**聲控飲料系統**

指導老師: 曾士桓 教授

參賽組員(姓名):黃柏皓 尤譽蒼 謝嘉侑

**摘要**

本專題欲做出一套辨識語音後利用機器產出飲料後送至取餐口的系統。在速食店中飲品通常都是由店員所裝取後再送至取餐口給予消費者取用。

在繁忙時刻，站在飲料機前等待飲料裝滿在走至櫃檯給予消費者享用，這之中所花費的時間及人力，若能利用機器代替完成，不僅更加有效率，且可以減少多餘的人力成本。

欲做出的聲控飲料系統包含了以下四個部分，第一部分需使系統紀錄使用者聲紋，以便在吵雜環境中仍可辨識。第二部分為機器利用分杯器將飲料杯平穩地放置履帶上，使杯子能夠在履帶上穩定的行進。第三部分為系統控制飲料機在正確的時機倒出飲料至杯中。第四部分，杯裝飲料經由履帶送出取餐口供消費者飲用，且可由螢幕確認是否取餐完成。

整套系統可簡易的完成製做飲品的流程，店員只需說出飲品名稱即可完成剩餘的所有流程，在忙碌的廚房中可減少相對的負擔，十分方便且快捷。

1. **前言**
   1. **研究動機**

飲食是人們每天必不可少的一環，雖然街頭巷尾也處處充滿了餐廳，但一到用餐時間，各式各樣的餐廳也因此忙碌了起來，常常可以看到每間餐廳都排滿了客人，趟若能更快速的將餐點完成收益不僅提升許多也能滿足更多客人的味蕾。

在各式各樣的餐廳中，強調速度快的素食店中，也常常因飲料的原因導致餐點耽擱的問題，由知名連鎖店X德基舉例，裝餐員需先倒入飲品再封膜後還需拿取消費者所需餐點裝入袋中再拿給消費者，油炸食品在30分中前都會提前炸至充足份量等待消費者購買，而飲品則為現點現做，為此浪費了許多寶貴的時間。

* 1. **研究目的**

在現今社會中，人們對時間的把握越來越注

重，在午間時刻更是如此，有許多人因等候時間過久，選擇了其他的店家；有許多人因等候時間過久，選擇不吃中餐亦或延後在吃，因此出餐速度在尖峰時刻對於餐廳來說更為重要了許多。

因此，此專題的目的就是利用機器取代人力同時機器需可選擇各式飲品，來滿足消費者的各種慾望，為了方便店員使用更是添加了語音系統，店員只需經由麥克風說出飲料名稱，而後機器將會將飲品做出並送至指定地點，不僅減少了人力更能快速的將所需餐點做出，效率相較以往也能因此大大提升。



　 圖1.1 系統簡易工作圖

1. **語音辨識**

**2.1降噪**

餐廳是人們用餐的環境，交談及廚房的整備等等都會發出許多的聲音，因此為了避免收到其他類似的聲音導致系統偵測錯誤，必須要將環境中的雜音給除去。

而通過系統產生與外界噪音相等的反向聲波，將噪音中和，從而實現降噪的效果。

由於耳機內之麥克風，主動偵測人耳可聽到之環境低頻噪音，然後將雜訊信號傳至控制電路，讓控制電路進行即時的運算，通過喇叭發射與噪音相位相反、振幅相同的聲波就能夠抵銷噪音，使得系統較不易發生錯誤。

