

辨識科技產品操弄評論之研究

陳隆昇
朝陽科技大學
資訊管理系
副教授
lschen@cyut.edu.tw

林瑞裕
朝陽科技大學
資訊管理系
研究生
s9914640@cyut.edu.tw

摘要

越來越多消費者在購買任何形式產品之前，都會根據部落格、社群網站或電子商務網站上他人使用經驗與評論意見，做為購買產品的主要參考。相關研究證明，比起現有網路廣告，消費者認為個人的評論更值得信賴。因此，越來越多的企業試圖寫出有利銷售量的操弄顧客評價，以影響消費者購買行為。近來，操弄評論的辨識已成為電子商務中熱門議題之一，唯辨識評論為真實亦或操弄有其難度。因此，本研究導入可讀性等特徵，利用支持向量機(Support Vector Machine)來提升操弄評論分類績效，並試圖找出評論操弄的重要特徵。並以智慧型手機產品 Samsung Galaxy Note II 使用者評論資料為範例，驗證方法之有效性。**關鍵詞**：評論操弄、語意分類、線上口碑、支持向量機、文字分類

Abstract

More and more customers purchase any kind of products or services depending on personal experiences and comments from blogs, virtual communities, and e-business websites. According to related researches, customers think these personal opinions are more reliable than the existing internet advertisements. Therefore, some enterprises attempt to manipulate comments to influence customer behaviors and increase their sales volume. Recently, to identify manipulated reviews has become one of hot issues in the area of e-business. But, it's a difficult task to effectively classify manipulated reviews. Therefore, this study employs Support Vector Machines (SVM) to improve the classification performance of review manipulation by introducing some additional attributes such as "readability". In addition, we tried to discover the important attributes of identifying manipulated reviews. Finally, we used a real users' comments data regarding the famous smart phone "Samsung Galaxy Note II" to testify the effectiveness of the proposed method.

Keywords: Review Manipulation, Sentiment Classification, Online Word-of-mouth, Support Vector Machines, Text Classification

1. 緒論

由於 Web2.0 的興起，許多新興媒體如部落格、社群網站、論壇或電子商務網站吸引大量使用者發表個人想法與評論文章，產生許多帶有語意情感的文本資訊，進一步影響了網路使用者的想法與思維[10]。近來，許多公司企業紛紛透過這些新興媒體來提高其新產品能見度，希望藉由網路使用者的評論回饋，來快速擴散其產品資訊。這些線上口碑式(Online Word of Mouth)分享的做法，一直被認為是一個影響消費者行為的重大因素[31]。倘若潛在的客戶透過這些評論產生好印象，那麼這些評論是非常有影響力的，因為許多有意願購買產品的消費者往往都會依賴這些評論來做為購買決策的依據[30]。

經由網路的便利性，許多使用者透過上述的平台進行心得與意見交流。而信譽是企業的寶貴財富，需要長時間努力與投入和重視與顧客之間的關係[31]，所以企業的正向評論與聲譽是很昂貴的。當使用者的正向評論非常多時，其負向評論就顯得更加重要，因為少量使用者的負向評論可能會對企業帶來致命傷害，也因此增加了語意分類(Sentiment Classification)的重要性[20]。然而，企業也注意到了這個問題的存在，於是試圖寫出有利銷售量的操弄顧客評價與不實評論，以影響消費者購買行為。相關研究已證明網路上的評論操作是存在的[26、28]。

對於網路使用者評論操弄的研究，目前相關研究[28、30]的資料收集都來自於數年前，而操弄的文章則會隨著時間增加逐漸減少，這歸咎於時間會帶來更多真實使用者加入回覆評論，而廠商對其操弄評論成本效益會因大幅增加而減少操弄[26、28]。然而消費者也會因產品的推陳出新，對其舊有產品也會減少購買

意願。再加上現今 3C 產品因科技的進步與互相競爭導致產品生命週期短[4、5]。也因產品生命週期的短暫，初期上市時間內，公司企業紛紛於產品發表前後，利用形象廣告塑造產品在消費者心中的美好憧憬。此時產品發表初期使用者評論將會影響到潛在消費者的購買決策。因此產品於發表上市不久之時真實產品使用者不多，公司企業便能利用此段時期，導入操弄評論以至增加產品正向評論。所以，如能盡快找出真實使用者的回覆評論或了解哪些為操弄評論，這對於潛在的消費者於購買決策時能有所幫助，可以減少許多不確定性與得知產品的真實情況。

