



## 3.2 Azure AI 視覺服務

Azure AI 視覺 API 服務提供進階的電腦視覺演算法，而這些演算法都是以機器學習模型為基礎，開發人員可以將影像檔案上傳或指定影像 URL，傳送給 Azure AI 視覺 API 進行影像分析，API 即可傳回影像分析的結果。

### Tips

- ① API：應用程式介面。
- ② 機器學習模型：一種已定型的服務，可辨識特定類型的模式(監督式學習、非監督式學習)。可以使用一組資料訓練模型，提供演算法，依演算法的規則以便模型用於推理這些資料，從中學習。

Azure AI 視覺可分析相機、影片或影像檔案，或是分析影像 URL 的視覺內容，如下為電腦視覺常見的功能：

### 一. 影像分類

「影像分類」是根據影像內容來分類影像，例如：在交通監視使用影像分類，根據車輛類型進行分類影像，將車輛類型分為計程車、公車、自行車、機車...等。



- ⓐ 交通工具影像分類識別出影像中的車輛為計程車 (圖片取自 Microsoft 技術文件網站)

### 二. 物件偵測

物件偵測可以預測影像中的物件位置與分類。物件偵測可取得物件的類型，還會偵測物件的頂端、左邊、寬度和高度的矩形框座標。在一個影像中可識別多種項目，例如：交通監視使用物件偵測，可以識別出影像中公車、計程車、單車騎士...等不同類別車輛的位置，並估計車與車之間的距離。



▲ 使用物件偵測取得公車、計程車、單車騎士在影像中的位置  
(圖片取自 Microsoft 技術文件網站)

### 三. 語意分割

語意分割屬於影像分割的一種，是進階的機器學習，做法是給一張影像，將影像中的所有「像素」點進行分類，把影像中的物件切割出來。例如：交通監視使用「遮罩」將交通中的影像進行切割，以不一樣的色彩區分不同的車輛。



▲ 使用語意分割將影像中的車輛加上遮罩  
(圖片取自 Microsoft 技術文件網站)

### 四. 影像描述 (說明影像)

Azure AI 視覺能夠分析影像、評估偵測到的物件，產生人類看得懂的片語或句子，來描述從影像中偵測到的內容。這些內容可以包含影像中「描述性標題」，或是描述性標題的「信賴分數」。

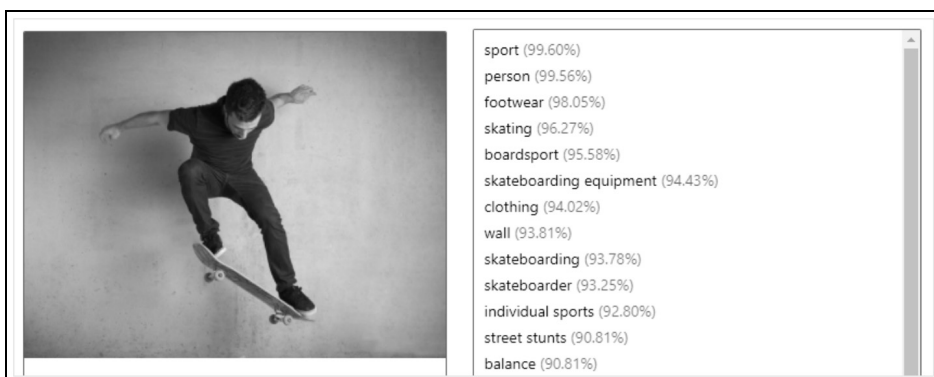


▲ 電腦視覺分析影像描述出在街道上遛狗的行人(圖片取自 Microsoft 技術文件網站)



## 五. 標記視覺特徵

Azure AI 視覺會根據可辨識的物件來產生影像描述，並傳回多個標籤，信度 (信賴分數) 最高的標籤會最先列出。而這些標籤除了包含影像主體，也會包含影像中的環境，如室內、室外、動物、工具或家具 ... 等。這些標籤可與影像建立關聯，作為摘要影像屬性的中繼資料；可用來搜尋具有特定屬性或內容的影像



⬆ 電腦視覺依信度高低順序傳出多個標籤。

**Tips**

標籤：描述物件的標題文字，或稱為標記。

## 六. 偵測品牌

偵測品牌功能提供識別商業品牌的能力。此服務已有一個資料庫，其中包含來自全球數千個可辨識產品標誌的商業品牌。例如：電腦視覺服務偵測有 Microsoft 品牌標誌的筆記型電腦。