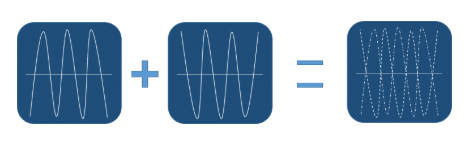
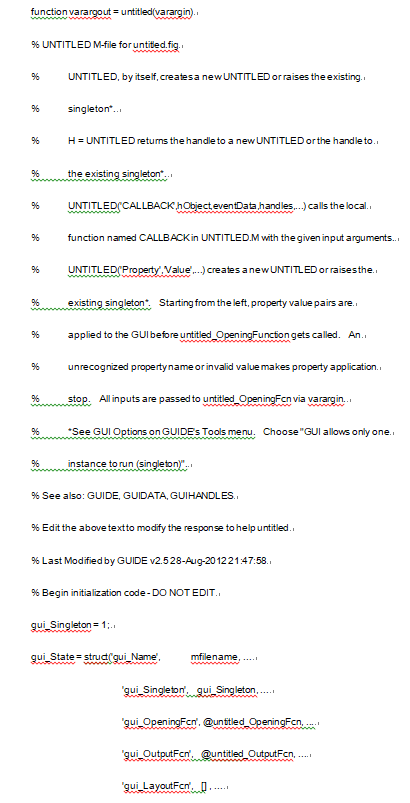
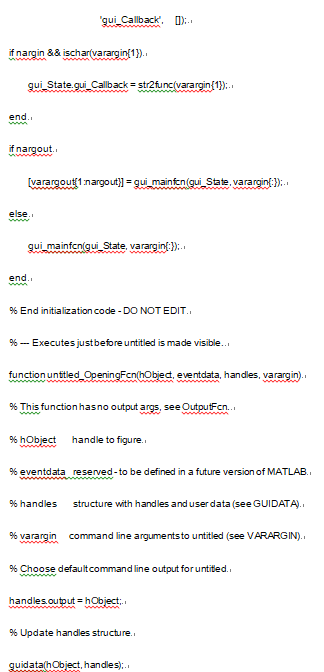
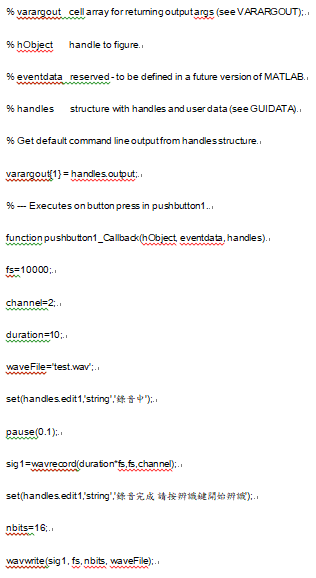


圖2.1產生之反向聲波

**2.2程式碼(語音辨識)**







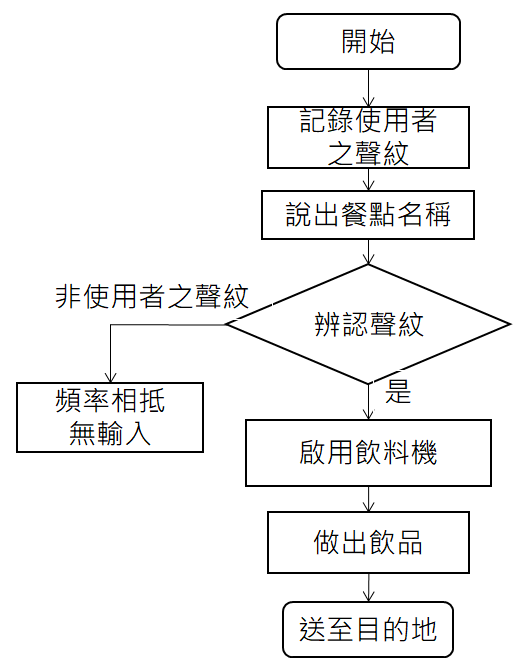
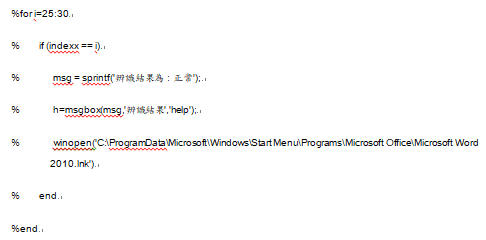


圖2-2系統流程圖

**2.3資料庫查詢**

填裝飲料機的時間也是十分寶貴的，感測器可得知當前飲品CC數，且回傳至螢幕，使得店員能通過螢幕得知飲品剩餘CC數，方便其及時補充飲品，且並不耽誤系統運作，可同時裝填飲品，同時運作系統。



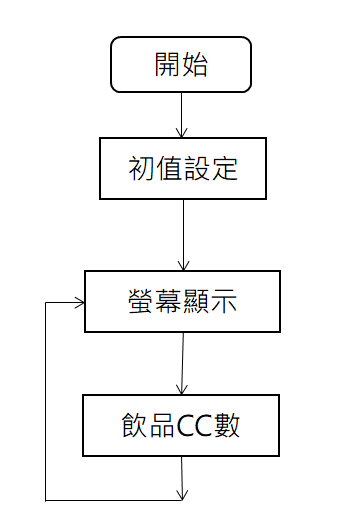
****

圖2-3 資料庫流程圖

1. **語言模型**

語言模型用來表示某一種語言的機率模型，目的是描述給定詞序列的機率分布，也就是計算一個句子/段落/文檔的機率，或者說，就是判斷一句話符合自然語言的表達習慣的可能性。