因此，本研究希望藉由導入可讀性等特徵，利用支持向量機(Support Vector Machine)來提升操弄評論分類績效，並試圖找出評論操弄的重要特徵，進而幫助到潛在消費者能獲得真正有用的評論資訊，減少預期購買產品不安定因素。並以智慧型手機產品 Samsung Galaxy Note II 使用者評論資料為範例，驗證方法之有效性。

2. 文獻探討

本章節將介紹操弄評論與找出操弄特徵之方法，依序為「操弄評論」、「產品操弄評論之要素」、「支持向量機」，分述如下。

2.1 操弄評論

透過網路的便利性，許多使用者透過網路平台進行心得與意見的交流。因此使用者的正向、負向評論會對企業形象影響很大，而信譽是企業的寶貴財富，需要長時間努力與投入和重視與顧客之間的關係，所以正向的評論與聲譽是很昂貴的。雖然參考網路上評論資訊購物的使用者越來越多，但許多消費者仍害怕擁有負面的經驗。因此負向的評論會對企業造成傷害，並帶來消費者對其公司服務或產品的負面觀感 [26、31]。所以當正向評論非常多時，其負向評論就顯得更加重要，因為少量的負向評論可能會對企業帶來致命傷害，也因此增加了語意分類(Sentiment Classification)的重要性 [20]。比較專家與顧客的推薦，發現比起現有網路廣告，消費者認為個人的評論更值得信賴。同時也有更多的證據顯示，消費者個人使用經驗的評論會影響潛在消費者對產品的選擇 [31]。

企業也注意到了這個問題的存在，於是開始尋找改善方法。越來越多的企業試圖寫出有

利銷售量的操弄顧客評價，以影響消費者購買行為。這些企業透過撰稿者操弄評論與網站內部管理配合機制，同時進行負向評論控制，並增加正向評論是目前最常見到的操弄方法。而相關研究也已證明網路上的評論操作是存在的 [26、28]。例如，在音樂產業中已被證實聘請專業人員撰寫有利於消費者閱讀的資訊，以促進銷售的新專輯 [14]。在臺灣 Mobile 01 電子商務網站上，Samsung 與 HTC 被證實是一個實際操弄案例 [2]。另一華擎科技公司徵工讀生寫手之案例 [3]，都可看出網路上的評論操弄確實存在。

此外，經相關研究顯示 Amazon、Barnes 與 Noble 網站內的使用者評論也都曾經過供應商、出版商與撰稿者的配合進行評論操弄。面對這些操弄的評論，網路使用者雖然了解操弄的存在，但缺乏正確的資訊來分辨真實評論與操弄評論。雖然一般消費者會試著去了解產品相關的服務與政策，以進一步減少不確定性。然而對於產品本身的不確定，則會依照網路上的評論來預期產品的質量，再降低其心中對產品未知部分的不確定性，最後做出購買決策。所以過往使用者回覆的評論將成為這些未來消費者重要的影響力之一 [26、28]。

Hu et al. [27]發現，網路上約有 10.3% 的產品操弄評論，因此 Hu et al. [28]提出了一個方法來確定操弄評論，並建立了一個自由給評的線上書評系統。透過簡單的統計方法來確定操弄評論，與評估消費者如何應對產品的操弄評論。操弄評論的辨識已成為電子商務中熱門議題之一，唯辨識評論為真實亦或操弄有其難度 [26]。

根據 Archak et al. [25]研究，從回覆討論的文章中會有那些特點、那些重要的特徵是消費者所重視與探討的，因此他們針對產品的回覆評論進行特徵探勘並找出其重要特徵。再根據 Hu et al. [27]研究，非操弄的評論會因每個人所受的教育、文化、職業等有所不同的呈現，而操弄評論會因操弄評論的人員持續觀測其他真實使用者的評論，並引用真實使用者的評論內容後再發布與回覆。另外也提到操弄評論的寫作風格會有兩項屬性，一是情緒 (sentiments)，一是可讀性 (readability)，因此我們可以透過上述方法從評論當中找出其操弄特徵。最後，我們使用了上述學者們所提到檢測操弄屬性的方法，在此做為本研究實驗檢測操弄評論的要素。

