- ▶ **Google**於2017年買下Kaggle

The Kaggle logo, featuring the word "kaggle" in a lowercase, blue, sans-serif font.

# “天池”

---

- ▶ 大陸版本之Kaggle

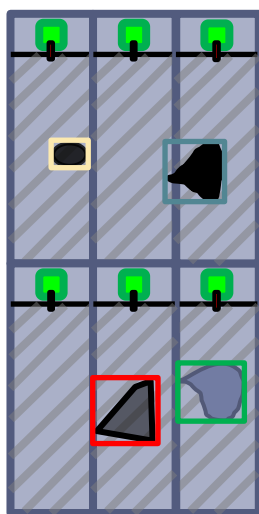
- ▶ <https://tianchi.aliyun.com/index.htm?spm=5176.100066.5610778.10.5198d780qaVmpq>

- ▶ 此平台由阿里巴巴所主導



# 案例分享一 面板製程缺陷辨識

- ▶ 面板在製造的過程中往往會產生不同種類缺陷/瑕疵
  - ▶ 不同種類的缺陷所產生的顏色、形狀會不太一樣
  - ▶ 面板廠希望能有一套自動辨識缺陷在哪裡以及缺陷種類為何的軟體



依據不同重要程度或不同種類之缺陷/瑕疵做出不同顏色之包圍框

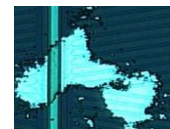
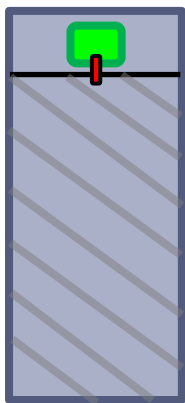
油汙

灰塵、微粒

玻璃碎片

塗料流出

# 案例分享一 面板製程缺陷辨識



成本低

以總數不到2000張的訓練樣本為例

高準確度

即可在一般PC上達到9X%的分類準確率

前置作業少且省時

且預測每張影像僅需4s內的運算時間

即時運行

# 案例分享一 面板製程缺陷辨識

---

- ▶ 傳統的解法
  - ▶ 用人工篩選的方式篩掉瑕疵品
  - ▶ 用傳統電腦影像方法做篩選 (準確度有待改善)





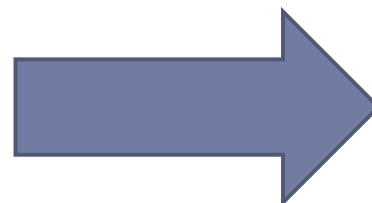
# 案例分享一 面板製程缺陷辨識

---



~70%  
(傳統電腦視覺方法)

導入AI 深度學習演算法



準確度大幅提升

~95%

# 案例分享一 面板製程缺陷辨識

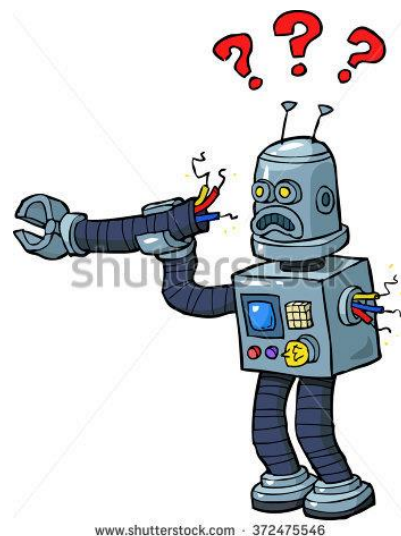
---

- ▶ 導入**AI**的好處
  - ▶ 大幅提高辨識的準確度(~70% → ~95%)
  - ▶ 同一套系統可用在面板廠其他影像辨識問題上面(不需要重新再開發)



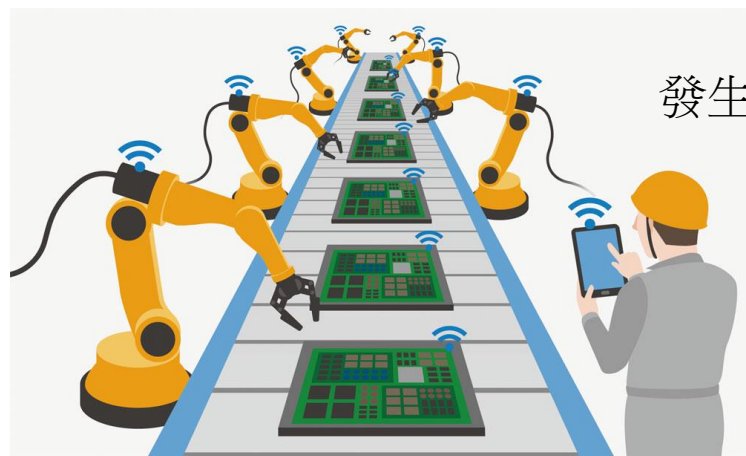
## 案例分享二 用AI來提升產線良率

- ▶ 良率往往是生產線當中一個重要的指標
  - ▶ 如何精準地控制產線上的設備以及監控製造過程是很重要的



## 案例分享二 用AI來提升產線良率

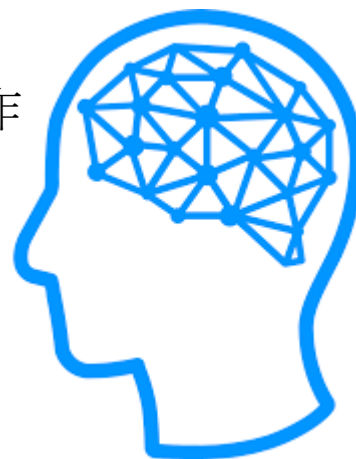
- ▶ 將工廠的設備加入感應器用來蒐集資料
  - ▶ 包括機器手臂訊號、產線影像等
- ▶ **AI**即時分析目前生產線狀況



