

以社群文字評論預測 O2O 商務之滿意度

陳隆昇

朝陽科技大學

資訊管理系

e-mail :

lschen@cyut.edu.tw

簡婉婷

朝陽科技大學

資訊管理系

e-mail :

st313149@gmail.com

摘要

離線商務模式 O2O(Online to Offline)已成為許多企業與商人們關注的焦點，企業爭相進入 O2O 市場。有學者認為信任度是 O2O 成功的關鍵，多項研究結果指出消費者滿意度會直接影響消費者的信任度。因此，企業有必要了解在 O2O 環境中影響消費者滿意度的重要因子。以往研究都使用問卷調查，無法提供企業即時結果。此外，近年來社群的影響力不容小覷，客戶生成內容如消費者評論、評價，也已經成為提供其他消費者與廠商重要的消費資訊來源。因此本研究試圖從社群媒體文字評論中利用決策樹(Decision Tree)來預測顧客滿意度，並找出影響滿意度之重要因子。研究結果望能成為 O2O 市場中的企業與研究人員作為提升滿意度與信任度之參考依據。

關鍵詞：O2O、特徵選取、滿意度、決策樹。

Abstract

O2O (Online to Offline) business model has attracted lots of attentions of enterprisers and business people. Companies attempt to enter the O2O market. Some scholars considered that trust is the key factor to the success of O2O. A number of studies have shown that consumer satisfaction will directly affect the consumer's trust. Therefore, enterprises need to understand the crucial factors of influencing consumer satisfaction in O2O environment. In previous studies, researchers have used questionnaires to survey customer's satisfaction, but, it cannot provide instant information. In addition, the influence of social communities can not be underestimated in recent years. Online user-generated content such as consumer reviews, evaluations, have also

become one of important information for consumers and manufacturers. Consequently, this study attempts to predict customer satisfaction and discover the important factors using decision tree (DT) from textual reviews on social media. Hopefully, the results can be served as a reference for enterprises and researchers in the O2O market to improve their satisfaction and trust.

Keywords: O2O, Feature selection, Satisfaction, Decision Tree.

1. 前言

隨著資通訊科技的發展與資訊產物如網際網路、智慧型手機、物聯網和手機 APP 的普及，逐漸的改變了購物(He et al., 2016)、生活和社交的方式。電子商務成為生活中不可或缺的一環，而電子商務模式由 B2B、B2C、C2B 發展到現在的離線商務(Online to Offline, O2O)模式。與以往的電子商務模式不同，O2O 的模式可從線上消費或預訂到線下實體店面享受服務或消費。

O2O 現在已經成為許多企業關注的電商模式。O2O 為人們帶來需多的方便，讓消費者逐漸改變了消費方式和付款方式。相關文獻指出在中國市場上有許多企業關注於 O2O 發展，也有許多各式各樣的企業正努力進入 O2O 市場，因為 O2O 被認為是電子商務中的一塊大餅(Carsten et al., 2016; Zhong et al., 2015)。O2O 模式也促進了線上金流產業的發展(Zhang et al., 2015)。在美國與中國等地區都有發展出相關知名的 O2O 服務，如知名的旅遊網站 TripAdvisor 和 Airbnb，餐飲服務 OpenTable、Yelp 和 Dianping，運輸業則有 Uber 和 Zipcar。而其中 OpenTable 在 2016 年全球的餐廳覆蓋量已經成長到 3,700 多家餐廳 (OpenTable n.d.)。

