蕭瑞祥、姜青山、曹金豐、陳柏翰（2015），『基於中文語法規則的情感評價單元抽取方法之研究』，*中華民國資訊管理學報*，第二十二卷，第三期，頁 243-272。

基於中文語法規則的情感評價單元抽取方法之研究

蕭瑞祥
淡江大學資訊管理學系

姜青山
中國科學研究院深圳先進技術研究院

曹金豐
育學資訊
淡江大學資訊管理學系

陳柏翰*
淡江大學資訊管理學系

摘要

隨著 Web 2.0 的概念被提出，加上近年來社群媒體興起，情感分析（sentiment analysis）逐漸成為新興研究的趨勢，其相關研究與應用的價值也越來越重要。意見單元（或稱情感評價單元）是評價語句中的評價對象及其對應的意見詞的組合，由於意見單元決定了此評價的意見傾向，因此意見單元的抽取是為情感分析領域的重要任務之一。

本研究採用系統發展研究法建置一套基於語句層級中文語法規則的意見單元抽取方法之雛型系統，並使用資料探勘技術歸納出意見單元的抽取規則，以建立意見單元抽取模式。研究以「智慧型手機」產品的評論文章驗證方法架構，實驗結果發現，同時使用語句結構與句法路徑結構作特徵屬性，有助於本系統意見單元抽取模式品質的提升，且語句結構在意見單元抽取較句法路徑結構具影響性。研究結果顯示，本研究所建立的意見單元抽取模式，與相關研究的意見單元抽取方法比較，具有較佳的 F-Measure 值。

關鍵詞：情感分析、意見單元、句法路徑、類神經網路

* 本文通訊作者。電子郵件信箱：abkobe8@gmail.com
2013/05/14 投稿；2014/07/25 修訂；2014/09/15 接受

Shaw, R.S., Jiang, Q.S., Tsao, C.F. and Chen, P.H. (2015), 'A study of opinion unit extraction based on Chinese syntactic rules', *Journal of Information Management*, Vol. 22, No. 3, pp. 243-272.

A Study of Opinion Unit Extraction Based on Chinese Syntactic Rules

Ruey-Shiang Shaw
Department of Information Management, Tamkang University

Qing-Shan Jiang
Shenzhen Institute of Advanced Technology Research Institute

Chin-Feng Tsao
Technical Seed Corporation
Department of Information Management, Tamkang University

Po-Han Chen*
Department of Information Management, Tamkang University

Abstract

Purpose－Through Web 2.0 concepts being advocated to bring about internet opinion groups growing in recent years, the field of Chinese Sentiment Analysis related research has expected more attention and value. Opinion Unit (or Appraisal Entity) is to define the association of opinion words and their corresponding subjects. Because of the opinion unit regulates the polarity of comments, extracting and analyzing opinion units is significant task for the field study of Chinese Sentiment Analysis.

Design/methodology/approach － This paper used the systems development process in information systems research to build a prototype system of a method of opinion unit recognition based on the syntactic rules in Chinese we proposed, and used the techniques of data mining to summarize opinion unit recognition rules to establish an opinion unit extracting mode.

* Corresponding author. Email: abkobe8@gmail.com
2013/05/14 received; 2014/07/25 revised; 2014/9/15 accepted

Findings－The study subjected to smartphonediscussing comments was used to test our method of opinion unit recognition and the experiment indicated using the sentence structure and syntactic path structures as feature attributes would contribute to opinion unit extracting mode, and the statement structure was more influential in the opinion unit recognition rules. Results showed that our opinion unit extraction mode is better than correlation studies in F-Measure.

Research limitations/implications－This paper focuses on the discussing comments related to smartphone appraisal group. Hence, it is suggested that future research may apply our method of opinion unit extracting mode to other areas, such as computer, car or food. Also, future research is recommended to compare with using other data mining classification, such as SVM, Decision Tree, K-NN or Bayesian Statistics.

Practical implications－This paper proposes the extraction principle of opinion unit. In commerce, it may apply to the sentiment analysis of products usage discussing comments. Also, future research may use the proposed attribute of statement structure and attribute of syntactic path structures to make an extensive study.

Originality/value－This paper proposes a method of opinion unit extraction based on statement level that can be applied to the discussing comments about smartphone appraisal group. Also, it implement the method and use data mining classification with attribute of statement structure and attribute of syntactic path structures to fulfill a rule of opinion unit extraction.

Keywords: sentiment analysis, opinion unit, syntactic path, artificial neural network

壹、導論

自從 2004 年 Web 2.0 的概念被提出，人們在網際網路上的行為發生了重大改變，互動參與和資訊分享變得即時且容易，越來越多的人在網際網路上表達個人觀點，由被動的接受資訊轉變為參與創建網際網路資源，現代人受網際網路的影響也越來越深，同樣地，每個人都可以在網際網路上發表意見，個人的影響力也越來越大，在強調互動的 Web 2.0 環境之下，部落格（Blog）、論壇、微網誌（Micro Blog）成為了產品有力的行銷工具。

根據 ACNielsen 在 2009 年針對全球網路消費者所做的調查報告（ACNielsen 2009），全球平均有 70%的消費者相信網友在網路上發表的意見與評價。Liu、Hu 與 Cheng（2005）認為分析網友在網際網路的評論，可以分析出整個社會的趨勢，尤其是網友對產品的評論，能夠經由分析轉換成具有商業價值的資訊，因此提出利用資料探勘（data mining）的方法對評論進行分析，試圖找出評論中對產品各方面特徵的正、負面評論，這些產品評論提供消費者及製造商重要的參考資訊，但這一類產品評論內容組織雜亂且數量龐大，單純依靠人工收集和整理資訊已不能滿足需求，因此，如何快速從海量的網際網路內容，自動化有效取得與整理，並快速分析網路上的評論成為重要的研究議題。在商業應用上，透過自動從網際網路大量的評價文章中定義其中所含有的意見，並進一步歸納、分析其所表達出的意見傾向，能夠讓人們快速檢視評價文章的評價目標在網路上的整體評價，有鑑於此，情感分析相關研究應運而生。

情感分析（sentiment analysis），又有稱為意見探勘（opinion mining），是指透過自動化的方式搜索網際網路上相關的資源，從中取得含有評價意見的內容，此類意見主要描述對於人物、事件或事物等的想法及觀點，並進一步讀取、總結、最後將其轉換成可利用的格式（Liu 2010）。目前情感分析領域的相關方法，目的皆是期望能瞭解文章作者在文章中的意見或評價等，透過文獻探討與觀察，情感分析在整體系統架構上已有相似的固定模式，本研究將情感分析一般性做法彙整成情感分析架構如圖 1。本研究所探討的範圍著重在定義意見單元，即意見單元的抽取。

意見單元抽取的目的是要識別評價語句中意見詞與其所修飾之評價對象的組合，即為意見單元，此一組合為評價語句的意見傾向的評判關鍵，且意見傾向不僅取決於意見詞本身，也取決於意見詞所評價的對象（Huang et al. 2011），例如：「價錢很高」、「螢幕的解析度很高」，這兩句話中的「高」對於所修飾的評價對象所表達的意見傾向並不相同，若沒有準確的抽取出評價對象與其對應的意見詞，就會導致情感分析結果出現偏差，因此如何準確地抽取意見詞與其修飾的評

