

國立臺灣大學管理學院資訊管理學系

碩士論文

Department of Information Management

College of Management

National Taiwan University

Master Thesis

以 Instagram 資料試做一套餐飲評分機制

On the Use of Instagram Data for Building  
a Rating System in Catering Industry

林瑋鴻

Wei-Hong Lin

指導教授：莊裕澤 博士

Advisor: Yuh-Jzer Joung, Ph.D.

中華民國 109 年 7 月

June 2020

## 論文摘要

作者：林瑋鴻

民國一百零九年七月

指導教授：莊裕澤 博士

### 以 Instagram 資料試做一套餐飲評分機制

隨著網際網路的發達，各種社群平台如雨後春筍般的興起，如：facebook、instagram、ptt、twitter、dcard...等，大眾儘需透過手機連上網即可隨時隨地發表自己的評論意見，討論著每一件事、每一個事物，因此對於美食愛好者而言，能夠透過大眾的意見找到真正的美食更是省下了自行搜尋的麻煩，也因此關於美食的評分與評論機制變的更加普及，如：google 評論、愛食記、大重點評網、yelp、menu 美食誌...等，然而，在眾多被推薦的店家中，該如何快速選擇最想造訪的店家儼然變成了一道難題，此時綜合大眾評價的店家排名即顯得相當重要。

因此本篇論文欲以公認排名推薦較準的社群平台為標準，設計餐飲評分機制學習此社群平台排名的計算方式，日後不再需要透過大眾投票的問卷調查或是美食社群的調查便能找出符合大眾心中的最佳餐廳排行。

本研究利用社群聆聽(social listening)的概念結合情緒分析技術，分析 Instagram 社群上用戶對於 100 間店家的評論資料，以公認推薦排名較準的 menu 美食誌 top10 餐廳排行榜，作為本研究餐飲評分機制的模型訓練資料，嘗試各種方式預測出趨近於 menu 美食誌 top10 的餐廳排名，並進一步探究預測排名準確或不準確之原因，最後透過三種方式驗證排名預測的準確度。

關鍵字：情緒分析、社群聆聽、線上評分機制、評論探勘

## 目錄

中文摘要.....	i
第一章、緒論 .....	1
1.1 研究背景.....	1
1.2 研究動機.....	2
1.3 研究目的.....	3
1.4 論文架構.....	4
第二章、文獻探討.....	5
2.1 線上評分機制.....	5
2.1.1 Google map 評分 .....	6
2.1.2 Yelp.....	7
2.1.3 大眾點評網 .....	8
2.1.4 評分機制比較 .....	9
2.1.5 MENU 美食誌 .....	10
2.2 評論探勘.....	11
2.2.1 情緒分析相關研究.....	12
2.2.2 現有評分機制問題.....	14
2.3 Instagram 社群行銷 .....	16
2.3.1 主題標籤行銷 .....	16
2.3.2 網路打卡行銷 .....	17
2.4 影響力分析.....	17
2.4.1 社群影響力計算方式.....	18
2.4.2 社群影響力與排名之關係 .....	20
2.4.3 社群影響力結合情緒分數與排名之關係 .....	20
2.5 小結 .....	21
第三章、研究方法.....	23
3.1 研究問題.....	23

3.2 研究架構.....	23
3.2.1 資料蒐集.....	26
3.2.2 線性規劃法與機器學習分類演算法.....	29
3.3 研究驗證.....	33
第四章、研究結果.....	35
4.1 資料爬取與分析.....	35
4.2 排名預測與驗證.....	39
4.2.1 線性規劃之餐飲評分機制.....	39
4.2.2 機器學習分類演算法.....	54
4.2.3 線性規劃法與機器學習分類法之結果比較.....	62
4.3 各餐飲種類的排名差異分析.....	63
4.3.1 線性規劃法的排名結果分析.....	63
4.3.2 機器學習分類演算法的排名結果分析.....	64
4.4 小結.....	65
第五章、結論.....	67
5.1 研究成果.....	67
5.2 研究貢獻.....	68
5.3 研究限制.....	69
5.4 未來研究方向.....	70
中文參考文獻.....	71
西文參考文獻.....	74
附錄.....	1
附錄一、線性規劃法排名結果.....	1
一、義大利麵.....	1
二、牛肉麵.....	1
三、滷肉飯.....	2
四、咖哩料理.....	2
五、港式餐廳.....	3
六、韓式炸雞.....	3

七、牛排 .....	4
八、鐵板燒.....	4
九、冰品 .....	5
十、韓式料理.....	5
附錄二、 擴增情緒詞典 .....	1

# 第一章、緒論

## 1.1 研究背景

隨著網際網路發展的日益進步，任何人都能不受時間、地點的限制在網路上留下自己對於某項商品或服務的評論與意見，人們也開始習慣把自己的體驗與感受透過網路評分與評論機制傳達出去，舉凡 google map 評分，任何人都能在任何一個商家地址或是景點位址，根據自己在此地的經驗撰寫評論並加入評分，也因此網路上充斥著愈來愈多使用者原創的評論內容(User Generated Content)。

除此之外，評分與評論機制的普及更可以從各式電子商務平台中看出，如：旅遊類型的 TripAdvisor、購物類型的 eBay、Amazon、餐廳類型的 Yelp，愈來愈多人會參考他人在網路平台上給出的意見進行購買決策，根據美國市調公司尼爾森於「2009 年全球網路消費者調查」中早已發現近約七成的消費者都會相信網路上的評論(簡之文，2012)。又如：共享經濟平台 Uber，利用用戶的評價篩選不良的司機，以提供更優質的服務，或是 Airbnb 仰賴房客的評論建立房東的信譽以提供他人作為選擇房間的參考，由此可見，評分與評論機制的重要性亦是不言而喻，此時評分機制的好與壞就顯得非常重要，因為好的評分機制才能給予參考者最正確的資訊。

再者，消費者除了透過特定網站或應用程式進行評分與評論之外，亦常見於社群平台上抒發個人意見，如：Facebook 或是 Instagram，其中又以圖像為主的分享平台 Instagram 用戶對於餐飲業的評論內容較多，根據英國連鎖餐廳 Zizzi 的調查(2017)，18-35 歲的用戶平均一年會花五天的時間瀏覽 Instagram 上的食物照片，當中 3 成的用戶會參考這些照片決定是否消費，可見消費者已逐漸將 Instagram 視為選擇餐廳的媒介之一，因此本篇論文將以 Instagram 為研究對象探討如何建立新的餐廳評分機制，讓消費者可直接透過評分結果選擇

欲造訪的商家。

## 1.2 研究動機

在這個人人都持有手機的世代裡，隨手拍照已不再是件難事，也因此愈來愈多人喜歡透過 Instagram 拍照分享美食，在美食上桌後第一件事就是讓手機“先食”，同時也吸引愈來愈多人想透過 Instagram 來找美食，因此餐廳店家的名聲度將是消費者選擇商家的重要指標之一，因此若能提供消費者各種餐飲類別的排名，消費者將更能夠快速判斷欲造訪的商家。而本文所設計的餐飲評分機制即是希望透過 Instagram 意見領袖(Key opinion leader)的影響力因素計算出最符合大眾心目中的餐廳排行榜。

然而目前現有的評分機制，如：google 評分(如圖 1.)，單純從所有評分統計出平均值，可能會因為被“灌水”加入假評分，加上每則評分權重一樣，導致評分結果不如預期，高評分的店家實質上卻提供低品質的服務或商品，更無法確實反映店家的知名度，此外，長久累計評分的結果也無法即時反映近期店家的品質狀況及知名度，倘若店家因為更換老闆導致品質嚴重下降，新加入的低評分卻仍不足以讓整體評分明顯下降，亦無法反映知名度下降的趨勢。



圖 1. Google 評分-新馬辣火鍋店

(資料來源：Google Maps)

根據 Dai 等人(2018)的研究針對“灌水”問題將用戶的每一則評分進行比

重加權，評論字數較少的評分予以較低權重，反之亦然，因此，本研究將延續加權的概念，有鑑於 Instagram 用戶相較於 Google 用戶多具備粉絲數(追蹤者)的因素，一般而言粉絲數多的用戶能夠影響愈多人的購買決策，但是根據 Lee (2015)和 Peters 等人(2013)關於影響力行銷(Influencer marketing)的研究指出粉絲數多的用戶(Influencer)影響力不一定大，因粉絲數不一定代表實際粉絲互動率，而 Ling Hang Yew(2018)加入用戶發文平均按讚數以及平均評論數等因子，整理出較能代表用戶真正影響力的計算方式，故本研究參考 Ling Hang Yew(2018)的影響力計算方式後，嘗試將用戶的影響力作為用戶評論的加權比重，以排除既有評分機制的”灌水”問題。

然而，Instagram 用戶評論並未包含用戶評分，因此本篇論文會採用目前發展成熟的情緒分析技術(Sentiment analysis)，先進行用戶評論的情緒分數計算，將評論轉為評分，再以此評分進行加權處理，影響力大的用戶擁有較高的評分權重，以此結果驗證是否能計算出符合大眾喜好的餐廳排名結果。

### 1.3 研究目的

本研究希望建立新的評分機制，以 MENU 美食誌所調查出的排行榜作為符合大眾心中排行的標準代表，透過本研究餐飲評分機制計算出的新排名，預測 MENU 美食誌的排名，目的是為了解決以問卷調查大眾意見時造成的耗時與耗人力等低成本問題，以及不須透過 MENU 美食誌的排名計算方式，即可找出符合大眾心中的餐廳排名。

至於以 MENU 美食誌的排行榜為標準排名，是因為 MENU 美食誌決定出的餐廳排行榜較具有公信力，其排名標準都是來自於 MENU 美食誌 app 上面常在發表美食心得的美食部落客，相較網路溫度計(2019)也是用情緒分析方法進行排名，但卻未經美食客認證，所以本研究採用 MENU 美食誌作為餐廳排名的標準。



首先，本研究根據 MENU 美食誌(2020)於 2019 年、2020 年對於義大利麵、牛肉麵、滷肉飯、咖哩料理、港式餐廳、韓式炸雞、韓式料理、牛排、鐵板燒、冰品十大類調查出的前十名店家，代表美食客們推薦的前十名，作為本研究分析的店家資料，從 Instagram 上蒐集這些店家於指定時段的相關評論，包括以 hashtag 方式提及店家名稱時所提供的評論或是直接到店家地址打卡時所寫下的評論，再利用情緒分析技術(Sentiment analysis)將評論轉為評分，接著參考 Ling Hang Yew(2018)的影響力計算方式嘗試將影響力大的用戶(Influencer)評分進行加權，加總結果即為改善後的評分，最後，根據調整後的評分計算出餐廳排名前十名。

總而言之，本研究主要目的為，讓消費者可以根據本研究的餐飲評分機制排名結果更快速地選擇欲造訪的餐廳店家，而不須透過任何其他方式，如：問卷調查或是只能等待 MENU 美食誌的排行榜公布。

## **1.4 論文架構**

在接下來的論文架構，將於第二章探討過去相關的文獻，包含社群影響力分析、情緒分析以及相關技術文獻，於第三章詳述研究過程及實驗設計，第四章提出實驗研究成果及驗證，最後於第五章說明結論與未來研究方向。

## 第二章、文獻探討

本篇論文主要目的可分為兩個方向，第一個為從 Instagram 上蒐集指定時間區段內關於欲分析店家的所有 Instagram 用戶評論，藉由情緒分析技術 (Sentiment analysis)、意見探勘技術 (Opinion mining) 計算出評論分數，再進一步藉由 Instagram 用戶的影響力進行分數加權，計算出店家在特定時間區段的評分。第二個目的則是結合店家的打卡數、hashtag 數與用戶影響力利用機器學習分類演算法預測排名。目前已有許多發展成熟的技術可供本研究使用，像是情緒分析技術 (Sentiment analysis) 以及意見探勘技術 (Opinion mining)，亦有相關的研究可供本研究進一步優化與使用，如：Ling Hang Yew (2018) 研究中對於 Instagram 用戶影響力的計算方式以及陳佳雯對於社群影響力的分析研究 (2017)。

於本章第一節，本研究將探討既有線上評分機制的現況，而過去研究中已針對評論提出新的評分機制，因此，第二節將探討評論探勘 (Review mining) 如何利用情緒分析 (Sentiment analysis) 以及意見探勘技術 (Opinion mining) 分析評論，並基於這些研究結果延續本研究之實驗。接著，因為考慮到本研究以 Instagram 社群作為分析資料的來源，因此第三節將進一步探討社群行銷 (Social Media Marketing) 的相關研究，以利後續更加了解如何使用 Instagram 資料進行排名預測。最後，於第四節加入影響力分析，探討影響力會如何影響排名結果，故本研究嘗試加入影響力的因素計算排名。

### 2.1 線上評分機制

網路上充斥著各式關於餐飲業的大眾評分網站，如：於 google 搜尋引擎搜尋店家名稱時最先跳出的 Google 評分、美國最大評分網站 Yelp、中國最大的生活消費指南”大眾點評網”、數位美食雜誌”愛食記”、香港餐廳指南”OpenRice 開飯喇”等網站，以下內容將進一步探討用戶數眾多的三個線上評分

系統，包括 google、yelp、大眾點評網，有何特色以及分別存在哪些大眾評分機制問題，並且介紹本研究作為標準排行的 MENU 美食誌如何讓自有的排行榜具有影響力與公信力。

### 2.1.1 Google map 評分

Google Maps 的軟體工程師 Jonathan Goldman 於 2007 年宣布大眾可以開始在 Google Maps 上發表自己的評論，不論是餐廳、醫院、美髮店、旅館、旅遊景點甚至是學校，希望大眾能透過 Maps 告訴全世界在當地的感受(Pitman, 2019)。根據 BrightLocal 行銷公司研究(如圖 2.)指出 Google 蒐集到的評論數量在 2015 年到 2016 年之間成長快速，成長率達到 278%，遠超出其他美國當地知名的評論網站，如：Facebook、yelp、TripAdvisor 等網站(Murphy, 2018)。值得一提的是，根據 Google 官方的說明，用戶評論的豐富度及完整性、評論數量更是會大大影響搜尋排名，如同 Murphy(2018)研究所述 google 排名 1-3 名的店家相較 4-6 名會多出平均 472% 的評論數量。

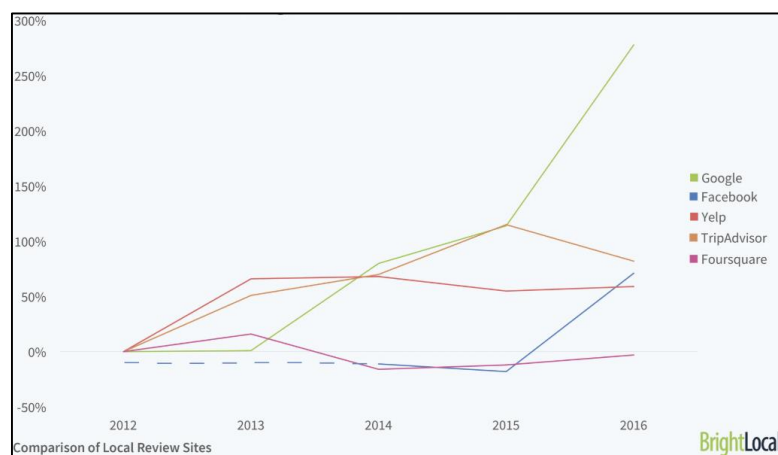


圖 2. 美國當地評論網站的新評論成長率

(資料來源：BrightLocal)

然而，隨著評論數量的增長，卻無法代表評論可靠度的增加，尤其很多店家考量到高評分會為他們帶來更多顧客，因而雇用專門“洗評價”的

行銷公司(如：FANSDOOR、Forestracker、WPR2)(Kaviya, 2017)，在短時間內衝高自己的評分與評論數量，因此本研究的目的之一即在探討如何確實避免遭受”洗評價”的威脅。

### 2.1.2 Yelp

Yelp 由任職 PayPal 的前員工 Russel Simmons 和 Jeremy Stoppelman 於 2004 年所創建，目前在全世界收錄一億則以上的商家評論，其經營理念始終保持著評論者至上，消費者其次，商家最後的理念，抱持” Real People Real Review” 的初衷，為了不讓商家影響評論者最真實的評論，更嚴禁商家於 Yelp 網站宣傳廣告，而為了鼓勵評論者提供更高品質、更細節的評語，Yelp 創建 Elite 的會員概念，只要符合基本準則(如圖 3.)即能享有特別的待遇以及獎勵(如：專屬 Elite 的 VIP 派對)(黃郁珊，2015)，也因為其相當重視評論者，直至今日才能累積如此驚人的評論數量。

成為 Elite 的基本準則
• 必須使用真人大頭照、與真實姓名成為會員。
• 積極寫評論(尤其是成為第一個評論的人的評論數量)。
• 評論要囊括從酒吧、餐廳、水電服務、當地診所、政府機構、指甲沙龍、乾洗店、到精品店等任何類型的店家評論，想成為 Elite，評論必須同時具備廣度和深度。
• 積極與社群互動，如和其他 Yelper 成為朋友、瀏覽且回覆其他 Yelper 的評論、FUCs 投票、贈送讚美等。

圖 3. 成為 Elite 的基本準則

(資料來源：Yelp Elite Squad)

但是這麼多的評論卻仍舊面臨假評論的威脅(Luca, 2016)，因此，Yelp 為了因應假評論，設計新演算法並依據內容品質高的評論進行評分加權，舉例而言，寫愈多細節的評論愈容易被排在商店評論前幾則顯眼的位置，而粗劣的評價則容易被排在後幾則不容易被看見，如此一來即可成功隱藏假評論，不影響使用者觀看評論所獲得的商家資訊，並且不將假評論的評分計入整體評分中，然而此方法卻因誤刪使用者真正的留言引起使用者反

彈(彭博社網站, 2015), 因此本研究欲探討是否能轉為利用用戶影響力當作評分加權以降低假評論的影響程度, 避免假評論的威脅。

### 2.1.3 大眾點評網

大眾點評網由張濤於 2003 年創立, 目前為中國最大且全世界第一個建立的大眾評論網站, 截至 2014 年, 網站已有超過 1.9 億的用戶數量, 評論數量也達到 6000 多萬則, 至於其商家評分方式則是提供用戶 0-4 分的選擇, 代表差、一般、好、很好、非常好, 可在口味、環境、服務三個指標上打分數, 有別於其他一般的評分機制只提供整體評分的選擇(如圖 4.), 同時我們也可以從圖中發現大眾點評網除了希望用戶提供評分之外, 也希望能夠一次獲得其他的細節資訊, 包括喜歡的菜色、圖片、平均消費價格、餐廳特色等資訊, 如此一來可讓其他消費者有更多關於店家的資訊參考。此外, 在評分的計算方式上, 不單單計算所有用戶評分的平均, 大眾點評網多加考慮用戶的評論總數、評論時間以及信譽度, 以調整此用戶對評分計算結果的貢獻(廖敏惠, 2015)。

The screenshot displays the Dianping.com (大眾點評) review interface for the restaurant '鹿港小鎮(環貿廣場店)' (Lugang Town (Huanjiao Plaza Branch)). The page includes a header with the logo and navigation links. The main content area features a star rating system for the overall score, with sub-ratings for '口味' (Taste), '環境' (Environment), and '服務' (Service). Below these are sections for '喜歡的菜' (Favorite Dishes) and '評價' (Review). The '評價' section includes a text input field and a character count. The '上傳圖片' (Upload Photos) section allows users to upload photos, with a note about file size and format. The '人均' (Per Person) section shows a price input field. The '餐廳特色' (Restaurant Features) section includes checkboxes for '可以刷卡' (Can Use Card), '無線上網' (Free Internet), '免費停車' (Free Parking), '有露天位' (Outdoor Seating), '有表演' (Performance), and '可送外賣' (Takeout Available). The '停車信息' (Parking Information) section includes a text input field for '车停在哪里了, 收费么?' (Where is the car parked, is there a fee?). The bottom of the page has a '提交點評' (Submit Review) button and a '且同步到:' (Also sync to:) section with social media icons.