從廠商的角度來看，要在 O2O 市場中吸引客戶將越來越困難，因為越來越多競爭者紛

紛進入 O2O 市場中(He et al., 2016)。以台灣著名的 EzTable 訂位平台為例,自從 2008 年開始到現在餐廳數量已經達到 1 萬間中高級餐廳。而全球著名的 Groupon 團購網,從 2008 年至 2015 年,特色商家已經超過了 950,000 家;而這些成長現象說明了 O2O 正在快速的發展,同時也會出現非常多的競爭者,這樣快速發展的環境下,Zhong 等學者認為混亂和信任不足將為 O2O 帶來很大的問題(Zhong et al., 2015),由此可知信任度將是 O2O 的一大問題;另外在 2016 年時 Castaldo 等學者提出在零售的藥品店中,消費者的滿意度與對於藥品店的信任度呈現正相關(Castaldo et al., 2016)。Wang 等學者也指出,在團購網站上消費者的滿意度會影響消費者對於團購網站的信任度與黏著度(Wang et al., 2016)。因此本研究認為消費者滿意度將會對消費者信任度有相當的影響,所以了解滿意度之影響因素對於 O2O 平台與廠商來說都是非常重要的。

而目前相關的文獻(Zhong et al., 2015; Li et al., 2016) 幾乎都使用問卷調查的方式進行研究,而問卷調查的方式,無法提供立即的顧客意見,且須曠日廢時花費大量人力物力進行資料收集。此外,也會因為抽樣方法的問題,造成樣本的誤差。

近來,社群媒體(Social media)對電子商務消費者的影響力日漸上升。客戶生成內容如消費者文字評論、評價,也已經成為提供其他消費者與廠商重要的消費資訊來源之一(Guo et al., 2017)。故本研究試圖利用客戶生成內容,也就是客戶評論,利用決策樹(Decision Tree)來預測顧客滿意度,並找出影響滿意度之重要因子。研究結果望能成為 O2O 市場中的企業與研究人員作為提升滿意度與信任度之參考依據。此外,OpenTable 在 O2O 電子商務環境中,是一個經典代表。因此本研究將以 OpenTable 中之任一餐廳之評論為例,找出在 O2O 中影響客戶滿意度的重要因子。廠商和網路平台便可以了解到影響客戶滿意度的重要因子,並且調整未來營運的目標和策略。

2. 文獻探討

2.1 O2O(Online to Offline)

O2O 電子商模式可以被定義為任何設備所提供的線上商務和線下實體之間的無縫購物體驗。換句話說,也就是消費者可以透過線

上網路平台的方式預訂或訂購想要的服務或餐廳訂位等,再到線下實體店面享受服務或用餐。這樣的消費模式吸引許多企業進入 O2O 市場(Carsten et al., 2016; Zhong et al., 2015)。雖然 O2O 市場帶來了許多新的機會,但市場的混亂和信任不足將會是 O2O 的一大問題(Zhong et al., 2015)。

目前 O2O 相關文獻幾乎都使用問卷的方式來作資料收集,其中可能隱含了一些問題,其中包含了問卷無法確定填答者是否真正了解 O2O 議題,以及無法確定填答者是否真正使用過 O2O 服務等。故本研究欲利用客戶生成內容—評論,來作為研究資料,如此一來本研究便可以確定評論填答者為真正有使用過 O2O 服務的人,也可以避免一些因問卷填答者對於問項意思表達不清而使研究者無法得到準確答案等問題。

2.2 星等與滿意度

根據以往的文獻,線上產品的評級與評論,是消費者、零售商和製造商商品相關訊息的來源之一(Guo et al., 2017),有許多文獻一致提出,在各種情況下,評論與星等都會影響到銷售情況(e.g. Lin et al., 2011; Park et al., 2007),然而 Engler 學者認為產品星等代表的意思是客戶對產品的滿意度(He et al., 2016),因此本研究認為線上的評論星等,可以當作是顧客滿意度指標。在 OpenTable 的評論中可以看到消費者填寫星等評級,故本研究將星等視為顧客滿意度指標。

2.3 決策樹(Decision Tree, DT)

決策樹是資料探勘中常見的特徵選取方法之一,主要的特色是容易理解和實現,並且能夠在較短的時間內對大量的數據建立可行的決策樹。

根據以往的文獻研究,決策樹成功地運用在許多地方,如 Everaert 學者將決策樹用於建立偵測水質適用性之模型(Everaert et al., 2016),而在 2014 年 Zhanga 學者提到特徵選取的方法中決策樹的準確率高於其他方法(Zhang et al., 2014),因此本研究將決策樹作為特徵選取的方法,並使用 C5.0 演算法來建構決策樹。