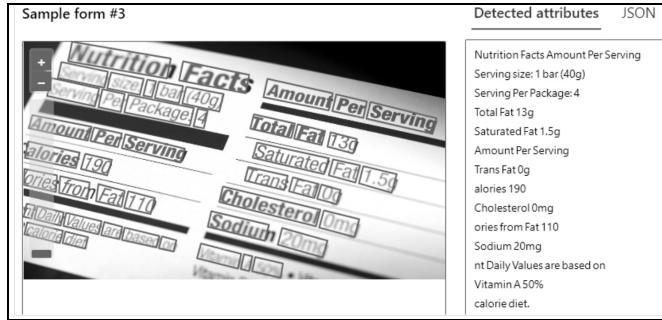


⬆ 偵測筆電有 Microsoft 品牌標誌  
(圖片取自 Microsoft 技術文件網站)



## 七. 光學字元辨識

Azure AI 視覺服務可使用光學字元辨識 (OCR) 功能，來偵測影像中的列印和手寫文字 (如：道路標誌或店面招牌)。



- ⊙ 光學字元辨識可偵測影像中的文字資料  
(圖片取自 Microsoft 技術文件網站)

## 八. 臉部偵測

臉部偵測可找出影像中的人臉，是一種特殊形式的物件偵測。能夠判斷年齡、性別以及代表臉部位置的矩形方塊，還能進行臉部識別 (判斷影像中的兩個臉部是否為同一人) 以及臉部表情偵測。(隱私權因素，年齡、性別以及表情偵測預設無法使用)



- ⊙ 透過臉部服務識別影像中人員的年齡和性別  
(圖片取自 Microsoft 技術文件網站)



## 九. 偵測特定領域內容

影像分類時，Azure AI 視覺服務支援兩個特製化領域模型：

1. **名人**：包含已定型的模型，用來識別來自政治、運動、娛樂和商業界數以千計的知名人物。(隱私權因素，預設無法使用)
2. **地標**：可識別知名地標，例如：艾菲爾鐵塔、台北 101 大樓 ... 等。



⊕ 電腦視覺服務可識別地標 Taipei 101，並具有 98.59% 的信度。

## 十. 其他進階功能

電腦視覺服務另提供如下進階分析功能：

1. **偵測影像類型**：影像內容類型分為「美工圖案」與「線條繪圖」。
2. **偵測影像色彩配置**：影像中的色彩有三個屬性，分別為主要前景色彩、主要背景色彩、整體主要色彩集合。
3. **產生縮圖**：所謂縮圖就是壓縮圖片，將容量太大的圖片建立成小型的影像版本。
4. **內容仲裁**：可以掃描文字、影像、影片內容，偵測出是否含有冒犯意味、有風險或不當的資料。例如：具冒犯性或不恰當的文字、管制影片中的血腥暴力場景或限制級影像、種族歧視內容。也可自訂字詞或影像仲裁清單。



## 3.3 Azure AI 服務開發環境與必要條件

Azure AI 服務(Azure AI Services)是使用 REST API 和用戶端程式庫 SDK 的雲端式服務，可協助開發人員在應用程式中建立認知智慧。微軟 Azure AI 服務提供 Azure AI 視覺、Azure AI 語音、Azure AI 語言與 Azure OpenAI 等多種 AI 服務。開發人員無需具備 AI 人工智慧專業的技術，只要能將要分析的資料 (含圖檔、文字)，傳送到 Azure AI 服務的演算法模型進行運算處理，即可傳回用戶端所需要的資訊，如此可協助開發人員進行建立具 AI 智慧功能的應用程式。

### Tips

- ① REST：是一種軟體架構風格，目的是幫助在世界各地不同軟體，其程式在網際網路中能夠互相傳遞訊息。
- ② REST API：是一種 Web API 的設計規範。
- ③ SDK：軟體開發工具套件。

在使用 Azure AI 服務前，您需要先註冊 Azure 帳戶並建立服務資源。建立完成後，系統將提供您「金鑰 (Key)」和「端點 (EndPoint)」這兩項重要資訊，這兩項資訊是開發 AI 應用程式時必須使用的，因為它們用於身份驗證並連接到您的 Azure 資源。