價對象組合，是為情感分析中重要的任務之一。

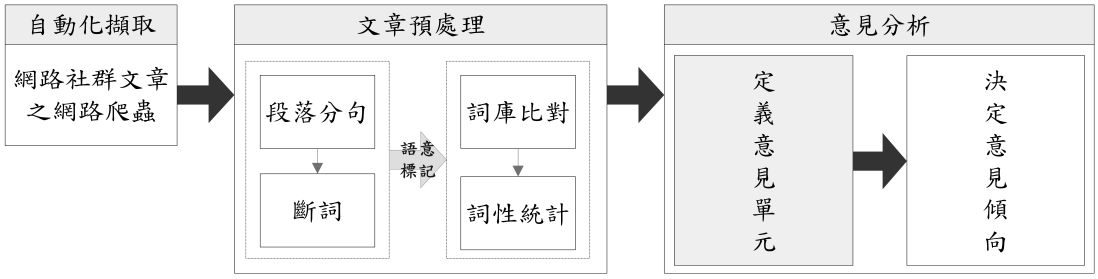


圖 1：情感分析架構圖

目前意見單元抽取的相關研究中，有些學者以字詞彼此之間的距離為基礎抽取意見單元，透過選取字詞距離與評價對象最靠近的形容詞做為對應的意見詞（Hu & Liu 2004），或設定評價對象與意見詞之間的字詞距離不超過一個範圍，在此距離中尋找形容詞做為其所對應的意見詞（Kim & Hovy 2004），也有學者使用 SVM 分類工具，歸納出主觀情緒語句在文字距離上的分類模式（簡之文 2012）。但此類方法忽略語句本身的結構，應用於長度較長的語句，可能發生評價對象與意見詞之間的字詞距離較遠，導致抽取出的意見單元組合並非正確的配對。

再則，也有透過預先建立評價對象詞庫與意見詞庫，進而將語句與詞庫中的字詞逐一比對，以抽取意見單元定義之各元素的方法（Hu & Liu 2004; Kobayashi et al. 2007; 楊盛帆 2009），不過由於不同領域有其特定使用的評價對象詞語或意見詞語，實際上不容易建立全面且適用於各領域的詞庫，此外，這類方法在建立詞庫時需要花費時間及人力成本，在對於新用詞的維護上也勢必需花費時間及人力成本，且意見單元的準確率也受到詞庫侷限。

相較於經由詞庫比對的方法以抽取意見單元，有研究將文章中各字詞進行詞性標註後，篩選語句中特定詞性，以做為抽取意見單元各元素的方法（Hu & Liu 2004; Qiu et al. 2011），也有學者透過計算與評價主題同時出現機率最高的名詞做為評價對象（Popescu & Etzioni 2005）或比較名詞短語在評價文章與普通文章中出現的頻率來識別評價對象（Scaffidi et al. 2007）。此類方法通用性高且不需耗費時間建立相關詞庫，但由於此類方法已經抽離字詞本身的意義，且很難將所有詞性的組合狀況列出，以致於意見單元抽取的準確率低於詞庫比對方法。

透過句法路徑或句法規則的意見單元抽取方法（Bloom et al. 2007; Huang et al. 2011; Zhao et al. 2011），由於忽略字詞本身的意義，若面對長句較多且結構複雜的中文語句，若其中包含的評價對象與意見詞數目較多，會產生多組評價對象與意見詞組成的意見單元，經由比對常見的意見單元句法路徑的方法，能夠將這些

意見單元都識別出來，但並非所有組合都是正確的評價對象與意見詞組合，導致抽取的準確率降低。

此外，中文「一字多義」、「一詞多義」的現象，使得有些字詞雖然存在於意見詞典，但在語句中不做為意見詞的意義使用（李林琳 2008），在此情況發生時，不能與評價對象組成意見單元。但在透過詞庫比對的方式或自動構建句法路徑模式時，這些字詞也會被標註成意見詞並與評價對象組合，產生錯誤的意見單元。

綜合相關研究探討，本研究目的旨在建立一套基於語句層級中文語法規則的意見單元抽取方法，期望能夠較準確且快速地從候選意見單元中，識別出正確的評價對象與其所對應之意見詞的組合。本研究主要研究流程歸納如下：

1. 以中文部落格及論壇關於「智慧型手機」類型產品評價文章為例，建立雛型系統實作建立意見單元抽取模式的流程，並利用資料探勘的類神經網路分類技術幫助建立意見單元抽取模式。
2. 使用特徵選取方法進行類神經網路的特徵選取，並從評估結果探討特徵的選擇對於意見單元抽取的影響，進而提高所建立的意見單元抽取模式的準確率。
3. 以本研究建立的意見單元抽取模式，與現有相關研究的抽取方法進行比較，並將比較結果加以探討。

貳、文獻探討

本研究主要為探討意見單元的抽取方法。本節探討與本研究領域相關的文獻，包括意見單元、句法路徑、中文剖析系統、類神經網路等相關議題。

一、意見單元

（一）意見單元的定義

意見單元的擷取為網路上產品意見文章探勘的任務之一（Morinaga et al. 2002），學者探討意見單元的研究文獻甚多，楊盛帆（2009）在研究中提到，意見探勘當中最關鍵的步驟在如何以結構化的方式表達網路使用者的意見；Hu 與 Liu（2004）定義了評論模型，每一個被評論的對象可能是產品、服務、主題、個人、組織或事件，並且對應到一組意見詞；Hu 與 Liu（2004）認為網路使用者在表達意見時，是對於某項人、事、物等特徵（Feature），發表看法或評價（Evaluation），因此符合特徵（Feature）與評價（Evaluation）的組合，即可組成意見單元；Kobayashi、Inui 與 Matsumoto（2004）定義了一個評價單元三元組，即「Evaluated Subject」、「Focused Attribute」、「Value」，其中「Focused Attribute」對應意見單元

中的評價對象，「Value」則對應意見單元中的意見詞；也有學者提出將意見詞與鄰近的程度詞和否定詞，組合成一個意見單元（王正豪&李啟菁 2010）。

本研究歸納相關研究的方法，將「評價對象、意見詞」視為一組意見單元。由於意見詞的意見傾向不僅取決於意見詞本身，也取決於意見詞所評價的對象，因此，準確地抽取意見詞與其修飾的評價對象，為判斷意見單元意見傾向的重要關鍵。

（二）意見單元的抽取方法

不少學者提出各種意見單元的抽取方法，包括基於詞庫比對擷取（Hu & Liu 2004; Kobayashi et al. 2007; 楊盛帆 2009）、基於詞性擷取（Hu & Liu 2004; Popescu & Etzioni 2005; Qiu et al. 2011; Scaffidi et al. 2007）、基於字詞距離擷取（Hu & Liu 2005; Kim & Hovy 2004; 簡之文 2012）、句法規則模式擷取（Bloom et al. 2007）或句法路徑模式擷取（Huang et al. 2011; Zhao et al. 2011）等。

楊盛帆（2009）預先建立了筆記型電腦領域的各項中文意見單元元素詞庫，做為分析網路文章中對於某品牌筆電情感分析之依據，其中的評價對象詞庫是從筆電官方網站提供的筆電規格表建立，而意見詞則是採用台大意見詞典 NTU Sentiment Dictionary（台灣大學自然語言處理實驗室 2007），再以人工方式標記出詞庫中未包含的字詞；Hu 與 Liu（2004）使用 WordNet 做為英文詞庫來源，接著對語句各個字詞逐一比對，並針對詞頻、詞性及相似詞集合等進行分析，找出與意見相關的組合，以做為產生意見單元之方法；Kobayashi、Inui 與 Matsumoto（2007）先以人工的方式建立評價對象詞庫與意見詞庫，接著將語句中的字詞逐一與詞庫比對，以做為抽取意見單元各元素的方法。但此類方法實際上不容易建立全面且適用於各領域的詞庫，建立詞庫時也須耗費時間及人力成本，且意見單元的準確率也會受到詞庫所侷限。

Hu 與 Liu（2004）先將文章中各個字詞進行詞性標註後，比對語句中特定詞性做為抽取意見單元的各元素的方法，研究中歸納評價對象的詞性多為名詞，意見詞的詞性多為形容詞或副詞；Popescu 與 Etzioni（2005）的研究中計算 PMI（point-wise mutual information）值來擷取與主題同時出現機率最高的名詞做為評價對象；Qiu 等（2011）使用消費者評論資料集裡的評論，將 POS tagging 標註詞性結果為名詞的做為評價對象，標註詞性結果為形容詞的做為意見詞；Scaffidi 等（2007）透過比較名詞短語在某一評價文章中出現的頻率與在普通文章中出現的頻率的不同來識別出有價值的評價對象。此類方法與透過詞庫比對抽取意見單元比較，通用性高且不需耗費時間及人力建立相關詞庫，但由於不論是在中文語料或英文語料，此類方法皆已經抽離字詞本身意義，且很難將所有的詞性組合狀況列出，以致於意見單元抽取的準確率低於詞庫比對方式。