3. 研究方法

本研究使用網站中使用者評論和評價進行資料分析，實驗流程如圖 1，詳細說明如下：

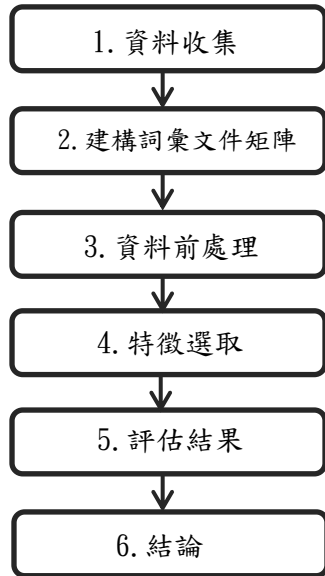


圖 1. 研究流程圖

步驟 1：資料收集

本研究以 OpenTable 網站作為主要資料收集對象，並使用爬蟲工具進行網頁爬蟲，將網頁中的評論資料例如會員帳號、日期、評論、標籤評論、細項星等及總星等，其中評論指的是一整文字敘述，而標籤評論為使用幾個精確的詞來評論用餐與服務。而本研究將資料分為兩個部分，評論與標籤評論進行實驗。

步驟 2：建構詞彙文件矩陣

本研究收集之文本資料後，將每一份評論資料以特徵向量的形式表示(Zhang et al., 2008)，也就是建置詞彙文件矩陣 (Term-document matrix, TDM) 並且計算出每一個詞彙特徵之權重 TF-IDF (Term Frequency - Inverse Document Frequency) 如來分析單詞頻率。

TF-IDF 計算步驟如下：

Step 2.1：算出詞頻 TF (Term Frequency) 是指算出某個詞語在該文件中出現的頻率，如公式(1)。

$$tf_{ij} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (1)$$

上式子中分子 $n_{i,j}$ 為該單詞在該文件中使用的次數，而分母則是在文件中所有字詞出現次數之合。

Step 2.2：接著算出逆向文件頻率 IDF (Inverse document frequency) 如公式(2)。

$$idf_i = \log \frac{|D|}{|\{j: t_i \in d_j\}|} \quad (2)$$

而以上的式子中分子 D 為文件總數，分母中的 t_i 就是該單詞在文件總數 D 中所出現的文件數。

Step 2.3：最後再算出 TF-IDF 如公式(3)，將 TF 與 IDF 相乘便可算出 TF-IDF 之權重值。

$$tfidf_{i,j} = tf_{i,j} \times idf_i \quad (3)$$

步驟 3：資料前處理

此階段根據本研究所採用的特徵選取方法來對資料進行格式轉換以及數值化處理並將資料進行二元分類，而本研究將類別(class label)資料分為消費者滿意度高於平均與低於平均兩類。最後再使用五折交叉驗證法(5-fold cross-validation)，將資料分為五等份，讓實驗結果更具穩健性。

步驟 4：特徵選取

決策樹(Decision Tree, DT)

本研究將決策樹作為特徵選取的方法並使用 C5.0 演算法來建構決策樹。另外，因考量到資料類別分布不平均因素，避免類別失衡問題。故使用考量分類錯誤成本並幾何平均值(G-mean, GM)以及準確率(Accuracy)為評估指標。最後將每個特徵之利用率進行累加，而其特徵利用率單一屬性在決策樹規中所利用的比率。例如屬性 A 在 5 調規則中出現 3 次，此屬性的利用率則為 60%(3/5)。

建立決策樹的步驟如下：

Step 4.1：決定決策樹之自變數和因變數。

Step 4.2：給定分類錯誤之成本。

Step 4.3：採用 C5.0 演算法建構決策樹。

Step 4.4：將每個 fold 之特徵利用率進行累加，並據以挑選重要特徵。

步驟 5：實驗結果

經過決策樹分類後計算出 5 次實驗中，找出利用率較高的決策節點之變數，並分析其詞意，便可得知影響滿意度之重要因素。並且比較出評論與標籤評論之分類準確率及分類效能何者較佳。