圖 4. 大眾點評網用戶評分頁面

(資料來源：大眾點評網)

由大眾點評網的評分計算方式可見，整體評分確實無法單由全部評分的平均來判定，還需加入用戶的評論總數、評論時間以及信譽度等權重，而大眾點評網的信譽度若轉換為 Instagram 可利用分析的資料，即為每篇貼文的按讚數，以直觀而言按讚數在某種程度上可以解釋為用戶對於貼文的同意程度或喜好程度，因此本研究嘗試採用用戶的按讚數作為權重基礎，試著用不同方式改善”票票等值”而導致”洗評分”的問題。

#### 2.1.4 評分機制比較

根據以上三種大眾評論網站的背景介紹以及發展現況，此章節將針對網站賦予商家的評鑑指標和提供給消費者的商家參考資訊，於以下表格進行統整，且此三種大眾評論網站的評分計算流程仍存在一些具爭議的評分問題，因此，也將於以下表格逐一列出。

	評鑑指標	參考資訊	評分問題
Google map 評分	商家整體評分 評論數量 好感度、價格	距離 熱門時段 菜單、相片 餐廳特色(無障礙空間、可停車...)	假評論 評論與評分不符 評分無法即時反應最新評論
Yelp	商家整體評分 評論數量 價格	距離 菜單、相片	假評論 評論與評分不符
大眾點評網	商家整體評分 評論數量	菜單、相片 餐廳特色(可刷	假評論 評論與評分不符

	口味、環境	卡、無線上網...)	
	服務、價格	推薦菜	

表 1. 大眾評論網站之評鑑指標比較

(資料來源：本研究整理)

由上述表格可以觀察到目前的大眾評分機制仍無法脫離假評論以及相關的衍生問題，導致以評分大小計算出的餐廳排名，不一定能夠代表符合大眾心目中的排行，故接下來的章節將進一步探討過去文獻對於評分機制問題所提出的改善方針，作為發想本研究餐飲評分機制的參考。

### 2.1.5 MENU 美食誌

MENU 美食誌是一款提供美食客(用戶)簡單紀錄並可快速搜尋好吃店家的 app，期望用戶在享用完每一間好吃的店家之後，也可以隨手將食記心得紀錄在 app 上，成為最懂吃的美食客(如圖 5)，且提供各種徽章象徵美食客的”吃貨等級”，以及各自擅長、專業的美食領域。每隔一段時間，MENU 美食誌亦會計算商家於某段時間內獲得的美食客打卡數，以打卡數最多的商家賦予排行榜中的第一名，進而計算出前十名的排行榜，其排名計算方式因以打卡數為計算基準，因此不受長期累計評分的評分匯總問題或假評論問題所影響，且一旦排名發布後，在網路上可見時常會有媒體報導其餐飲排行榜，可見其公信力，故本研究將 MENU 美食誌的排行榜作為最符合大眾喜好的標準餐廳排行。



圖 5. MENU 美食誌的美食客範例(2020)

## 2.2 評論探勘

評論探勘(Review mining)一詞根據 Jin 等人(2014)著作中的定義，將之定義為透過分析評論且總結轉換成有用的資訊之流程，此流程能夠帶給分析者或商家增加收入或是減少成本的利益，此外，亦可被稱為意見探勘(Opinion mining)(Kamal, 2015)，意見探勘隸屬於網頁內容探勘(Web content mining)的子類別，被定義為資料探勘技術的其中一種應用(圖 6.)，目的是利用自然語言處理(Natural Language Processing)追蹤大眾對於一產品或服務的意見(Perera & Caldera, 2017)。



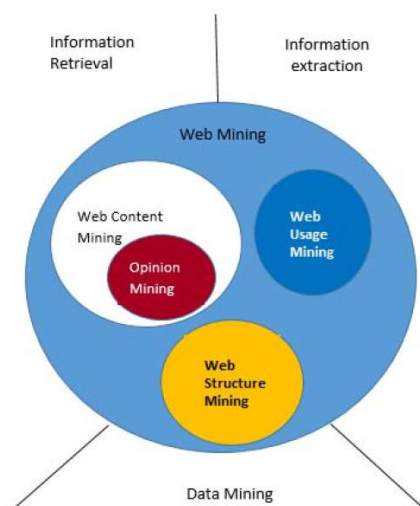


圖 6. 意見探勘以及相關領域

(Perera & Caldera, 2017)

至於意見探勘一詞亦有不少研究將之視為情緒分析(Sentiment Analysis)(林彩雯, 2015)，目的均為對一主觀之資料進行分析，而目前主要的情緒分析方法可分為兩種，分別是機器學習法(machine learning approaches)和字典法(lexicon-based methods)。前者又分為監督式(Supervised)以及非監督式(Non-Supervised)兩類，監督式即為利用標記好正負面極性的語料讓機器學習，常用於文件分類，如支援向量機(Support Vector Machine)，非監督式則不需人工標記資料，而是透過資料之間的相似度進行分群，然而監督式的準確率高於非監督式(Chaovalit & Zhou, 2005)，至於後者字典法則是使用現有的辭典(如：台大中文情感極性情緒辭典 NTUSD、知網 HowNet)對照文中的情緒詞即可計算情緒分數(廖敏惠, 2015)。

### 2.2.1 情緒分析相關研究

而過往對於評論探勘的研究多著重於利用情緒分析技術將評論分類為正面或負面，然而僅僅知道正負面意見並無法有效改善商家所提供的商品或服務(Jin, 2014)。在 Perera 和 Caldera (2017)的研究中提到，若欲從評論中人工擷取關於商家各方面的評價(如：價格、環境、地點、口味...等主題

資訊)，幾乎是不可能之事，因為有太大量的評論數目難以逐一分析，因此這份研究採用自然語言處理方式，先從 zomato 餐廳搜尋網站上蒐集 100 個國家近一萬則評論，利用詞性標記器（part-of-speech tagger）將每個字詞標上詞性，再利用相依性剖析器(dependency parser)將每個主題的意見詞擷取出來，最後計算出每則評論對於每個主題的情緒分數。

其他關於餐廳評分機制的過往研究則趨於少數，但相較電影評分機制或是書籍銷售評分機制卻有較多的研究探討，因為評分機制是否改善需透過商家的利益收入是否增加得知，而書籍銷售量(Chevalier & Mayzlin, 2006)和電影收入(Moon, Bergey, & Iacobucci, 2010)具有公開的紀錄，但私人餐廳收入較難以取得，故較少關於餐廳評分機制的研究(Jia, 2018)。根據 Jia 的研究(2018)，作者利用多重線性回歸篩選出影響口味、服務、環境評價最劇烈的因素。Gan(2016)認為過去研究普遍只找出評論中關於食物、服務、環境、價格四個面向如何影響最終評分，因此利用情緒分析從評論中找出第五個面向情境，研究結果顯示食物、服務、情境影響最終評分最劇，其次為價格和環境。

而在廖敏惠(2015)的研究中，作者將大眾點評網的評論資料利用中研院的 CKIP 中文斷詞系統斷詞後，進行人工標記情緒詞、搭配詞、程度詞、否定詞、未知詞，並建立餐飲業專用詞庫，而後利用情緒分析計算出每則評論的情緒分數，再與原評論的評分比對正確率，研究發現，對極性強烈(情緒分數較高或較低)的評論可以達到 88-96%的準確率。李啟菁

(2010) 則以字典法比對數位相機評論的情緒詞以及程度詞計算評論的情緒分數，分析結果與人工給予的評分相似。Yang Liu 等人的研究(2017)則是透過情緒分析與模糊理論(fuzzy set theory)設計演算法分析用戶對於替代性商品的線上評論，利用線上評論將商品進行推薦排名。因此本研究嘗試將評論透過情緒分析轉換為情緒分數，再進行後續的餐廳推薦排名計

算。

在本研究中，斷詞系統可採用發展至目前為止具有高精準度且被廣泛應用於各研究中的 CKIP 中文斷詞系統(An, 2011)或是結疤斷詞系統(Jieba)(2016)，而後利用台大擴增情緒辭典(ANTUSD) (Shih-Ming, 2016) 搭配人工標記的餐飲領域擴增情緒詞彙，對於餐飲業的評論進行情緒分數的計算，相較單純使用台大擴增情緒辭典(ANTUSD) (Shih-Ming, 2016)，加入擴充的餐飲領域情緒詞彙較能夠在餐飲評論中識別出特殊的情緒詞。

### 2.2.2 現有評分機制問題

而本研究主要目的之一在於利用情緒分析嘗試建立餐飲評分機制，計算出餐廳排名結果，並排除現有評分機制所潛藏的問題，包括虛假評論、長期累計評分的評分匯總（rating aggregation）問題以及評論與評分不符問題，導致排名結果無法反應近期使用者的喜好，以下將分別進一步進行文獻探討。

#### 2.2.2.1 假評論問題

針對假評論的問題，亦有相關研究持續探討，在 Luca(2016)對於 Yelp 假評論的研究中有四項發現指出，第一，大約有 16%的評論被篩選為有假造的嫌疑，這些評論通常具有較極端的喜好，可能為極度支持或是極度反對，第二，當餐廳有較少的評論或是有較多負面評論時，較容易引起餐廳自行提交假評語，第三，對於受益於 Yelp 的連鎖店較不會自行假造假評語，第四，當店家面臨的競爭提高時較容易收到負面的假評語。而 Akoglu 等人(2013)則提出一個可快速且有效辨識假評論的框架，FRAUDEAGLE，有別於一般利用文字探勘或行為分析方法偵測假評論，作者利用評論者和產品的網路效應，將評論者和評論賦予一分數以偵測假評論及假評論製造者，對於大量的評論內容也相

當適用。

由過去文獻可發現假評論確實存在，且為數不少，而過去研究解決假評論問題的方式都是透過文字探勘、行為分析或是 Akoglu 等人 (2013) 提出的網路效應篩選假評論，但是本研究將以不同方式處理假評論問題，希望透過評論者影響力高低調整評論權重，藉此降低假評論的權重，而非直接偵測假評論。

#### **2.2.2.2 評分匯總問題**

除了假評論問題之外，評分匯總 (rating aggregation) 亦是許多研究探討的議題之一，評分匯總意指長久累計評分導致評分結果無法反映店家的真正品質，如 Dai 等人 (2018) 研究所述，舉例而言，A 餐廳在前 12 個月獲得 2 顆星評價，後 12 個月獲得 4 顆星，而 B 餐廳在前 12 個月獲得 4 顆星評價，後 12 個月獲得 2 顆星，在整體評分結果卻是相同的，造成參考者無法判斷 B 餐廳實際上品質正在下降，而誤選 B 餐廳，因此作者根據評論的資訊含量、評論的時間、評論者是否為 Yelp 認證的精英賦予評分不同權重，設計新的演算法計算整體評分。此外，在現有的評分網站中，yelp 根據認證的菁英賦予較高的評分權重，大眾點評網則根據時間因素，愈近期的評論給予較高的權重，反之亦然，以期望改善評分匯總的問題。

過去文獻當中透過不同因素調整評分的權重，包括評論者個人特質、時間因素、資訊含量等，而本研究也將考慮進這些因素並且在個人特質的部分加入影響力的因素，透過不同方式解決評分匯總問題。

#### **2.2.2.3 評論與評分不符問題**

除了上述評分問題之外，不可否認的，並非所有的評論者都會給予一致的評論與評分，倘若評分無法反映評論內容，則可能導致整體

評分的偏差，無法真正反映商品或服務的品質好壞，在 Riyanul Islam(2014)的研究中提出一個新的評價系統，為了移除使用者評分與評價之間的模稜兩可，藉由整合原星級評分和評論的數值極性(情緒分數)，產生調整後的評分，驗證結果顯示，由調整後的評分總結計算的整體評分更能夠反映產品或服務的品質。

過去關於評論與評分不符的文獻專注在評論的情緒分數計算，而本研究亦會實作情緒分析技術計算評論的情緒分數，但是不同之處在於本研究不考慮使用者的原始評分，而是直接利用評論轉成的情緒分數代表評論者的真正意見，接著直接考慮進評論者影響力作為評分權重調整，以去除可能的潛藏問題，包括假評論帶來的”洗評分”問題以及無法即時反應用戶意見的評分匯總問題，而接下來將針對 Instagram 社群行銷的文獻進行探討。

## **2.3 Instagram 社群行銷**

至今已有愈來愈多的用戶透過 Instagram 搜尋美食，同時也愈來愈多的美食部落客在 Instagram 社群上發展，由金蛋網路數位行銷公司(2018)的調查中亦指出，許多店家因為經過粉絲數多、影響力大的美食相關媒體或粉絲專業(如：ETToday 餓勢力、波波發胖)介紹後，知名度大漲，業績翻倍成長，因此本研究選擇將 Instagram 用戶影響力作為本研究建立餐飲評分機制的權重依據，此外，Instagram 打卡數、hashtag 數是否對排名結果有所影響，也將在此章節進行文獻探討。

### **2.3.1 主題標籤行銷**

主題標籤(Hashtag)是一個在各大社群媒體(如:Facebook, Instagram, twitter)上被廣為使用的標記方式，任何人都可以在自己的文章裡寫上欲標記的詞彙，而他人可以透過搜尋主題標籤(Hashtag)，篩選出有附上此標籤的文章，因此在社群媒體上常被商家用來增加品牌認知度(Brand

awareness)，宣傳自家的商品或服務(Bhasin, 2019; Vedantam, 2017)。而主題標籤行銷則是商家透過主題標籤在社群網路上宣傳活動，並找到真正目標客群(Target audience)的行銷方式之一(Bhasin, 2019)。而許多商家也開始尋求社群媒體上的網路紅人(Social media influencer)或是意見領袖(Key opinion leader)的幫助，在他們的介紹文章中附上商家產品、活動等…相關主題標籤，透過他們的影響力提升品牌知名度(Bhasin, 2019)。

因此本研究欲進行的機器學習分類演算法，即是希望將指定時間區段中店家被用戶提及的主題標籤數，作為排名依據的特徵值，並測試是否能夠幫助提升預測排名的準確度。

### 2.3.2 網路打卡行銷

除了主題標籤數之外，商家獲得的打卡數亦可代表商家名氣的高低。打卡意旨社群媒體上透過 GPS 系統定位，在一地點上進行登錄，並可附上相片、文字或標記他人，隨後社群媒體便會將這些資訊發布在個人的動態上。在馬千惠(2012)對於餐飲業網路打卡(Check-in)的實證結果顯示，網路打卡具有口碑傳播效果，閱讀者會受到打卡訊息的影響增加造訪商家的興趣，甚至會優先選擇親友打卡過的餐廳進行消費。

因此，在本研究計算排名的過程中，也會嘗試指定一個時間區段，計入商家在 Instagram 上獲得的打卡數，因為打卡數愈多，即代表著商家的名氣也可能隨之升高。

## 2.4 影響力分析

至今已有許多針對影響力進行分析的研究，而本研究選擇嘗試將 Instagram 中每則評論轉換後的情緒分數，透過 Instagram 的用戶影響力進行調整後計算餐廳排名，因此將在此章節進行影響力的相關文獻探討。

### 2.4.1 社群影響力計算方式

在過去已有許多針對社群網路中用戶之間的訊息傳播與影響力計算之研究，最早在此領域提出的研究之一是 Domingos 與 Richardson(2002)利用機率方式探討社群網路使用者之間的互動關係，並透過此方式將影響力最大化問題視為演算法之問題。Kempe 等人的研究(2003)則是最早將影響力最大化問題視為最佳化之問題，並提出 independent cascade model(IC)和 linear threshold model(LT)兩種 model 作為評估影響力程度的依據，independent cascade model(IC)模型即是探討如何模擬社群網路中影響力的擴散，此外，Kempe 等人(2003)亦利用貪婪演算法(Greedy algorithm)，計算出各節點在社群網路中的影響力後，找出社群網路中具有影響力的使用者。Tian Zhu 等人的研究(2009)利用 pagerank(Page, Lawrence 等人, 1999)計算社群中各成員的 pagerank 值，分為社群內影響力的 InnerPagerank 以及對其他社群的影響力 OutterPagerank，根據計算結果定義每個節點代表的角色，包括 Normal node、Bridge node、Core node、Huge influence node。

至於現今社群中的影響力計算亦可由社群自行定義，例如：痞客邦部落格依據用戶產生的文章內容(user generated content)與用戶之間的互動過程定義出五種用戶的社群影響力指標，分別為搜尋力(Search)、擴散力(Share)、人氣力(Publicity)、互動力(Engagement)與號召力(Drive)(痞客邦，2020)，各指標意涵如下表：

➤ <b>搜尋力 Search</b>
內容資訊透過搜尋引擎被搜尋出現的頻率，包含文字闡述、圖文品質、以期達到質量符合瀏覽者期待的程度。
➤ <b>人氣力 Publicity</b>

強調內容資訊在網友閱讀過程中發生的忠誠度行為，包含被造訪次數、閱讀停留時間、被訂閱關注數、不重複訪客數...等。
➤ <b>擴散力 Share</b>
因社群經營者的內容資訊於站內外被引用的廣泛程度，包含站內外來源流量...等。
➤ <b>互動力 Engagement</b>
因著內容資訊產生與會員或粉絲的各項互動行為，包含推、給分、留言...等。
➤ <b>號召力 Drive</b>
透過閱讀內容資訊或是互動，引發會員產生實際行動，達成實際成效的引導。

#### 痞客邦社群影響力指標(痞客邦官方網站, 2020)

透過這五種指標，用戶可以觀察到自己的文章內容是否對社群中其他會員或是在整個網際網路中造成實際影響。

依據陳佳雯於探討 Instagram 美食圖片內容力及社群影響力對按讚行為之影響的研究(2017)中分析，社群平台的影響力可在不同面向中呈現，包括讚數、留言、分享，然而研究發現 Instagram 的留言現象較不盛行，而分享功能亦僅限將 Instagram 的貼文分享到其他的社群上(如：facebook、twitter)，無法達到如同 facebook 的分享功能可在同一個社群中分享各個用戶的貼文，進而達到擴散的效果，故此篇研究將按讚數作為衡量社群影響力的因素，而 Ling Hang Yew 等人(2018)亦考慮貼文按讚數以及貼文觸及數作為 instagram 用戶影響力的計算因素，但是受限於貼文觸及數僅限使用者本人觀看，本研究無法透過爬蟲爬取此項資訊，因此綜合上述，本研究以直觀而言，一則貼文獲得的按讚數在一定程度上可視為成功影響用戶對其進行按讚的動作，故將按讚數設計為用戶的影響力。



## 2.4.2 社群影響力與排名之關係

上一小節中提到各種社群影響力的計算方式，然而，若以 pagerank 而言，pagerank 可用來計算網頁的排名(Page, Lawrence 等人, 1999)，而 pagerank 的概念則說明了一個網頁被其他網頁鏈結的數量，代表著此網頁的重要性，愈多網頁鏈結到的網頁(pagerank 高)，其重要程度愈高，意即影響力愈大的網頁在大眾心中的搜索排名就愈前面，同理，pagerank 高的網頁所主動鏈結到的網頁，其重要程度也會變高，因此本研究嘗試以此概念設計餐飲評分機制，按讚數多的貼文代表其重要性愈高，影響力愈大，其評論內容受到愈多用戶的認可，此篇貼文所介紹的餐廳也就具有一定程度的重要性，在大眾心中的排名也可能隨之上升，反之，若貼文內容為負面情緒，則會降低在大眾心中的排名，因此本研究嘗試將按讚數視為用戶貼文所具有的影響力，並作為後續計算餐廳排名的依據之一。

此外，吳美慧也在文章強度、部落客評價及推薦產品取得方式對購買意圖之影響的研究(2013)中，以美妝保養部落格為例，調查有閱讀美妝部落格習慣的女學生共 120 人，利用 SPSS 統計分析，證實部落格文章的強度愈強，讀者對其推薦之產品所抱持的態度會愈好，並證實若讀者對產品的態度愈好，也會增加讀者對產品的購買意願。范錚強等人在線上口碑對消費者購買意圖之影響－網路論壇的實驗研究(2012)中亦證實，網路論壇中產品評論的內容豐富度愈高且推薦此評論的人數愈多，資訊搜尋者就會認為此文章內容具有參考價值。因此，綜合上述，若將餐廳視作上述研究中的產品，本研究推測社群影響力與餐廳排名之間將具有一定程度的關係，因此在計算餐廳排名的過程中加入社群影響力之因素。

## 2.4.3 社群影響力結合情緒分數與排名之關係

在 2.2.1 節中提及，過去研究證實透過情緒分析將評論轉為評分可代表

用戶對於產品的評分，因此本研究的餐廳排名計算方式即是將評論透過情緒分析轉換為情緒分數，且在前一小節提及社群影響力與餐廳排名之間具有一定程度的關係，故本研究將再加入用戶影響力的因素設計排名演算法。

K. Zhang 等人的研究(2011)即是利用 Amazon 的顧客評論資料設計產品排名的模型，調整評論中各個因素的權重，為每一個產品計算排名分數，且最終模型產生的產品排名與產品實際的銷售排行類似。現有評分機制的大重點評網亦是將用戶的信譽程度作為用戶評分的權重調整依據，此外，網路溫度計(2020)的網路口碑排行榜亦是將網路聲量與網友評論的語意情緒進行綜合評比，最終計算出各領域的口碑排行榜。因此，綜合上述，本研究將社群影響力與情緒分數結合，將評論透過情緒分析轉換為情緒分數後，再將評論者的影響力作為評分的權重調整依據，計算出餐廳的排名分數，最終產生餐廳的排行結果。

## 2.5 小結

回顧此章節各類別的文獻探討，在 2.1 節分別對目前現有的評分機制，包括 Google Maps 評分、Yelp、大眾點評網的發展現況做摘要介紹，並進一步比較各評分機制提供的評鑑指標以及仍待解決的計分問題。在 2.2 節指出雖然較少針對餐飲評論的相關研究，但過去文獻除了利用情緒分析技術將評論進行情緒極性分類之外，仍有研究利用情緒分析於評論中，建立新的評分機制；此外，針對假評論問題的研究，大部分就文字探勘、行為分析方面著手，但漸漸地亦有其他研究利用不同方式，如網路效應等方式偵測假評論；針對評分匯總（rating aggregation）問題，多數研究會就評論的資訊含量或是使用者因素給予評分不同權重，線上評分機制(Yelp、大眾評分網)亦根據使用者因素調整使用者的評分權重；至於評論與評分不符的問題則有研究利用評論的情緒分數

調整原始評分，讓評分與評論更趨於一致。在 2.3 節中，則是探討 Instagram 社群對於排名計算可能的影響，包括 Instagram 用戶影響力、Instagram 餐廳帳號所獲得的主題標籤數、打卡數。而 2.4 節則是依據過去研究與現有評分機制探討影響力與排名之間的關係。

根據以上章節所盤整的過去文獻中可觀察到，據我們所知尚未有人利用 Instagram 影響力當權重建立餐飲評分機制，並排除潛藏的長久累計評分問題，即時反映消費者意見，以及假帳號“灌水”、評論與評分不一致的問題，使得排名結果更符合 menu 美食誌所調查的餐廳排行，同時考慮 Instagram 的打卡數、主題標籤(Hashtag)數、加上時間因素，測試是否能更精準預測排名，以提供使用者快速選擇餐廳的依據，以下章節將針對如何利用 Instagram 影響力與情緒分析技術建立餐飲評分機制的研究方法進一步探討。

## 第三章、研究方法

### 3.1 研究問題

憑藉著網路社群媒體的發達，愈來愈多人希望透過其他人在網路上分享的意見，了解一間餐廳的各種資訊，包括知名度、食物味道、服務品質、環境品質…等面向，作為選擇餐廳的依據，而除了參考部落格對於某間餐廳的評論之外，最快的方式莫過於參考已經整合眾多意見的評分機制或是餐廳排名，然而現存的評分機制卻仍存在一些問題，如：假評論、評分無法即時反應最新意見、評分與評論不一致…等現象，導致參考者無法得到關於餐廳最正確的資訊，而過去許多研究指出可根據評論者因素或是時間因素調整評分的權重解決上述問題，因此本研究欲結合情緒分析方法將評論轉為評分，與 Instagram 使用者的影響力作為評分的權重，計算出餐廳排名，解決現有評分機制無法提供最符合大眾心中餐廳排名之問題，此外，亦利用機器學習分類演算法，考慮進評論者的個人資訊、餐廳打卡數(Check-in)和主題標籤數(Hashtag)等資訊，提供更多資訊給機器參考，預測餐廳的排名，並在第四章研究結果中比較兩方法的排名預測準確度。

### 3.2 研究架構

本實驗分為三個階段，分別為：(1)資料蒐集、(2) 線性規劃之評分模型與機器學習分類演算法的排名預測模型之實作、(3)模型驗證與排名結果分析。整體的實驗流程架構如圖 7 所示，第一階段為圖 7 中的步驟一先進行資料蒐集，根據 MENU 美食誌的各餐飲種類的商家排名調查(詳見附錄一)，蒐集網友推薦的前十名店家名稱，於 Instagram 擷取指定時段的店家評論，作為本研究餐飲評分機制模型的訓練資料來源，並取得評論者的相關用戶資訊，進行第二階段的步驟二至六。在步驟二中，整理出各評論者的發文數、粉絲數、追蹤中帳戶數