在[語音辨識](https://zh.wikipedia.org/wiki/%E8%AA%9E%E9%9F%B3%E8%BE%A8%E8%AD%98)和在[資料壓縮](https://zh.wikipedia.org/wiki/%E8%B3%87%E6%96%99%E5%A3%93%E7%B8%AE)的領域中，這種模式試圖捕捉語言的特性，並預測在語音串列中的下一個字。我們在語音辨識程式中也多次使用了語言模型的公式來計算使用者所述之描述，使程式能夠更加精確的判斷。

**3.1語言模型公式**

統計語音辨識的最基本問題是，給定輸入訊號或特徵序列， O={O1,O2,… On}，符號集（詞典）{W}={W1,W2…,Wn}，求解符號串W=W1,W2 ,…Wk使得：

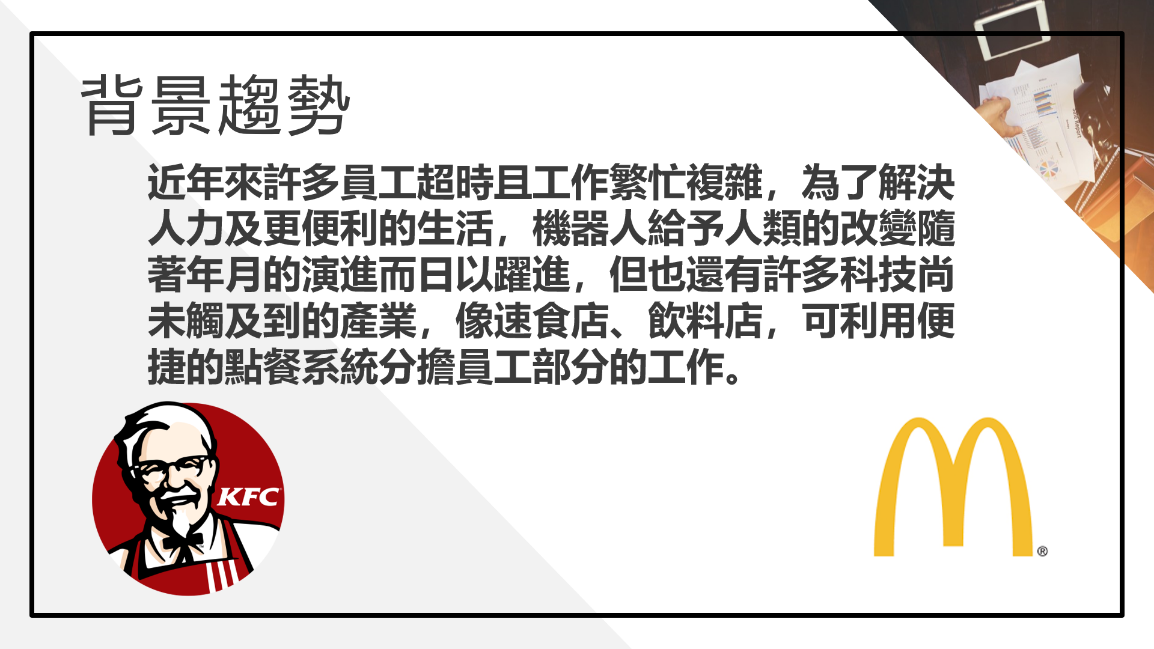


由於對於確定的輸入串O， P(O)是確定的，因此省略它並不會影響上式的最終結果，因此，一般來說語音辨識所討論的問題可以用下面的公式來表示，可以將它稱為語音辨識的基本公式。 {\displaystyle W=\arg \max P(O|W)P(W)}