4. 實驗結果

4.1 衡量指標

在文本實驗中，最常使用的二元分類表也就是 Confusion matrix 來評估分類效能如表 1。表中 TP、FP、FN、TN 分別定義如下：

- (1) TP 指的是樣本為正向(高於平均)且分類也是正向(高於平均)。
- (2) FP 指的是樣本為負向(低於平均)但分類成正向(高於平均)。
- (3) FN 指的是樣本為正向(高於平均)但分類成負向(低於平均)。
- (4) TN 指的是樣本為負向(低於平均)且分類也是負向(低於平均)。

表 1.二元分類表

Predicted Actual	Positive (滿意度高於平均)	Negative (滿意度低於平均)
Positive (滿意度高於平均)	TP	FN
Negative (滿意度低於平均)	FP	TN

而本研究中以 PA、NA 來計算 G-Mean(GM)，並運用 GM、準確率 Accuracy 與訓練時間來評估分類效能。其中 PA、NA 與 GM 計算方式與定義如下：

- (5) PA 指的是正確判斷滿意度高於平均之樣本，公式如下：

$$PA = \frac{TP}{TP + FN} \quad (4)$$

- (6) NA 指的是正確判斷滿意度低於平均之樣本，公式如下：

$$NA = \frac{TN}{FP + TN} \quad (5)$$

- (7) GM 則表示判斷該分類之 PA 與 NA 之間的幾何平均數來評估分類效能，其公式如下：

$$GM = \sqrt{PA \times NA} \quad (6)$$

4.2 資料收集及分析

本研究使用 import.io 爬蟲工具對 OpenTable 進行爬蟲，並選取其中位於倫敦的“The Five Fields”餐廳以及評論時間為 2013 年 1 月到 2016 年 9 月 28 日來作為實驗資料，並且本實驗蒐集之評論資料如圖 2，圖中方框 1 為總星等，方框 2 為評論資料，方框 3 為標籤評論。



圖 2.網站評論示意圖

而網路評論共收集 363 筆文字評論以及標籤評論，本研究將資料分為兩個部分，第一部份評論資料如表 2，而第二部分的標籤評論刪除空白資料後，剩餘 324 筆資料如表 3 所示。

表 2. 評論資料之樣本統計

項目	分布
評論時間	2016 年(20%) 2015 年(29%) 2014 年(34%) 2013 年 (16%)
評論者身分	一般使用者 (86%) Vip 使用者 (14%)
滿意度	5 等星(88%) 4 等星(10%) 3 等星(2%) 2 等星(0.2%) 1 等星(0%) 高於平均星等(88%) 低於平均星等(12%)

表 3. 標籤評論資料之樣本統計

項目	分布
評論時間	2016 年(15.74%) 2015 年(30.55%) 2014 年(36.41%) 2013 年 (17.28%)
評論者身分	一般使用者 (87.34%) Vip 使用者 (12.65%)
滿意度	5 等星(89.19%) 4 等星(9.87%) 3 等星(0.92%) 2 等星(0%) 1 等星(0%) 高於平均星等(89.19%) 低於平均星等(10.8%)

4.3 建立詞彙文件矩陣

收集到文本資料後，本研究使用 QDA Miner 分別對評論資料以及標籤評論進行詞語

切割，抽取出詞彙特徵，並對文本資料進行詞彙文件矩陣的建構 (TDM) 如圖 3 及圖 4 分別為評論之詞彙文件矩陣示意圖與標籤評論之詞彙文件矩陣式意圖。評論之詞彙文件舉陣中共有 635 個文字特徵，而標籤評論之詞彙文件矩陣共有 59 個特徵。

共 635 個文字特徵

	ABSOLUTE	ABSOLUTELY	ACCOMMODATING	ACCOMPANIED	ADDED	ADDITION	ADDITIONAL	ADVENTUROUS	AHEAD	AMAZING	...
1	0	0	0	0	0	0	0	0	0	0	...
2	0	0	0	0	0	0	0	0	0	0	...
3	0	0	0	0	0	0	0	0	0	0	...
4	0	0	0	0	0	0	0	0	0	0.819544	...
5	0	0	0	0	0	0	0	0	0	0	...