量...等個人資訊，接著執行步驟三參考 Ling Hang Yew 等人(2018)、陳佳雯(2017)的影響力計算方式，以直觀而言將各篇貼文的按讚數作為每個評論者的貼文影響力，結合由步驟四、五計算出每則評論的情緒分數，步驟六將情緒分數依據影響力進行加權，即為改善後的評分。最後第三階段模型驗證與排名結果分析將在 3.3 節研究驗證說明，驗證排除假評論、長久累積評分以及評分與評論不符之問題的可能之後，是否可計算出趨近於 menu 美食誌的餐廳排名，意即新排名能夠符合大眾的喜好，以及步驟九利用十折交叉驗證和加入其他知名店家再選出前十名，兩種方式驗證線性規劃模型的準確度。另外，本研究預測排名的另一方法機器學習分類法，即為步驟七結合評論的情緒詞彙數量、店家打卡數、主題標籤數、評論者資訊進行排名預測。最後比較線性規劃法與機器學習分類演算法的排名預測準確度。

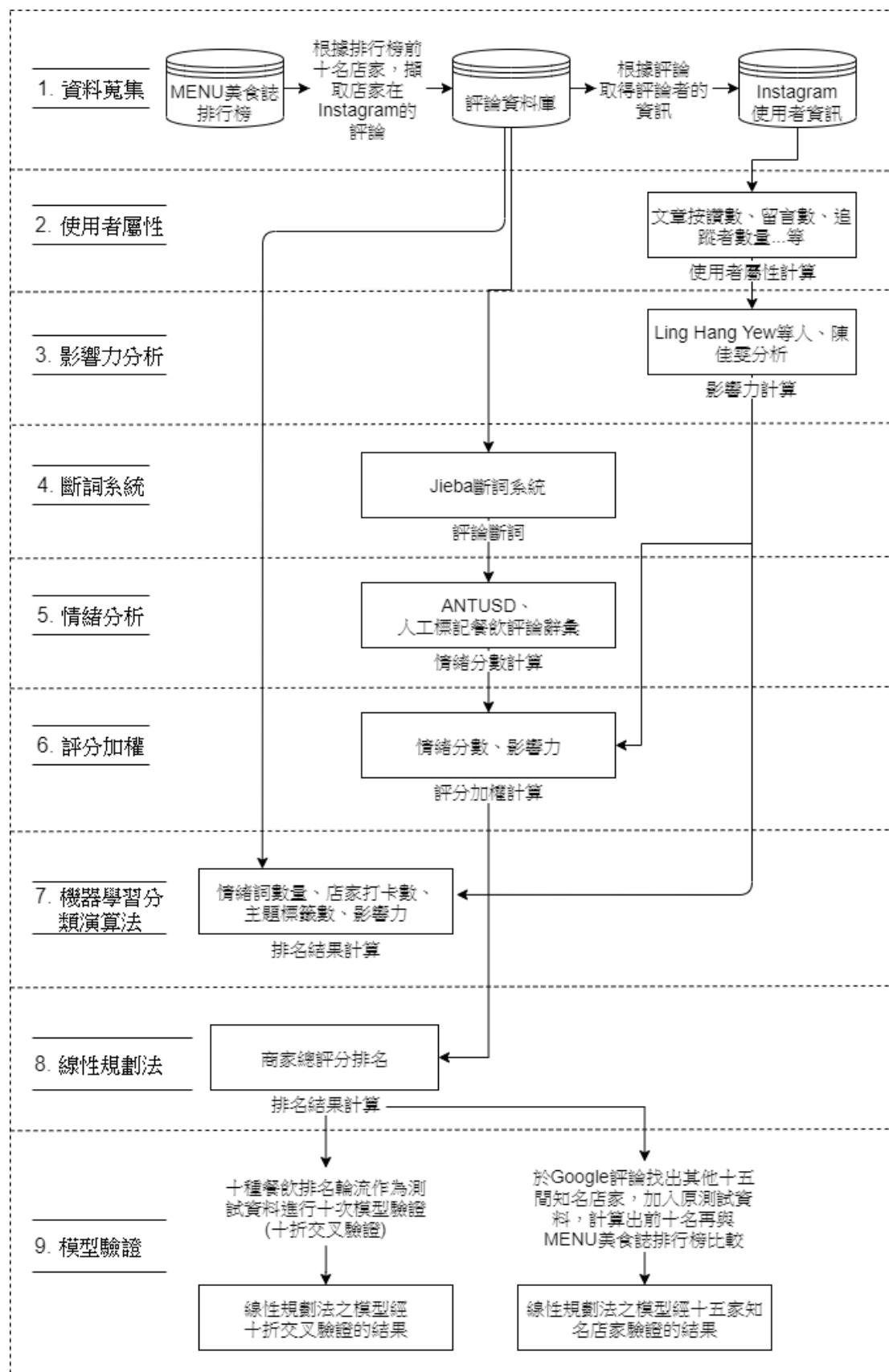


圖 7.研究架構圖

### 3.2.1 資料蒐集

#### 3.2.1.1 店家選取

本研究欲針對 Instagram 店家的評論進行情緒分析計算，根據 MENU 美食誌的排名計算方式(如表 2)，選出十種近兩年的 Top10 餐飲排行榜(圖 8 & 附錄一)，包括義大利麵、牛肉麵、滷肉飯、咖哩料理、港式餐廳、韓式炸雞、韓式料理、牛排、鐵板燒、冰品，作為本研究的店家評論分析資料來源。



圖 8. MENU 美食誌牛肉麵排行榜前三名範例(2019)

➤ 排名標準	於某一時間區段蒐集 MENU 美食誌 App 內的用戶打卡數，打卡數愈多的店家，排名愈前面，以此標準排出第一名至第十名的商家。依照不同餐飲主題，平均獲得的打卡數有所不同。
--------	---

表 2. MENU 美食誌的排行計算方式(2019, 2020)

#### 3.2.1.1 Instagram 評論者資訊擷取

在 Instagram 社群中可分為一般帳戶以及商業帳戶，一般帳戶可於自己的個人資訊頁面(如圖 9)中查看總共發布過幾篇貼文、有幾位追蹤

者(粉絲數)、正在追蹤其他帳戶的追蹤中數量、個人自我簡介、精選動態(將影片匯集成一個短影片)以及曾經發布過的貼文資訊或是被別人標記的貼文資訊，至於商業帳戶則是需要另外申請將一般帳戶轉換為商業帳戶，相較一般帳戶新增查看洞察報告的功能，一般帳戶的貼文僅可看見多少人按讚過此則貼文(若為影片則為觀看數量)、貼文內文、留言內文、留言數量(如圖 10)，商業帳戶的洞察報告則可以觀察到每篇貼文的按讚數、留言數、分享數、多少人收藏這則貼文的收藏數、多少人因為觀看這則貼文而點進個人資訊頁面查看的商業檔案瀏覽次數、多少不重複的 Instagram 用戶看見此則貼文的觸及人數(如圖 11)，然而因為商業帳戶的數據僅限發文者本人觀看，因此無法作為本研究的資料分析來源。

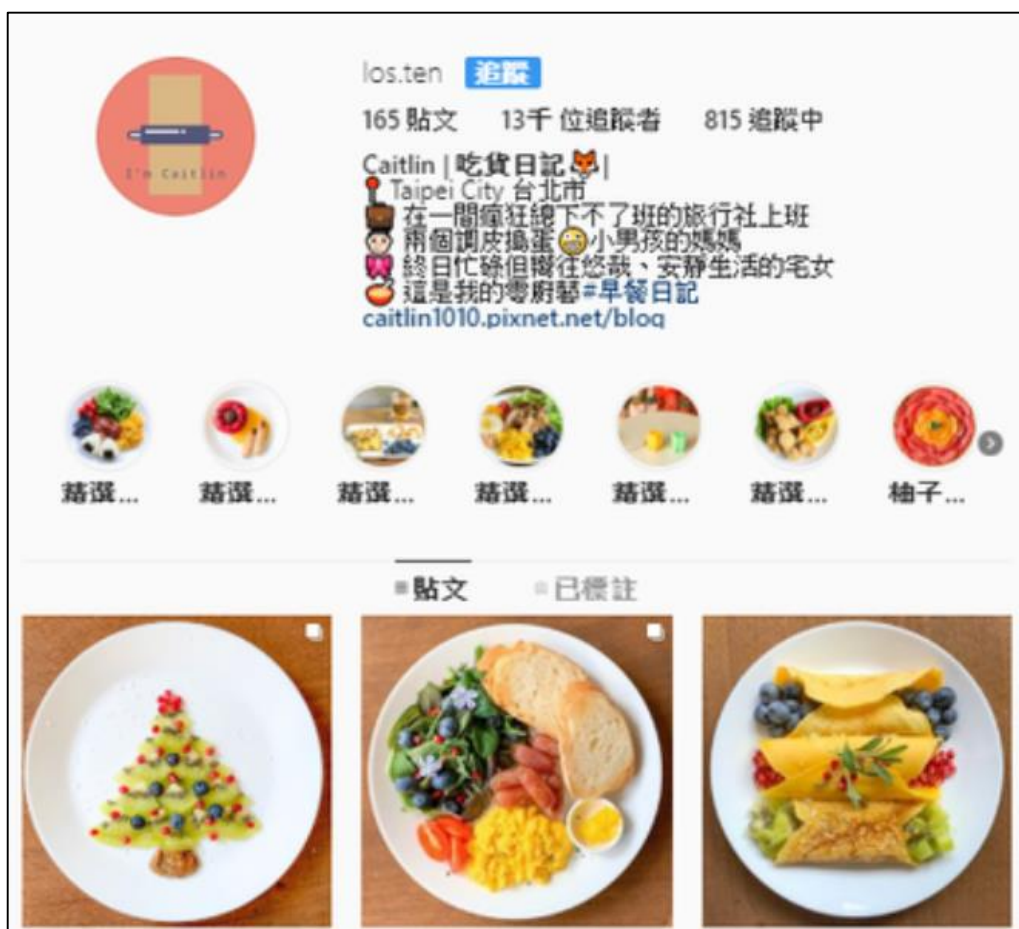


圖 9. Instagram 個人資訊頁面(Caitlin,2019)





圖 10. 貼文資訊(Caitlin,2019)



圖 11. 貼文洞察報告(Caitlin,2019)

因此本研究依據網路爬蟲套件 selenium 以及瀏覽器擴充套件 webscraper 爬取評論者貼文內容，以及評論者個人資訊(表 3.)，利用這些數值作為線性規劃法之餐飲評分機制的權重以及機器學習分類法的特徵值。

評論者個人資訊			貼文資訊		
1	Account	Instagram 帳號	1	Contents	內文
2	Posts	貼文數	2	Likes	按讚數
3	Followers	粉絲數			
4	Following	追蹤中數量			

表 3. 本研究擷取評論者與貼文資訊

### 3.2.2 線性規劃法與機器學習分類演算法

本研究提出的餐飲評分機制，分為線性規劃法與機器學習分類演算法。在文獻探討中提及社群影響力結合情緒分數對於排名預測存在一定程度上之關係，因此本研究嘗試先以現有情緒辭典 antusd 辭典(Shih-Ming, 2016)、emoji 辭典(Kralj Novak 等人, 2015)，透過影響力權重調整情緒分析後的評分計算排名，測試排名結果是否擬合 menu 美食誌排名，再以線性規劃法結合現有情緒辭典 antusd 辭典(Shih-Ming, 2016)、emoji 辭典(Kralj Novak 等人, 2015)以及本研究人工標記情緒詞彙，設計排名計算公式後測試排名結果。機器學習分類演算法則是以同為 Instagram 資料的評論者個人資料與貼文資訊作為特徵值，預測餐廳排名。因此此小節將先在 3.2.2.1 討論影響力權重制定方式、3.2.2.2 說明情緒分析方法如何將用戶評論轉為用戶評分、3.2.2.3 說明線性規劃法的排名計算過程，接著在 3.2.2.4 敘述機器學習分類演算法的排名預測過程。

#### 3.2.2.1 影響力權重

本研究參考 Ling Hang Yew 等人(2018)、陳佳雯(2017)的影響力計算方式之後，由於 Ling Hang Yew 等人(2018)的 Instagram 用戶影響力計算方式需考慮用戶的貼文觸及數(Reach)，然而本研究無法取得此觸及數(Reach)，因觸及數只有用戶自己將帳戶轉為商業帳號後才可自行

查看，並不公開給其他人觀看(2020, INSTAGRAM)，無法於本研究中實作，因此本研究以直觀而言，按讚數可直接代表一篇貼文被多少人審閱過，而被正面影響表達贊同之意，因此本研究將按讚數視為影響力大小的代表，作為後續情緒分數的調整權重。

### 3.2.2.2 評分計算

在過去的情緒分析方法中程序大致相同，而本研究將參考 Kaviya 等人(2017)的研究將評論轉成情緒分數，以下為 Kaviya 等人(2017)的計算過程：

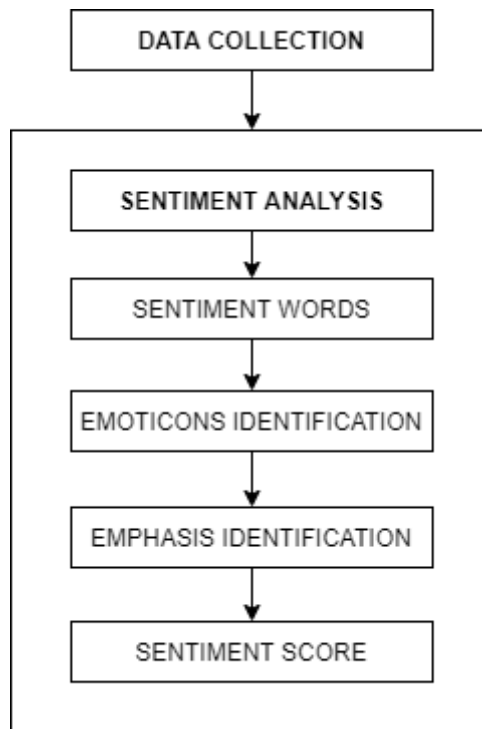


圖 12. 情緒分數計算流程 (Kaviya 等人(2017))

Kaviya 等人(2017)依照字典法選取情緒詞，再依照情緒詞的情緒分數(如圖 13)加總，搭配情緒程度副詞加權情緒分數，最終每則評論會產生一個情緒分數(如圖 14)，而依據游和正、黃挺豪、陳信希的研究(2012)中提及，部分詞彙會因為領域的不同而由非情緒詞彙轉變成情緒詞彙，例如，”鼎泰豐”詞彙原本不帶有情緒程度，但是在餐飲領

域卻可能帶有正面情緒，因此本研究將會測試是否額外加入人工標記的餐飲領域詞彙，可以更加準確的預測評論的情緒分數。

<i>Term</i>	<i>Score</i>
Positive terms	
Great	3.1
Good	1.9
Negative terms	
Bad	-2.5
Negative	-2.7

圖 13. 情緒詞的情緒分數(Kaviya 等人(2017))

<i>Sentence</i>	<i>Score</i>
AWESOME WAFFLES. LOVED IT !!!!!!!	0.9124
Really GOOD. BEST ambience !!!!	0.9063

圖 14. 評論的情緒分數(Kaviya 等人(2017))

具備評論的情緒分數以及評論者的影響力之後，即可依照評論者的影響力調整評論者的評分權重，實作結果將於第四章詳細描述。

### 3.2.2.3 線性規劃法

本研究欲先嘗試僅利用兩個現有情緒辭典進行情緒分析法搭配權重概念調整評分，測試排名結果是否趨近於 menu 美食誌排行榜，然而考慮到單一則評論情緒分數加總可能過大，故亦測試正規化後的結果，且考慮評論數過多可能會造成總評分過高，亦需測試不同平均方式，而僅利用情緒分析方法搭配權重計算出店家總評分的公式如下：

$$\text{店家總評分} = \left( \sum_i^n \left( \left( \text{各情緒詞彙的情緒分數加總} \right)_i * \left( \text{評論按讚數} \right)_i \right) \right) \div n$$

註：n: 評論數量, i: 評論 ID

公式一、情緒分析結合權重

再嘗試以線性規劃法，結合兩個現有情緒辭典與人工標記情緒詞彙，找出彼此結合可讓排名預測效果最好的最佳參數組合，而因為 antusd 辭典(Shih-Ming, 2016)之情緒詞彙的情緒分數數值區間落於[-1,1]、emoji 辭典(Kralj Novak 等人, 2015)又是以另一種方式計算出情緒詞的情緒分數，以及本研究人工標記情緒詞嘗試將其情緒分數數值區間設置於[-10,10]，擴增情緒詞之間的情緒程度差異，故利用線性規劃方式結合三種不同型態的情緒詞之情緒分數，計算店家總評分，其公式如下：

$$\text{店家總評分} = \left( \sum_i^n ((w_1 * \text{antusd 辭典} + w_2 * \text{emoji 辭典} + w_3 * \text{本研究標記情緒詞})_i * (\text{評論按讚數})_i) \right) \div n$$

註：n: 評論數量，i: 評論 ID， $w_1, w_2, w_3$ : 模型參數

#### 公式二、線性規劃法

最後依據各間店家的總評分進行排序，計算排名結果，而線性規劃的目標函式即是找到一組最佳模型參數，使各店家的預測排名與 menu 美食誌實際排名差距最小化。

#### 3.2.2.4 機器學習分類演算法

根據過去研究可用於預測排名的演算法，包括深度學習模型 tf-ranking、分類演算法：線性回歸、決策樹、隨機森林等，其中分類演算法曾被廣州帷策智能科技有限公司執行長江穎，實作於第五屆《歌手》電視節目的決賽排名預測(江穎, 2017)八名預測結果中冠軍預測正確，四位排名預測誤差一名，其餘預測誤差皆在四名內，而三種分類演算法中又以隨機森林的預測準確率最高。因此，本研究將參考江穎

(2017)的做法，利用隨機森林演算法預測餐廳排名。

在下一節研究驗證中，將說明採用哪些驗證指標驗證本研究餐飲評分機制，包括線性規劃法以及機器學習分類法的實驗排名結果，而關於詳細的實驗結果將於下一章深入探討。

### 3.3 研究驗證

為了驗證利用本研究餐飲評分機制計算出來的排名是否能趨向 MENU 美食誌的排名，意即符合大眾心中的排名，本研究採用均方誤差 MSE(Mean Square Error)(公式一)與一致序列對比率評分 FCP(Fraction of Concordant Pairs) (公式二)，計算餐飲評分機制的新排名與 MENU 美食誌排名之間的 MSE 與 FCP，MSE 代表新排名與標準排名之間的距離，故值趨近於 0(預測完全正確)，代表排名預測愈準確；而 FCP 代表相對排名正確的組數比例，組數愈多，即值愈大愈趨近於 1(預測完全正確)，代表排名預測愈準確。若單純以 MSE 計算排名的準確度無法完全判定排名演算法的好壞，因 MSE 可能因為少數排名的偏離值導致 MSE 過大(如：僅第一名預測錯誤為第十名，第十名預測錯誤為第一名，但其餘排名預測正確，則 MSE=16.2，較 baseline=16 差)，但在此情況下，FCP 仍有 0.62(較 baseline=0.5 優)，因此本研究加入 FCP 指標幫助辨識排名演算法的準確度。

$$\text{MSE} = \sum_{i=1}^n (y_i - y_i^p)^2$$

註： $y_i$ : 實際值,  $y_i^p$ : 預測值, n: 預測排名的數目

公式一、均方誤差(Mean Square Error)

$$n_c^u = |\{(i, j) | (\hat{r}_{u,i} > \hat{r}_{u,j} \ \& \ r_{u,i} > r_{u,j}) \text{ or } (\hat{r}_{u,i} == \hat{r}_{u,j} \ \& \ r_{u,i} == r_{u,j})\}|$$

$$n_d^u = |\{(i, j) | (\hat{r}_{u,i} - \hat{r}_{u,j})(r_{u,i} - r_{u,j}) \leq 0 \ \& \ i > j\}|$$

$$n_c = \sum_u n_c^u \quad n_d = \sum_u n_d^u$$

$$\text{FCP} = \frac{n_c}{n_c + n_d}$$

註：u: 排名各名次的所有pair組合, i, j: 預測排名的數目

$n_c^u$ : 相對排名一致的pair組合,  $n_d^u$ : 相對排名不一致的pair組合

$\hat{r}_{u,i}$ : 第i個項目的預測排名,  $r_{u,i}$ : 第i個項目的實際排名

$n_c$ : 相對排名一致的pair數量,  $n_d$ : 相對排名不一致的pair數量

#### 公式二、一致序列對比率評分(Fraction of Concordant Pairs)

此外，機器學習分類演算法的排名驗證方式，除了將預測排名結果與實際排名計算 MSE 之外，也採用 Macro Average，將第一名至第十名視為第一類至第十類別，分別計算第一類別至第十類別的精度(precision)、召回率(recall)、F1 分數，將十類別的精度(precision)、召回率(recall)、F1 分數個別平均後，意即計算十次 F1/precision/recall，再分別取平均，即為 Macro Average 下的精度(precision)、召回率(recall)、F1 分數。(Rachid, 2018)

## 第四章、研究結果

本章將逐步說明利用線性規劃法和機器學習分類演算法預測排名的實作過程，並進一步分析與探討造成各式餐飲種類排名差異的可能原因，在第一節，將對抓取的資料進行初步分析介紹，在第二節敘述線性規劃法和機器學習分類法預測排名的過程與排名結果驗證，在第三節中，針對不同餐飲種類的排名結果進行排名差異分析，最後於第四節中進行本章小結。

### 4.1 資料爬取與分析

為了爬取 Instagram 用戶對商家的評論與用戶個人資訊，首先，根據欲分析商家的打卡地點，蒐集用戶在此打卡地點的文章，利用網路爬蟲套件 Selenium 先行爬取此打卡地點的文章網址(如範例圖 15)，存於 HackMD(2020)，再利用 google chrome 擴充套件 webscraper 前往各網址內爬取貼文文章與用戶個人資訊，然而，考慮到時間因素，目標是從過去的文章預測 MENU 美食誌的商家排名，因此爬取的貼文時間，限制於 MENU 美食誌公開發布商家排名之日期(詳見附錄一)前的 Instagram 文章，最後爬取的範例結果呈現於圖 16 中。