有學者是以字詞距離為基礎抽取意見單元，Hu 與 Liu（2004）選取字詞距離與評價對象最靠近的形容詞做為對應的意見詞；Kim 與 Hovy（2004）設定評價對象與意見詞之間的字詞距離不超過研究中定義的值 k；簡之文（2012）將文字距離、程度詞分級等屬性，做為 SVM 分類工具的訓練資料與測試資料，歸納出主觀情緒語句在文字距離上的分類模式，且透過實驗發現若誤差範圍在 3 以內則將視為主觀情緒評論的意見單元，即以評價對象的位置起算，距離誤差範圍 3 以內可以找到對應的評價詞。由於此類方法忽略語句本身的結構，應用於長度較長的語句，可能發生評價對象與意見詞之間的字詞距離較遠，導致抽取錯誤的意見單元組合，且在實際的評價語句中，動詞、名詞、形容詞都可以做為意見詞使用。

在基於句法規則的意見單元抽取研究中，Bloom、Garg 與 Argamon（2007）使用 Stanford Parser（http://nlp.stanford.edu:8080/parser），以人工方式制定了 31 條句法規則，以表示評價對象與意見詞之間的關聯，例如：target（評價對象）$\xrightarrow{nsubj}$ x $\xrightarrow{dobj}$ y $\xrightarrow{amod}$ attitude（意見詞），句法規則中的有向箭頭顯示了箭頭兩端的依存關係，該方法更為深層的探討評價對象與意見詞之間的關係，但建立句法規則需耗費過多的人工，且透過人工方式所建立的規則可能不夠全面。

Zhao 等（2011）探討透過句法路徑以抽取意見單元，是使用意見詞庫從大量評價語句中尋找意見詞，再透過詞性分析出評價對象，進而自動獲取單一語句中包含的所有評價對象與意見詞之間的句法路徑，接著透過統計句法路徑出現的次數，建立了句法路徑模式庫，並藉由人工制定的兩條泛化規則進行句法路徑的泛化，最後將語句的句法樹與句法路徑模式庫進行比對，從句法樹中抽取符合模式庫的句法路徑，將路徑頭尾節點的字詞分別做為評價對象與意見詞，以組成意見單元。研究實驗結果證明，此方法應用在英文評價語料結果良好，但是處理長句較多且結構複雜的中文語句，由於單一評價語句包含的評價對象與意見詞的數目較多，導致準確率較低，例如：有一評價語句中含有 m 個評價對象、n 個意見詞，則會產生 m * n 條候選意見單元組合，透過比對常見的意見單元句法路徑，皆能夠將這些組合都識別出來，但在這 m * n 條候選意見單元組合中，並非所有組合都是正確的評價對象與意見詞組合，以 iPhone 5 的評價意見語句為例：「iPhone 5 不僅螢幕的畫質細膩且色彩準確」這句話包含兩個評價對象「畫質」、「色彩」及兩個意見詞「細膩」、「準確」，透過排列組合會產生四種候選意見單元組合，其中只有「畫質＋細膩」、「色彩＋準確」，這兩種組合是為正確的意見單元。

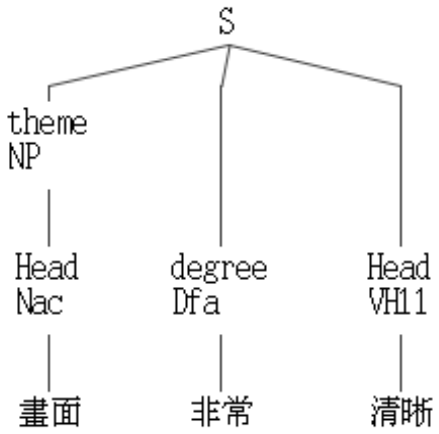
Huang 等（2011）的研究與 Zhao 等（2011）透過句法路徑以抽取意見單元的方法類似，提出了基於句法樹結構的意見單元抽取方法，是以一種新的基於句法樹的結構 GMCT（generalized minimum complete tree）來表示意見單元，該方法保留了句法樹的結構資訊，藉此區分出錯誤路徑，同時使用了一種基於 CTK（convolution tree kernel）的相似度計算方法 ACTK（approximate convolution tree

kernel），對 GMCT 進行模糊比對，從而提高泛化性。研究實驗結果證明，在中文評價語料中，該方法提升意見單元抽取的準確率及召回率。

二、句法路徑

透過文獻探討與觀察，句法路徑是指在中文句法樹上連接任意兩個節點之間的句法結構（Zhao et al. 2011）。本研究參考 Zhao 等（2011），將句法路徑特指在短語句結構中文句法樹上，連結評價對象和意見詞的有向句法結構（如圖 2），亦即透過句法路徑來描述評價對象與意見詞之間的修飾關係，例如：「畫面 ↑Nac↑NP↑S↓VH11↓ 清晰」就是一條意見單元的句法路徑，其中「畫面」、「Nac」等表示中文句法樹的各節點，「↓」符號表示其左邊節點是其右邊節點的父節點，「↑」符號表示其左邊節點是其右邊節點的子節點。

句法路徑表示了意見單元中的評價對象與意見詞，在評價語句中兩者之間的關係，本研究利用此特性，自動化抽取評價對象與意見詞之間的句法路徑，並進一步將其轉換後的各項特徵屬性值做為類神經網路的訓練資料，藉此歸納出意見單元的抽取模式。



資料來源：CKIP 中文剖析系統（http://parser.iis.sinica.edu.tw/）

圖 2：句法路徑示意圖

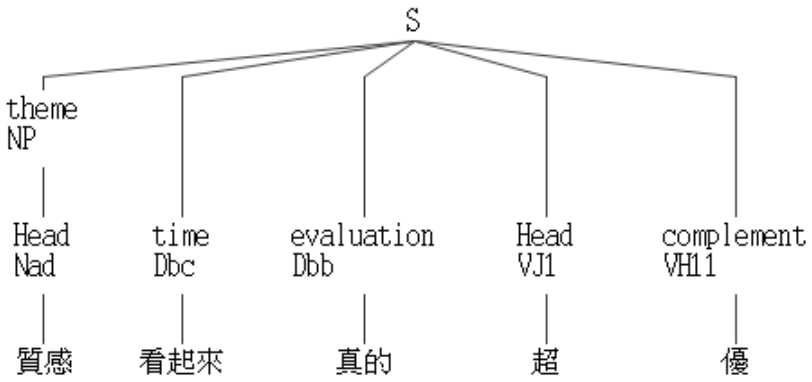
三、中文剖析系統

為順利取得評價語句中評價對象與意見詞之間的句法路徑，本研究採用 CKIP 中文剖析系統（http://parser.iis.sinica.edu.tw/）將語句進行剖析處理，將語句進行剖析處理的目的在產生語句的句法樹（如圖 3），句法樹表示了語句結構，而語句結

構是語義分析及瞭解的必要訊息，要使電腦具有智慧的語言處理能力，電腦系統都必須先能夠分析語句結構，因此中文語句自動剖析的工作便成為語言理解不可或缺的技術。產生出句法樹之後，就能夠從句法樹中抽取評價對象與意見詞所對應的句法路徑。

CKIP 中文剖析系統為中央研究院 CKIP 中文詞知識庫小組所建立，採用機率式無語境規律的模型（probabilistic context-free grammar）為基本剖析架構並加入結構中詞彙搭配關係機率解決結構歧義，系統工作包含斷詞、斷詞標記、中文剖析、角色指派。以下為評價語句「質感看起來真的超優」的句法樹表示式及句法路徑範例：

- 評價語句：質感看起來真的超優
- 句法樹表示式：S（theme:NP（Head:Nad:質感）| time:Dbc:看起來 | evaluation:Dbb:真的 | Head:VJ1:超 | complement:VH11:優）
- 句法路徑：質感 ↑Nad ↑NP ↑S ↓VHl1 ↓ 優