2.2 產品操弄評論之潛在要素

本小節試圖由文獻中歸納潛在評論操弄之因素。根據 Archak et al. [25]的發現，「產品特徵」可能是一個潛在因素。依其討論可得出，操弄評論會針對其探討主題之設備與物品之特有名詞和規格等相關特徵，會於無形中增加這些關鍵字詞出現次數。

第二個潛在因素為「語意特徵」（在此即正、負向語意評論）。在網頁探勘領域裡辨識文章中，網路使用者所表達的語意評論，是語意分類中重要議題之一[11]。普遍來講，語意分類的主要目的是在於擷取文章中，消費者對於特定服務或產品的評論，並辨識其語意情感[29]。因此，語意分類可以辨識出部落客撰文時所表達的情緒，以協助公司企業更加謹慎的回應消費者意見或加以改進產品與服務。透過研究發現，許多操弄文章會使用情感語意來引導讀者激昂憤慨或投以感性來說服引導讀者[27]。

其餘潛在的因素，本文納入「適讀性(Readability)」，主要的想法是假設操弄的評論，應該非常容易閱讀。根據 Dale 與 Chall's [13]的定義，係指『在給定一個印刷品(如書籍、雜誌)中的要素總和，會影響到一群閱讀人。在何種程度上他們認為「最佳的閱讀速度」和「覺得很有意思」才是成功』。Klare [19]提出的可讀性可判斷因素包括：句子的平均長度、句子中不同詞彙的多寡、詞彙語法變化的複雜性。

在文本資料中包含了句長與詞彙，句長與詞彙是衡量文本資料難度的重要參考依據。因此，閱讀教學領域相關的研究者便利用這兩項主要的指標，進而發展適讀性公式。根據文獻記載，適讀性公式已創建超過 200 多種[15]。其中著名的適讀性公式包括如：Flesch Reading Ease Formula (1948), Flesch-Kincaid Grade Formula (1976), the Fry Readability Graph (1977), Dale and Chall's Formula (1944), the SMOG formula, etc. [7]，本研究使用 Flesch Reading Ease formula 適讀性公式[16]。此公式使用句子平均長度，與每一個單字的平均音節數。分數從 1 至 100 分，30 分為“非常困難”和 70 分“簡單”。一般標準文件的分數落在 60 至 70 分。

另一個適讀性量測特性為「學校年級水平(Flesch-Kincaid Grade Level)」又稱為“Flesch-Kincaid”[17]。利用其計算公式將文章等級轉換為美國學校年級程度，表示文章為該年級學生所能閱讀與理解。主要以平均句子的長度與每一個單字的平均音節數做計算。一般

標準文件分數落在 7.0 到 8.0 等級。

最後一個適讀性屬性為「相異詞比率(Type-Token Ratio)」[21]。一篇能廣為人們所閱讀的文章，除了要考慮到使用字詞的頻率，也要考慮字詞在文章中所佔有的比率。舉例來說，作者 A 寫了一篇 100 字的文章，當中用到了 40 個不同的字詞，那麼這篇文章的 TTR 為 0.4；作者 B 也寫了一篇 100 字的文章，當中用到了 50 個不同的字詞，那麼這篇文章的 TTR 為 0.5。兩者相較下，我們可以說作者 A 的文章用字變化小，閱讀難度比作者 B 的文章簡單[6]。

2.3 支持向量機

根據現有文獻，機器學習已被視為語意分類的一個有效方法。例如：Dave et al. [18]發展了一種自動分類正向、負向評論和數個實驗特徵選取與評分的功能。Whitelaw et al. [12]提出了一個語意分析在 SVM 上提取分析評估組(extracting and analyzing appraisal groups)。Abbasi et al. [9]提出了一個 EWGA 在 SVM 上使用語意探測來選取電影評論的特徵分類法。因此，本研究採用 SVM 分類器來提高操弄評論的辨識效能。

機器學習方法(Machine Learning Approaches)像統計方法一樣，為尋求一個最佳模式來符合測試資料。不同於之處在於機器學習方法大多數利用資料自動化學習過程，自動歸納出分類規則及建立模組。SVM 是眾多機器學習方法中，最為多人使用。支撐向量機(Support Vector Machines, SVM)於 1995 年由 Vapnik 所提出，主要核心來自於統計學習理論中結構風險最小化原理(Structural risk minimization, SRM)。SVM 主要是使用一個區分超平面(Separating Hyperplane)以最大的邊界超平面作為決策面，用來處理兩類別或類別資料問題的一種二元分類器(Binary classifier)。