圖 15. Instagram 用戶於富宏牛肉麵打卡的文章網址



1	article link	text	user info page	post count	fans count	follow others count	likes
2	https://www.instagram.com/	Best Beef Noodle Soup of Taiwan, walking distance of Taipei	bier.sommelier	26 貼文	121	304	26
3	https://www.instagram.com/	夜寒くて暖かいもの食べたかったので、 宿近くの富宏牛肉麵へ🍜 #台湾好きな人と繋がりたい 餓了～	shining_610	218 貼文	638	660	103
4	https://www.instagram.com/	每個禮拜必報到☺☺ #氣球人生 #我有蜜大腿	mickey_zeng77	856 貼文	925	395	78
5	https://www.instagram.com/	女朋友帶我去吃牛肉麵。很好吃，再來一碗！	kinguup	158 貼文	43	134	4
6	https://www.instagram.com/	哥吃的是回憶， 姐吃的是初めまして。 #lanmihsiehseatlife #富宏牛肉麵 #大學林先生很常來吃 #牛雜麵90元 #taipei #beefsoupnoodle	carolismehere	1981 貼文	236	359	12
7	https://www.instagram.com/	下午四點半想說睡個午覺 結果一個睡死 睡醒就在這了... #豪便宜啊 #加個牛油94👉	gn01495788	380 貼文	254	68	6
8	https://www.instagram.com/	Again ☺	shunshuncc	521 貼文	1518	600	46
9	https://www.instagram.com/	吃一半才想到要拍照，打卡才發現是原「建宏牛肉麵	min.yung	1081 貼文	386	1557	25

圖 16. Instagram 用戶對於富宏牛肉麵的評論與個人資訊

有鑑於 Dai 等人(2018)研究中提及評分匯總（rating aggregation）問題，長久累計的評分會導致評分結果無法反映店家的真正品質，因此本研究爬取的貼文時間，僅限 MENU 美食誌公開發布商家排名的前數個月內之文章，且因為本研究之目的為預測 MENU 美食誌的排行榜，故僅考慮 MENU 美食誌發布排行榜之前的 Instagram 用戶評論資料。

然而，如何挑選前數個月內之文章，本研究嘗試三種方式，第一種方法為將挑選評論的時間區段(time window)設為一致六個月並蒐集六個月內的全部評論，根據各店家於每月獲得的文章數量分布(圖 17)來看，每個店家六個月內獲得的評論總數不一，甚至相差極大，有些店家僅獲得 100 篇內的評論，有些店家卻獲得 600 篇以上的評論數量；第二種方法則是將挑選評論的時間區段(time window)設為一致六個月，但每間店家隨機挑選固定數量的評論；第三種方式一樣是將挑選評論的時間區段(time window)設為一致，但依時間順序挑選最接近 menu 美食誌公布排行榜之時間點前的數篇評論，且各店家挑選固定數量的評論。最終測試結果以第三種方法的排名結果較為準確，測試方式則是依據 4.2.1.2 節中僅使用台大擴增情緒辭典 antusd (Shih-Ming, 2016)的方式進行前測。

然而，進一步研究發現，第三種方式中，部分店家於六個月內所蒐集到的

評論數量僅在 50 篇以內，若嘗試將蒐集評論的時間區段擴大為一年，增加各店家的評論數量進行排名預測，將會增加排名的預測準確度，而各店家如何挑選固定數量的評論，本研究分別測試 50 篇、100 篇、150 篇、200 篇左右的排名結果，以 150 篇左右的排名結果較準，因此以 150 篇為基準繼續進行後續的實驗。

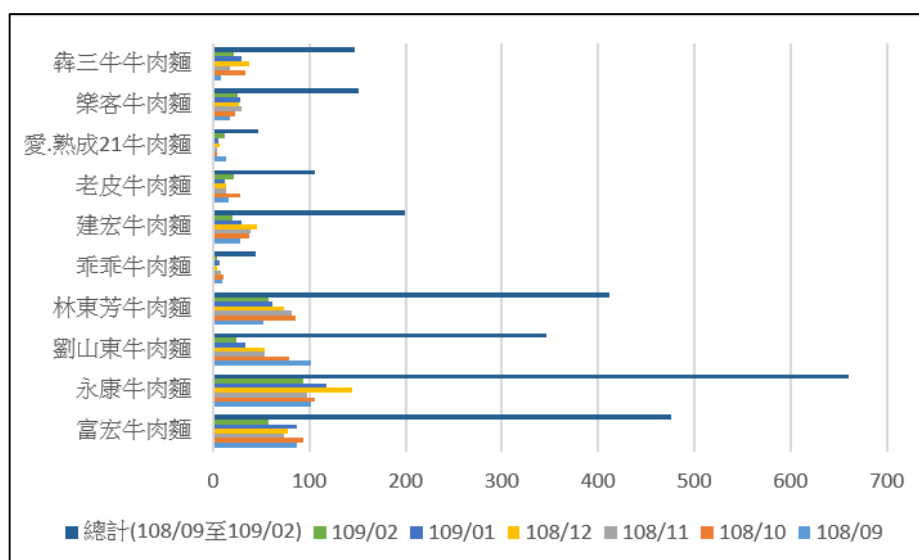


圖 17. 各店家於每月獲得的文章數量分布(以牛肉麵六個月為例)

此外，亦可選擇根據標記有商家名稱的 hashtag 蒐集商家的貼文，但是由圖 18(左)可看出標有 hashtag 的貼文內容較容易偏離對於商家的評論，因其可能包含其他主題的 hashtag，意即文章內容並非以評論餐廳為主軸，甚至完全沒有提及餐廳評論相關的字詞，而位於商家地標打卡的貼文則大部分會談論關於餐廳的相關評論(如圖 18(右))，因此本研究選擇用戶的打卡貼文，作為排名預測的資料來源。



圖 18. 標記有店名的 hashtag 貼文(左)對照打卡貼文(右)(以林東芳牛肉麵為例)

然而，並非每一種餐飲種類的資料都適合進行排名預測，以手搖飲為例，因為本研究預計從打卡貼文取得評論內容，但經過觀察 Instagram 上的手搖飲打卡貼文，不僅評論少且雜訊多，部分貼文內容都非跟評論相關，且手搖飲的分店過多，資訊過於分散，不易取得，而若是提及評論相關的內容，也大部分都屬於外文，可推測為外國觀光客所發布，如圖 19 以五十嵐為例，可見連續十一天的貼文內容都是外文，因此不適合本研究以中文為分析目標的排名預測，因此本研究從 MENU 美食誌已發布的餐飲排行榜中，最後整理出 2019 年之後的十種排行榜，包括義大利麵、牛肉麵、滷肉飯、咖哩料理、港式餐廳、韓式炸雞、韓式料理、牛排、鐵板燒、冰品(詳見附錄一)。

1	text	user_info	post_count	fans_count	follow_others	likes
2	50 Lan ร้านชามกราคาถูกอีกแล้ว 40-50 บาท เองว่า ละคือเป็นชามกรเนื้	tdredline	343 貼文	43	5	null
3	You know you ' re surrounded by people that care about you, when they give into your need for food photos. ☺ #love #food #foodporn #foodphotography #yum #yummy #nomnom #foodie #yvr #vancouver #vancity #dailyhive #vancitybuzz #eeeeeats #instafood #foodstagram #foodgasm #hungry #buzzfeedfoods #instagood #feedfeed #f4f #foodiesofinstagram #taiwan #bubbletea	melceats	196 貼文	264	422	23
4	年末年始は台湾で過ごします。 早速タピオカミルクティーの50嵐へ。毎年海外にいくと、暑い、食べ物辛いなど不評もちらほらあるけど、既然大満足 ^ ^ ☆	taisou_s_c	17 貼文	48	55	16
5	#波霸奶茶 #bigbubblemilktea #50嵐 #koihe	kaeoooo_	147 貼文	7	8	7
6	#台湾 #台北 リビタピ〇〇大粒ホットミルクティ 甘さ70%。タピオカっておいしいんですね〇気づくの遅かったです。40元(約140円) #台湾 #台北美食 #タピオカミルクティー #タピオカ #50嵐	koutan36	460 貼文	116	108	19
7	A little sunshine for the first day of 2020 #mùa mưa tại tỉnh bản	nammeob	782 貼文	1217	574	18
8	Ôi hoàng tử xin người hãy tha thứ cho những người em gái bị dính lò	tiênphuc_	679 貼文	6698	867	134
9	#手デブ #タピ活 #邪道	sakatter	4244 貼文	1016	5669	33
10	やっぱり現地はタピオカの味も量も違うよねー。 #台湾旅行 #台北旅行 #タピオカ #50嵐 #台北グルメ	gammmmm	635 貼文	123	94	26
11	ไปอ่านเจอข้อมูลนี้มาว่า ชามมไข่มุกเจ้าเนี้ย "อู่สือหลัน (50嵐)" ก็คือแบรนด์เดียวกับ koihe แหะละ แต่เป็นชื่อที่ขายในประเทศเท่านั้น แต่ส่วนตัวคิดว่ารสชาติชามม, รสชาติไข่มุกมันไม่เห็นจะเหมือนกันเท่าไรเลย ราคาก็ต่างกันเยอะ ☺ #pearlmilktea #珍珠奶茶	nut_seed	756 貼文	488	634	20

圖 19. 五十嵐手搖飲打卡貼文(108.12.29 至 109.1.8)

## 4.2 排名預測與驗證

本節將分別敘述實作線性規劃之餐飲評分機制和機器學習分類演算法的過程與排名結果驗證。

### 4.2.1 線性規劃之餐飲評分機制

首先，先將蒐集到的評論資料進行資料前處理後，以 3.2.2.3 節的公式一測試僅利用兩個現有情緒辭典進行情緒分析法搭配權重概念調整評分後的排名結果，並測試正規化情緒分數以及各種調降排名 mse 的方式，再以線性規劃方式結合兩個情緒辭典(antusd 辭典(Shih-Ming, 2016)、emoji 辭典(Kralj Novak 等人, 2015))與本研究人工標記之情緒詞彙，利用 3.2.2.3 節的

公式二找出讓排名 mse 最小化的參數組合，之後將此參數組合套用於測試資料驗證模型準確度。

#### 4.2.1.1 資料前處理

本研究抓取的資料如下圖 20，每一列為一位 Instagram 用戶的貼文，其中 text 為評論內容、user info 為使用者 id、post count 為此使用者曾經發布在 Instagram 上的總發文數、fans count 為使用者擁有的粉絲數量、follow others count 為使用者追蹤其他使用者的數量、likes 為此則評論貼文獲得其他用戶按讚的總按讚數(此用戶對於其他用戶按讚的數量因 Instagram 官方未公布，故在本研究中則不討論)，此外本研究分析的資料量如下表 4 所示，表中列出十種餐飲種類各名次商家(menu 美食誌排名)所收集的 Instagram 評論數量以及各種類的評論總數，依據 4.1 節前測的結果，擷取 menu 美食誌公布排行榜之時間點前一年的評論，各店家收集至多 150 篇的評論。

1	text	user_info	post_count	fans_count	follow_others_count	likes
1	意外找到一間網美餐廳 來到這人人都是網美 網美聖地 每個角落都很好拍 餐點也好吃 玩到被店員提醒餐點涼掉就不好吃囉 #台中美食#taichung#jai宅#吃貨動起來#食 在好好吃#非吃不可#美好食光#潔妮 日 常#台中景點#網美景點	jenny_rui	181 貼文	364	1231	27
2	第一次來是粉色頭髮 第二次來還是粉色 他們的青醬爆炸蛤蜊義大利麵真的推 #網美當起來	muamuac	1355 貼文	741	613	32
3	・橘醬爆炸蛤蜊義大利麵 ・明太子炭烤雞肉串 ・黑松露爆炸蛤蜊飯（加燻烤） ・JA現炸脆薯 義大利麵超級好吃 那個醬太猛了 吃完還會想再點一份 蛤蜊也真的爆炸多的 雞肉串是店員推薦的 很令人驚艷的雞肉串 咬下去 裡面超級超級嫩 又有滿滿的肉汁 加上外面的明太子 太好 吃了 燉飯跟薯條 我就覺得普通了 來這裡點義大利麵就對了 而且他號稱台中最美的餐廳 裡面空間很大 各個地方都有精心佈置 絕 對是會回訪的店	abby.0512	155 貼文	136	297	21
4	不過幾歲都要開心健康快樂 🍰🎂🍷🍷 #happybirthdaytome #thanks #店員送我 🍷🍷	yenxin7	1174 貼文	1033	1243	106
5						

圖 20. 本研究欲分析的評論資料範例

	評論總數	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10
義大利麵	1281	150	109	150	103	87	85	147	150	150	150
牛肉麵	1228	147	150	150	150	101	92	116	90	82	150
滷肉飯	1271	150	150	150	111	116	119	90	150	123	112
咖哩料理	1219	132	104	132	150	122	108	132	138	76	125
港式餐廳	1289	150	150	150	124	127	126	134	87	146	95
韓式炸雞	1370	150	129	145	150	70	147	150	129	150	150
鐵板燒	1376	150	150	150	144	132	60	140	150	150	150
牛排	1481	137	148	150	150	150	150	150	150	146	150
冰品	1500	150	150	150	150	150	150	150	150	150	150

韓式料理	1321	150	145	150	150	150	96	92	147	150	91
------	------	-----	-----	-----	-----	-----	----	----	-----	-----	----

表 4. 本研究蒐集的各名次商家 Instagram 評論數量

在文字資料前處理的階段，通常會考慮到重覆值、空值、標點符號、雜訊以及停用詞問題。本研究爬取的資料在利用瀏覽器擴充套件 web scraper 爬取時，會直接將同一個 Instagram 用戶的貼文視為重覆資料，因而只會爬取到不重覆用戶的貼文，故爬取的資料不存在重覆的問題。但是用戶打卡的貼文可能不存在任何留言內容，僅留下照片，故必須移除此類空值(Nan)資料，並且移除與商家用餐評論完全不相關的貼文，本研究發現在 Instagram 社群中，美髮業者常會透過在商家地址打卡並發布廣告貼文，因此利用關鍵字(如：染髮、燙髮、美甲等...美髮相關詞彙)篩選出這些貼文並移除，另外，在此階段不需特別移除停用詞(stop words)，因為在下個階段計算評論分數時，可以直接依據情緒詞彙計算整篇文章的情緒分數，以達到排名計算的目標。

#### 4.2.1.2 計算商家總評分與排名

本研究先將評論透過 jieba 斷詞系統斷詞後，對照情緒辭典中的情緒詞與對應的情緒分數，挑出評論中的所有的情緒詞彙，並將這些情緒詞的情緒分數加總，即為此則評論的情緒分數，計算過程如下圖 21，圖中 comment 為評論，評論下方為各情緒詞與對應的情緒分數，sent\_score 為各則評論的情緒總分。

[0]comment: 這個時間最適合放食物照了🤔🤔, #MINT #學生最愛 #平價美食 , #米蘭茄汁海鮮義大利麵 🍝 175元, 海鮮份量頗多, 義大利麵本人也蠻入味, 重點是CP值很高啊啊🤔🤔, , #台北美食 #台灣 #義大利麵 #PASTA #LIFESTYLE #FOODLOVER	
時間	0.0550611
適合	0.775
🤔	0.701
🤔	0.701
最愛	0.5801090999999999
美食	0.16
🍝	0.251
份量頗多	0.1
入味	0.0
重點	0.1391674
CP值很高	0.8
🤔	0.701
美食	0.16
sent_score:	5.1233376
[1]comment: 思思念念♥, 終於吃到你😊😊	
♥	0.657
終於吃到	0.3
😊	0.25
😊	0.25
sent_score:	1.457

圖 21. 兩則評論的情緒分數分析範例

本研究隨機取八種餐飲種類(義大利麵、牛肉麵、滷肉飯、咖哩料理、港式餐廳、韓式炸雞、牛排、鐵板燒)設為訓練資料，其餘兩種(冰品、韓式料理)設為測試資料(詳見附錄一)，利用訓練資料對 3.2.2.3 節的公式一做適當調整，並調降排名 mse，並且透過線性規劃方式找出 3.2.2.3 節的公式二中讓排名 mse 最小化的參數組合，之後將此參數組合套用於此兩種測試資料中驗證模型準確度。而調降排名 mse 的過程如下(表 5 至表 8)：

表 5、僅用台大擴增情緒辭典 antusd (情緒詞：27221)

	義大利麵	牛肉麵	滷肉飯	咖哩料理	港式餐廳	韓式炸雞	牛排	鐵板燒
MSE	17.8	23.8	19.2	16.6	18	12	25	17.6
FCP	0.467	0.311	0.467	0.511	0.489	0.6	0.289	0.467

表 6、加入 emoji 情緒分數(Kralj Novak 等人, 2015) (情緒詞：27970)

	義大利麵	牛肉麵	滷肉飯	咖哩料理	港式餐廳	韓式炸雞	牛排	鐵板燒
MSE	22.4	19.6	20.2	16.6	19.2	11	26.2	18
FCP	0.378	0.422	0.4	0.533	0.444	0.622	0.244	0.444

表 7、情緒分數與按讚數(權重)正規化

	義大利麵	牛肉麵	滷肉飯	咖哩料理	港式餐廳	韓式炸雞	牛排	鐵板燒
MSE	11.8	20.6	9.6	18.8	11	17.6	15	21.8
FCP	0.622	0.422	0.644	0.467	0.622	0.422	0.511	0.356



表 8、加入本研究人工標記情緒詞(情緒詞：28884)

	義大利麵	牛肉麵	滷肉飯	咖哩料理	港式餐廳	韓式炸雞	牛排	鐵板燒
MSE	5.4	2.2	3.8	5.6	9	5.4	5.2	7.2
FCP	0.778	0.844	0.8	0.778	0.644	0.778	0.733	0.733

若僅僅用台大擴增情緒辭典 antusd，進行 3.2.2.3 的公式一計算排名，排名結果如表 5，八種餐飲類別中有七種都較 Baseline (mse=16.4，詳見 4.2.1.3) 差，其原始 menu 美食誌排行詳見附錄一，經過觀察原始評論後，研究發現 Instagram 用戶經常使用 emoji 作為情緒表達的方式之一，因此考慮將 emoji 的情緒分數(Kralj Novak 等人, 2015) (如下圖 22) 加入 3.2.2.3 的公式一計算，排名預測效果如表 6，有五種餐飲類別的 mse 不減反增，觀察 Novak 等人研究的 emoji 情緒分數後發現，其研究共收錄 751 個 emoji 及相對應的情緒分數，但是卻有一些 emoji 直觀上不屬於餐飲評價的範疇，對於分析餐廳排名沒有幫助，甚至在 Instagram 評論中被大量使用，造成本研究的排名結果誤判第一名的店家為最後一名，使得 mse 大幅增加，例如：'❄️', '💩', '\$'... 等，故本研究人工篩選並移除這些字詞後，也確實能夠幫助校正排名的準確度，因此，使用 Novak 等人研究的 emoji 情緒詞彙的同時，亦須經過人工篩選，從排名結果判斷哪些 emoji 為對餐飲排名有幫助的情緒詞彙。













Char	Image [twemoji]	Unicode codepoint	Occurrences [5...max]	Position [0...1]	Neg [0...1]	Neut [0...1]	Pos [0...1]	Sentiment score [-1...+1]	Sentiment bar (c.i. 95%)	Unicode name	Unicode block
😭		0x1f602	14622	0.805	0.247	0.285	0.468	0.221		FACE WITH TEARS OF JOY	Emoticons
♥️		0x2764	8050	0.747	0.044	0.166	0.790	0.746		HEAVY BLACK HEART	Dingbats
♥️		0x2665	7144	0.754	0.035	0.272	0.693	0.657		BLACK HEART SUIT	Miscellaneous Symbols
😍		0x1f60d	6359	0.765	0.052	0.219	0.729	0.678		SMILING FACE WITH HEART-SHAPED EYES	Emoticons
😭		0x1f62d	5526	0.803	0.436	0.220	0.343	-0.093		LOUDLY CRYING FACE	Emoticons
😘		0x1f618	3648	0.854	0.053	0.193	0.754	0.701		FACE THROWING A KISS	Emoticons

圖 22. Novak 等人研究的 emoji 情緒分數(2015)

然而，加入 emoji 辭典後，仍有七種餐飲類別的排名結果無法趨

近 menu 美食誌排行(表 6)，且研究發現因為每則評論提到的情緒詞彙數量不一，評論內容較長的文章通常被偵測出的情緒詞彙也較多，其情緒分數總分會較其他短評論高出許多，將會有不同評論分數之間差距過大的問題，因此本研究發現若將情緒分數與按讚數(權重)正規化後，有四種餐飲類別(義大利麵、滷肉飯、港式餐廳、牛排)的 mse 有大幅下降的趨勢(如表 7)，本研究亦多方嘗試各種正規化方法(Normalization)，包括以下三種正規化方式(如下圖 23)：

方法一(norm\_score\_1)：

將數據限制於 $[-1,1]$ ，若數據大於 1，則設為 1，反之亦同，若小於 -1，則設為 -1，其餘數值不變。

方法二(norm\_score\_2)：

將數據按比例縮放到 -1 到 1 的區間。

方法三(max\_min\_normalization)：

將數據按比例縮放到 0 到 1 的區間。

測試結果以將評論分數與按讚數進行方法三的正規化後的表現最優，

因此對 3.2.2.3 節的公式一做正規化調整。

```
1 def norm_score_1(all_comments_scores):
2     '''bigger or less than 1/-1 => 1/-1'''
3     norm_all_comments_scores = []
4     for one_score in all_comments_scores:
5         if one_score >= 1:
6             norm_all_comments_scores.append(1)
7         elif one_score <= -1:
8             norm_all_comments_scores.append(-1)
9         else:
10            norm_all_comments_scores.append(one_score)
11
12    return norm_all_comments_scores
13
14 def norm_score_2(x):
15     '''max_min norm to [-1,1]'''
16     return [(float(i)-np.mean(x))/(max(x)-min(x)) for i in x]
17
18 def max_min_normalization(x):
19     '''max_min norm to [0,1]'''
20     return [(float(i)-min(x))/float(max(x)-min(x)) for i in x]
```

圖 23. 三種正規化方法

此外，本研究觀察到些許評論文章中，仍有些許情緒詞彙未被辨

識出來，因此本研究另外從 google 評論中找出更多情緒詞彙，在加入這些人工標記的情緒詞彙以及透過本研究之情緒分數計算方式之後，可發現各餐飲種類的 mse 均能夠成功調降(如表 8)，意味著計算出的排名也愈趨於 MENU 美食誌的標準排名(大眾心目中的排名)。