共 363 筆評論實

圖 3.評論之詞彙矩陣

共 59 個文字特徵

	ACCESS	AFTERNOON	ANNIVERSARIES	AUTHENTIC	BIRTHDAYS	BRUNCH	BUSINESS	COCKTAILS	CONVERSATION	CREATIVE	CUISINE	DATE
1	0	0	1.860937	0	1.860937	0	0	0	0	1.860937	1.860937	0
2	0	0	0	0	1.860937	0	0	0	0	0	0	1.860937
3	0	0	0	2.258877	1.860937	0	0	0	2.258877	0	0	1.860937
4	0	0	1.860937	0	1.860937	0	0.547049	0	0	0	0	1.860937
5	0	0	0	0	0	0	0.547049	0	0	0	0	0

共 324 筆評論實

圖 4.標籤評論之詞彙矩陣

4.4 特徵選取

4.4.1 考量分類錯誤成本的決策樹

在此階段由表 4 可以看出在不設定分類錯誤成本(cost 為 1)時 NA 與 GM 值很低，表示有資料不平衡的問題，故本實驗找出分類錯誤成本設定值較佳者，作為實驗設定值；由表 4 評論之分類錯誤成本表中可以看到當分類錯誤成本設定為 5 的時候，其平均 GM 值最高而平均準確率平均也高達 80.42%，因此本實驗將評論實驗之分類錯誤成本設定為 5；而表 5 標籤評論之分類錯誤成本表中可以看到當成本質設定為 6 時，其平均 GM 值最高，而平均準確率位則有 70.4%。因此本實驗將標籤評論實驗之分類錯誤成本設定為 6。

表 4. 評論之分類錯誤成本表

	cost1	cost2	cost3	cost4	cost5	cost6	cost7
	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)
PA(%)	95.90 (0.01)	93.02 (0.02)	91.46 (0.05)	85.32 (0.04)	86.01 (0.07)	85.33 (0.05)	74.76 (0.12)
NA(%)	10.61 (0.08)	20.33 (0.15)	12.56 (0.09)	34.06 (0.18)	42.39 (0.13)	27.94 (0.08)	35.11 (0.24)

GM(%)	27.96 (0.17)	37.81 (0.23)	29.75 (0.18)	52.31 (0.16)	59.73 (0.08)	48.43 (0.08)	48.50 (0.14)
Accuracy (%)	85.38 (2.69)	84.3 (4.42)	81.84 (6.1)	78.98 (4.57)	80.42 (4.7)	78.24 (3.32)	69.68 (8.49)
Time	0.2	0.14	0.2	0.2	0.16	0.14	0.16

表 5.標籤評論之分類錯誤成本表

	cost1	cost2	cost3	cost4	cost5	cost6	cost7
	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)	平均值 (標準差)
PA (%)	100.00 (0)	97.24 (0.03)	95.84 (0.04)	87.15 (0.12)	78.15 (0.13)	73.30 (0.09)	71.23 (0.11)
NA (%)	0.00 (0)	2.50 (0.06)	5.71 (0.08)	11.43 (0.12)	28.57 (0.2)	47.14 (0.06)	44.29 (0.18)
GM (%)	0.00 (0)	6.89 (0.15)	14.52 (0.2)	22.98 (0.22)	39.98 (0.25)	58.66 (0.06)	54.38 (0.13)
Accuracy (%)	88.92 (0.57)	86.78 (2.72)	85.86 (2.31)	82.20 (8.14)	72.54 (9.48)	70.40 (8.28)	68.24 (8.84)
Time (%)	0.02	0.02	0.06	0.04	0.02	0.02	0.04