資料來源：CKIP 中文剖析系統（http://parser.iis.sinica.edu.tw/）

圖 3：句法樹示意圖

四、類神經網路（Artificial Neural Network; ANN）

類神經網路是以電腦系統模擬生物腦神經網路構造所建構的一套資訊處理方式，國外學者 Aleksander、Morton 與 Myers（1990）對類神經網路所下的定義，「類神經網路是一種以自然特性儲存並運用經驗知識的平行分散處理器。」

類神經網路系統透過模擬生物神經元的構造，由許多的人工神經元（artificial neurons）或稱處理單元組成，這些人工神經元具有簡單的函數運算功能，並由神經鍵（synapses）加以連接，每一個人工神經元可以接收其他神經元所傳來的訊號，當某一個神經元經由神經鍵接收到其他神經元傳來的訊號之後，會先將訊號經由

轉換函數（transfer function）處理運算，運算完成後再把轉換後的訊號傳遞到下一個神經元，當訊號要傳入下一個神經元之前，會經由神經鍵乘以一個加權值（weights），稱為 W_{ij}，表示第 i 處理單元對第 j 處理單元之影響強度，再傳遞給其他神經元做為輸入訊號，每一個神經元的輸出以扇形的方式送出，傳送到其他處理單元的輸入。

類神經網路的學習演算法可分為監督式學習網路、無監督式學習網路、聯想式學習網路、最適化應用網路（葉怡成 2009），其中監督式學習是由訓練資料中學習或建立一個模式，並依此模式推測新的案例。類神經網路的網路型態有許多不同的種類，倒傳遞類神經網路模式（back-propagation neural network）為目前最重要且應用最廣的一種監督式學習演算法之一（Fish et al. 1995; 葉怡成 2009），倒傳遞類神經網路模式增加了隱藏層，使其具備建構非線性模型之能力，且改用平滑可微分的轉換函數，使網路可應用最陡坡降法（the steepest descent method）觀念將誤差函數予以最小化，導出修正加權值的公式，圖 4 為倒傳遞類神經網路模式示意。

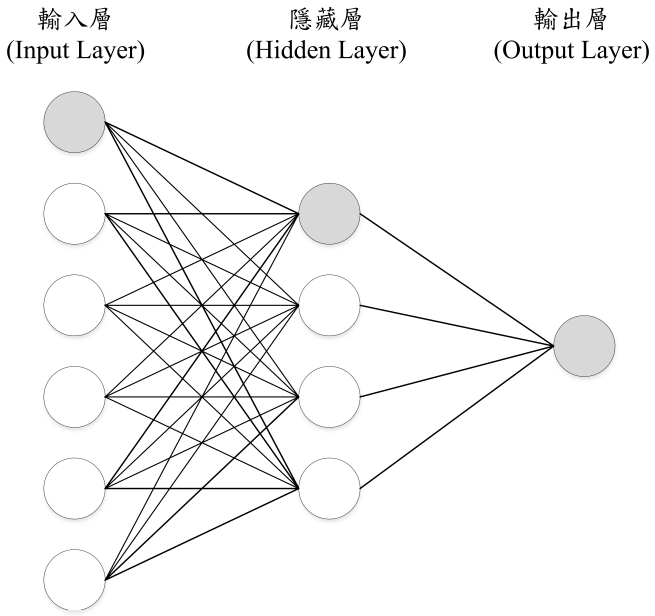


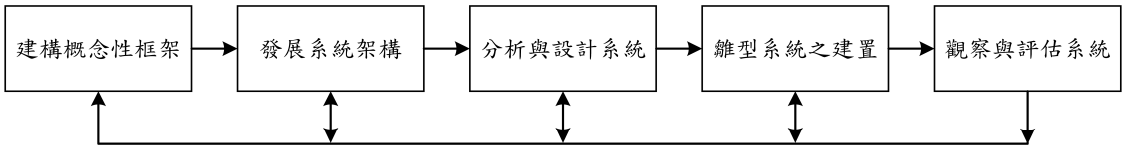
圖 4：倒傳遞類神經網路模式示意圖

監督式學習是透過訓練資料中學習或建立一個模式，並依該模式推測新的案例，本研究從中文語句特徵及句法特徵中獲取影響意見單元的重要因素，進而建構本研究模式，亦即從問題的領域中取得輸入與輸出的變數值做為訓練資料，並從中學習輸入與輸出變數的內在對應規則，最後再將此規則應用於新的案例，輸

入新的輸入變數值，以推論出輸出變數值，故本研究適合以監督式學習網路的倒傳遞類神經網路模式加以檢驗研究模型。

參、研究方法

本研究採用 Nunamaker、Chen 與 Purdin（1990）所提出的系統發展研究方法（systems development in information dystems research），做為本研究系統的發展步驟。研究流程包含了下列五個主要步驟，分別為建構概念性框架、發展系統架構、分析與設計系統、雛型系統之建置及觀察與評估系統，如圖 5 所示：



資料來源：Nunamaker et al.（1990）

圖 5：系統發展研究流程

本研究依據圖 5 之步驟及流程，在確認研究動機與目的後，深入瞭解相關知識，並依據蒐集到的文獻及資料，著手建置雛型系統，發展系統各元件之功能，在雛型系統完成之後，該系統即為本研究核心進行評估，最後再對評估結果作歸納討論。

本節先就系統發展研究流程之前三項步驟做說明，亦即建構概念性框架，再依此概念框架發展為系統架構，最後再進行系統分析與設計。後兩項的雛型系統之建置及觀察與評估系統，將在後面分成兩節做更詳細的探討與說明。

1. 建構概念性框架：本步驟是以問題為導向，先收集整理意見單元抽取的方法與可能問題，再基於解決問題之方向，擬定意見單元抽取之功能與需求，以做為本研究概念性框架之建構。
   (1) 意見單元抽取的問題：依據目前在意見單元抽取上所使用的方法及所遇到的問題彙整如下：
      - 以字詞距離抽取意見單元，由於忽略語句本身的結構，導致可能抽取錯誤意見單元組合。
      - 以比對評價對象詞庫與意見詞庫，抽取意見單元之方法，在建立與維護詞庫之人力成本較高，且應用於不同領域之準確率也有所侷限。
      - 以詞性標註為基礎的意見單元抽取方法，由於已抽離字詞本身意義，且較難列出所有詞性組合狀況，以致準確率較低。
      - 以句法規則抽取意見單元，由於句法規則的建立需耗費過多的人力

成本，且透過人工方式所建立的規則不夠全面。

- 以句法路徑規則抽取意見單元，由於忽略字詞本身的意義，若面對長句較多且結構複雜的中文語句，可能造成準確率較低。
- 對於中文「一字多義」、「一詞多義」的情形，各類方法較無法準確處理。

(2) 意見單元抽取之功能與需求：依據上述的意見單元抽取問題尋找相關的文獻資料，經由整理、研讀、分析後統整出一個概念性框架（如圖 6），本研究所研究發展之意見單元抽取模式的功能與需求為透過將網際網路的評價文章，經由資料探勘方法訓練及學習，建立一套意見單元的抽取模式，並可依此模式推測出新的案例，藉以判斷正確之意見單元的組合。