至於本研究從 google 評論中找出情緒詞彙的方式，是先蒐集 google 一星評論 1712 篇與五星評論 8732 篇，再將評論透過 jieba 套件斷詞後，從一星評論中人工標記出負面情緒詞彙 19218 個詞彙，五星評論中標記出正面情緒詞彙 70511 個詞彙，然而，過去研究(謝佩庭, 2014)會將詞頻(term frequency)過低的詞彙視為不具文件鑑別力，詞頻過高的詞彙視為停用詞(stop words)，而將這些詞彙先移除再人工篩選出情緒詞彙，但是本研究認為此作法仍有可能遺漏些許情緒詞彙，因此為了蒐集到更完整的情緒詞彙，將直接進行人工標記的程序。

而過去研究(謝佩庭, 2014)進行情緒分數計算時，普遍會多加檢驗情緒詞在句子中的前後是否包含程度副詞或否定詞，並調整其情緒分數，但在本研究中，則是嘗試將所有程度副詞或是否定詞合併進情緒詞中，並依據本研究情緒詞的計算方式，賦予其新的情緒分數。而本研究人工標記情緒詞之情緒分數給定方式則是，透過調整情緒詞的情緒分數使各項排名的 mse 降低，若為正面情緒詞則將情緒分數從 0 開始累計到 10，負面情緒詞則是情緒分數從 0 開始累計到負 10，觀測 mse 的變化，若平均 mse 有變小且每個餐飲種類的 mse 增加或降低的幅度不超過 0.1，則把當下的分數指定為此情緒詞彙的情緒分數，並加入人工判斷各情緒詞情緒程度，例如：賦予高強度的情緒詞彙較高的分數。

上述的排名結果都是經過 3.2.2.3 節的公式一計算商家總評分而來，依照評分大小給予適當排名，完整的商家總評分計算程序可參考

下圖：

輸入：經前處理後的評論、按讚數

輸出：商家總評分

FOR(商家所有評論)

    一則評論的評分 = 此評論中所有情緒詞的情緒分數加總

FOR(商家所有評論的按讚數)

    一則評論的按讚數 = 此評論按讚數 + 1

商家總評分 =

加總(正規化(所有評論評分) \* 正規化(所有評論的按讚數))/所有按讚數

Return 商家總評分

圖 24. 商家總評分計算程序

首先，將經前處理後的評論進行情緒分數轉換，計算出每一則評論的評分，再分析每則評論的按讚數，經研究發現些許評論的按讚數為空值，意即此評論沒有獲得任何按讚數，但是經過研究測試過後，發現這些按讚數為空值的評論若為 0 則對最終評分沒有任何貢獻，若為 1 則可以影響總評分，且有助於提升最終排名結果的準確度，因此，本研究嘗試兩種方式，分別為僅將按讚數為空值的評論賦予其值為 1，以及利用平移概念將每則評論的按讚數加 1，測試後的排名結果相差不多，相較未調整前均有所改善，故本研究採用平移概念。之後，再分別將評分與按讚數(權重)進行正規化後相乘，把所有評分加總，並且測試以所有按讚數或是以所有評論數量將總評分平均的排名效果最佳，測得所有按讚數最佳，因此最後將總評分再除以商家所有評論獲得的所有按讚數平均，即為此間店家的最終評分。

再利用線性規劃方式結合兩個情緒辭典(antusd 辭典(Shih-Ming,

2016)、emoji 辭典(Kralj Novak 等人, 2015))與本研究人工標記之情緒詞彙，利用 3.2.2.3 節的公式二找出讓排名 mse 最小化的參數組合，之後將此參數組合套用於測試資料驗證模型準確度。測試過各組參數後， $w_1, w_2, w_3$  以趨近於 1 的值進行排名預測可得最佳結果(如圖 25)，若以圖 25 左上圖  $w_1$  為例，固定另外兩個  $w_2, w_3$  參數為 1 後觀察調整  $w_1$  對八種訓練資料平均的 mse 變化，可觀察到當  $w_1$  趨近於 1 時平均 mse 為最小，而因  $w_1$  為 0 時會將情緒詞彙分數都轉為 0， $w_1$  為負值時會將正向情緒詞分數轉為負向分數，故  $w_1$  僅從 0.1 開始調整，以此方法反覆測得  $w_2, w_3$  趨近於 1 時平均 mse 為最小，因此以此組參數對未分析過評論的測試資料，以 3.2.2.3 節的公式二預測排名並測試預測準確度，由下表 9 可見測試資料冰品和韓式料理的評估指標均優於 Baseline(詳見 4.2.1.3 表 10)，而十種餐飲種類經線性規劃法預測排名的結果則可參考附錄一。

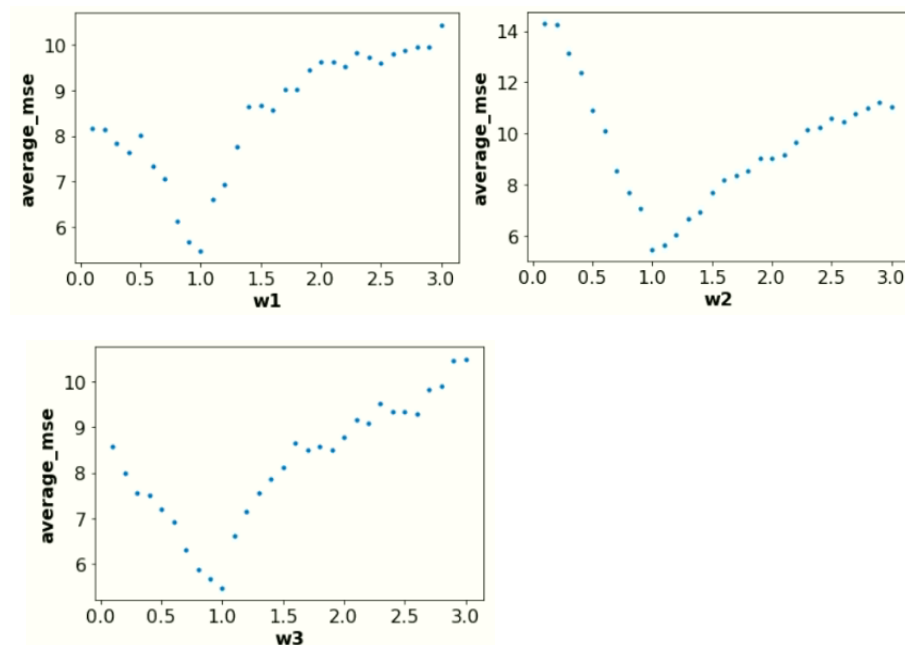


圖 25. 分別調整  $w_1, w_2, w_3$  後，平均 mse 的變化

	冰品	韓式料理
MSE	4.6	5.8
FCP	0.8	0.733

表 9、測試資料

#### 4.2.1.3 排名結果驗證

本研究為了驗證排名結果是否至少優於 baseline，本研究將一萬次的隨機排名與正確排名(第一名至第十名)計算 mse 之後，取平均得到  $16.4 \pm 0.2$ ，因此將 mse 的隨機排名 baseline 設為 16.4，另一評估指標 fcp (Fraction of Concordant Pairs)亦採用相同作法，取一萬次的平均為  $0.5 \pm 0.02$ ，因此隨機排名的 baseline 如下表：

	MSE	FCP
Baseline	16.4	0.5

表 10. 隨機排名的 baseline

由下表 11.可見線性規劃法的排名都優於 baseline，十種餐飲種類的 mse 均低於 baseline=16.4，fcp 均高於 baseline=0.5。

	Baseline(mse/fcp)	線性規劃法(mse/fcp)
義大利麵	16.4/0.5	<b>5.4/0.778</b>
牛肉麵	16.4/0.5	<b>2.2/0.844</b>
滷肉飯	16.4/0.5	<b>3.8/0.8</b>
咖哩料理	16.4/0.5	<b>5.6/0.778</b>
港式餐廳	16.4/0.5	<b>9/0.644</b>
韓式炸雞	16.4/0.5	<b>5.4/0.778</b>
牛排	16.4/0.5	<b>5.2/0.733</b>
鐵板燒	16.4/0.5	<b>7.2/0.733</b>
冰品	16.4/0.5	<b>4.6/0.8</b>
韓式料理	16.4/0.5	<b>5.8/0.733</b>

表 11. 線性規劃法之評分機制與 baseline 的比較

此外，為了讓本研究的模型更加 robust，本研究利用十折交叉驗證(10-Fold Cross Validation)進一步驗證模型，本研究共收集十種餐飲

種類的資料，所以可以輪流將每一種資料當作測試資料，其餘九種資料作為訓練資料，最後將十次的驗證結果取平均，即為十折交叉驗證的模型驗證結果。在訓練模型中不會使用到測試評論的情緒詞彙，所以在不考慮測試資料集的情緒詞彙的情況下，驗證模型是否還是能計算出趨近於 menu 美食誌的排名。總計進行十次的模型驗證，每一種資料輪流作為測試資料，其模型驗證結果與預測排名(實際排名：[1,2,3,4,5,6,7,8,9,10])如下表 12，最終十折交叉驗證的平均結果亦優於隨機猜測排名的 baseline(表 10)。

Test data	MSE	FCP	預測排名
義大利麵	6.2	0.756	[1, 2, 4, 5, 9, 6, 10, 3, 8, 7]
牛肉麵	3	0.822	[5, 1, 3, 2, 4, 7, 6, 10, 8, 9]
滷肉飯	5	0.778	[2, 3, 6, 1, 9, 4, 7, 5, 8, 10]
咖哩料理	5.6	0.778	[1, 2, 3, 5, 7, 9, 10, 4, 8, 6]
港式餐廳	9.4	0.644	[6, 1, 4, 7, 3, 10, 2, 5, 9, 8]
韓式炸雞	5.6	0.756	[2, 3, 4, 7, 6, 5, 1, 10, 8, 9]
鐵板燒	8.4	0.667	[4, 3, 5, 2, 7, 9, 1, 8, 10, 6]
牛排	5.2	0.733	[4, 1, 6, 5, 2, 3, 10, 8, 7, 9]
冰品	6.6	0.733	[1, 5, 4, 2, 9, 6, 7, 8, 3, 10]
韓式料理	6	0.711	[2, 1, 3, 9, 5, 8, 7, 6, 4, 10]
10-fold cv	<b>6.1</b>	<b>0.7378</b>	

表 12. 十折交叉驗證結果

而本研究的目的是要在眾多餐廳中，預測出符合大眾心中喜好的餐廳排行，如：menu 美食誌餐廳排行，因此除了預測出與 menu 美食誌 top10 餐廳排行榜類似的相對排名之外，應能夠在十間以上的餐廳中，選中 menu 美食誌的 top10 餐廳，且排名需與 menu 美食誌類似，故本研究進行第二種實驗驗證方式，將 menu 美食誌 top10 餐廳以外且同樣知名的其他店家加入預測排名的計算過程，而依據本研究觀察 menu 美食誌 top10 餐廳在 google 評論中通常都具有 1000 則以上的評論以及 4.0 以上的評分，故本研究以此標準(1000 則左右的評論且 4.0

以上的評分)從 google 評論中挑選知名的十五間不重複店家，與 menu 美食誌 top10 餐廳共 25 間餐廳計算排名。

因此本研究以前一小節所述隨機選取的測試資料(冰品、韓式料理)為例，進行第二種實驗驗證。冰品的驗證結果如下表 13，加入的 google 知名店家其評分與評論數量如下表 14，由結果中可發現雖然第五名、第八名、第十名被預測錯誤，由其他知名店家所取代，但是其他七名的排名預測仍能夠趨近於 menu 美食誌的排行，且研究發現本研究預測為第一名的店家，其大部分評論內容都較豐富，情緒詞使用量較多，可見其在 Instagram 上的人氣度固然很高，因此雖不在 menu 美食誌的排行之中，但在一定程度上仍屬大眾心中十分喜好的店家之一。

Menu 美食誌排名	餐廳名稱	排名	排名分數	評論數量
第 1 名	懷舊小棧 - 五妃廟口豆腐冰	2	0.36249	150
第 2 名	清水堂	3	0.35231	150
第 3 名	春美冰菓室	5	0.28869	150
第 4 名	朝日夫婦	4	0.30571	150
第 5 名	<b>Kinber 金帛手製</b>	<b>20</b>	<b>0.15646</b>	<b>150</b>
第 6 名	Lindaogebia 你家隔壁	7	0.27153	150
第 7 名	御品元冰火湯圓	8	0.25955	150
第 8 名	<b>阿宗芋冰城</b>	<b>15</b>	<b>0.2113</b>	<b>150</b>
第 9 名	有春冰菓室	6	0.28495	150
第 10 名	<b>桑原商店</b>	<b>22</b>	<b>0.14217</b>	<b>150</b>
-	<b>阿爸の芋圓</b>	<b>1</b>	<b>0.38321</b>	<b>150</b>
-	好想吃冰	18	0.17878	150
-	大方冰品	13	0.22469	137
-	冰島雪花妹	24	0.11007	81
-	建中黑砂糖刨冰	19	0.16485	139
-	<b>建華冰店</b>	<b>10</b>	<b>0.25015</b>	<b>150</b>
-	雪王冰淇淋	11	0.22982	150



-	涼心冰店(黑店冰店)	17	0.1973	148
-	<b>溝壩清涼冰店</b>	<b>9</b>	<b>0.2585</b>	<b>146</b>
-	慢食堂	14	0.22078	149
-	新建利冰店	12	0.22515	148
-	剛好水果室	16	0.20883	150
-	西門町芒菓冰	23	0.11508	150
-	船來雪淇淋	25	0.08057	140
-	豐春冰菓店	21	0.15032	150

表 13. 線性規劃法-冰品類別加入其他知名店家後的驗證結果

google 知名店家	google 評分	評論數
阿爸の芋圓	4.3	3,500
好想吃冰	4.3	2,241
大方冰品	4.3	1,264
建中黑砂糖刨冰	4.2	1,599
慢食堂	4.5	1,487
雪王冰淇淋	4.2	1,435
涼心冰店(黑店冰店)	4.4	1,551
溝壩清涼冰店	4.4	1,895
建華冰店	4.1	3,458
冰島雪花妹	4.7	1,160
新建利冰店	4.6	954
剛好水果室	4.2	1,599
西門町芒菓冰	3.9	1,533
船來雪淇淋	4.4	1,115
豐春冰菓店	4.3	2,938

表 14. google 評論上其他知名冰品店家(109.6.30)

韓式料理的驗證結果如下表 15，加入的 google 知名店家其評分與評論數量如下表 16，由結果中可發現雖然第四名、第十名被預測錯誤，由其他知名店家所取代，而第九名的預測誤差較大，但是其他七名的排名預測仍能夠趨近於 menu 美食誌的排行，預測誤差都僅在一至兩名內。

Menu 美食誌排名	餐廳名稱	排名	排名分數	評論數量
第 1 名	韓虎嘯	2	0.4328	150
第 2 名	豬對有韓式烤肉吃到飽	1	0.53897	145
第 3 名	I' m Kimchi	3	0.38206	150
第 4 名	Broccoli Beer 韓國餐酒食堂	14	0.21039	150
第 5 名	NeNe Chicken Taiwan	5	0.29905	150
第 6 名	韓石食堂	8	0.27771	96
第 7 名	奔跑吧年糕鍋	6	0.28183	92
第 8 名	米豐屋	7	0.28154	147
第 9 名	韓湘辣年糕	4	0.33791	150
第 10 名	咚咚家韓式豬肉	15	0.21036	91
-	一桶 tone 韓式新食	19	0.18745	138
-	台中高麗屋	21	0.15941	136
-	花小豬	9	0.25281	134
-	菜豚屋	13	0.22648	131
-	好好吃肉	20	0.16822	150
-	韓太閣	17	0.19137	130
-	韓鄉韓國料理	10	0.24326	136
-	囍韓燒	11	0.23502	148
-	FASHION PIG 韓式熟成五花肉	24	0.13326	150
-	GangNam 首爾韓鍋	18	0.18752	114
-	肉鮮生	23	0.15521	150
-	Meat Love 橡木炭火韓國烤肉	25	0.10652	150
-	小飯館兒	16	0.19859	94
-	韓川館	12	0.23144	113
-	順風韓國料理	22	0.15573	141

表 15. 線性規劃法-韓式料理類別加入其他知名店家後的驗證結果

google 知名店家	google 評分	評論數
韓鄉韓國料理	4.4	2,220
台中高麗屋	4.1	1,791

一桶 tone 韓式新食	4.5	1,076
GangNam 首爾韓鍋	4.5	<a href="#">1,201</a>
FASHION PIG 韓式熟成五花肉	4.8	1,432
好好吃肉	4.5	4,821
囍韓燒	4.5	1,415
菜豚屋	4.5	1,214
韓太閣	4.3	958
花小豬	4.2	1,261
肉鮮生	4.6	1,501
Meat Love 橡木炭火韓國烤肉	4.3	1,692
小飯館兒	4.1	1,767
韓川館	4.4	1,753
順風韓國料理	4.1	940

表 16. google 評論上其他知名韓式料理店家(109.6.30)

總結以上驗證線性規劃法的三種方式，第一種為與隨機猜測排名的 baseline 比較排名預測結果，第二種為十種餐飲類別輪流進行模型測試的十折交叉驗證，第三種則是加入十五間其他知名店家測試是否能挑選中 menu 美食誌的 top10 餐廳，最後測試結果雖然在少數餐飲種類或店家上無法完全精準預測排名，但是以整體排名而言，均有趨近於 menu 美食誌排行的現象。

#### 4.2.2 機器學習分類演算法

此章節實作的機器學習分類法為沿用前一節線性規劃之餐飲評分機制的資料前處理結果進行特徵工程(feature engineering)，再以特徵工程整理出的特徵進行模型訓練與排名結果預測，且以一則評論作為單筆訓練資料，此評論所屬的店家排名作為訓練資料的標籤(label)，共 10408 筆資料，進行多分類(multiclass classification)的監督式學習，模型將預測出一則評論所屬的排名，再利用 random forest 解決分類問題時，按照多棵樹分類器投票決

定最終分類結果的投票法概念((Breiman, L.,(2001)), (戰爭熱誠,(2019)))，計算此店家中所有評論的所屬排名，選取眾數的排名作為此店家的排名預測結果，並且沿用線性規劃法隨機挑選的八種餐飲種類(義大利麵、牛肉麵、滷肉飯、咖哩料理、港式餐廳、韓式炸雞、牛排、鐵板燒)作為訓練模型的資料，再以此模型預測其他兩種餐飲種類(冰品、韓式料理)的排名，最後在下一節比較線性規劃法和機器學習分類法的排名結果差異。

依據資料前處理後的原始資料進行特徵工程，欲找出各名次店家的評論通常都具有哪些特徵，例如：直觀而言，用戶對於第一名店家的評論通常具有較多的情緒詞彙，因此本研究整理出九項特徵，分別為各則評論的情緒詞數量、各間店家獲得所有評論中，各則評論的情緒詞數量最大值、平均值、中位數以及各間店家獲得的平均主題標籤數(hashtag)、粉絲數、按讚數、打卡頻率、距離 menu 美食誌排行發布的時間，以下分別描述此九項特徵的生成方式(表 17)，以及特徵範例(如圖 26)：

特徵	生成方式
情緒詞數量 (senti_word_count)	依據情緒辭典篩選出評論中包含的情緒詞，並計算此則評論中情緒詞總數
情緒詞數量最大值 (senti_word_count_max)	於一間店家中，比較所有評論的情緒詞總數，計算出情緒詞總數的最大值，代表此店家的所有評論於此特徵的值
情緒詞數量平均值 (senti_word_count_mean)	於一間店家中，比較所有評論的情緒詞總數，計算出情緒詞總數的平均值，代表此店家的所有評論於此特徵的值
情緒詞數量中位數 (senti_word_count_median)	於一間店家中，比較所有評論的情緒詞總數，計算出情緒詞總數的中位數，代表此店家的所有評論於此特徵的值

平均主題標籤數 (hashtag_count_mean)	於一間店家中，計算所有評論平均獲得的主題標籤數量，代表此店家的所有評論於此特徵的值
打卡頻率(check_list)	於一間店家中，計算店家於指定時間區段內獲得的平均貼文數，代表此店家的所有評論於此特徵的值
粉絲數 (fans_count)	每一則評論的發文者之粉絲數
按讚數(like_list)	每一則評論貼文獲得其他人按讚的總數量
距離 menu 排行發布的時間(time_list)	每一則評論時間距離 menu 美食誌發布排行榜之時間的天數差距

表 17. 九項特徵的生成方式

以直觀的角度而言，一則評論中的正向情緒詞被提及愈多次，即代表此間店家愈受歡迎，在大眾心中的排行應當愈前面，而本研究發現 Instagram 的用戶評論多為正面，負面情緒詞鮮少被提及，因此在情緒詞數量此項特徵中，因為居多為正面情緒詞，故可以直觀的假設評論的情緒詞數量愈多，店家可能愈受歡迎。另外，經研究發現各間店家獲得所有評論中，各則評論的情緒詞數量的最小值都為 0，對預測沒有幫助，故不加入此項特徵。

	senti_word_cnt	senti_word_cnt_max	senti_word_cnt_mean	senti_word_cnt_median	hashtag_count_mean	fans_count
0	19	74	9.793333	5.0	4.8	364.0
1	8	74	9.793333	5.0	4.8	741.0
2	1	74	9.793333	5.0	4.8	349.0
3	4	74	9.793333	5.0	4.8	247.0
4	22	74	9.793333	5.0	4.8	136.0

check_list	like_list	time_list	Label
2.888	8	-2	1
3.466	38	-7	1
2.644	52	-7	1
2.466	33	-8	1
0.366	5	-8	1

圖 26. 第一名店家的特徵範例(接續上圖)

若計算各特徵值與排名(label)的相關係數(Correlation coefficient)(如下表 18), 可篩選出情緒詞數量平均值、中位數對預測準確度較有幫助, 可見平均每則評論提及的情緒詞數量確實能夠一定程度地代表店家在大眾心中的排名, 且情緒詞數量愈多, 排名可能愈前面, 然而, 研究發現平均主題標籤數以及用戶個人資訊(粉絲數)對於排名結果卻不會有較直接的影響。

特徵值	相關係數
senti_word_count_median	0.206708
senti_word_count_mean	0.194224
senti_word_count_max	0.119092
hashtag_count_mean	0.086899
senti_word_count	0.045236
check_list	0.001197
fans_count	0.001075
time_list	0.001039
likes_list	0.000068