### Tips

- ① 金鑰 (Key)：驗證用戶端應用程式的管理帳戶。
- ② 端點 (EndPoint)：提供存取資源的 HTTPS 位址。

設計 Azure AI 服務應用程式的開發語言，可使用 Java、Python、C#... 等，本書使用 C# 配合 Visual Studio 環境開發 Azure AI 服務應用程式。

進行 Azure AI 視覺分析的影像必須符合下列需求：

- ❖ 必須是 JPEG、PNG、GIF 或 BMP 格式的影像
- ❖ 檔案大小必須小於 4 MB
- ❖ 像素必須大於 50 × 50 像素



## 3.4.4 影像分析範例實作

### 範例：cv02.sln

練習製作影像分析程式。程式執行時按下 **開檔** 鈕開啟開檔對話方塊，並指定要分析影像的圖檔。接著會將影像的影像描述與信度、人數、品牌、成人資訊以及標籤和標籤信度，顯示於多行文字方塊中。

#### 執行結果

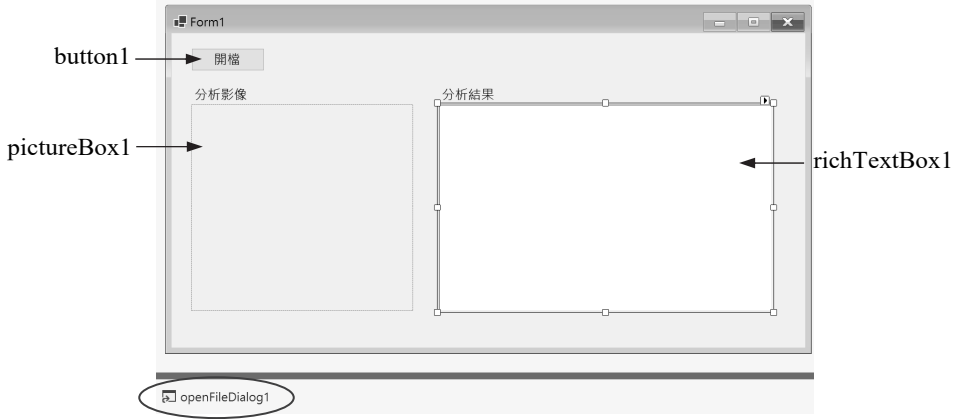


#### 操作步驟

**Step 01** 進入 Visual Studio 開發環境並執行功能表的【檔案(F) / 新增(N) / 專案(P)...】指令建立「Windows Forms 應用程式」，使用 C#程式語言開發，專案名稱為 cv02。

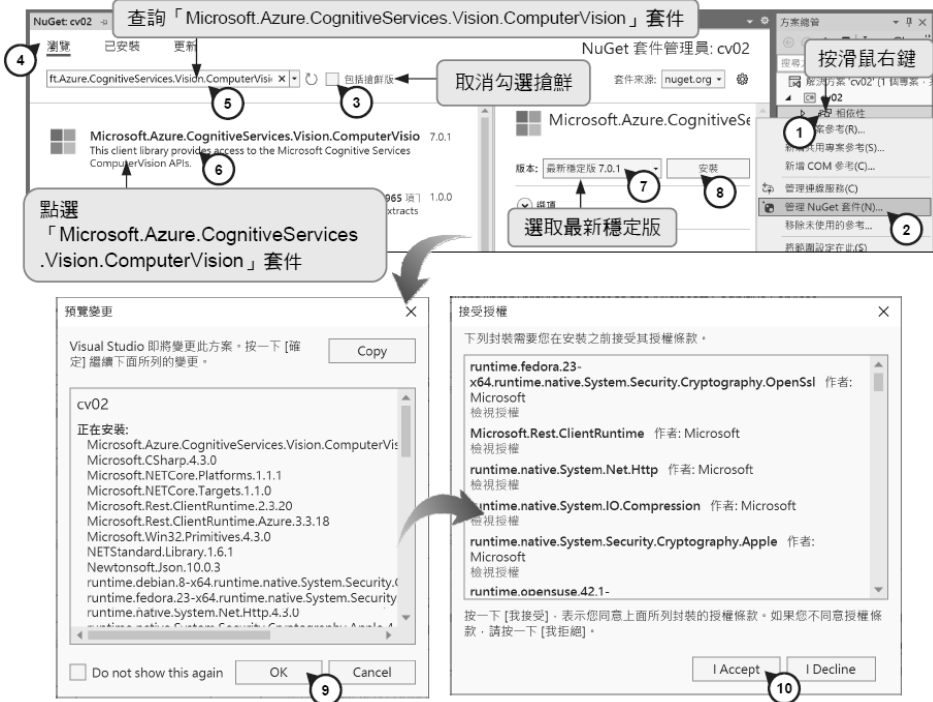


Step 02 建立如下表單控制項輸入介面：



Step 03 安裝 Computer Vision 電腦視覺套件：

在方案總管視窗的「相依性」按滑鼠右鍵執行【管理 NuGet 套件 (N)】，接著依圖示操作安裝「Microsoft.Azure.CognitiveServices.Vision.ComputerVision」套件。







## Step 04 撰寫程式碼：

## 程式碼 FileName:Form1.cs

```
01 using Microsoft.Azure.CognitiveServices.Vision.ComputerVision;
02 using Microsoft.Azure.CognitiveServices.Vision.ComputerVision.Models;
03
04 namespace cv02
05 {
06     public partial class Form1 : Form
07     {
08         public Form1()
09         {
10             InitializeComponent();
11         }
12
13         // 表單載入時執行
14         private void Form1_Load(object sender, EventArgs e)
15         {
16             // pictureBox1 的邊框樣式設定為固定單線
17             pictureBox1.BorderStyle = BorderStyle.FixedSingle;
18             // 使圖片會依據 pictureBox1 的大小進行拉伸或縮小，填滿整個 PictureBox
19             pictureBox1.SizeMode = PictureBoxSizeMode.StretchImage;
20         }
21
22         // 按下 [開檔] 鈕執行
23         private async void button1_Click(object sender, EventArgs e)
24         {
25             if (openFileDialog1.ShowDialog() == DialogResult.OK)