圖 6：系統概念性框架

2. 發展系統架構：由圖 6 之系統概念框架需求，本研究參考文獻探討與商用相關系統的架構，發展之雛型系統初步架構，如圖 7 所示。

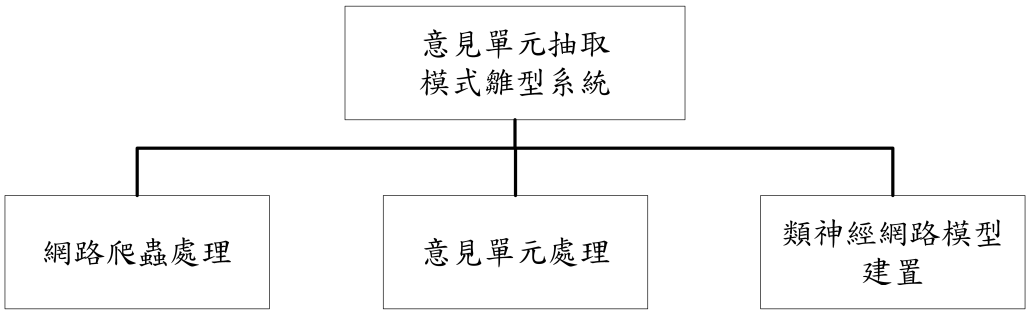


圖 7：雛型系統初步架構圖

(1) 網路爬蟲處理：具備自動化依據符合相關領域主題或關鍵字，搜尋擷取網際網路的相關網路評論文章之功能。

(2) 意見單元處理：具備本研究所提以語句層級中文語法規則的意見單元抽取方法之功能，包含：抽取意見單元及產生句法路徑、產生語句及句法特徵屬性值等功能。

(3) 類神經網路模型建置：具備以意見單元組合訓練資料建立類神經網路模型，篩選較佳的類神經網路模型之功能。

3. 分析與設計系統：本研究使用 Visual C#程式語言進行系統開發工作，於網際網路搜索相關的評價文章頁面，轉換為標準的 HTML 語言之後，進一步解析以取得所需要的內容，儲存在 Microsoft SQL Server 資料庫中，提供進行分析處理的資料。此外，本研究使用 Microsoft SQL Server Analysis Services 做為分析建模工具，對演算法相關參數進行設定調整，以產生訓練資料的最佳模型，進一步產生意見單元抽取模式。

肆、雛型系統之建置

本步驟是根據上述三項步驟所設計發展之系統架構與分析，細部規劃系統架構與流程，並實作建置雛型系統。為驗證本研究所提出的意見單元抽取方法之可行性，本研究參考 Zhao 等（2011）提出的系統發展前提：評價詞與其真正具有搭配關係的評價物件間滿足一定的句法關係，且這些句法關係是有規律可循的、可總結的，而非雜亂無章的。本研究將雛型初步系統架構（如圖 7）依據前述假設與參考 Zhao 等（2011）和 Huang 等（2011）的模式，並以建立意見單元抽取模式為目的，發展雛型系統細部架構與流程（如圖 8），並藉以建置雛型系統。

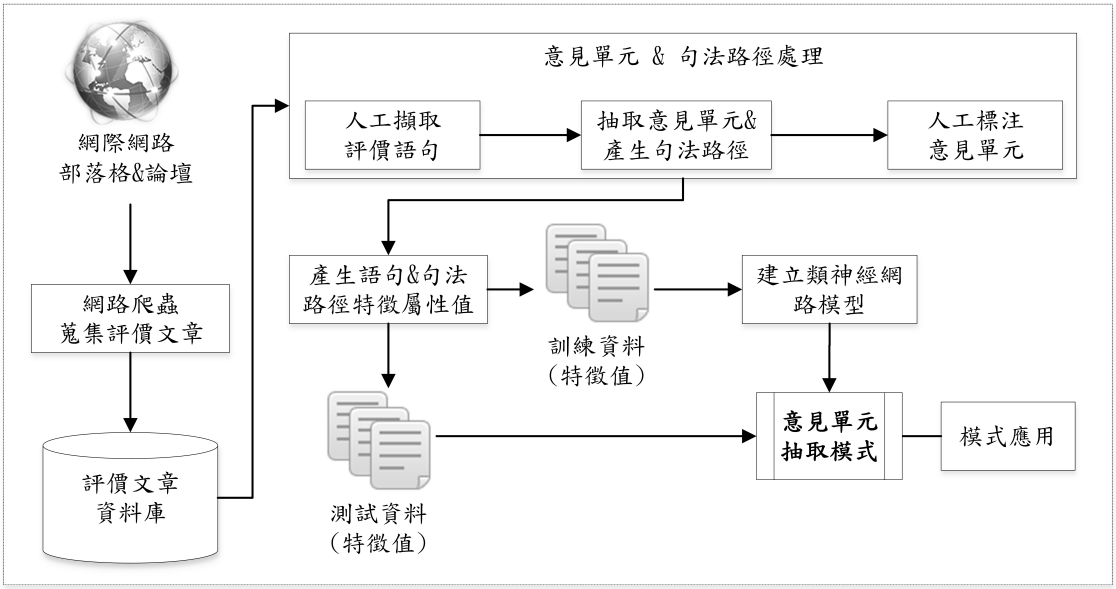


圖 8：雛型系統細部架構與流程圖

一、網路爬蟲蒐集評價文章

網路爬蟲（web crawler）為網頁資訊擷取技術，能夠自動於網際網路中解析並取得網頁內容，本系統使用網路爬蟲，自動搜索網際網路的相關資源，以取得符合相關領域或關鍵字的網路評論文章。

本系統透過網路爬蟲於痞客邦（http://www.pixnet.net/blog）、天空部落（http://blog.yam.com/index.php?op=blog）等部落格及 Mobile01 論壇（http://www.mobile01.com/category.php?id=4），自動擷取符合「智慧型手機」產品相關的中文網路評論文章，儲存於關聯式資料庫中，提供進行分析處理的資料。

二、人工擷取評價語句

本系統在自動抽取候選意見單元時，是以字詞的詞性做為評價對象與意見詞的抽取依據，由於透過詞性的特性並未考慮到字詞本身的意義，為確保不會出現非本系統定義相關領域的評價語句，系統預先以人工方式從已擷取的評論文章，篩選出同時具有相關領域評價對象與意見詞之語句，做為一評價語句。

本系統經由人工方式從透過網路爬蟲取得的 250 篇「智慧型手機」類型產品相關評論文章中，篩選共 552 條同時具有評價對象與意見詞的評價語句。

三、抽取意見單元&產生句法路徑

Hu 與 Liu（2004）在抽取意見單元的研究中，歸納出評價對象或稱產品特徵（feature）大部分的詞性為名詞，及意見詞或稱產品特徵的評價（evaluation）的詞性通常是形容詞或副詞的形式，在情感分析中，某些詞性的詞語具有更大的概率包含情感傾向性（唐都鈺 2012），表示將字詞進行詞性標註是有必要的，詞性標註的功用在於對語句內每一字詞，判斷在語句中表示何種詞性並加以標註，本系統將依其抽取意見單元方法，擷取詞性為名詞的字詞做為評價對象。另外，由於本系統使用 CKIP 中文剖析系統建立句法樹，在句法樹中被標註的詞性會依此剖析系統為標準，本系統將台大意見詞典 NTU Sentiment Dictionary（台灣大學自然語言處理實驗室 2007）的正、負面意見詞，經過 CKIP 中文剖析系統進行詞性標註，並統計各詞類數量（如表 1），統計結果發現詞性為 Vi 的字詞數量最多，約 3,688 筆，因此本系統將參考語句中詞性標註為 Vi 的字詞做為意見詞。

系統將語句轉換成句法樹的格式後，自動化解析此語句的句法樹，以抽取評價對象與意見詞節點間（即意見單元）之有向路徑做為此語句的一條句法路徑，本系統參考 Zhao 等（2011）基於句法路徑的意見單元抽取方法為基礎。在句法樹的產生，使用 CKIP 中文剖析系統將語句剖析處理，轉換成句法樹的表示式，例如：

「S（theme:NP（Head:Nad:質感）｜time:Dbc:看起來 ｜evaluation:Dbb:真的 ｜Head:VJ1:超 ｜complement:VH11:優）」，接著自動化解析句法樹，抽取評價對象與意見詞節點間的有向路徑做為一意見單元的句法路徑，例如：「質感 ↑Nad ↑NP ↑S ↓VH11 ↓ 優」。

表 1：NTUSD 意見詞之詞性統計表

詞性	數量
Vi	3688
Vt	2404
N	2302
ADV	214
T	150

註：依照數量排序列出前五項

由於單一評價語句的評價對象與意見詞的數量不限，因此一條評價語句可能含有一至多條評價對象與意見詞的組合，本系統從 462 條評價語句自動化抽取 1,128 條候選意見單元，並產生各意見單元在句法樹中的句法路徑。

四、人工標註意見單元

由於一條評價語句可能包含一至多組評價對象與意見詞的組合，且這些候選的意見單元包含正確及錯誤的意見單元組合，系統以人工方式標註出正確與錯誤的意見單元組合，做為後續建立意見單元抽取模式的依據。