表 18. 各特徵值與排名的相關係數

在實驗結果的部分, 利用不同的分類模型, 包含 SVM、KNN、Random Forest 與各模型的降維(PCA), 進行模型訓練與預測的排名結果(標準排名為: [1,2,3,4,5,6,7,8,9,10])如下表 19, 可發現不論以何種模型預測冰品的排名, 其結果都不如隨機猜測排名(mse=16), 但在韓式料理中, 各模型的排名預測均優於 baseline(mse=16), 又以隨機森林(Random Forest)的效過相對最好。

	冰品	韓式料理
SVM	27.1/[5, 8, 7, 8, 7, 3, 2, 2, 2, 2]	11.6/[7, 4, 1, 7, 4, 2, 4, 8, 8, 4]

PCA + SVM	28/[5, 8, 8, 8, 7, 3, 2, 2, 2, 2]	11.6/[7, 4, 1, 7, 4, 2, 4, 8, 8, 4]
KNN	22.5/[7, 9, 2, 10, 7, 3, 8, 3, 9, 2]	9.1/[2, 10, 4, 7, 4, 9, 8, 8, 10, 8]
PCA + KNN	26.1/[5, 9, 2, 10, 7, 2, 8, 3, 2, 2]	8.9/[2, 10, 4, 7, 4, 8, 9, 8, 10, 8]
Random Forest	26.8/[8, 9, 5, 10, 7, 3, 5, 1, 9, 2]	3.3/[2, 1, 1, 4, 4, 5, 4, 8, 5, 10]
PCA + Random Forest	19.8/[5, 9, 1, 3, 7, 3, 8, 1, 10, 2]	11.7/[4, 8, 1, 7, 8, 3, 9, 8, 8, 4]

表 19. 不同模型下的 MSE/排名結果

而若以 macro average 來驗證模型準確度，以 PCA 搭配 SVM 的 macro average 為例(如下圖 27)，可得知第五名的評論無法被正確辨識，而第八名的評論則相較之下較能被正確預測，然而，macro average 的評估標準是以完全命中排名為基準，但本研究主要目的為預測排名，意即若預測排名與實際排名差距一名之內，仍為可接受的結果範圍，若為 macro average 則會判定為預測錯誤，因此本研究認為以 MSE 作為本研究的評估指標較能符合實際需求，而 macro average 則只作為參考的評估指標。

	precision	recall	f1-score	support
1	0.49	0.26	0.34	247
2	0.32	0.32	0.32	226
3	0.19	0.51	0.28	235
4	0.11	0.19	0.14	221
5	0.00	0.00	0.00	159
6	0.16	0.23	0.19	171
7	0.44	0.35	0.39	195
8	0.69	0.31	0.43	201
9	1.00	0.14	0.25	197
10	0.40	0.33	0.36	230
accuracy			0.27	2082
macro avg	0.38	0.26	0.27	2082
weighted avg	0.39	0.27	0.28	2082

圖 27. 以 PCA 搭配 SVM 的 macro average 為例

而為了從眾多店家中挑選出符合 menu 美食誌的餐廳排行，故延用線性規劃法的其他 15 間 google 知名店家，加入分類模型進行測試，然而上

述投票法的分類方法僅能將店家分類為一至十名(如下表 20，以冰品為例對 25 間店家進行投票法的排名結果)，無法有效計算出 25 名，因此利用 random forest 解決回歸問題時，由多顆樹預測值的均值決定最終預測結果的均值概念((Breiman, L.(2001)), (戰爭熱誠,(2019)), (江穎, 2017))，將一店家所有評論的預測排名進行平均，(例如：該店家 150 則評論的預測排名皆為 1，則此店家的排名分數為 1)，再將所有店家的排名分數排序，排名分數愈低，代表排名愈前面，最終可得所有店家的排名(如表 21、22)。

Menu 美食誌排名	餐廳名稱	排名	排名分數	評論數量
第 1 名	懷舊小棧 - 五妃廟口豆腐冰	8	-	150
第 2 名	清水堂	9	-	150
第 3 名	春美冰菓室	5	-	150
第 4 名	朝日夫婦	10	-	150
第 5 名	Kinber 金帛手製	7	-	150
第 6 名	Lindaogebia 你家隔壁	3	-	150
第 7 名	御品元冰火湯圓	5	-	150
第 8 名	阿宗芋冰城	1	-	150
第 9 名	有春冰菓室	9	-	150
第 10 名	桑原商店	2	-	150
-	阿爸の芋圓	2	-	150
-	好想吃冰	9	-	150
-	大方冰品	3	-	137
-	冰島雪花妹	5	-	81
-	建中黑砂糖刨冰	8	-	139
-	建華冰店	5	-	150
-	雪王冰淇淋	5	-	150
-	涼心冰店(黑店冰店)	8	-	148
-	溝壩清涼冰店	2	-	146
-	慢食堂	7	-	149
-	新建利冰店	9	-	148
-	剛好冰果室	7	-	150



-	西門町芒菓冰	3	-	150
-	船來雪淇淋	4	-	140
-	豐春冰菓店	4	-	150

表 20. 機器學習分類法-冰品類別以眾數決定 25 間店家的排名

Menu 美食誌排名	餐廳名稱	排名	排名分數	評論數量
第 1 名	懷舊小棧 - 五妃廟口豆腐冰	2	1.857	150
第 2 名	清水堂	13	4.043	150
第 3 名	春美冰菓室	2	1.857	150
第 4 名	朝日夫婦	2	1.857	150
第 5 名	Kinber 金帛手製	10	3.012	150
第 6 名	Lindaogebia 你家隔壁	10	3.012	150
第 7 名	御品元冰火湯圓	7	2.646	150
第 8 名	阿宗芋冰城	18	5.289	150
第 9 名	有春冰菓室	5	1.925	150
第 10 名	桑原商店	17	5.124	150
-	阿爸の芋圓	9	2.709	150
-	好想吃冰	12	3.507	150
-	大方冰品	15	4.709	137
-	冰島雪花妹	20	6.395	81
-	建中黑砂糖刨冰	23	7.443	139
-	建華冰店	21	7.22	150
-	雪王冰淇淋	24	7.633	150
-	涼心冰店(黑店冰店)	6	2.428	148
-	溝壩清涼冰店	22	7.441	146
-	慢食堂	16	4.745	149
-	新建利冰店	25	9.283	148
-	剛好水果室	14	4.332	150
-	西門町芒菓冰	1	1.444	150
-	船來雪淇淋	19	6.077	140
-	豐春冰菓店	8	2.649	150

表 21. 機器學習分類法-冰品類別以平均方式決定 25 間店家的排名

Menu 美食誌排名	餐廳名稱	排名	排名分數	評論數量
第 1 名	韓虎嘯	1	1.415	150

第 2 名	豬對有韓式烤肉吃到飽	12	3.436	145
第 3 名	I' m Kimchi	11	3.21	150
第 4 名	Broccoli Beer 韓國餐酒食堂	16	3.84	150
第 5 名	NeNe Chicken Taiwan	21	4.945	150
第 6 名	韓石食堂	7	2.316	96
第 7 名	奔跑吧年糕鍋	19	3.972	92
第 8 名	米豐屋	16	3.84	147
第 9 名	韓湘辣年糕	3	1.625	150
第 10 名	咚咚家韓式豬肉	8	2.463	91
-	一桶 tone 韓式新食	13	3.445	138
-	台中高麗屋	10	3.189	136
-	花小豬	22	5.861	134
-	菜豚屋	23	6.265	131
-	好好吃肉	24	6.67	150
-	韓太閣	4	1.738	130
-	韓鄉韓國料理	5	2.223	136
-	囍韓燒	18	3.947	148
-	FASHION PIG 韓式熟成五花肉	20	4.244	150
-	GangNam 首爾韓鍋	15	3.645	114
-	肉鮮生	25	9.095	150
-	Meat Love 橡木炭火韓國烤肉	1	1.415	150
-	小飯館兒	14	3.604	94
-	韓川館	9	2.663	113
-	順風韓國料理	5	2.223	141

表 22. 機器學習分類法-韓式料理類別以平均方式決定 25 間店家的排名

從上表 21、22 可以發現以平均方式計算出的排名較無法有效預測

menu 美食誌的排名，在冰品的預測結果中，前十名僅預測出七間 menu 美食誌所選的餐廳，且當中僅有三間較類似於 menu 美食誌的排行；在韓式料理的預測結果中，則僅有三間較類似於 menu 美食誌的排行，故探討預測不

夠準確的原因，研究發現以同樣方式進行歌手決賽排名預測的江穎(2017)，利用至少 16 個特徵值，且成功找到與決賽排名之相關係數超過 0.7 的兩項特徵值，分別為排名平均值和中位數，故在八名參賽者中準確預測冠軍以及四位預測誤差僅一名的成績，然而，反觀本研究的機器學習分類法，因 Instagram 可提供的分析資料有限，僅能找出九個特徵值，且未能找到與排名非常相關的特徵值，僅有低度相關(Modestly correlated)的(Wikipedia, (2020))的情緒詞數量平均值、中位數、最大值，故較難以準確預測 menu 美食誌的排行榜。

#### 4.2.3 線性規劃法與機器學習分類法之結果比較

本研究實作兩種不同方法預測大眾心目中的餐廳排行，利用的資料來源都為 Instagram 的用戶貼文與用戶個人相關資訊，然而利用不同的方法會產生不同的排名效果，因此，此小節簡單統整本研究餐飲評分機制之線性規劃法與機器學習分類法之最佳結果如下(以隨機選取的冰品與韓式料理之 mse 為例)：

	冰品	韓式料理
線性規劃法	4.6	5.8
機器學習分類演算法之 Random Forest	26.8	3.3

由上表可發現機器學習分類法對於韓式料理的排行預測能力優於線性規劃法，但是在冰品中的預測能力卻大幅降低，而線性規劃法對於冰品的排行預測則顯然準確許多，且以整體的排名而言，線性規劃法也優於機器學習分類法，造成此現象的可能原因推測為兩方法各自採納不同的排名標準，以線性規劃法而言，主要以情緒分數計算不同排名之間的情緒程度差異，而機器學習分類法則是不考慮情緒程度，直接以情緒詞彙的多寡進行判斷，因此，可以歸納出以情緒程度進行排名比起單純以情緒詞彙量進行

排名，更適合用於各種類的排名，而以情緒詞彙量為基準的機器學習分類法僅適用於特定的餐飲種類。總而言之，若以相同資料進行模型預測，仍以線性規劃法之餐飲評分機制的預測表現較現有分類演算法穩定。

### 4.3 各餐飲種類的排名差異分析

本節將依據線性規劃法和機器學習分類演算法的排名結果，分別探討不同方式計算下的排名，造成各餐飲種類產生排名差異的可能原因。

#### 4.3.1 線性規劃法的排名結果分析

研究發現增加情緒詞彙且情緒詞彙之間的情緒程度差異愈大，意即將不同程度的情緒詞彙之間的情緒分數差距擴大，愈可增加排名準確度，由 4.2.1.2 小節的表五與表八可見，加入本研究人工標記的情緒詞彙後，排名準確度優於僅利用 antusd 擴充詞典的表現，而本研究人工標記之情緒詞的情緒分數即為擴大至 $[-10,10]$ 之間，相較於 antusd 情緒詞彙的情緒分數僅限制於 $[-1,1]$ 之間。

此外，另一研究發現新排名可以趨於 MENU 美食誌的排名，可能原因在於 MENU 美食誌的排名是根據其用戶的打卡數決定，而新排名的分析資料來源亦來自於 Instagram 用戶的打卡貼文，共通點都是用戶實際到過商家用餐並打卡所留下的紀錄，因此本研究可以計算出類似於 menu 美食誌所調查出的餐廳排行。

若進一步分析造成牛肉麵的排名準確，但港式餐廳排名卻相較不準的可能原因，經過分析原始評論之後，發現港式餐廳的第一名(詳見附錄一)在 Instagram 社群上，被認為 cp 值不高，雖然非常好吃但是導致人潮很多需要花很多時間排隊，因此在分析結果中被分類為第五名，無法趨於 menu 美食誌所判定的第一名，而第七名的評論確實都含有大量的正向情緒詞彙，導致被線性規劃法排名為第一名，無法趨於 menu 美食誌所判定的第七名，但排

除第一名與第七名之外，其餘排名均能趨於 menu 美食誌的排名。且觀察新排名結果(詳見附錄一)可發現，若以前五名與後五名兩區段的角度觀看排名，在每個種類中，線性規劃法幾乎都能正確地將前五名的店家分類至前五名的區段中，反之亦然，屬於前段排名的店家較不會誤判為後段排名。

另外，本研究也發現線性規劃法計算出的第一名，普遍每篇評論的內容都很豐富，使用到的情緒詞彙也隨之增加，僅有少數評論包含較少的文字，所以最後排名總分較高而列為第一名，反之，排名愈後面的店家平均每則評論的內容豐富度則隨之減少。故本研究推測第一名在某種程度上而言可算是人氣度較高，可以用許多正向情緒詞描述此店家所提供的餐點、環境或服務，因為觀察排名結果為一至三名的店家後，研究發現許多評論都會在餐點、環境或服務等面向上詳細描述自身的體驗與感受，同時亦呼應了 Menu 美食誌的排名計算方式，因為 Menu 美食誌是依據打卡數排名，所以某種程度而言亦算是根據人氣度來進行排名。

最後，研究亦發現若是來自各國的外文評論亦會造成排名的失準，因為線性規劃法的評分機制只透過中文的情緒詞彙計算排名，無法判斷各種外文的情緒詞彙，因此會造成排名的落差。

#### 4.3.2 機器學習分類演算法的排名結果分析

利用機器學習分類法透過八種餐飲類別建立模型之後，套用在冰品以及韓式料理上進行排名預測，結果顯示韓式料理的排名預測準確，但冰品卻嚴重失準，究其原因，4.2.2 章節中提及機器學習分類法主要是以各店家所有評論的平均情緒詞數量作為預測的特徵值，而觀察原始評論資料後發現，韓式料理的評論中，情緒詞使用的頻率較高，且每一則評論的使用量較平均，然而，在冰品的評論中，各篇評論的情緒詞使用量差異大，一部分評論幾乎沒有使用到情緒詞彙，而一部份評論則使用到非常多的情緒詞

彙，因此造成機器無法以平均情緒詞彙使用量準確判斷排名結果。

#### 4.4 小結

本研究根據 MENU 美食誌的十種餐飲種類排行榜進行實驗，每一種餐飲種類共排出前十名，依據這十種類別的前十名，共 100 間店家，蒐集這些店家在 Instagram 社群上的用戶評論，各店家收集 150 篇左右的評論，且因為現有評分機制潛藏著長期累計評分的評分匯總問題，因此考慮進時間因素，僅蒐集 MENU 美食誌於網路上發布排名前一年以內的用戶評論，但是研究發現，少數店家一年內所獲得的評論過少，甚至不足 50 篇，導致計算出的商家總評分較高，嚴重影響最終排名結果，因此在這些少數店家中，蒐集不限於一年內的評論，擴張為蒐集兩年左右內的評論，直到蒐集至 150 篇左右，再經由本研究的線性規劃法之評分機制計算過後，也確實能夠獲得較準確的排名。

本研究經過一連串測試不同方法下的排名精準度後，調降 mse，並利用線性規劃法結合現有情緒辭典和人工標記情緒詞彙，找到一組參數  $w_1, w_2, w_3$  在趨近於 1 的情況下讓排名 mse 最小化，且在兩種測試資料當中亦能正確計算出符合 MENU 美食誌的排名結果，測試過程與最終排名結果可參考 4.2.1.2 的表 5 至表 9 與附錄一，測試過程中，本研究亦證實加入用戶影響力作為權重，也確實能夠增加排名的精準度。最後，本研究共以三種驗證方式，分別為計算隨機猜測排名之 baseline、十折交叉驗證、加入其他知名店家，驗證是否達到準確預測 menu 美食誌排行的目的。

進一步分析發現，造成某些餐飲種類的新排名結果較無法符合標準排名 (MENU 美食誌排名) 的可能原因，推測是因為某些餐廳在 Instagram 社群上的人氣與 MENU 美食誌所計算出的人氣固然不同，導致排名結果的表現，會因為少數幾間店家的預測與實際的排名差距過大而產生排名較不準的情況，此外，亦有可能是因為，各國的外文評論無法經過本研究餐飲評分機制正確判斷出情緒

詞彙，然而，雖然線性規劃法之評分機制在某些餐飲種類排行中，較無法完全精準預測排名，但是以整體而言，大部分的餐飲種類都能經由線性規劃法較準確地預測出符合大眾心中的排行，而不須透過大眾投票或是 MENU 美食誌的特別計算過程得知。

最後，進行第二種方式，以機器學習分類演算法進行排名預測，由 4.2.2 的排名結果可見，機器學習分類法僅適用於情緒詞使用頻率較高，且每一則評論的使用量較平均的狀況(如韓式料理)，其排名結果較為準確，但是在加入其他 google 知名店家後進行排名，卻證實無法有效從眾多商家中找出符合 MENU 美食誌的排行，故以整體排名而言，還是以線性規劃法之評分機制的排名結果較為穩定。

## 第五章、結論

### 5.1 研究成果

本研究根據過去文獻觀察後發現，目前鮮少利用社群資料，針對評論者影響力進行餐飲評分機制設計以辨識出符合大眾心目中餐廳排行的相關研究，利用社群聆聽(social listening)的概念自動調查出最佳的餐廳排行榜，因此，本研究利用 Instagram 用戶具有影響力的特性，結合情緒分析技術分析餐廳評論，將評論轉為評分後，經由用戶影響力的加權轉換，依照最終分數高低進行餐廳排名，不須透過大眾投票的問卷調查或是美食社群的調查便能找出符合大眾心中的最佳餐廳排行。

為了找到符合大眾心中餐廳排行的標準，以作為實驗的參照對象，本研究比較大眾評分網站(如：google 評分、愛食記評分、yelp 評分)與相關美食社群後，發現 menu 美食誌最為合適，因為大眾評分網站存在著假評論、評論與評分不符、評分無法即時反應最新評論等潛藏的問題，相較之下，menu 美食誌的排行榜為根據近幾個月其社群平台上，美食客到店用餐的打卡數(check in)作為排名依據，且發布後會吸引各網路媒體報導，引起大眾前往用餐，故相對具有公信力，足以代表大眾心目中的最佳餐廳排行，而實驗結果亦顯示可直接根據 Instagram 的用戶打卡評論推算 menu 美食誌的排名，不須透過 menu 美食客的打卡資料即可得知大眾心中的最佳餐廳排行。

經本研究觀察後發現，單純以台大擴增情緒辭典 antusd 進行評論的情緒分數計算後，依據分數高低進行排名的結果與 menu 美食誌的實際排名仍存有一些差距，八種餐飲類別中僅有一種排名結果優於隨機排名 baseline，然而，研究發現 Instagram 用戶通常會用 emoji 表達他們的情緒，評論中含有較多的 emoji，因此本研究也嘗試將 emoji 加入情緒分數的計算，但排名效過沒有太大的變化，而若是加入權重概念以及將各則評論的情緒分數正規化後，可見八種



餐飲中有四種(義大利麵、滷肉飯、港式餐廳、牛排)的預測誤差都有顯著下降，且都優於隨機排名 baseline，而再進一步加入本研究人工標記情緒詞，則八種作為訓練資料之餐飲類別的排名結果都可優於隨機排名 baseline，可見情緒詞的增加更能夠幫助排名準確度的提升，最後利用線性規劃法，結合現有情緒辭典與本研究人工標記情緒詞，找到最佳組合參數讓排名 mse 最小化，並測試於兩種餐飲類別的測試資料中，冰品與韓式料理的預測誤差均在 7 以下(隨機排名 baseline=16)，且冰品的一致序列對比率評分(FCP)可達八成以上，韓式料理則為七成。另外，本研究嘗試利用第二種方式，機器學習分類演算法進行排名預測，然而研究發現僅有情緒詞彙數量與排名結果較相關，用戶個人資訊、主題標籤數、評論發布時間對結果預測較無幫助，故在兩種餐飲類別之測試資料排名預測結果中，僅有韓式料理的排名預測較為準確。

最後根據研究結果，進行各餐飲種類的排名差異分析，在線性規劃法中，研究發現增加情緒詞彙且情緒詞彙之間的情緒程度差異愈大，愈可增加排名準確度，且較能穩定的預測出十種餐飲類別較接近 menu 美食誌的排名，其原因可能來自於，線性規劃法與 menu 美食誌排名都是利用實際到過商家用餐並打卡所留下的資料進行排名計算。另外，導致線性規劃法在特定種類中無法完全精準預測的可能原因為，店家在 Instagram 社群與 MENU 美食誌中的人氣固然不相同，造成排名預測的不一致，以及線性規劃法較無法從各國外文評論中分析情緒程度高低，若店家的評論多為各國外文，亦會導致排名預測誤差。而在機器學習分類演算法中，因考慮的特徵值不多，僅適用於預測情緒詞使用頻率較高，且每一則評論的使用量較平均的餐飲種類，因此仍以線性規劃法的預測表現較為穩定。

## 5.2 研究貢獻

本研究的主要貢獻如下：

1. 第一個以經過大眾驗證的餐廳排名為標準，進行排名預測的研究：現行評分機制(如：google、愛食記、網路溫度計)的排名未經過大眾驗證，僅可作為參考，但不一定符合大眾心目中的最佳餐廳排名，而據我們所知本研究為第一個利用情緒分析方法並加入權重概念透過社群聆聽(social listening)分析 instagram 資料，計算出符合大眾心目中的餐廳排名。
2. 餐飲方面的情緒辭典擴充：本研究除了利用過往研究的情緒詞典進行實驗之外，另外結合本研究人工標記出的情緒詞彙，除了可以增進本研究的排名預測精準度之外，更可以作為日後其他利用情緒分析方法進行實驗之研究的幫助，以因應持續變化的網路語言，達到更精準地分析文字中所含有的情緒。
3. 不需花費大量成本，可隨時計算出趨於大眾心目中的餐廳排行：本研究的價值之處是透過本研究就可以隨時計算出趨於大眾心中的餐廳排行，不再需要 menu 美食誌透過蒐集美食客的資料，花費大量人力與時間調查排行，只需借助 Instagram 上面眾多美食部落客自然產生的貼文內容，即可產出排名，日後亦可提供 menu 美食誌作為排名參考。