3. 研究方法

本節將介紹本研究所提出方法的實施流程。首先，本研究將分成二個階段。第一階段，根據可讀性、語意特徵、產品特徵，定義出六項潛在的評論操弄屬性。第二階段，根據所找出的產品操弄特徵建構文件詞彙矩陣進行實驗。

3.1 定義產品操弄文章之潛在屬性

可讀性特徵

利用 Readability Analyzer 分析本實驗資料集，並找出可讀性分數、學校年級水平與閱讀難易度之數據。適讀性分析程式(Readability Analyzer)是由 Xu, Jiabin & Yunlong Jia. [33]學者，引出微軟公司 Microsoft Office 其內建“拼字及文法檢查”功能撰寫而成。其撰寫目的在於可進行批次文章“拼字及文法檢查”功能，快速計算出本實驗所需三項數據：Flesch Reading Ease、Flesch-Kincaid Grade Level、Type-Token Ratio，其公式依序如公式(1)~(3)。

適讀性(Flesch Reading Ease)

$$206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}} \right) \quad (1)$$

學校年級水平(Flesch-Kincaid Grade Level)

$$0.39 \left(\frac{\text{total words}}{\text{total sentences}} \right) + 11.8 \left(\frac{\text{total syllables}}{\text{total words}} \right) - 15.59 \quad (2)$$

相異詞比率(Type-Token Ratio)

$$TTR = \frac{\text{Types}}{\text{Token}} \quad (3)$$

其中，Types 詞彙的總類，如 10 個“do”的出現則歸為一個類別。Token 詞彙的總數，如“How are you”，即為三種詞彙類別。

語意特徵

使用 Hu and Liu [22]學者的語意特徵辭典，找出本研究資料集之特徵。語意特徵，透過 Hu et al. [27]的定義，我們使用上述學者所整理出的正、負向語意辭典透過 QDA miner 軟體處理，並做為一項自定義屬性值。

產品特徵

根據 Archak et al. [25]學者發現，依其討論可得出，操弄評論會針對其探討主題之設備與物品之特有名詞和規格等相關特徵，會於無形中增加這些關鍵字詞出現次數。因此本研究使用 QDA miner 軟體進行資料處理，找出頻率高的產品特徵詞彙，並做為一項屬性。