為避免人工介入時，因個人主觀或疏忽影響結果，故安排三位人員進行相關工作，且預先對人員進行說明，若發現其中兩位人員的結果不同，則邀請第三人進行評估決定結果，系統採用 Kappa 一致性係數做為前二位人員標註結果一致性評估的量測方法，計算公式如式(1)：

$$\text{Kappa 值} = \frac{P_0 - P_c}{1 - P_c} \tag{1}$$

P_0：為觀測一致性（observed agreement），表示前後兩位人員標註結果一致的百分比。

P_c：為期望一致性（chance agreement），表示前後兩位人員標註結果預期相同的機率。

Kappa 值的計算結果介於-1.00 至+1.00 間，由於目前對於 Kappa 值一致性表現沒有絕對的標準，本系統參考 Landis 與 Koch（1997）提出的 Kappa 值一致性解釋（如表 2 所示）。

表 2：Kappa 值一致性解釋

Kappa 值	一致性強度	Kappa 值	一致性強度	Kappa 值	一致性強度
低於 0.00	弱	0.21 – 0.40	尚好	0.61 – 0.80	高度
0.00 – 0.20	輕	0.41 – 0.60	中度	0.81 – 1.00	最高

資料來源：Landis 與 Koch（1977）

表 3 為前二位人員 A 與人員 B 對 1,128 筆意見單元組合標註的結果，Kappa 值計算結果約 0.9688，由計算結果得知，前二位人員的意見單元標註結果具有高度的一致性，因此本系統將標註結果進一步使用。

表 3：人員 A 與人員 B 標註候選意見單元組合結果表

人員 B ＼ 人員 A	正確	不正確
正確	452	6
不正確	11	659

五、產生語句&句法特徵屬性值

本系統自動將前述經人工標註的正確與錯誤的意見單元結果，其語句本身及其表示的句法路徑正規化，轉換成特徵屬性值，做為類神經網路模型的訓練資料及測試資料。

Qu 等（2008）在判斷語句是否為評價語句及其意見傾向的子任務中，利用語句結構特徵做為 SVM 分類的各項特徵屬性；Kobayashi 等（2007）以語句結構特徵做為 SVM 訓練及測試的特徵屬性，以識別出評價語句；Wu 等（2009）使用 Stanford parser（http://nlp.stanford.edu:8080/parser）將語句轉換成句法樹的格式，並定義了句法路徑的特徵做為 SVM 訓練及測試的特徵屬性，藉此抽取出評價語句中的意見單元。本系統引用上述研究，定義各特徵（如表 4、表 5）及意見單元識別結果（如表 6），並且將詞性分別轉換以編號表示（如表 7），各特徵分述如下：

表 4：語句結構特徵表

編號	特徵名稱	特徵描述	值域
S1	語句長度	表示語句的長度	＞0
S2	評價對象與意見詞的距離	表示評價對象與意見詞之間的字詞距離	＞0
S3	評價對象前一個詞的詞性	表示評價對象其前一個字詞所屬的詞性	≥0
S4	評價對象後一個詞的詞性	表示評價對象其後一個字詞所屬的詞性	≥0
S5	意見詞前一個詞的詞性	表示意見詞其前一個字詞所屬的詞性	≥0
S6	意見詞後一個詞的詞性	表示意見詞其後一個字詞所屬的詞性	≥0

表 5：句法路徑結構特徵表

編號	特徵名稱	特徵描述	值域
P1	意見詞位於評價對象之前（後）	表示意見詞相對於評價對象的位置，若意見詞位於評價對象之前為 0，反之為 1	1,0
P2	↑方向總數	表示句法路徑中，子節點指向父節點的方向總數	＞0
P3	↓方向總數	表示句法路徑中，父節點指向子節點的方向總數	＞0
P4	↑方向總數 -↓方向總數	表示特徵編號 P2 與特徵編號 P3 的差值	整數
P5	節點總數	表示句法路徑中，所有節點的總數	＞0

表 6：意見單元識別結果規則表

編號	結果名稱	結果描述	值域
R	意見單元識別結果	表示意見評價對象與意見詞組成的意見單元組合是否正確，需人工標註。若正確為 1，反之為 0。	1,0

表 7：詞性編號對照表

編號	詞性	精簡詞類標記	編號	詞性	精簡詞類標記
0	空白	空白	10	Vi	動作不及物動詞
1	A	非謂形容詞	11	T	感嘆詞
2	ADV	動詞前(後)程度、副詞、數量副詞	12	P	介詞
3	DET	指代定詞、數量定詞、特指定詞	13	COMMACATEGORY	逗號
4	FW	外文標記	14	EXCLAMATIONCATEGORY	驚嘆號

5	M	量詞	15	PAUSECATEGORY	頓號
6	N	代名詞、名化物動詞	16	PERIODCATEGORY	句號
7	POST	後置詞	17	QUESTIONCATEGORY	問號
8	Vt	動作使動動詞、動作及物動詞	18	ASP	時態標記
9	C	對等連接詞			

資料來源：中研院平衡語料庫詞類標記集（中央研究院詞庫小組 1998）

本系統從 462 條評價語句中隨機抽取 100 條評價語句，其中共包含 289 條候選意見單元組合，同時透過自動化解析句法樹，共取得 289 條各候選意見單元所對應的句法路徑，系統將其做為測試資料，並使用另外 362 條評價語句包含的 839 條候選意見單元（即句法路徑）做為訓練資料，訓練資料與測試資料相關的統計結果如表 8。本系統將語句本身的結構及意見單元的句法路徑結構正規化，並將正規化結果做為類神經網路的各項特徵屬性。

表 8：訓練資料與測試資料相關的統計結果表

統計項目	訓練資料數量	測試資料數量
所含評價句語總數	362	100
所含意見單元(句法路徑)總數	839	289
正確意見評價單元數	368	91
錯誤意見評價單元數	471	198
意見單元總數/句	2.32	2.89

以「打算換 iphone5 因為聽說電池挺耐用」之評價語句為例，表 9 列出此語句包含 2 組候選意見單元組合：編號 1 與編號 2，以及各組候選意見單元與其在語句的句法樹中所對應的句法路徑。

表 9：候選意見單元及其對應之句法路徑表

編號	候選意見單元	句法路徑
1	iphone5 + 耐用	N ↑ NP ↑ VP ↑ VP ↓ VP ↓ Vi
2	電池 + 耐用	N ↑ NP ↑ VP ↓ VP ↓ Vi

本系統接著透過自動化方式將評價語句依照所定義的語句結構特徵屬性（如表 4），將各候選意見單元的來源語句正規化，產生對應的語句結構特徵屬性值（如表 10 中的 S1～S6），例如：表 10 中編號 1 的 S3 為 8，透過表 4 的定義表示評價對象其前一個字詞所屬的詞性為 Vt，這裡需透過表 7 將 8 轉換成對應的詞性 Vt；並且依照本研究所定義的句法路徑結構特徵屬性（如表 5），將各候選意見單元所對應的句法路徑正規化，產生對應的句法路徑結構特徵屬性值（如表 10 中的 P1～P5），例如：表 10 中編號 1 的 P1 為 1，透過表 4.5 的定義表示意見詞位於評價對象之後，最後透過人工的方式判斷各候選意見單元的組合是否正確，組合正確則將識別結果值 R 輸入 1，反之為 0。

表 10：候選意見單元之特徵屬性值列表

編號	S1	S2	S3	S4	S5	S6	P1	P2	P3	P4	P5	R
1	19	5	8	3	2	0	1	3	2	1	6	0
2	19	1	8	2	2	0	1	2	2	0	5	1

系統將各評價語句之候選意見單元轉換後的特徵屬性值列，做為建立類神經網路模型的訓練資料，以建立意見單元的抽取模式。

六、建立類神經網路模型

本系統採用監督式學習網路的倒傳遞類神經網路演算法，並使用 Microsoft SQL Server Analysis Services 做為建模工具，對演算法相關參數進行設定調整，以產生訓練資料的最佳模型，模型的訓練資料是使用語句與句法路徑正規化後的特徵屬性。