26             {
27                 try
28                 {
29                     string cvApiUrl = "電腦視覺服務端點";
30                     string cvApiKey = "電腦視覺服務金鑰";
31                     string imagePath = openFileDialog1.FileName;
32                     //建立 FileStream 物件 fs 開啟圖檔
33                     FileStream fs = File.Open(imagePath, FileMode.Open);
34
35                     //建立電腦視覺物件，同時指定電腦視覺的雲端服務金鑰
```



```
36         ComputerVisionClient visionClient =
37             new ComputerVisionClient(
38                 new ApiKeyServiceClientCredentials(cvApiKey),
39                 new System.Net.Http.DelegatingHandler[] { });
40
41         //電腦視覺物件指定雲端服務端點
42         visionClient.Endpoint = cvApiUrl;
43
44         // 指定要分析的列舉項目(視覺特徵)，並將分析的列舉存入 visualFeatures 陣列
45         VisualFeatureTypes?[] visualFeatures = new VisualFeatureTypes?[]
46         {
47             VisualFeatureTypes.ImageType, //影像類型
48             VisualFeatureTypes.Color, //顏色資訊
49             VisualFeatureTypes.Faces, //臉部資訊
50             VisualFeatureTypes.Adult, //成人資訊
51             VisualFeatureTypes.Categories, //影像分類
52             VisualFeatureTypes.Tags, //影像中的項目
53             VisualFeatureTypes.Objects, //影像中的物件
54             VisualFeatureTypes.Brands, //影像中的品牌
55             VisualFeatureTypes.Description //影像描述
56         };
57
```

## Q 說明

1. 第 23,60 行：ComputerVisionClient 物件的 AnalyzeImageInStreamAsync() 為非同步方法，故呼叫時必須加上 await 關鍵字，使用的事件處理函式也要定義為 async。
2. 第 29,30 行：請填入自行申請的電腦視覺服務的金鑰與端點。(可參考 3.4.2 節 cv01 範例的 Step01 步驟)
3. 第 45~56 行：建立分析影像的視覺特徵物件 visualFeatures。臉部資訊可取得臉部方框，至於年齡和性別因隱私權因素預設無法使用。
4. 第 60~61 行：使用 AnalyzeImageInStreamAsync()方法指定要分析影像與視覺特徵，並傳回影像分析結果 ImageAnalysis 類別物件 res。