3.2 產品操弄評論實驗流程

本文研究的流程圖如圖一所示，詳細步驟說明如下：

步驟一：資料收集與前處理

由產品評論網站收集正常與操弄評論資料，再經過刪除冗詞贅字和標點符號等前處理步驟。

步驟二：篩選候選屬性集合

本研究透過相關文獻特徵[25、27]，區分文字特徵集合與操弄特徵集合。並利用這些屬性來描述所收集之文字評論資料。

步驟三：建構實驗數據

本研究依照使用的屬性集合區分為三種實驗數據，分別為

實驗一：僅使用文字特徵

使用 TF-IDF 權重，並取詞彙頻率 5 以上為候選屬性集合，構建文件詞彙矩陣。

實驗二：使用文字特徵與潛在操弄特徵

將實驗一文件詞彙矩陣加入操弄特徵值矩陣，形成實驗二數據。

實驗三：僅使用潛在操弄特徵

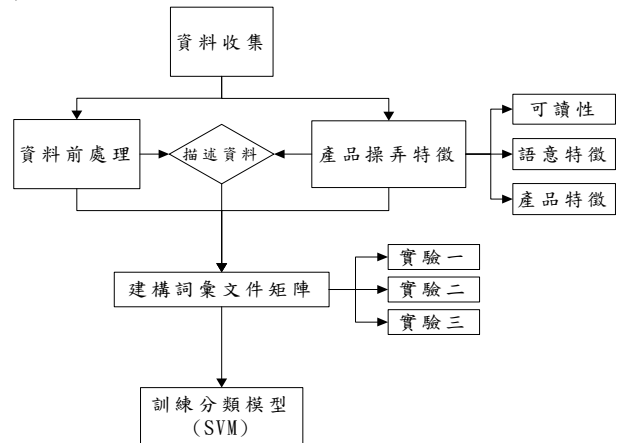
僅利用階段一所定義之前操弄屬性，建構實驗三數據。

此外，我們使用五摺交叉驗證以證實所提出方法的有效性。

步驟四：訓練 SVM 分類模型與實驗結果

此步驟，本研究使用訓練資料來建構 SVM 分類模型，並得出實驗數據

步驟五：討論與結論



圖一 實驗流程圖

4. 實驗結果

本研究使用 QDA miner 來構建出文件詞彙矩陣。在第二階段中，我們使用 LIBSVM 與 Minitab 進行實驗。

4.1 資料收集與前處理

如先前所述，本研究資料來自於 Engadget 商務網站中的 Samsung Galaxy Note II 使用者回覆評論¹進行實驗資料收集，再根據 3.1 章節產品操弄特徵將收集 200 筆語意資料進行分類出 145 筆非操弄評論與 55 筆操弄評論。Samsung Galaxy Note 於 2012 年 8 前後於各個國家、地區發表上市。現今一款智慧型手機平均生命週期約在 1~2 季，少數熱銷如蘋果

¹<http://www.engadget.com/2012/10/08/samsung-galaxy-note-ii-reviews/>

(Apple) iPhone、三星電子(Samsung Electronics) Galaxy 系列機種的生命週期較有可能大過於 3~4 季以上。

資料前處理的部分，使用 QDA miner 做為本研究資料探勘工具，首先刪除冗詞贅字、標點符號與低詞頻詞彙，並同時建構為實驗一文件詞彙文件矩陣。我們將所提取的詞彙以 TF-IDF 權重來表示向量。此外，我們使用五摺交叉驗證以證實所提出方法的有效性，資料切割如表 1。

表 1 收集評論資料

Fold 編號	資料量	類別分佈
Fold #1	200 篇	非操弄評論:145 操弄評論:55
Fold #2		
Fold #3		
Fold #4		
Fold #5		

表 2 評論操弄屬性

No.	屬性	定義	支持	定義屬性值
1	適讀性: Flesch Reading Ease	適讀性是另一種文本分級的研究取向，由閱讀材料與讀者適配程度得出結果。	Hu et al.(2012)	1: 70-100 0: 60-69 -1: 0-59
2	學校年級水平: Flesch-Kincaid Grade Level	以美國學年級作為分級，分數越低代表其閱讀性越適合低年級。	Hu et al.(2012)	1: 1-5(含 1 以下) 0: 6-8 -1: 9-12(含 12 以上)
3	文章閱讀難易度: Word TTR(TTR)	文章中使用的詞彙的變化影響閱讀的難易程度。	Hu et al.(2012)	1: 0.5 以下 0: 0.6-0.7 -1: 0.8 以上
4	語意特徵	以正向或負向詞彙引導讀者情感趨向。	Hu et al.(2012)	1: 帶有正向情緒字詞 0: 無情感 -1: 帶有負向情緒字詞
5	語意特徵-1	使用情感字詞影響讀者情緒。	Hu et al.(2012)	1: 有情感詞彙 0: 無
6	產品特徵	提到品牌、款式、規格。	Archsak et al.(2007)	1: 有提到 0: 無

接著我們使用可讀性分析(Readability Analyzer)找出產品操弄特徵並計算出本實驗所需三項數據：Flesch Reading Ease、Flesch-Kincaid Grade Level、Type-Token Ratio。另外使用 Hu and Liu [22]學者所整理出正、負向語意辭典，並使用 QDA miner 偵測出使用這些語意評論資料，接著再次使用 QDA miner 尋找出高詞頻產品特徵詞彙。依表 2 彙整出上述定義屬性值，並加入實驗一文件詞彙矩陣成為實驗二文件詞彙矩陣。最後，將表 2 六項產品操特徵值，自主為實驗三文件詞彙矩陣。

4.2 衡量指標

在傳統文本分類中，使用一般最為常見

的二元分類表來評估分類效能。如下表 3 所示。TP 表示為正面(非操弄評論)與 TN 表示為負面(操弄評論)樣本被分類正確的數量，FP 表示為正面與 FN 表示為負面樣本被分類錯誤的數量。本研究中，PA 與 NA 分別表示為正、負面類別準確率，其定義如公式(4)與(5)所示。再藉由總準確率(Overall Accuracy, OA)、G-mean(GM)、F-measure(F1)等指標分別驗證。