最後，語言模型則提供了 P(W)。這樣，基本公式就可以更加具體的寫成：



本專題程式特過短分時析，擷取特定影格來分辨使用者是否為本人的聲學特徵，並結合Google語音訓練軟體來進行訓練，增加數據量以防偵測錯誤。

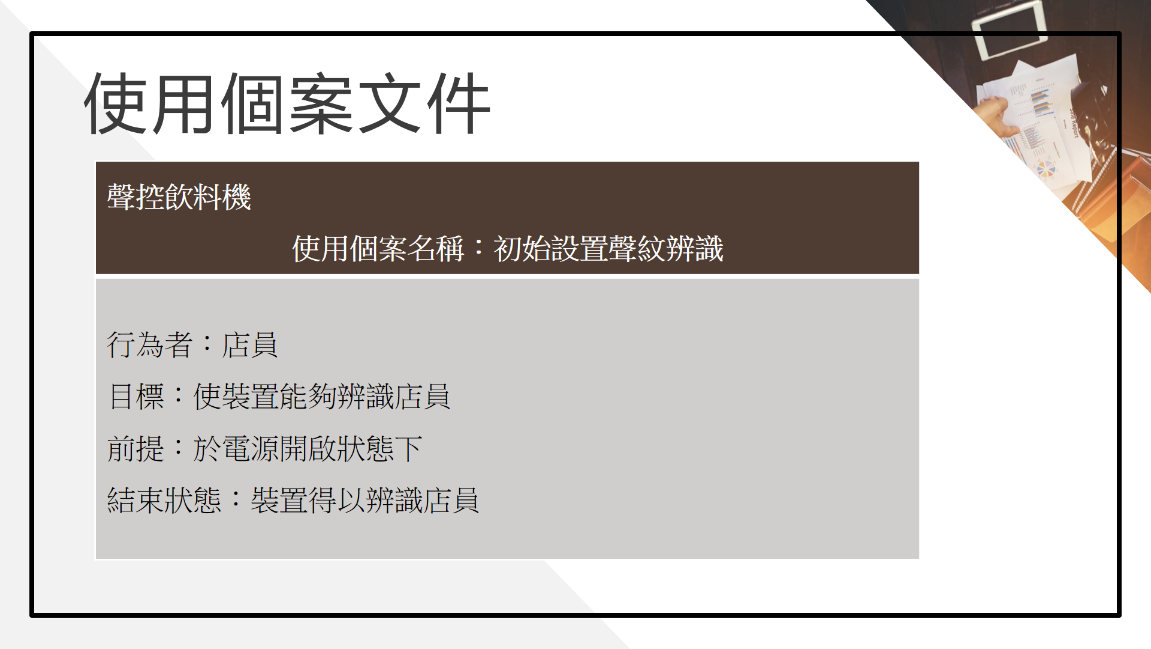
**背景趨勢:**

**動機目的:**

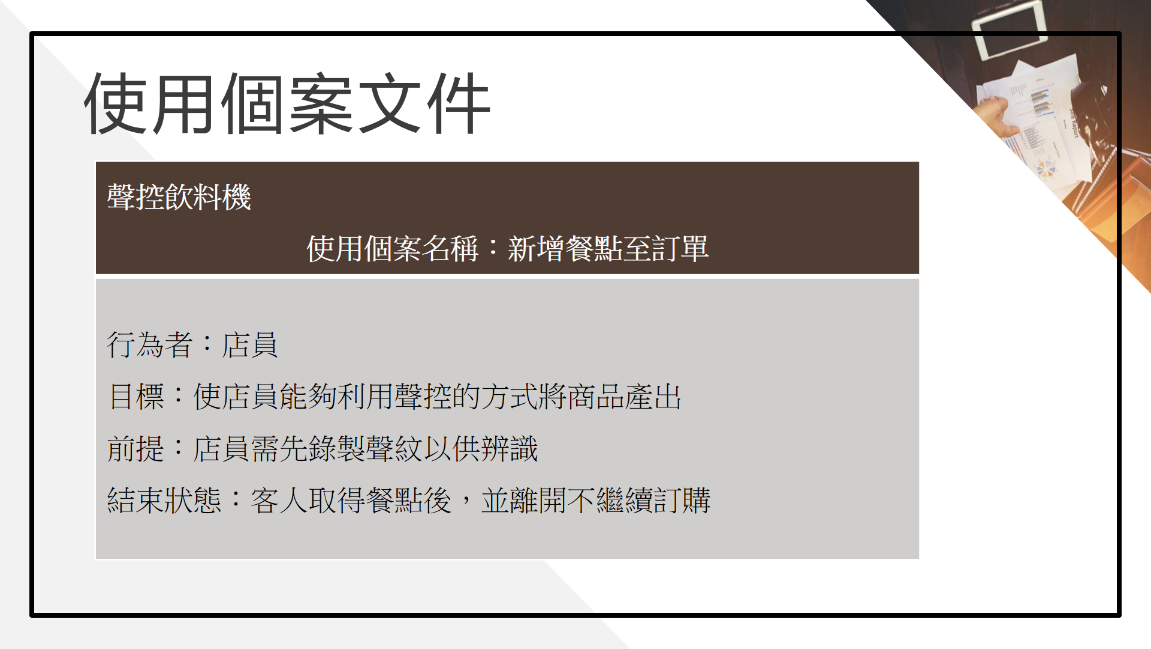