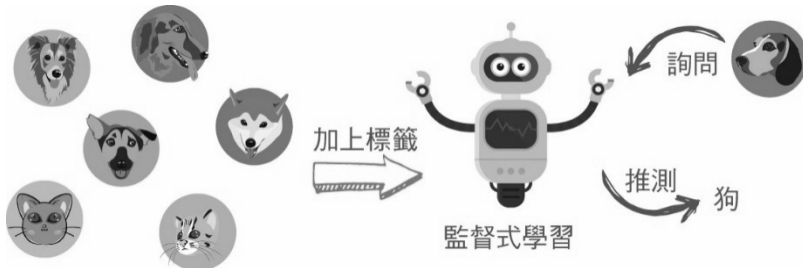


## 10.3 機器學習的模型

機器學習模型是一種電腦演算法，可以使用資料來進行評估或決策。簡單來說可以將模型視為接受輸入資料然後產生輸出的函式。機器學習模型與傳統演算法的設計方式不同，傳統演算法需要改善時要使用人力進行編輯；而機器學習會運用資料讓指定工作得到更好的效能。例如使用機器學習的股價預測方案，會隨著股市資料的增加，機器學習模型可以累積更多經驗而提升預測能力。如果是採傳統演算法，則必須由工程師修改股價預測公式。模型是機器學習服務的核心元件，常用的機器學習模型大致分成「監督式學習」、「非監督式學習」、「半監督式學習」和「增強學習」四種類別。

### 10.3.1 監督式學習

監督式學習 (supervised learning) 是給含有標籤 (label) 的許多資料，也就是附有答案的資料，透過機器學習模型計算來找出最佳解答。提供給機器學習的資料集欄位值稱為特徵值，而作為答案 (標籤) 的欄位值稱為目標值，為資料加上標籤的動作就稱為標記 (或稱預定義)。例如輸入 1000 張附有「貓」或「狗」標籤的照片，作為模型的定型資料後，再輸入新的沒有標籤照片，檢驗模型是否能正確識別出是貓還是狗。

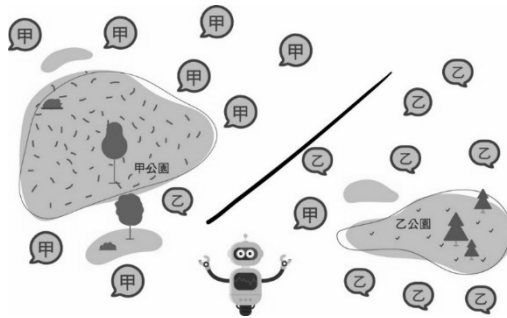


⊕ 監督式學習過程示意圖



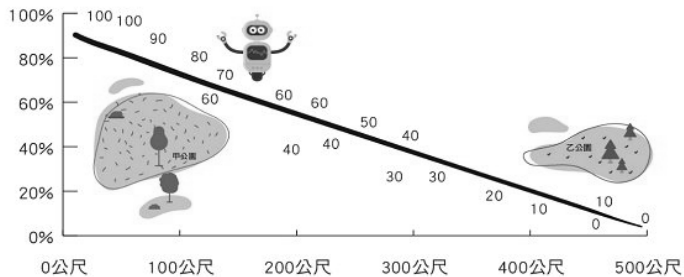
監督式學習主要分成分類 (classification) 和迴歸 (regression) 兩大類。分類的目標是找到資料所屬的種類，將資料分派到不同的群組，會忽略同一群組內資料間的些微差異。迴歸的目標則是找到資料的趨勢，將所有資料視為一個群組，分析資料的差異而找到整體的傾向。

例如某個小鎮有甲、乙兩個公園，抽樣調查各家戶常去的公園，並在地圖上標示出來。此時可以在地圖畫出一條直線，來大致區分出兩個群組，這就是分類的方法。利用這條直線就可以預測出，其它未調查的家戶可能常去的公園。



▲ 以分類方法預測家戶常去的公園

上例如果改用迴歸方式來處理，在調查時改詢問到甲公園的機率 (100 ~ 0%)，並在地圖上依距離和機率標示出來。此時也可以在地圖畫出一條直線，來表示家戶和甲公園距離與到公園機率的趨勢，這就是迴歸的方法。利用這條直線就可以根據距離預測出，其它未調查的家戶到甲公園的機率。