### 5.3 研究限制

本研究的研究限制主要包括兩種限制，分別為可取得的模型訓練資料量有限以及 Instagram 可運用之資料有限。以下將以條列方式進行說明：

1. 可取得的模型訓練資料量有限：menu 美食誌的排行榜也有 2018 年調查的排名，但是 Instagram 資料不一定有留存 2018 年的資料，且太過久的資料可能會隨時間被使用者刪除、設為不開放等變數影響，導致可分析的資料變很少，因此本研究僅利用 2019、2020 年的 menu 美食誌排行榜進行模型訓練。
2. 難以解釋線性規劃法之餐飲評分機制中各因素與排名結果之關係：本研究實作的線性規劃法是以情緒詞彙作為模型參數，找出結合各情緒詞的最佳參數，調校這些參數預測排名結果，無法清楚知道模型中考慮的變數如何影響排名結

果，若是能夠釐清更多其他變數對排名結果的影響，即可改善演算法，使得排名結果更趨於 menu 美食誌的排行。

## 5.4 未來研究方向

本研究之目的為預測出與 menu 美食誌相近的排名，以節省額外人力與時間成本調查其排名，轉為僅需利用社群評論資料進行社群聆聽(Social Listening)即可計算出類似於 menu 美食誌的排名，因此以下將描述未來可進一步研究的方向以改善排名預測結果：

1. 蒐集更多訓練資料進行分析：目前僅用八種 menu 美食誌的餐廳排行榜作為訓練資料，如果能夠蒐集更多 menu 美食誌的訓練資料才能更精準地篩選出不合適的非情緒詞彙 emoji，以更精準的情緒詞彙改善排名結果，或是結合更多社群的評論資料進行分析，測試是否能增加預測精準度。
2. 機器學習分類演算法加入更多特徵值：本研究中亦實作 random forest 等監督式機器學習演算法預測排名，但是本研究限於可從 Instagram 評論中取得的特徵值不多，因此未來若能增加更多相關特徵值，勢必能對預測結果有所幫助。
3. 試做其他具公信力的餐廳排行標準：未來研究可試著尋找與 menu 美食誌類似具有公信力的其他知名機構之排名結果，與其比較 mse 或是其他排名校正指標，測試是否能更趨於其排名結果。
4. 翻譯評論增加準確度：若以中文的情緒詞彙進行情緒分析，將無法對各國語言進行分析，而某些餐飲種類的 Instagram 評論其中一部份為來自各國包括日文、韓文、泰文等語言，因此若能將其語言進行翻譯，再進行情緒分析，將有可能讓排名結果更趨向 menu 美食誌的排行。

## 中文參考文獻

- 馬千惠 (2012)。《網路打卡的口碑傳播效果對消費者購買決策影響之研究——以餐飲業為例》。國立中山大學傳播管理研究所碩士論文。
- 廖敏惠 (2015)。《網路美食評論情緒分析之研究》。國立高雄餐旅大學台灣飲食文化所產業研究所碩士論文。
- 吳美慧 (2013)。《文章強度、部落客評價及推薦產品取得方式對購買意圖之影響》。國立中央大學資訊管理研究所碩士論文。
- 陳美如、蔡精育、宋鎧、范錚強 (2012)。《線上口碑對消費者購買意圖之影響——網路論壇的實驗研究》。中山管理評論，第 20 卷，第 2 期。
- 陳佳雯 (2017)。《探討美食圖片內容力及社群影響力對按讚行為之影響——以 Instagram 為例》。國立中山大學資訊管理學系研究所碩士論文。
- 鉅亨網新聞中心 (2015)。〈商戶點評網站 Yelp 遇困境：評分有時並不公正〉。檢自 <https://news.cnyes.com/news/id/483830> (Dec. 19, 2019)
- 簡之文 (2012)。《部落格文章情感分析之研究》。淡江大學資訊管理學系碩士論文。
- 謝佩庭 (2014)。《基於使用者情緒關鍵詞彙之臉書粉絲專頁評論分類與評分系統》。國立交通大學多媒體工程研究所碩士論文。
- 游和正、黃挺豪、陳信希 (2012)。《領域相關詞彙極性分析及文件情緒分類之研究》。Computational Linguistics and Chinese Language Processing, 17(4), 33-48。
- 金蛋網路數位行銷 (2018)。〈口碑行銷 | 餐廳美食爆紅關鍵：IG 打卡 +GOOGLE 評論〉。檢自 <https://www.gemarketing.com.tw/article/wom/food-ig-google/> (Dec. 23, 2019)
- 江穎 (2017)。〈機器學習預測《歌手》決賽全排名〉。檢自

<https://zhuanlan.zhihu.com/p/26883727> (Jun. 29, 2020)

結巴中文斷詞 (2016)。〈結巴中文斷詞台灣繁體版本〉。檢自

[https://github.com/ldkrsl/jieba-zh\\_TW#readme](https://github.com/ldkrsl/jieba-zh_TW#readme) (Jun. 29, 2020)

網路溫度計 (2019 年 12 月 19 日)。〈美食網路口碑排行〉。檢自

<https://dailyview.tw/Top100/Topic/90?volumn=0> (Dec. 19, 2019)

網路溫度計 (2020)。〈網路溫度計網路口碑排行榜資訊〉。檢自

<https://dailyview.tw/Top100/RankInfo> (Jun. 29, 2020)

痞客邦 (2020)。〈痞客邦社群影響力〉。檢自

<https://pixranking.events.pixnet.net/influence> (Jun. 29, 2020)

愛食記 (2020)。〈從超過 50,000 家精選餐廳中，探索您不知道的熱門美食〉。

檢自 <https://ifoodie.tw/> (Jun. 29, 2020)

戰爭熱誠 (2019)。〈Python 機器學習筆記：隨機森林算法〉。檢自

<https://www.cnblogs.com/wj-1314/p/9628303.html> (Jun. 29, 2020)

HackMD (2020)。〈用 Markdown 即時協作知識庫〉。檢自

<https://hackmd.io/> (Jun. 29, 2020)

INSTAGRAM (2020)。〈在 Instagram 查看洞察報告〉。檢自

[https://www.facebook.com/help/instagram/1533933820244654?help\\_ref=uf\\_permalink](https://www.facebook.com/help/instagram/1533933820244654?help_ref=uf_permalink) (Jun. 29, 2020)

MENU 美食誌 (2020)。〈吃貨必備，簡單記錄，搜尋好吃店家的美食神器〉。

檢自 <https://menutaiwan.com/tw/about> (Jun. 29, 2020)

Caitlin (2019)。〈INSTAGRAM 新手入門：怎麼看後台的數據分析〉。檢自

<https://caitlin1010.pixnet.net/blog/post/311663027>  
(Jun. 29, 2020)

Wikipedia (2020)。〈皮爾遜積差相關係數〉。檢自

<https://zh.wikipedia.org/wiki/%E7%9A%AE%E5%B0%94%E9%80%8A%E7%>

[A7%AF%E7%9F%A9%E7%9B%B8%E5%85%B3%E7%B3%BB%E6%95%B0](#)

(Jun. 29, 2020)

## 西文參考文獻

- An, J.X., Huang, J., Yu, W., Akoglu, L., Chandy, R., Faloutsos, C.(2011). Algorithm of Disambiguation and Matching of Chinese Word Segmentation in Connected Strategies Research. *Advanced Materials Research (Volumes 219-220)*, 1702-1706.
- Bhasin, H. (2019). Retrieved from What is Hashtag Marketing? Importance Of Hashtag Marketing <https://www.marketing91.com/what-is-hashtag-marketing/> (Dec 23,2019)
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32.
- Brown, D., Hayes, N. (2008). *Influencer Marketing: Who Really Influences Your Customers?*, Routledge.
- Chaovalit, P., Zhou, L. (2005). Movie review mining: a compareson between supervised and unsupervised. In *Proceedings of the 38th Hawaii International Conference on System Sciences*.
- Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43, 345–354
- Dai, W., Jin, G.Z., Lee, J., Luca, M. (2018). Aggregation of Consumer Ratings: An Application to Yelp.com. *Quantitative Marketing and Economics*, 16(3), 289-339
- Dewey, J., Wheeler, J., (2009). *Interest and Effort in Education*. Southern Illinois University Press: eBook Academic Collection.
- Domingos, P., Richardson, M. (2002). Mining the network value of customers. In *Proceeding of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, 57–66.
- Gan, Q., Ferns, B.H., Yu, Y., Jin, L. (2017) A Text Mining and Multidimensional Sentiment Analysis of Online Restaurant Reviews. *Journal of Quality Assurance*

*in Hospitality & Tourism*, 465-492

- GONG, X. (2014). Strategic Customer Engagement on Instagram- A Case of Global Business to Customer (B2C) Brands. Master's Thesis in Media Management, Media Management Master Program KTH Royal Institute of Technology.
- Jia, S. (2018). Behind the ratings: Text mining of restaurant customers' online reviews, *International Journal of Market Research*, 60(6), 561–572.
- Jin, J., Ji, P., Liu, Y. (2014). Recommending Rating Values on Reviews for Designers. *Encyclopedia of Business Analytics and Optimization*.
- Kamal, A. (2015). Review Mining for Feature Based Opinion Summarization and Visualization. *International Journal of Computer Applications*. 119(17)
- Kaviya, K., Roshini, C., Vaidhehi, V., Dhalia Sweetlin, J. (2017) Sentiment Analysis for Restaurant Rating. *IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)*, 140-145.
- Kempe, D., Kleinberg, J., Tardos, E. (2003). Maximizing the spread of influence through a social network. In the *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 137-146.
- Kralj Novak, P., Smailovic, J., Sluban, B., Mozetic, I. (2015). Sentiment of Emojis. *PLoS ONE* 10(12): e0144296.
- Lee K, D. (2015). Analytics, Goals, and Strategy for Social Media. *Library Technology Reports*, 51(1), 26–32.
- Ling Hang Yew, R., Binti Suhaidi, S., Seewoosurn, P., Kumar Sevamalai, V. (2018). Social Network Influencers' Engagement Rate Algorithm Using Instagram Data. *2018 Fourth International Conference on Advances in Computing*,



Communication & Automation (ICACCA).

Luca, M., Zervas, G. (2016). Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud. *Management Science*, 62(12), 3412-

3427.<https://doi.org/10.1287/mnsc.2015.2304>

Luca, M. (2016). Reviews, Reputation, and Revenue: The Case of Yelp.com. Harvard Business School Working Paper, No. 12-016

Mir Riyanul Islam. (2014). Numeric Rating of Apps on Google Play Store by Sentiment Analysis on User Reviews. International Conference on Electrical Engineering and Information & Communication Technology, pp.1-4.

Moon, S., Bergey, P. K., & Iacobucci, D. (2010). Dynamic effects among movie ratings, movie revenues, and viewer satisfaction. *Journal of Marketing*, 74, 108–121.

Murphy, R. (2018). Comparison of Local Review Sites: Which Platform is Growing the Fastest? Retrieved from

<https://www.brightlocal.com/research/comparison-of-local-review-sites/> (Dec 23,2019)

Page, L., Brin, S., Motwani, R., Winograd, T. (1999). The PageRank Citation Ranking: Bringing Order to the Web. Technical Report. Stanford InfoLab.

Perera, I.K.C.U., Caldera, H.A. (2017). Aspect Based Opinion Mining on Restaurant Reviews. 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), 542-546.

Peters, K., Chen, Y., Kaplan, A., Ognibeni, B., Pauwels, K. (2013). Social Media Metrics—A Framework and Guidelines for Managing Social Media. *Journal of Interactive Marketing*, 27, 281–298.

Pitman, J.(2019).The Ultimate Guide to Google My Business Reviews. Retrieved

- from <https://www.brightlocal.com/learn/how-do-google-reviews-work/> (Dec 23,2019)
- Rachid, A.D., Abdellah, A., Belaid, B., Rachid, L. (2018). Clustering Prediction Techniques in Defining and Predicting Customers Defection: The Case of E-Commerce Context. International Journal of Electrical and Computer Engineering (IJECE).
- Richardson, M., Domingos, P. (2002). Mining Knowledge-Sharing Sites for Viral Marketing. In Proceeding of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining.
- Shih-Ming, W., Lun-Wei, K. (2016). ANTUSD : A Large Chinese Sentiment Dictionary. Language Resources and Evaluation Conference(LREC).
- Yang, L., Jian-Wu, B., Zhi-Ping, F. (2017). Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory. Information Fusion, 36, 149-161.
- Yelp Elite Squad. (2014). Yelp Elite Squad. Retrieved from <http://www.yelp.com/elite> (Dec 23,2019)
- Zizzi, Hosie, R. (2017). HOW INSTAGRAM HAS TRANSFORMED THE RESTAURANT INDUSTRY FOR MILLENNIALS. Retrieved from <https://www.independent.co.uk/life-style/food-and-drink/millenials-restaurant-how-choose-instagram-social-media-where-eat-a7677786.html> (Dec 23,2019)
- Zhang, K., Cheng, Y., Liao, W.K., Choudhary, A. (2011). Mining millions of reviews: a technique to rank products based on importance of reviews. Proceedings of the 13th International Conference on Electronic Commerce, ACM.
- Zhu, T., Wu, B., Wang, B. (2009). Social Influence and Role Analysis Based on Community Structure in Social Network. Advanced Data Mining and

Applications, 788-795.

## 附錄

### 附錄一、線性規劃法排名結果

#### 一、義大利麵

名次	MENU 美食誌排名 (2019-09-23)	新評分排名
第一名	JAI 宅	1
第二名	MINT PASTA 義大利麵館	2
第三名	HUN 混	4
第四名	LA PASTA 義大利麵屋	5
第五名	MIGA KITCHEN PASTA . 米家廚房 義大利麵	8
第六名	DOR,留手工義大利麵	6
第七名	PETIT DOUX 微兜 CAFÉ BISTRO	10
第八名	<a href="#">KAFFE@HOME 我在家咖啡</a>	3
第九名	螺絲瑪莉義麵坊	9
第十名	村口微光	7

註：寫於 MENU 美食誌排名後的日期為其公開發布此種餐飲排行榜的時間

來源：<https://findlife.com.tw/menu/blog/2019/09/23/201909italytop10/>

#### 二、牛肉麵

名次	MENU 美食誌 (2019.12.1)	新評分排名
第一名	富宏牛肉麵	3
第二名	永康牛肉麵	1
第三名	劉山東牛肉麵	5
第四名	林東芳牛肉麵	2
第五名	乖乖牛肉麵	4
第六名	建宏牛肉麵	7
第七名	老皮牛肉麵	6
第八名	愛.熟成 21 牛肉麵	10
第九名	樂客牛肉麵	8
第十名	犇三牛牛肉麵	9

來源：<https://findlife.com.tw/menu/blog/2019/12/01/2019beefnoodles/>

### 三、滷肉飯

名次	MENU 美食誌排名 (2019-12-15)	新評分排名
第一名	今大魯肉飯	2
第二名	金峰魯肉飯	3
第三名	阿義魯肉飯	4
第四名	山河魯肉飯	1
第五名	金澤魯肉飯	9
第六名	禾日香古早味魯肉飯	5
第七名	痣男	8
第八名	南豐魯肉飯	6
第九名	小王清湯瓜仔肉	7
第十名	矮仔財滷肉飯	10

來源：<https://findlife.com.tw/menu/blog/2019/12/15/2019lowbahbang/>

### 四、咖哩料理

名次	MENU 美食誌排名 (2019-12-28)	新評分排名
第一名	十巷咖哩	1
第二名	小吉咖哩	2
第三名	佐藤咖哩	3
第四名	別噲大叔	5
第五名	小倉館 CAFE & RICE	7
第六名	富士咖哩 FUJI CURRY	9
第七名	N.N. THAI THAI	10
第八名	貴一郎 S・R・T 咖哩部	4
第九名	找地方坐 HAVE A SEAT	8
第十名	NONAME 咖哩カレーライス専門店	6

來源：<https://findlife.com.tw/menu/blog/2019/12/28/2019curry/>

## 五、港式餐廳

名次	MENU 美食誌排名 (2020-03-20)	新評分排名
第一名	華嫂冰室	5
第二名	茗香園冰室	2
第三名	美天餐室 DAY DAY	4
第四名	檀島香港茶餐廳	7
第五名	復興咖啡交易所_FXCE	3
第六名	吉星港式飲茶	10
第七名	崔記小餐館	1
第八名	香港茶水攤	6
第九名	港味高師傅港式點心	9
第十名	鴨寮街-烤鴨點心坊	8

來源：<https://findlife.com.tw/menu/blog/2020/03/20/2019hongkongstyle/>

## 六、韓式炸雞

名次	MENU 美食誌排名 (2020-05-11)	新評分排名
第一名	起家雞韓式炸雞	2
第二名	CHIMAC175 TAIPEI	3
第三名	KATZ 卡司複合式餐廳	4
第四名	朴大哥的韓式炸雞	6
第五名	努娜炸雞 누나치킨	7
第六名	娘子炸雞	5
第七名	OPPADAK 歐巴答韓式炸雞	1
第八名	TAEBAK 大發韓式特色料理	10
第九名	NENE CHICKEN	8
第十名	木槿燒韓國料理	9

來源：<https://findlife.com.tw/menu/blog/2020/05/11/2020koreanchicken/>

## 七、牛排

名次	MENU 美食誌排名 (2020-02-27)	新評分排名
第一名	赤鬼炙燒牛排	4
第二名	德州鮮切牛排 TEXAS ROADHOUSE	1
第三名	莫爾頓牛排館 MORTON' S THE STEAKHOUSE	6
第四名	教父牛排 DANNY' S STEAKHOUSE	5
第五名	N°168 牛排館	2
第六名	來來牛排館	3
第七名	A CUT STEAKHOUSE 牛排館	10
第八名	沾美西餐廳	8
第九名	牛排教父 CAPSTONE STEAKHOUSE	7
第十名	凱恩斯岩燒餐廳	9

來源：<https://findlife.com.tw/menu/blog/2020/02/27/2020steak/>

## 八、鐵板燒

名次	MENU 美食誌排名 (2019-10-26)	新評分排名
第一名	TBS 易邦饅 TEPPANYAKI	2
第二名	犇鐵板燒	3
第三名	米半 鐵板料理	5
第四名	明水然無菜單鐵板燒	4
第五名	潼精緻鐵板料理	7
第六名	合榭 HOSIC 精緻鐵板料理	9
第七名	陶板屋和風創作料理	1
第八名	三鉄鐵板燒	8
第九名	大初鐵板燒	10
第十名	當代鐵板燒	6

來源：<https://findlife.com.tw/menu/blog/2019/10/26/2019tepanyaki/>

## 九、冰品

名次	MENU 美食誌排名(2020-05-06)	新評分排名
第一名	五妃街豆腐冰 懷舊小棧	1
第二名	清水堂愛玉專賣店	2
第三名	春美冰菓室	5
第四名	朝日夫婦	3
第五名	KINBER 金帛手製	9
第六名	你家隔壁	6
第七名	御品元傳統手工元宵-冰火湯圓	7
第八名	阿宗芋冰城	8
第九名	有春冰菓室	4
第十名	桑原商店	10

來源：<https://findlife.com.tw/menu/blog/2020/05/06/2020icetop10/>

## 十、韓式料理

名次	MENU 美食誌排名(2020-04-19)	新評分排名
第一名	TIGERROAR 韓虎嘯	2
第二名	豬對有韓式烤肉吃到飽	1
第三名	I' M KIMCHI	3
第四名	BROCCOLI BEER 韓國餐酒食堂	9
第五名	NENE CHICKEN	5
第六名	韓石食堂	8
第七名	奔跑吧！年糕鍋	6
第八名	米豐屋	7
第九名	韓湘辣年糕	4
第十名	咚咚家 DONDONGA 韓式豬肉專賣	10

來源：<https://findlife.com.tw/menu/blog/2020/04/19/2020koreantop10/>



## 附錄二、擴增情緒詞典

編號	word	score
0	送	0.15
1	服務	0.1
2	肉麻	0.05
3	好ㄗ	0.2
4	敲好	0.4
5	豪ㄗ	0.2
6	豪豪ㄗ	0.5
7	好爽	0.5
8	推推	0.3
9	推推推	0.4
10	激推	0.7
11	讚	0.25
12	過癮	0.35
13	超棒	0.35
14	超級棒	0.5
15	👍	0.1
16	👍	0.1
17	👍👍	0.35
18	👍👍👍	0.45
19	愛店	0.3
20	棒	0.13
21	美食	0.16
22	特制	0.2
23	難波萬	0.3
24	濃	0.1
25	愛	0.2
26	yummy	0.1
27	優	0.15
28	膩害	0.15
29	想	0.005
30	失望	-0.2
31	三	0.005
32	□	0.25

33	軟嫩	0.1
34	入口即化	0.3
35	特有	0.05
36	吸滿	0.05
37	鶴立	0.2
38	XD	0.1
39	濃厚	0.3
40	疑問	-0.1
41	吃重	0.01
42	☺	0.1
43	值得	0.23
44	二訪	0.2
45	鬼金棒	0.25
46	特製	0.05
47	拉麵	0.15
48	第一名	0.5
49	第一	0.2
50	超好吃	0.6
51	肥肉	-0.2
52	拉麵控	0.1
53	久沒吃	0.25
54	🙄	0.1
55	這麼能	0.1
56	雙倍	0.2
57	達成	0.1
58	飛舞	0.2
59	很讚	0.3
60	很濃	0.1
61	鬼金是天	0.65
62	又來吃	0.2
63	yummyfood	0.1
64	good	0.055061
65	instagood	0.2
66	體驗	0.2
67	增量	0.2
68	豪豪	0.25
69	ㄗ	0.15

70	膩	-0.2
71	傻眼	-0.2
72	嗚嗚	-0.1
73	最特別	0.4
74	撐	0.2
75	很好吃	0.4
76	下次	0.1
77	再去	0.1
78	加點	0.1
79	真的很優	0.5
80	很優	0.3
81	很膩害	0.3
82	排了	-0.1
83	山嵐拉麵	0.25
84	山嵐	0.25
85	解憂	0.35
86	沾麵	0.15
87	白湯	0.15
88	foodlover	0.1
89	instalikes	0.1
90	noodlelover	0.1
91	Recommend	0.1
92	ramen	0.15
93	Ramen	0.15
94	lamen	0.15
95	taipeicuisine	0.1
96	🍜	0.15
97	🍲	0.05
98	🍽	0.05
99	🍴	0.1
100	🐱	0.1
101	□	0.15
102	❤️	0.1
103	豐盛	0.2
104	終於吃到	0.3
105	普通	-0.02
106	不錯	0.2