在 Analysis Services 中，增益圖被應用於評估資料採礦模型（圖 9），增益圖曲線越向上，表示模型的效果越好。在圖 9 中的對角直線，為「理想模型」，因此預測模型的曲線越接近理想模型，代表預測效果越好。本系統透過增益圖做為評估使用訓練資料建立類神經網路模型的依據，以篩選出較佳的類神經網路模型，進而與相關研究的意見單元抽取方法比較。

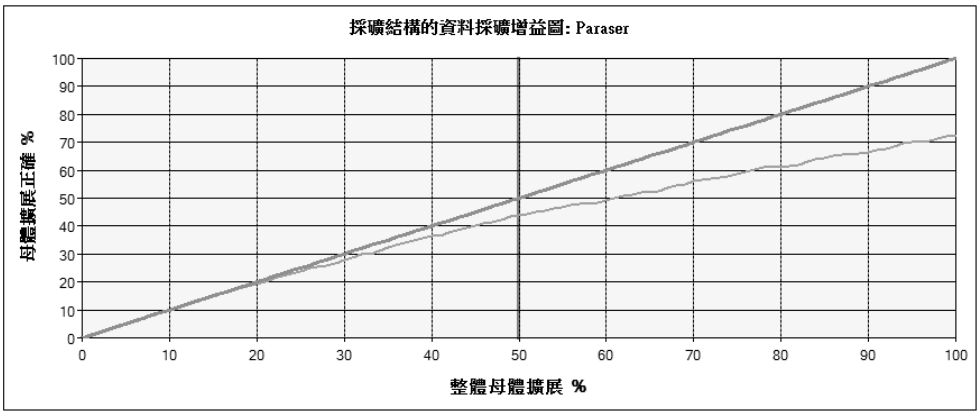


圖 9：Analysis Services 增益圖範例

七、模式應用

本系統旨在建立一套適用於產品中文評論文章的意見單元抽取模式（以「智慧型手機」類型為例），此模式可應用於語句層級的意見單元抽取。首先，系統將一篇評價文章中的語句，透過字詞的詞性特性，自動化抽取語句中被標註為相關詞性的字詞，以做為評價對象與意見詞，並將之排列組合以組成候選意見單元，接著解析這些候選意見單元在語句句法樹中所對應的組成路徑，然後透過自動化方式將此語句的語句結構及句法結構轉換成特徵屬性，最後將特徵屬性的值輸入本系統所建立的意見單元抽取模式，即可識別候選意見單元的組合是否為正確的意見單元，亦即判斷評價對象與意見詞的組合是否正確，進一步做為判斷語句意見傾向的依據。

以一篇中文「智慧型手機」類型產品評論文章中，「iPhone 不僅畫質細膩且色彩準確」之評價語句為例，系統首先透過字詞的詞性特性，抽取出所有 N（評價對象）與 Vi（意見詞）之組合做為候選意見單元組合，此例句透過排列組合方式，共包含了六種意見單元可能的組合（如表 11），以表 11 中編號 3 的候選意見單元組合為例，整體運作流程如圖 10，系統自動化產生此候選意見單元的特徵屬性值列，接著將這些特徵屬性值輸入本系統建立的意見單元抽取模式，即可產生識別結果 R 為 1，即表示編號 3 的評價對象與意見詞組合，在此語句中為正確組成意見單元。

表 11：範例之意見單元組合表

編號	評價對象（N）	意見詞（Vi）	句法路徑	組合結果
1	iPhone	細膩	N↑NP↑S↓S↓S↓Vi	錯誤
2	iPhone	準確	N↑NP↑S↓S↓S↓Vi	錯誤
3	畫質	細膩	N↑NP↑S↓Vi	正確
4	畫質	準確	N↑NP↑S↑S↓S↓Vi	錯誤
5	色彩	細膩	N↑NP↑S↑S↓S↓Vi	錯誤
6	色彩	準確	N↑NP↑S↓Vi	正確

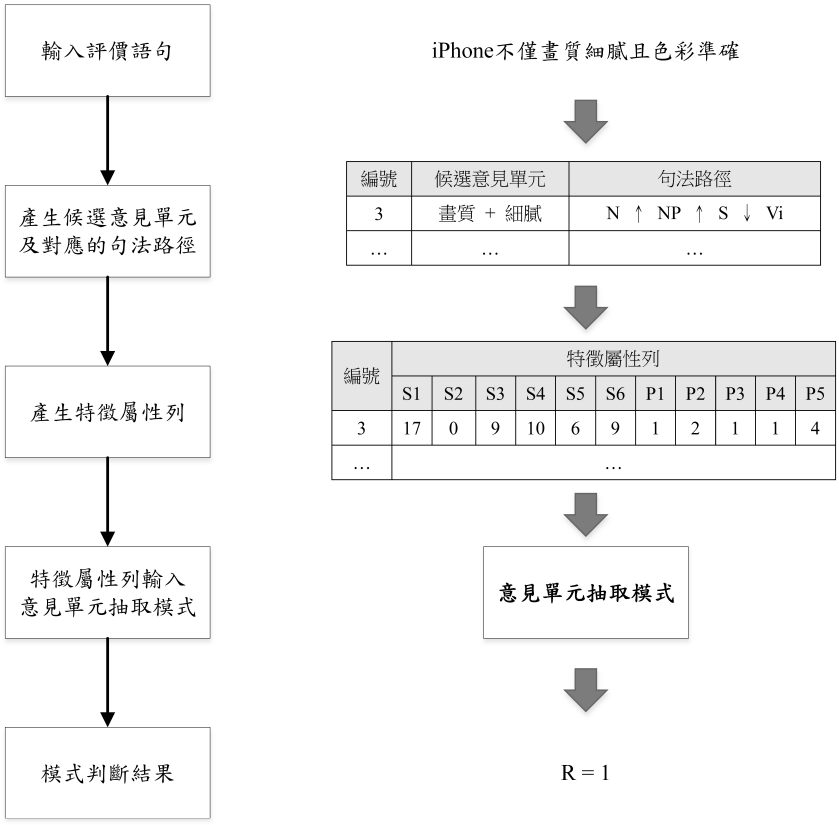


圖 10：模式應用運作流程示意圖

伍、觀察與評估系統

本步驟主要是設計與實施測試實驗流程與準備測試資料，以驗證本研究所建立意見單元抽取規則及方法實際應用之雛型系統的績效與效益。

一、系統變項 Pearson 相關分析

Pearson 相關分析主要在於分析兩個連續變項之間的相關程度，本系統在使用類神經網路建立模型前，以 Pearson 相關分析探討特徵屬性各別與意見單元識別結果（R）之間的相關性。依據統計學的定義，P-Value < 0.01 顯示兩變數之間有顯著相關（邱皓政 2010），本研究計算各特徵屬性的分析結果如表 12，透過分析結果篩選出 P-Value < 0.01 的特徵做為本系統的輸入特徵。

表 12：Pearson 相關分析結果

編號	Pearson 相關係數值	P-Value	編號	Pearson 相關係數值	P-Value
S1	-.326(**)	0.000	P1	.275(**)	0.000
S2	-.371(**)	0.000	P2	-.182(**)	0.000
S3	.123(**)	0.000	P3	-.180(**)	0.000
S4	-.242(**)	0.000	P4	.019	0.586
S5	-.060	0.081	P5	-.255(**)	0.000
S6	-.014	0.696			

**表示在顯著水準為 0.01 時（雙尾），相關顯著。

* 表示在顯著水準為 0.05 時（雙尾），相關顯著。

二、系統評估結果與討論

在雛型系統建置完成之後，本研究從經過人工方式篩選後的 462 條評價語句當中，隨機抽取 100 條評價語句，經過自動化方式抽取出候選意見單元，共抽取出 289 條候選意見單元，接著分別抽取候選意見單元在評價語句中所對應的句法路徑，系統將語句結構及句法路徑結構正規化後產生特徵屬性，做為系統評估的測試資料，以評估本雛型系統的有效性，本系統設計之評估方法說明如下：