**描述性項目與事件條列式:**



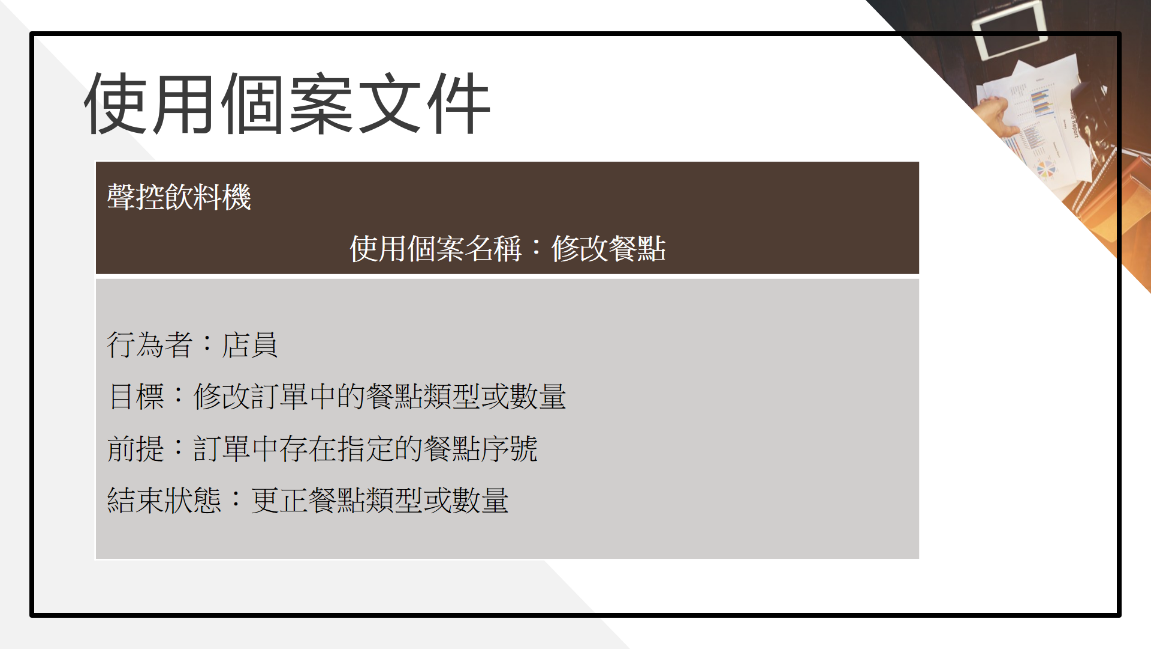
**使用個案圖-初始設置聲紋辨識:**

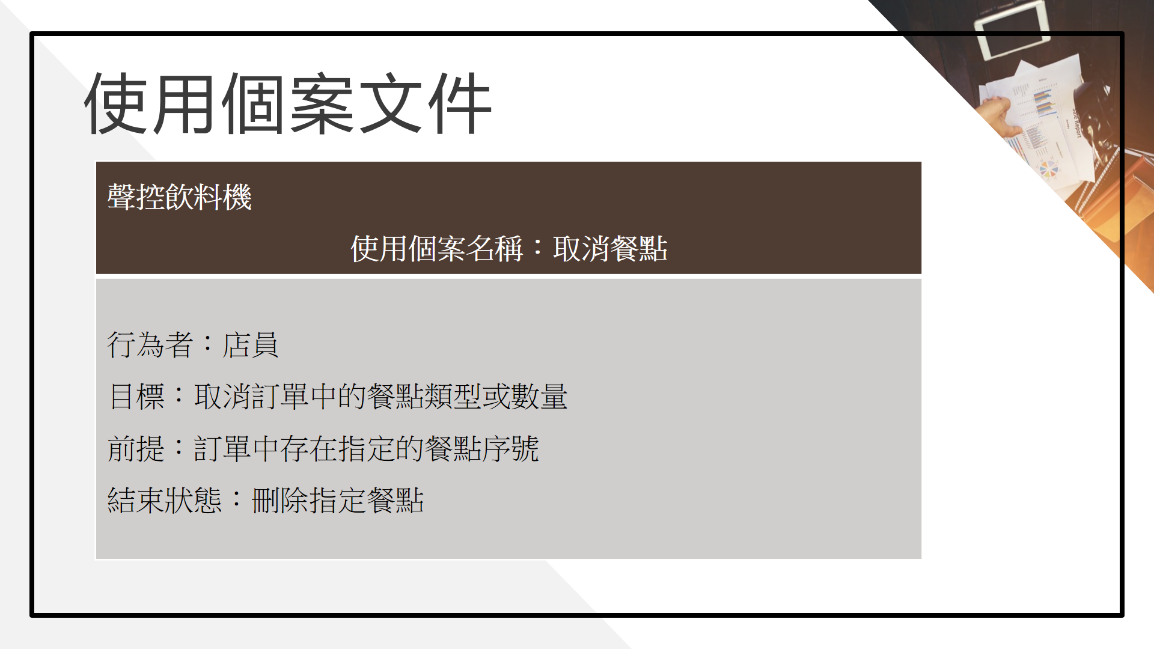


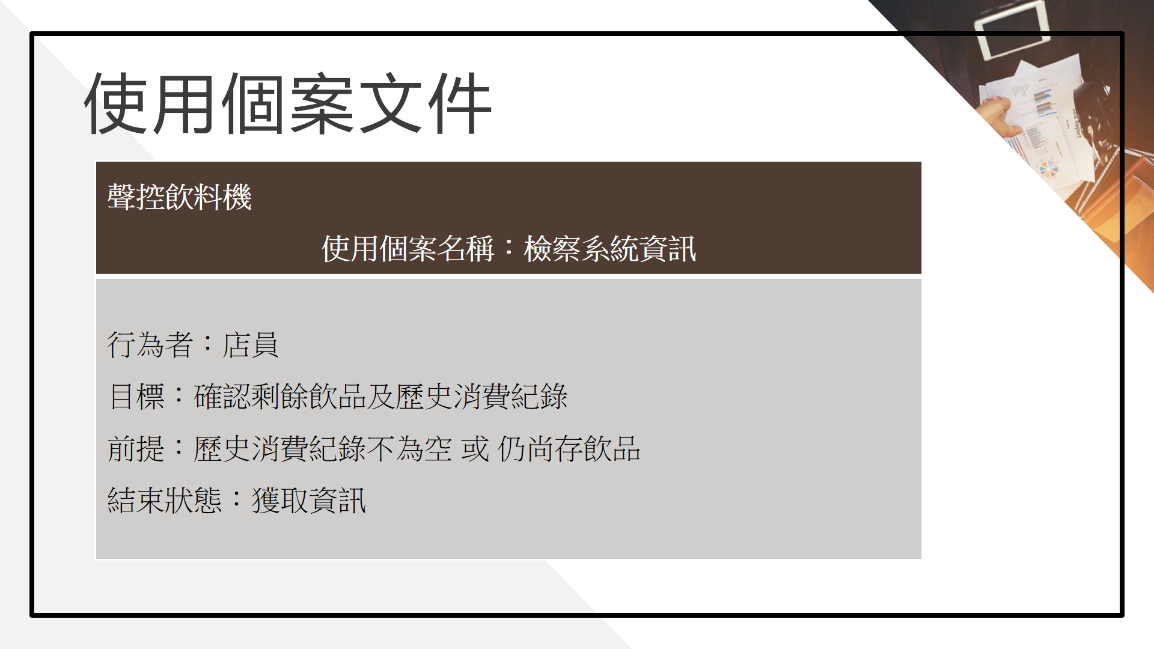
**使用個案圖-新增餐點至訂單:**



**使用個案圖-修改餐點:**



**使用個案圖-取消餐點**:

**使用個案圖-檢察系統資訊**:

**使用個案圖:** 

**分工表**:

**參考文獻**

[1] 賴玟杏，語音辨識相關資料檔

[2]Alex Lee.ANC Principle and analysis.

TechTeller.(2017)

**誌謝**

謝謝每個組員的努力才可使這次專題順利完成，也感謝賴玟杏教授的程式指導及曾士桓教授的報告