107	只剩	-0.1
108	築夢	0.3
109	厲害	0.2
110	愛吃	0.2
111	真的很好吃	0.8
112	拉麵是人生	0.3
113	第 N 次吃	0.6
114	很日本	0.2
115	肉不肥	0.1
116	香氣	0.2
117	香濃	0.2
118	辣麻味噌拉麵	0.25
119	相見恨晚	0.3
120	僅次於	-0.1
121	中規中矩	-0.05
122	感覺	0.02
123	覺得	0.02
124	虛	-0.2
125	爛	-0.2
126	差	-0.2
127	好差	-0.3
128	好硬	-0.15
129	久	-0.1
130	太久	-0.4
131	熱死	-0.2
132	排隊	-0.1
133	不高	-0.3
134	鹹	-0.1
135	好鹹	-0.2
136	太鹹	-0.1
137	沒	-0.1
138	不	-0.05
139	==	-0.05
140	不值得	-0.2
141	超優	0.4
142	👍👍👍	0.3
143	#yummm	0.1

144	#yummy	0.1
145	#delicious	0.1
146	熱點	0.2
147	網美	0.1
148	聖地	0.2
149	老朋友	0.1
150	Happy	0.1
151	霸氣	0.2
152	爽爽粗	0.3
153	聚餐	0.1
154	♥	0.1
155	😊😊😊	0.3
156	坐滿	0.3
157	😊😊😊😊	0.4
158	😊😊😊😊😊	0.4
159	👏👏	0.2
160	團聚	0.1
161	慰勞	0.1
162	其實	0.2
163	還好	0.2
164	棒棒	0.1
165	撈個	0.1
166	撈起來	0.1
167	😊	0.2
168	#foodie	0.1
169	#sharefood	0.1
170	#foodheaven	0.1
171	好友	0.1
172	歷久彌新	0.1
173	😊	0.1
174	眼淚	-0.1
175	不要	-0.1
176	大推	0.3
177	#海底撈	0.1
178	想再來	0.2
179	不失水準	0.01
180	意外找到	0.3

181	很好拍	1
182	食在好好吃	0.8
183	美好食光	0.8
184	非吃不可	0.8
185	真的推	0.8
186	排隊名店	0.1
187	□	0.8
188	最美的	0.9
189	空間很大	0.8
190	王美	0.8
191	約會	1
192	JAI 橘醬	1
193	CP 值高	0.2
194	CP 值夠	0.8
195	CP 值很高	0.8
196	CP 值超高	0.8
197	CP 值蠻高	0.6
198	CP 值算高	0.2
199	CP 值頗高	1
200	高 CP 值	0.2
201	CP 值尚可	0.05
202	CP 值中上	0.1
203	CP 值中下	0.01
204	CP 值中等	0.05
205	還會想再點一份	5
206	很方便食用	0.005
207	絕對是會回訪	2
208	超熱情	2
209	名不虛傳	0.3
210	下次再找其他人來	5
211	沒有空包彈	2
212	特別的好拍	6
213	蠻好入口的	2
214	吃得超飽	2
215	相機先食	0.5
216	相機食先	0.5
217	深受學生喜愛	5

218	竟然多了新餐點	2
219	最代表性的餐廳	0.95
220	吃的是回憶	0.8
221	極有名的	3
222	必點	0.1
223	有自己的味道	5
224	真的有夠讚	8
225	麵不加錢	8
226	超喜翻	8
227	整個韓風捏	6
228	拍起來就是好看	8
229	N 訪	8
230	蠻常來吃的	1
231	感動痛哭流涕	8
232	一定要試試	5
233	更加網美店了	6
234	撒花	5
235	太好拍了	6
236	燈光好氣氛佳	6
237	融入許多設計	5
238	留下不錯的印象	2
239	狂誇超好吃	8
240	獨家醬	5
241	光看就覺得很滿足	3
242	五燈獎	3
243	裝潢真的太美	6.2
244	名不虛傳	3
245	人氣店	2
246	耐嚼而柔軟	3
247	有層次更香更好喝	0.01
248	免費更換	6
249	表示敬意	6
250	超值破表	9
251	超級賣點	7
252	榜上有名	0.01
253	米其林必比登	0.1
254	台灣首創	3

255	慕名而來	8
256	還不錯	0.05
257	福鴻新人	4
258	最對味	6.5
259	俗擱大碗	5
260	油花很香	5
261	DELICIOUS	0.1
262	YUMMY	0.1
263	喝飽吃足	3
264	吃光光	1
265	飽到囊囊的	8
266	□	0.1
267	每個禮拜必報到	8
268	再來一碗	2
269	很常來吃	9
270	PERFECT	0.1
271	味道鮮美	1
272	肉又大	1
273	不用排隊直接入座	5
274	INVINCIBLE	9
275	必定要排除萬難嚐看看	7
276	最強的	3
277	美食家的代表	5
278	從未吃過的人一定要去	5
279	朝聖	0.1
280	嚐鮮	0.1
281	香氣味十足	1
282	平價好吃	3
283	很好喝	0.1
284	吃了好滿足	3
285	貼心的店員	0.5
286	味道依然沒變	0.3
287	生意還是很好	5
288	撐死我	0.5
289	□	0.05

290	好地方	0.01
291	完美絕配	6
292	爽口不油膩	0.8
293	送啦	1
294	富具時尚感	4
295	常常這麼多人	1
296	吃到飽吃到吐	3
297	有設計感	0.01
298	不需訂位	0.2
299	很親民	0.1
300	有咬勁	0.1
301	真的很想吃	2
302	好多網友推薦	8
303	衝來嚐鮮	3
304	吃過後就一直惦著	8
305	忍住隔天不去二刷	3
306	更深得我心	6
307	長知識	0.5
308	平價又樸實的好味道	3
309	Q 彈好吃	0.6
310	香有韻味	1
311	必吃	0.1
312	料理王道	8
313	常常吃	0.2
314	擠擠滿滿人	8
315	絕對是榜上有名	1.5
316	光看就要流口水	1
317	好澎湃	0.5
318	好吃到犯規	8
319	只是為了吃這個	0.2
320	忘不掉的味道	1
321	必來嗑一碗	2
322	不介意一直到那裡去	9.5
323	很高興價格低	3
324	多兩倍的肉	6
325	直接來兩碗	2
326	太美味吃了兩碗	9

327	太美味	0.8
328	不用排隊	0.01
329	豪香	0.2
330	有第二次	1
331	絕對想去	6
332	太舒服	0.8
333	仍然很棒	0.5
334	別來無恙	0.2
335	必到之地	9
336	未必是壞事	0.6
337	著實欣慰	0.7
338	又厚又美味	0.8
339	鴨蛋	0
340	GOOD	0.1
341	EXCELLENT	1
342	独特	0.2
343	大人氣	1
344	通通免費	5
345	難怪生意這麼好	0.8
346	CP 值真的很高	1
347	不難找	2
348	●	0
349	滑順感	0.3
350	小資族吃飽飽	1
351	加麵不加價	0.3
352	超級懷念	3
353	味道非常豐富	3
354	軟嫩不柴	0.1
355	一點都不老	0.2
356	吃不夠都可以免費加	2
357	鮮嫩多汁	0.1
358	還是念念不忘	5
359	忘記你我做不到	0.5
360	再訪	0.01
361	好愛	0.01
362	別有用心	0.5
363	沒有排很久	1

364	喜歡整個店的佈置	3
365	份量也蠻夠	1
366	網路上說的沒騙人	0.5
367	一直蠻多人在候位	5
368	酷爆	0.6
369	□	0.2
370	音樂很可以	1
371	沒失望	0.1
372	超級多汁	2
373	很有特色又很隱密	1
374	不會過熟	0.5
375	完全不油膩	0.01
376	厚實不乾不柴	1
377	都不用錢	2
378	大胃王不用怕	2
379	又薄又脆	0.3
380	好懷念	0.1
381	並沒有不好	0.5
382	都可以在這邊看到	5
383	GET 到	0.2
384	帶點果酸味	0.5
385	哈哈笑	1
386	前三名	0.5
387	變得好潮	3
388	時髦	0.3
389	一開店就坐滿	5
390	想吃好久了	6
391	比想像中快	0.5
392	比想像中快非常多	2
393	很 LUCKY	1
394	真的是感動到不行	3
395	讓我印象超深刻	3
396	保有口感卻不乾柴	1.5
397	是不是超級誘人	6.5
398	超多人推薦	6
399	都很好粗	1
400	認真對待每份餐點	3.5

401	就是放鬆	1
402	認真好喝	2
403	真是療癒到了個極點	5
404	真是療癒到了極點	5
405	超適合拍照打卡	5
406	讓我比較驚豔	6
407	就是無法擋	5
408	一開店就一堆人	3
409	真滴很好吃	3
410	可以去第二次	3
411	肉很嫩汁也很多	2
412	很對味	0.2
413	敲雞好吃	8
414	皮薄肉嫩	0.2
415	毛起來吃	8
416	便宜又划算	0.5
417	大家一起分享	0.5
418	沒在開玩笑的好吃	3
419	全部都沒雷	3
420	我尬藝	2
421	吃到還打包回家	2
422	真的好好粗	8.5
423	蠻多人推薦	1
424	超級超級想吃	8
425	舒壓	0.05
426	真的超級好吃啊	5
427	超稀有	6
428	可遇不可求	3
429	想再來一碗	3
430	超興奮	2
431	小確幸	0.1
432	吃多了也不會膩	2
433	最舒服了	1
434	很 Q	0.1
435	絕品	0.1
436	收進口袋名單	2
437	終於是被我吃到了	2

438	真心推各種口味	6
439	放在清單好久	6
440	會常來	1
441	浮誇	0.01
442	增量倍增	6
443	超大份	0.5
444	心血來潮	0.15
445	紅了好一陣子	8
446	超出期待	6
447	很棒的回憶	3
448	IG 打卡	0.1
449	一直看別人 IG 打卡	8
450	吃起來真棒	3.5
451	還是一堆人去吃	3
452	真的滿划算的	4
453	讓用餐感受更好	1
454	用餐人潮不少	2
455	真的難能可貴	3
456	慶祝不歸路	2
457	還這麼包容	2
458	HAPPY	0.01
459	LOVEYOU3000	8
460	愛你 3000	8
461	瘋狂吃	9
462	服務滿分	3
463	覺得可以再訪	3
464	鐵板	0
465	我的寶寶	3
466	我的嘻嘻寶寶	3
467	怎有辣摸好吃	9
468	太佛心	0.5
469	也太佛心	1
470	超讚推	1
471	老闆人超 NICE	6
472	□	0.01
473	好飽	0.001
474	出餐很快	1

475	這家一直想來	6
476	你是我心中的太陽	8.5
477	是無價	2
478	不管再怎麼忙	1
479	環境非常舒服	6
480	吃完還能去	1.5
481	網上有點太毒舌	0.5
482	好多人問我	5
483	來這拍照滿好的	3
484	大家都超愛	6
485	看超久終於來了	2.5
486	有夠喜歡這家	9
487	好吃又好拍	1
488	好好吃給他推一個	3
489	之前一直想去吃	8
490	超對我的味	5
491	茶香濃厚	3
492	軟嫩夠味	3
493	心還在	2
494	該收心囉	1
495	回歸現實	1
496	這麼久了才來	3
497	無條件付出	3
498	多元層次表現	4.5
499	超級絕配組合	6
500	吃了會懷念	3.5
501	所有菜色最喜歡的	3
502	在寒冬中享有溫暖	2
503	非常的剛好	2
504	看著都會想吃	2
505	滿座的狀態	2
506	抽獎進行中	0
507	✓	0.05
508	二次造訪	0.02
509	超差	-0.3
510	太貴	-0.1
511	蟑螂	-0.3

512	死鹹	-0.2
513	拉肚子	-0.3
514	送錯	-0.1
515	不推	-0.1
516	再也不會	-0.3
517	軟爛	-0.1
518	味精	-0.1
519	沒吃過	-0.03
520	幹嘛	-0.03
521	不怎麼樣	-0.05
522	老鼠	-0.2
523	差到	-0.2
524	吃不飽	-0.03
525	CP 值低	-0.1
526	趕人	-0.1
527	態度惡劣	-0.3
528	高估	-0.2
529	這麼久	-0.05
530	CP 值超低	-0.3
531	無言	-0.1
532	收走	-0.05
533	很難吃	-0.3
534	很髒	-0.3
535	臉色	-0.02
536	不給	-0.02
537	零分	-0.02
538	濃鬱	-0.01
539	沒多好吃	-0.05
540	太爛	-0.15
541	再也	-0.02
542	太軟	-0.05
543	面對面	-0.02
544	要不要	-0.01
545	沒洗	-0.01
546	超爛	-0.15
547	沒來	-0.02
548	腥味	-0.1

549	要命	-0.2
550	不優	-0.15
551	不敢恭維	-0.2
552	一星	-0.4
553	咳嗽	-0.02
554	薄薄的	-0.02
555	很薄	-0.02
556	大吼大叫	-0.03
557	下嚥	-0.01
558	超貴	-0.1
559	歐巴桑	-0.01
560	不太好	-0.05
561	差太多	-0.01
562	敷衍	-0.05
563	不來	-0.02
564	漏單	-0.1
565	CP 值很低	-0.5
566	有點鹹	-0.02
567	聽不懂	-0.02
568	差差	-0.15
569	超久	-0.05
570	還被	-0.01
571	小聲	-0.01
572	地上	-0.01
573	到極點	-0.02
574	排很久	-0.05
575	超少	-0.02
576	吃不出	-0.01
577	太油	-0.15
578	好意思	-0.02
579	嚴重不足	-0.5
580	踩到	-0.01
581	稱不上	-0.05
582	一臉	-0.02
583	半個	-0.01
584	都還沒	-0.01
585	爛店	-0.5



586	太重	-0.02
587	負評	-0.2
588	怪味	-0.15
589	翻白眼	-0.1
590	白跑一趟	-0.2
591	好久	-0.05
592	沒拿到	-0.01
593	插隊	-0.05
594	再也不敢	-0.2
595	很擠	-0.15
596	就掛	-0.02
597	薄薄	-0.01
598	偏貴	-0.05
599	肉薄	-0.01
600	堪憂	-0.2
601	大呼小叫	-0.01
602	爛死	-0.3
603	越來越少	-0.05
604	拒絕往來	-0.35
605	換掉	-0.01
606	還不	-0.01
607	過硬	-0.15
608	笑話	-0.01
609	沒熟	-0.15
610	沒人點	-0.05
611	好不好	-0.01
612	點錯	-0.01
613	夠爛	-0.2
614	逃漏稅	-0.2
615	一肚子火	-0.5
616	不吃	-0.03
617	天荒地老	-0.01
618	很油	-0.2
619	味道不好	-0.25
620	貴到	-0.1
621	言過其實	-0.15
622	油煙味	-0.2

623	飯太硬	-0.25
624	黏在	-0.01
625	打噴嚏	-0.01
626	品頭論足	-0.05
627	太普通	-0.06
628	失落感	-0.2
629	休想	-0.05
630	烙賽	-0.25
631	SOSO	-0.05
632	口乾舌燥	-0.05
633	小強	-0.5
634	忙碌	-0.05
635	憤式	-0.2
636	越賣	-0.01
637	越貴	-0.01
638	平淡無奇	-0.05
639	爛透	-0.5
640	大喊大叫	-0.01
641	名不符實	-0.3
642	沒有	-0.01
643	還凹	-0.2
644	黃色	-0.01
645	限時	-0.01
646	心不甘	-0.05
647	情不願	-0.05
648	很慢	-0.05
649	人擠	-0.01
650	敢收	-0.01
651	又爛	-0.25
652	在問	-0.01
653	沒上	-0.02
654	說壞話	-0.05
655	好歹	-0.01
656	差一點	-0.01
657	怎麼樣	-0.01
658	晃來晃去	-0.01
659	好聲好氣	-0.01

660	說不上	-0.02
661	裝傻	-0.13
662	扣分	-0.3
663	還兇	-0.15
664	千萬別	-0.05
665	硬要	-0.05
666	差評	-0.2
667	差成	-0.2
668	不打緊	-0.01
669	小塊	-0.01
670	不怎麼	-0.01
671	可不可以	-0.005
672	NOTGOOD	-0.25
673	不言	-0.01
674	比較忙	-0.01
675	沒問	-0.01
676	慢到	-0.01
677	崩壞	-0.05
678	愛理不理	-0.25
679	還等	-0.01
680	冷冷的	-0.005
681	無表情	-0.01
682	冷掉	-0.1
683	沒辣	-0.05
684	很怪	-0.2
685	貴了點	-0.02
686	只顧	-0.05
687	很重	-0.05
688	算錯	-0.15
689	多收	-0.15
690	嘻笑	-0.05
691	太吵	-0.15
692	料少	-0.25
693	遮三醜	-0.05
694	人少	-0.05
695	再加強	-0.3
696	吃不到	-0.05

697	87%	-0.05
698	辣太多	-0.01
699	更爛	-0.3
700	淡味	-0.05
701	太冰	-0.05
702	只差	-0.01
703	過多人	-0.05
704	點硬	-0.01
705	少得	-0.05
706	質低	-0.15
707	大霉	-0.2
708	低成本	-0.1
709	貴超	-0.05
710	價位偏高	-0.05
711	腎臟病	-0.02
712	無法控制	-0.01
713	不帶位	-0.05
714	一肚子	-0.05
715	烏氣	-0.2
716	真烏	-0.5
717	腸胃炎	-0.1
718	散掉	-0.1
719	最扯	-0.3
720	不修	-0.05
721	偏硬	-0.05
722	太扯	-0.3
723	嫌貴	-0.2
724	超薄	-0.01
725	誣陷	-0.5
726	漏水	-0.2
727	受罪	-0.35
728	渾然不知	-0.2
729	爆低	-0.15
730	半成品	-0.1
731	從沒	-0.05
732	還敢	-0.1
733	很窮	-0.05

734	差爛	-0.5
735	最爛	-0.5
736	還問	-0.05
737	打翻	-0.05
738	卻連	-0.05
739	貴爆	-0.3
740	退款	-0.3
741	全不	-0.05
742	不出所料	-0.1
743	重重	-0.01
744	油膩膩	-0.2
745	拒吃	-0.8
746	倒掉	-0.2
747	奇差	-0.3
748	酸言酸語	-0.2
749	噴口	-0.1
750	不太懂	-0.1
751	極爛	-0.65
752	媽的	-0.5
753	丟在	-0.05
754	走掉	-0.05
755	打鬧	-0.05
756	走鐘	-0.05
757	肥油	-0.05
758	彎腰	-0.001
759	路邊攤	-0.01
760	腹瀉	-0.05
761	怒罵	-0.5
762	廚餘桶	-0.01
763	騷味	-0.3
764	還少給	-0.25
765	還濺	-0.3
766	嚇死	-0.1
767	食物中毒	-0.3
768	變質	-0.4
769	不值一提	-0.15
770	價格昂貴	-0.1

771	沒做好	-0.05
772	吃不完	-0.01
773	爛的	-0.05
774	敗筆	-0.1
775	排錯	-0.05
776	太辣	-0.05
777	不爛	-0.05
778	太硬	-0.1
779	爛還帶	-0.1
780	不香且	-0.1
781	嗅到	-0.01
782	臭臉	-0.1
783	超硬	-0.15
784	命賤	-0.2
785	蠻差	-0.3
786	來路不明	-0.15
787	打消念頭	-0.1
788	未送	-0.05
789	怕人	-0.05
790	幹什麼	-0.05
791	高價位	-0.05
792	幻滅	-0.25
793	沒本事	-0.1
794	還燙	-0.01
795	打槍	-0.15
796	食材少	-0.25
797	加錯	-0.05
798	拖時間	-0.05
799	奧客	-0.3
800	屁孩來	-0.2
801	上樑不正下樑歪	-0.2
802	空蕩蕩	-0.02
803	未處理	-0.05
804	嬉皮笑臉	-0.05
805	變超差	-0.3
806	價格上漲	-0.1
807	常漏	-0.1

808	差且	-0.1
809	無法回答	-0.05
810	催趕	-0.1
811	小幹	-0.1
812	變苦	-0.05
813	極慢	-0.1
814	乾癟	-0.1
815	吃爛	-0.02
816	到爛	-0.01
817	笑死人	-0.02
818	大眼瞪小眼	-0.05
819	誣賴	-0.1
820	攻擊	-0.05
821	居然	-0.09
822	天灰灰	-0.05
823	overrated	-0.65
824	滿苦的	-9
825	太甜了	-0.95
826	懶妹	-5
827	CP 值算低	-0.2
828	CP 值蠻低	-0.2
829	低 CP 值	-0.2
830	不適合	-0.1
831	太肥	-0.1
832	卻要候位半小時	-3
833	沒打算吃這間	-0.5
834	人神共憤	-1.5
835	原本很期待的	-0.8
836	只適合拍照的網美店	-0.2
837	臭臭的味道	-0.1
838	人潮爆多	-8
839	端上來的時候是冷的	-5
840	口感很普通	-2
841	中午還不開	-2
842	有點澀	-0.6
843	收得偏乾	-5
844	美中不足	-0.05

845	不飽滿比較乾扁	-3
846	非朋友介紹完全不知道它的存在	-1
847	醬汁給的略少	-1
848	份量也不是太多	-2
849	皮有點過硬	-2
850	口感稍稍差了一點	-2
851	店很小	-0.5
852	意外的不嫩	-1
853	整體記憶點不高	-2
854	個人不太能接受	-0.2
855	吃久會有點膩	-4
856	沒有給我太大的驚艷感	-3
857	還有點小失望	-2
858	跟預想中的差太多	-0.6
859	冷氣不太涼	-0.6
860	等很久有點不 OK	-0.6
861	不會再有第二次	-0.1
862	選擇搭配困難	-2
863	zzz	0
864	休息	0
865	👤	0
866	📞	0
867	📅	0
868	不定休	-1
869	時段不定休	-6
870	時段才有供應	-1
871	沒特別驚艷	-0.05
872	不合	-0.1
873	絕不再訪	-0.3
874	不會再訪	-0.2
875	蠻多變化的	3
876	怎麼吃都好吃	3
877	👉	0
878	👉	0.1
879	💡	0.1

880	🍷	0.1
881	🍷	0.2
882	✓	0.1
883	🍷	0.1
884	R	0
885	U	0
886	S	0
887	🍷	0.5
888	🍷	0.1
889	🍷	0.1
890	沒什麼味道	-0.1
891	整個氣勢輸一半	-0.3
892	🍷	0.3
893	還是比較喜歡	-0.1
894	不會回訪	-0.4
895	☀	0.05
896	略鹹了些	-1
897	·	0.05
898	沒有味道	-0.1
899	比較偏酸	-1
900	融化很快	-1
901	不是我喜歡的	-0.5
902	會比較酸	-3
903	感覺可以再	-0.8
904	如果可以再	-0.6