表 3 二元分類表

預測 \ 實際	預測為非操弄	預測為操弄
實際為非操弄	TP	FN
實際為操弄	FP	TN

一般而言，傳統的評估指標普遍都使用總準確率(Overall Accuracy, OA)來評估分類效能，其 OA 指標如公式(6)所示。此外，Kubat et al. [23]提出 G-mean(GM)指標，如公式(7)所示。以計算 PA 與 NA 之間的幾何平均數來評估分類效能，假設 PA 與 NA 都是高的，那麼 GM 值就會高。另一項重要的評估指標是使用 F1 指標，該指標結合了精確率(Precision)與召回率(Recall)，可以有效的評估分類效能，當 Precision 和 Recall 高時，則一樣會有較高的 F1 值。Precision、Recall 和 F1 定義如公式(8)~(10)所示。

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

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

$$OverallAccuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

$$GM = \sqrt{PA \times NA} \quad (7)$$

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (10)$$

最後，本研究使用 Minitab 進行皮爾森相關係數(Pearson correlation)分析，採絕對值來分級：當 r_{xy} 等於 1 時，完全相關； r_{xy} 大於

0.8，高度相關； r_{xy} 大於 0.4，中度相關； r_{xy} 小於 0.4，低度相關； r_{xy} 等於 0，完全無相關；再將 0 與 0.4 之間取 1/2 值，並做為“接近無相關”，依 0 至 1 絕對值一共分為五級如表 4 所示。

表 4 相關係數絕對值

相關係數絕對值	相關程度
R=1	完全相關
R=0.8~0.99	高度相關
R=0.4~0.79	中度相關
R=0.2~0.39	低度相關
R=0.01~0.19	接近無相關
R=約=0	無相關

4.3 SVM 分類績效比較

下表 5 列出實驗一至三的 SVM 分類績效。實驗一僅使用文字詞彙矩陣，未導入任何操弄屬性。實驗二則是原文件詞彙矩陣外，亦加入了潛在操弄特徵。由實驗數據看來，不管在 OA、GM、F1 分類績效並無顯著差異。而實驗三僅利用六個操弄屬性，其分類績效在 GM 與 F1 上，則有顯著的提升。亦即利用較少的操弄特徵，即可與傳統使用詞彙文件矩陣的效果還好，因此導入操弄特徵的確可以提升分類績效。

而由實驗二的結果看來，因為選取之操弄特徵相較於文字詞彙數目上太過懸殊，所以實驗二的績效並無顯著提升。因為 SVM 的績效對參數調整十分敏感，因此，本研究採用最佳的參數設定如表 6 所示。

表 5 各實驗 SVM 分類績效

	實驗一		實驗二		實驗三	
	平均數	標準差	平均數	標準差	平均數	標準差
PA	95.17%	3.93%	94.48%	5.23%	88.97%	7.86%
NA	29.09%	11.85%	30.91%	18.85%	47.27%	19.71%
OA	77.00%	3.26%	77.00%	2.09%	77.50%	4.68%
GM	51.65%	10.25%	51.40%	15.99%	63.12%	13.30%
F1	85.71%	2.05%	85.63%	0.99%	85.08%	3.30%

表 6 各實驗 SVM 優化參數

實驗 Fold	實驗 1	
	cost	gamma
Fold1	2	0.03125
Fold2	512	0.0001220703125
Fold3	32	0.00048828125
Fold4	32	0.001953125
Fold5	512	0.0001220703125
實驗 Fold	實驗 2	
	cost	gamma
Fold1	2	0.0078125
Fold2	512	0.0001220703125
Fold3	32	0.00048828125
Fold4	512	0.0001220703125
Fold5	128	0.0001220703125
實驗 Fold	實驗 3	
	cost	gamma
Fold1	512	0.0001220703125
Fold2	2	0.5
Fold3	0.5	0.5
Fold4	8	0.3125
Fold5	2	0.5

4.4 重要操弄特徵之選取

本小節利用相關係數，探討所用之特徵與類別（是否為操弄評論）之相關性。表七為實驗一、及三之相關係數。

表 7 實驗一、三皮爾森相關係數(取絕對值 R 值 0.2 以上之候選屬性)