4.4.2 提取重要特徵

本研究中使用決策樹 DT 以及使用五折交叉驗證法來篩選出重要因素。第一部份本研究以評論來探討影響影響使用者滿意度之因素，也就可以了解到使用者在餐廳用餐或接受服務時會重視部份如表 6；表 6 為 DT 所使用之決策節點，本研究依據 5 Fold 中所使用之節點使用率超過 200%之因素作為該部份之重要因素。而其平均 G-Mean 為 59.73 與平均準確率為 80.42%，且平均訓練時間為 0.16 秒。

表 6. 評論之 DT 決策節點

單位:百分比

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	累加使用率
LONDON	100	94		89	99	382
GOOD	100		100		100	300
REAL	100	100		97		297
FABULOUS	100			80	100	280
STARTER		100	100		80	280
EYES	97	100		81		278
STAFF	71	91		100		262
PREFER			78	100	80	258
DELICIOUS	98	2	100	7	2	209
SMALL		100		100		200
GRAS		100		100		200
PERFECT		100		100		200
⋮	⋮					
⋮	累加使用率<200 之特徵(共 59)					

GM	69.9	63.5	51.8	62.3	51.1	
Accuracy	76.4	87.7	76.6	82.2	79.2	
Time	0.1	0.2	0.2	0.2	0.1	

第二部份本研究以標籤評論作為探討影響影響使用者滿意度之因素，如表 7；表 7 為 DT 所使用之決策節點，本研究依據 5 Fold 中所使用之節點使用率超過 200%之因素作為該部份之重要因素。而其平均 G-Mean 為 58.66 與平均準確率為 70.40%，且平均訓練時間為 0.02 秒。

表 7. 標籤評論之 DT 決策節點

單位:百分比

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	累加使用率
INGREDIENTS	100	100	100	85	97	482
OPTIONS	68	100	100	100	100	468
LUNCH	100	72	71	100	100	443
LIST	79	95		100	100	374
CREATIVE		100		70	66	236
GASTRONOMY	100	14	5		15	134
PEOPLE		100		43	29	172
POST	100					100
KID		89				89
CREATIVE	62					62
LOCAL	21				13	34
COCKTAILS				27		27
FRIENDLY		4				4
GM	49.8	56.4	66.6	58.9	61.5	
Accuracy	56.2	70.8	75.4	76.9	72.7	
Time	0.0	0.0	0.0	0.0	0.0	

4.5 實驗結果

根據實驗結果顯示，本研究根據每一個部份篩選出累加使用頻率大於等於 200 之因素。而實驗的第一部份則篩選出 12 個特徵分別為，LONDON, GOOD, REAL, FABULOUS, STARTER, EYES, STAFF, PREFER, DELICIOUS, SMALL, GRAS, PERFECT 而根據這些文字特徵，可以歸納出 Good、Fabulous、Delicious、Small、Perfect 通常是體驗詞彙，而 Real、Starter、London、Eyes、Staff、GRAS 都是用餐經驗的描述，這部分較不容易被解釋；而第二部份標籤評論所篩選出的特徵共有二個，分別為 INGREDIENTS、OPTIONS、LUNCH、LIST、CREATIVE，而根據原始標籤評論，可以歸納出 Ingredients、Lunch、Creative 與 List 都屬於描述商品的標籤。由上述實驗結果便可

以整理出消費者體驗、與消費經驗以及產品都是影響滿意度的重要因子。

從兩個實驗中可以看出，若使用標籤評論作為找出影響滿意度之重要特徵之資料，會較評論資料來的準確；標籤評論較為具體、精確，且消費者表達方式較為一致，少了敘述中的冗詞贅字；而從維度的角度中也可以發現，評論的文字特徵維度較廣且混亂，訓練時間也比標籤評論長，而訓練出來的決策樹也與標籤評論差不多；因此本研究認為使用標籤評論作為探勘資料的效用較評論資料來的好。

5. 討論與結論

本研究的目的是為了探討在消費者在 O2O 環境中滿意度的影響因素，並利運用評論來作資料分析，並從中特徵選取出重要影響因素。結果指出消費者的用餐體驗以及實體餐點的好壞都是影響滿意度的重要因素，藉此餐廳及廠商人員即可作為參考依據，如多注意實際的餐點呈現，以及消費者消費經驗，以提高消費者滿意度。表 8 為本實驗探討平台的 Vip 會員使用者對餐廳的討論通常有這幾個特徵。