發生異常馬上做出對應之動作  
並通知相關人員

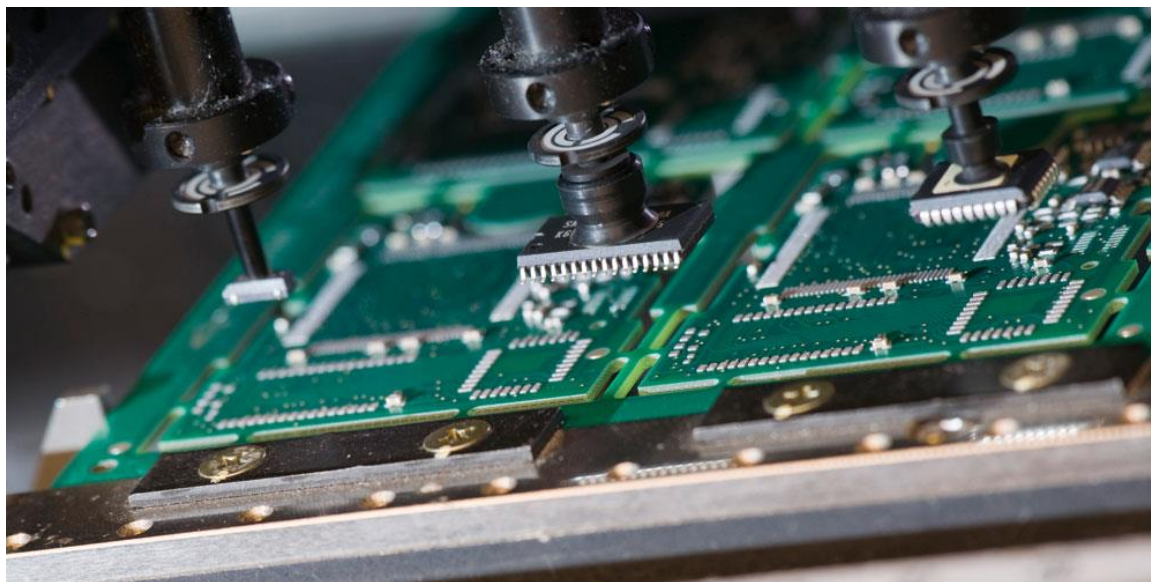


即時傳送各設備訊號



## 案例分享二 用AI來提升產線良率

---

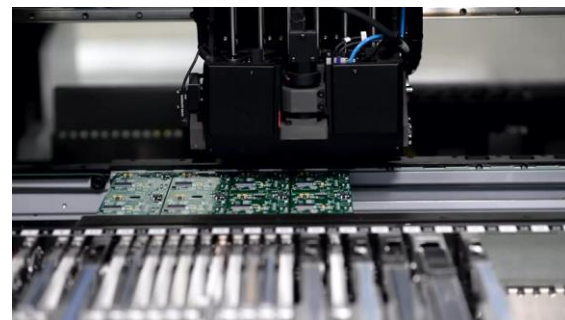
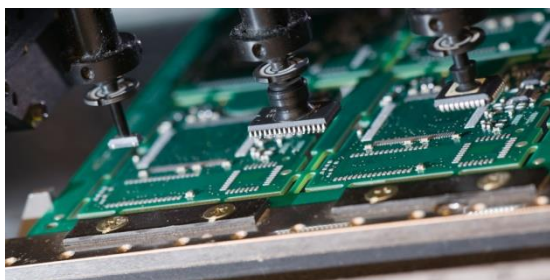


用AI來做影像定位，確保元件位置正確沒有太大誤差

## 案例分享二 用AI來提升產線良率

---

狀態隨時回饋給機台



用現有機台修正後的參數做更準確的控制

## 案例分享二 用AI來提升產線良率

---

- ▶ 導入AI的好處
  - ▶ 提升產品良率、減少人力檢查
  - ▶ 即時監控每個設備狀況
  - ▶ 提早預防生產線故障，減少損失



## 案例分享三 術後傷口症狀辨識

---

- ▶ 重大手術過後(如截肢)常常需要定期追蹤是否傷口發炎
  - ▶ 正常、紅腫、壞死、出血、化膿
- ▶ 傳統方法仰賴醫護人員定期觀看傷口影像
  - ▶ 非常花人力
  - ▶ 會有一定比率的人為出錯



# 案例分享三 術後傷口症狀辨識



傷口分割



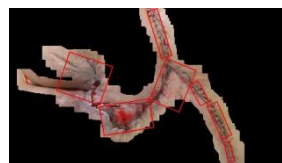
傷口縫合區定位



關注區域擷取



症狀評估



出血

## 案例分享三 術後傷口症狀辨識

---

- ▶ 導入AI的好處
  - ▶ 提升內部效率
  - ▶ 降低人為出錯率