實驗一		實驗三	
特徵詞彙	R 值	特徵候選值	R 值
PEN	-0.334	產品特徵	-0.416
NOTE	-0.3	Word TTR(TTR)	-0.34
SMARTPHONES	-0.286	正、負向語意特徵	-0.236
II	-0.27	Reading Ease	-0.166
IPHONE	-0.27	Grade Level	-0.074
SCREEN	-0.252	有無語意特徵	0.26
GALAXY	-0.239		
SMARTPHONE	-0.237		
YEAR	-0.221		
MIKEBPI	-0.22		
TINY	-0.22		
FIT	-0.217		
LIFE	-0.2		
USE	-0.2		

以上實驗數據，先從實驗一結果看來這些特徵值大多都提到與手機有相關聯。例如：PEN、NOTE、SMARTPHONES、II、IPHONE、SCREEN..等等，因此可以說明『產品特徵』是判斷操弄評論之重要特徵之一。

再看實驗三的產品操弄候選值，從高 R 值強的關係到弱的關係，可以看出，產品特徵相關係數最高，再次驗證『產品特徵』的確為判斷操弄特徵之重要特徵。此外，第二高的屬性 (Word TTR) 為適讀性特徵之一，然或許收集資料不足，因此後續可以再收集更多資料再行驗證。

5. 結論

在這項研究當中，我們使用了 SVM 來進行實驗。透過參數最佳化後，可以有效提高 SVM 分辨操弄資料的準確度。其中自行定義屬性“產品特徵”、“文章閱讀難易度(TTR)”、“語意特徵”，均有不錯的數據表現，更確定操弄評論均會使用到這些類別屬性。因此，這些類別屬性可確定於將來的檢測操弄評論所使用。但這項研究領域仍然是一項挑戰，歸咎於操弄資料的正確性無法真正百分之百的確認。例如，研究認為一筆資料是被操弄的，但其他人可能會認為這是一個真正消費者所撰寫的評論。因此，決定操弄與非操弄仍然是一個非常主觀的問題，所以未來研究如能取得真正的操弄與非操弄評論資料，讓研究人員可以使用這個基準建立各種模型來審查其它不同議題的評論。

致謝

本研究承蒙國科會計畫(契約編號 NSC 101-2628-E-324-004-MY3) 部分經費補助，作者在此表達感謝之意。

參考文獻

- [1] 104 人力銀行，鵬泰顧問有限公司，上網日：2013 年 1 月 5 日 取自：
<http://www.104.com.tw/jobbank/custjob/index.php?r=cust&j=403a43243836402256583a1d1d1d1d5f2443a363189j48>，2013。
- [2] Randy Walk，HTC 被刻意抹黑的品牌，上網日：2013 年 1 月 5 日 取自：
<http://forum.businessweekly.com.tw/topic.aspx?fid=74&tid=2244>，2012。
- [3] SinUpSexy，Fw:[爆卦]華擎在徵葉珮文工讀生，上網日：2013 年 1 月 5 日 取自：
http://www.ptt.cc/bbs/PC_Shopping/M.1357793890.A.47C.html，2013。
- [4] 低調空間，一款手機的生命周期，上網日：2013 年 1 月 5 日 取自：
http://www.tpst.com.cn/kexie/tjsinfo/index/web_main.action?id=114a4db8778041b895b89cb293c6432c，2012。
- [5] 沈勤譽，智慧型手機產品生命週期，上網日：2013 年 1 月 5 日 取自：
<http://www.digitimes.com.tw/tw/dt/n/shwnws.asp?CnID=10&Cat=25&Cat1=&id=288487>，2012。
- [6] 周中天，以學生為中心的英文教學—因材施教、循序漸進，取自：
<http://210.59.19.199/fdownload/getfile.asp?fn=621&fenable=1&id={A4DD4905-DADE-4673-88FD-930D4BA2C63C}>，2011。
- [7] 劉敏華、賴則中，從文本難度與特色看視譯之困難，碩士論文，國立臺灣師範大學翻譯研究所，台北，2010。
- [8] 鵬泰顧問有限公司，鵬泰顧問有限公司，上網日：2013 年 1 月 5 日 取自：
<http://www.opentide.com.tw/>，2007。
- [9] A. Abbasi, H. Chen, and A. Salem, "Sentiment analysis in multiple languages: feature selection for opinion classification in web forums," ACM Transactions on Information Systems, vol. 26, pp. 1-34, 2008.
- [10] Aixin Sun, Ee-Peng Lim, Ee-Peng Lim, Ee-Peng Lim, "On strategies for imbalanced text classification using SVM A comparative study", Decision Support Systems, vol. 48, no. 1, pp. 191-201., 2009.
- [11] B. Liu, M. Hu, and J. Cheng, "Opinion observer: analyzing and comparing opinions on the web," in 14th International conference on World Wide Web, 2005, pp. 342-351.
- [12] C. Whitelaw, N. Garg, and S. Argamon, "Using appraisal groups for sentiment analysis," in 14th ACM International conference on Information and knowledge management, Bremen, Germany, 2005, pp. 625-631.
- [13] Dale, E., & Chall, J. S., "The concept of readability. ", Elementary English, 26, 23. , 1949.
- [14] D. Mayzlin, Promotional chat on the internet, Marketing Science, 25, (2), 157-165., 2006.
- [15] DuBay, W. H., "Principles of readability. ", Retrieved November, 30th , 2009, from <http://www.impactinformation.com/impactinfo/readability02.pdf>., 2004.
- [16] F. Rudolph, "A new readability yardstick,"