表 8.Vip 使用者之標籤評論

	標籤評論特徵
Vip 使用者	ROMANTIC, OCCASION, BUSINESS
所有使用者	INGREDIENTS、OPTION、LUNCH

而 O2O 平台開發人員便可以了解到其 Vip 顧客對於餐廳所重視的因素，便可以作為推薦餐廳之依據，如 Vip 客戶都很重視用餐體驗，藉此網站便可以選出體驗滿意度較高之餐廳作為推薦。

致謝

本文受到行政院科技部專案計畫部分資助(契約編號 MOST 104-2410-H-324-009-M Y2)，作者在此表達感謝之意。

參考文獻

- [1] Castaldo, S., Grosso, M., Mallarini, E., Rindone, M., "The missing path to gain customers loyalty in pharmacy retail: The role of the store in developing satisfaction and trust," *Research in Social and*

- Administrative Pharmacy*, vol. 12, pp. 699–712, 2016.
- [2] Carsten, P., “China’s Wanda, Tencent, Baidu to set up \$814 million e-commerce company.”, 2016 from : <http://www.reuters.com/article/us-wanda-tencent-baidu-idUSKBN0GT04020140829>
- [3] Engler, T.-H., Winter, P., Schulz, M., “Understanding online product ratings: A customer satisfaction model,” *Journal of Retailing and Consumer Services*, vol. 27, pp. 113–120, 2015.
- [4] Everaert, G., Bennetsen, E., Goethals, P.L.M., “An applicability index for reliable and applicable decision trees in water quality modeling,” *Ecological Informatics*, vol. 32, pp. 1–6, 2016.
- [5] Guo, Y., Barnes, S.-J., Jia, Q., “Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation,” *Tourism Management*, vol. 59, pp. 467-483, 2017.
- [6] He, Z., Cheng, T. C. E., Dong, J., Wanf, S., “Evolutionary location and pricing strategies for service merchants in competitive O2O markets,” *European Journal of Operational Research*, vol. 254, no. 2, pp. 595-609, 2016.
- [7] Lin, C.-L., Lee, S.-H., Horng, D.-J., “The effects of online reviews on purchasing intention: the moderating role of need for cognition,” *Social Behavior and Personality an international journal*, vol. 39, no. 1, pp. 71–81, 2011.
- [8] Li, X., “Research on construction of takeout O2O platform service quality evaluation system,” *The 13th International Conference on Service Systems and Service Management (ICSSSM)*, 2016.
- [9] OpenTable from : <https://restaurant.opentable.com/>
- [10] Park, D.-H., Lee, J., Han, I., “The effect of on-line consumer reviews on consumer purchasing intention : the mode rating role of involvement,” *International Journal of Electronic Commerce*, vol. 11, no. 4, pp. 125–148, 2007.
- [11] Wang, W.-T., Wang, Y.-S., Liu, E.-R., “The stickiness intention of group-buying websites: The integration of the commitment–trust theory and e-commerce success model,” *Information & Management*, vol. 53, pp. 625–642, 2016.
- [12] Zhang, B., Huang, L., “The research status of O2O industry analysis Meituan, for example,” *International Conference on Logistics, Informatics and Service Sciences (LISS)*, 2015.
- [13] Zhong, W., “Service Recommendation Method on Multiple Dimension O2O,” *International Conference on Intelligent Transportation*, pp. 713-716, 2015.
- [14] Zhang, Y., Wang, S., Phillips, P., “Binary PSO with mutation operator for feature selection using decision tree applied to spam detection,” *Knowledge-Based Systems*, vol. 64, pp. 22–31, 2014.
- [15] Zhang, C., Zuo, W., Peng, T. and He, F., “Sentiment Classification for Chinese Reviews Using Machine Learning Methods Based on String Kernel,” *Proceedings of the Third International Conference on Convergence and Hybrid Information Technology*, vol. 2, pp. 909-914, 2008.