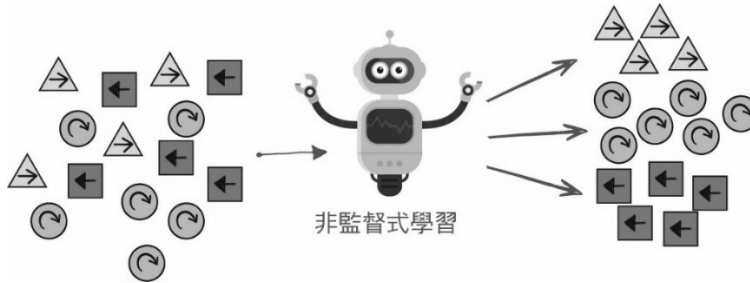


▲ 以迴歸方法預測家戶到甲公園的機率



### 10.3.2 非監督式學習

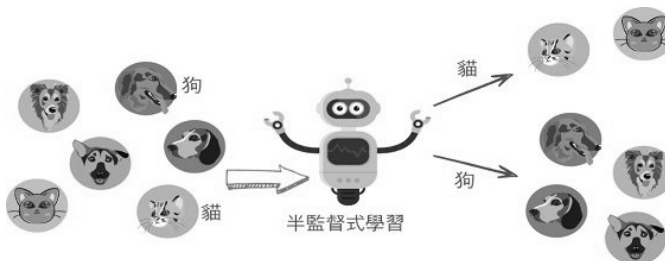
非監督式學習 (unsupervised learning) 的定型資料沒有標籤，讓機器學習模型自行摸索出資料規律，擷取出資料的特徵，達成解答未知資料的目標。非監督式學習最常見的方法就是叢集分析 (cluster analysis)，會根據特徵透過演算法計算資料間的相似程度，將資料樣本區分群組。



▲ 非監督式學習過程示意圖

### 10.3.3 半監督式學習

半監督式學習 (semi-supervised learning) 是介於監督式與非監督式之間的機器學習。因為現實中收集資料較簡單，但是完整標記的資料較少且耗費人力，所以開發出半監督式學習的技術。透過少量有標籤的資料找出特徵，就可以對其它大量無標籤的資料進行分類。這種方法可以減少標記資料的時間，又能讓預測結果比較精準。例如有 100 張照片，其中只有 10 張有加註「貓」或「狗」的標籤，半監督式機器學習模型可利用這 10 張照片的特徵去辨識及分類剩餘的照片。



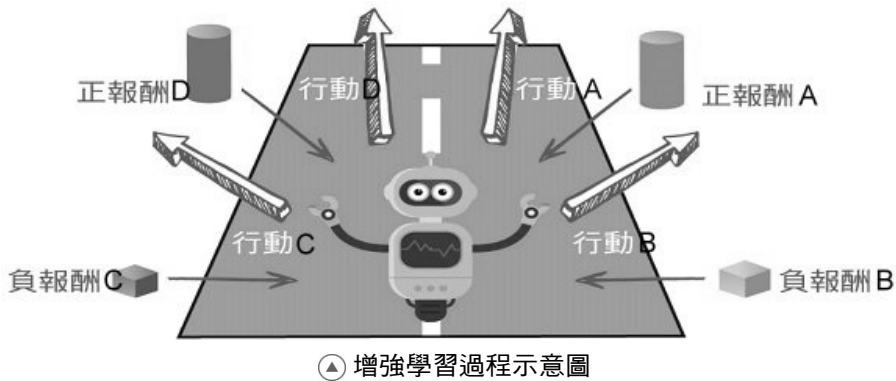
▲ 半監督式學習過程示意圖



## 10.3.4 增強學習

增強學習 (reinforcement learning) 是在指定環境中互動，模型會觀察環境而採取行動，並隨時根據回饋資料逐步修正，反覆嘗試錯誤來獲得最大利益。在過程中模型會進行一系列動作，隨著每個動作環境也會跟著變化。若環境的變化是接近目標就給予正報酬 (positive reward)；若遠離目標則給予負報酬 (negative reward)，增強學習模型的目標就是獲得最高的報酬。

例如模型開車時若保持在車道中就給正報酬，偏離跑道就給負報酬。雖然沒有給予標籤資料但根據報酬的多寡，增強學習模型會自行逐步修正最後得到正確的結果。如果問題是需要不斷做決策的情形，答案不能一次就解決時，增強學習模型就很適合，例如下棋時需要根據棋局變化不斷改變策略、或是開車隨時會遇到不同的路況。



## 10.4 分類模型

### 10.4.1 分類模型簡介

分類 (classification) 機器學習模型簡稱為分類模型，是屬於監督式學習。分類的目標是找到資料所屬的種類，將資料分派到不同的群組，會忽略同