- Applied Psychology, vol. 32, pp. 221-233, 1948.
- [17] J. P. Kincaid, J. R. P. Fishburne, R. L. Rogers, and B. S. Chissom, Variation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel, 1975.
- [18] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: opinion extraction and semantic classification of product reviews," in 12th International Conference on World Wide Web, New Jersey, 2003, pp. 519-528.
- [19] Klare, G. , "Readability. In Pearson P.D (Ed.) ", Handbook of Reading Research. (pp.681-744). USA: Longman., 1984.
- [20] K.-Y. Chen, L.-S. Chen, M.-C. Chen, and C.-L. Lee, "Using SVM based method for equipment fault detection in a thermal power plant," Computers in Industry, vol. 62, pp. 42-50, 2011.
- [21] M. C. Templin, Certain language skills in children: Their development and interrelationship. Minneapolis,: University of Minnesota Press, 1957.
- [22] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," in ACM SIGKDD International conference on knowledge discovery and data mining, Seattle, 2004, pp. 168-177.
- [23] M. Kubat and S. Matwin, "Addressing the curse of imbalanced data sets: One-sided sampling," in 14th International conference on Machine Learning, 1997, pp. 179-186.
- [24] M. Kubat, R. Holte, and S. Matwin, Learning when negative examples abound vol. 1224, 1997.
- [25] N. Archak, A. Ghose, and P. G. Ipeirotis, "Show me the money! deriving the pricing power of product features by mining consumer reviews," in 13th ACM SIGKDD International conference on Knowledge discovery and data mining, 2007, pp. 56-65.
- [26] N. Hu, I. Bose, Y. Gao, and L. Liu, "Manipulation in digital word-of-mouth: A reality check for book reviews," Decision Support Systems, vol. 50, pp. 627-635, 2011a.
- [27] N. Hu, I. Bose, N. S. Koh, and L. Liu, "Manipulation of online reviews: An analysis of ratings, readability, and sentiments," Decision Support Systems, vol. 52, pp. 674-684, 2012.
- [28] N. Hu, L. Liu, and V. Sambamurthy, "Fraud detection in online consumer reviews," Decision Support Systems, vol. 50, pp. 614-626, 2011b.
- [29] Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," Expert Systems with Applications, vol. 36, pp. 6527-6535, 2009.
- [30] S. Piramuthu, G. Kapoor, W. Zhou, and S. Mauw, "Input online review data and related bias in recommender systems," Decision Support Systems, vol. 53, pp. 418-424, 2012.
- [31] S. Utz, P. Kerkhof, and J. v. d. Bos, "Consumers rule: How consumer reviews influence perceived trustworthiness of online stores," Electronic Commerce Research and Applications, vol. 11, pp. 49-58, 2012.
- [32] V. N. Vapnik, The Nature of Statistical Learning Theory: Springer-Verlag, 1995.
- [33] Xu, Jiajin & Yunlong Jia., Readability Analyzer 1.0: A text difficulty analyzing tool. Beijing: The National Research Centre for Foreign Language Education, Beijing Foreign Studies University., 2009.