Precision（準確率）、Recall（召回率）和 F-Measure 為資料探勘領域中常見之評估指標，為了評估本系統所建立的意見單元抽取模式，在識別意見單元是否獲得良好的效果，本系統將採用此評估指標做為評估依據，並利用 Confusion Matrix（如表 13）說明此評估指標，準確率表示分類結果為正確的意見單元組合，而能準確辨識為正確的意見單元組合之比率；召回率表示實際為正確的意見單元組合樣本，而能被正確分類之比率。系統主要將以 F-Measure 做為衡量指標（Larsen & Aone 1999），使用這一項指標的優點在於，F-Measure 在計算時使用了 Precision 與 Recall 的計算結果，避免了在 Precision 或是 Recall 較高時而造成的假象，因而無法客觀的比較不同方法的優劣，當 Precision 與 Recall 的值越高，則 F-Measure 的

值也越高，表示其識別意見單元的正確率高、且正確的意見單元辨識錯誤率低，因此若是方法評估結果的 F-Measure 較大者，顯示其效果越佳，其計算公式如式(2)、(3)、(4)。

表 13：Confusion Matrix

		人工方式識別	
		Positive	Negative
意見單元抽取模式識別	Positive	True Positive Count (TP)	False Positive Count (FP)
	Negative	False Negative Count(FN)	True Negative Count (TN)

資料來源：Turban、Sharda 與 Delen（2011）

$$\text{Precision} = TP / (TP + FP) \tag{2}$$

$$\text{Recall} = TP / (TP + FN) \tag{3}$$

$$\text{F} - \text{Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

在雛型系統建置完成之後，本研究將測試資料的語句及句法路徑正規化後產生特徵屬性值，輸入本系統的意見單元抽取模式，以評估本系統的有效性，本系統設計以下方法進行評估，分述如下：

一、特徵屬性選擇

特徵屬性選擇主要是探討特徵屬性的選擇對於意見單元抽取模式各項評估值的影響，本研究分成三個層面做探討，如下所述：

1. 語句結構特徵：將特徵 S1 至 S6 做為類神經網路模型的訓練資料及測試資料，並參考 Pearson 相關分析結果（表 12），將特徵 S5 & S6 排除，接著將特徵輸入並調整神經網路相關參數，透過增益圖找出最佳的類神經網路模式，即為意見單元抽取模式，再將測試資料輸入此模式，模式將會自動計算測試資料透過此模式的預測結果，若預測與實際相同則為預測正確，不同則視為預測錯誤，最後計算此模式的 Precision、Recall 與 F-Measure。
2. 句法路徑結構特徵：重覆一、1.節的步驟，僅在特徵屬性的選擇，以特徵 P1 至 P5 做為類神經網路模型的訓練資料及測試資料，並參考 Pearson 相關分析結果（表 12），將特徵 P4 排除。
3. 結合語句與句法路徑結構特徵：嘗試結合語句本身與句法路徑結構特徵，

將 S1 至 S6、P1 至 P5 做為類神經網路模型的訓練資料及測試資料，並參考 Pearson 相關分析結果（如表 12），將特徵 S5 & S6 & P4 排除，接著重覆一、1.節的步驟。

本研究分別透過「語句結構」特徵、「句法路徑結構」特徵與同時使用兩種特徵，探討不同結構的特徵屬性選擇，對於意見單元抽取模式各項評估值的影響，由評估結果發現（如表 14），結合語句與句法路徑結構特徵所建立的意見單元抽取模式，其 F-Measure 皆較佳，顯示同時使用兩種特徵，有助於類神經網路模型品質提升。

此外，各別使用語句結構特徵與句法結構特徵所建立的意見單元抽取模式相比較，語句結構特徵所建立的模型獲得的 F-Measure 較佳，且準確率提升了 5%，顯示語句結構特徵意見單元的抽取較具影響。

表 14：評估結果比較

評估方法	Precision (%)	Recall (%)	F-Measure (%)
語句結構特徵（本研究）	51.92	59.34	55.38
句法結構特徵（本研究）	46.88	**65.93**	54.79
結合語句與句法路徑結構特徵（本研究）	**58.06**	59.43	**58.70**

二、字詞距離方法比較

依據第 1 節的評估結果，本研究將意見單元識別效果較佳的模式與其他做比較，即選擇結合語句與句法路徑結構特徵，做為其訓練資料，以建立意見單元抽取模式。

Kim 與 Hovy（2004）設定在距離評價對象 K 的範圍內尋找意見詞，以組成意見單元，此方法在英文語料中取得很好的結果，本研究模擬此方法，首先透過人工方式標註語句中的評價對象，接著抽取距離評價對象 K 的範圍內的意見詞，組合成意見單元，並透過人工方式標註正確及錯誤的意見單元，以評估此方法抽取意見單元的有效性。

由評估結果發現（如表 15），Kim 與 Hovy（2004）基於字詞距離的意見單元抽取方法雖然在英文語料中取得好的結果，但在中文語料中 F-Measure 值約 46%，且準確率僅有 35%，這可能因為中文語料在很多情況，意見詞與評價對象距離較遠，且中文在語意表達上較英文複雜，此方法在中文語料上存在者侷限性，也反映出，本系統取得評價對象與意見詞之間的關係比單純取得他們在評價語句中的位置有意義。

三、句法路徑方法比較

本研究模擬 Zhao 等（2011）基於句法路徑的意見單元抽取方法，評估結果發現（如表 16），本系統與該方法相比，準確率提升了 23%，F-Measure 值提升了 10%。Zhao 等（2011）在其研究顯示，基於句法路徑的方法用於英文語料時，獲得很好的結果，但 Huang 等（2011）將 Zhao 等（2011）的方法應用於中文語料，F-Measure 值約 39%，此外，本研究實際將 Zhao 等（2011）的方法應用本系統數據，F-Measure 值約 48%。由於透過句法路徑模式庫比對的方式，忽略字詞本身意義及語句結構，且中文字本身「一字多義」的現象，皆可能導致抽取錯誤的意見單元，使得 Zhao 等（2011）的方法應用於語句結構複雜的中文語料，結果不佳。

表 15：評估結果比較

評估方法	Precision (%)	Recall (%)	F-Measure (%)
字詞距離 (k=3)*（Kim & Hovy 2004）	34.63	**68.13**	45.93
結合語句與句法路徑結構特徵（本研究）	**58.06**	59.43	**58.70**

*設定評價對象與意見詞之間的字詞距離不超過 K

表 16：評估結果比較

評估方法	Precision (%)	Recall (%)	F-Measure (%)
句法路徑（Zhao et al. 2011）	35.35	**76.92**	48.44
結合語句與句法路徑結構特徵（本研究）	**58.06**	59.43	**58.70**

陸、結論與未來展望

意見單元的抽取在情感分析領域佔了重要的地位，尤其應用在產品的情感分析上，意見單元的正確抽取就顯得十分關鍵，影響了此產品的意見傾向，進而影響產品給予大眾的觀感，由此可見意見單元的重要性。近年來對於意見單元抽取的研究越來越多，學者在研究中提出各類意見單元的抽取方法，但仍存在一些需要解決的課題，特別在中文的意見單元抽取，由於中文並無文法且語句的結構複雜，相較英文的意見單元抽取所面臨到的課題更為廣泛，本研究發展一套中文意見單元的抽取模式，提供較準確且快速地識別候選意見單元中，正確的評價對象與其所對應之意見詞的組合，進而做為意見傾向的評判依據。

本研究的目的在於建立一套基於語句層級中文語法規則的意見單元抽取模式，透過建立一套雛型系統，並以「智慧型手機」產品評論文章為例，在語句層級上，將語句本身結構及其句法路徑結構轉換後的特徵屬性，做為使用類神經網

路演算法執行資料探勘分析的訓練資料，以歸納出意見單元的抽取模式，此模式可識別候選意見單元是否為正確的意見單元，亦即判斷評價對象與意見詞的組合是否正確，最後透過實驗進一步驗證抽取模式的有效性。本研究整理並歸納出結論如下：

1. 在特徵屬性的選擇上，同時使用「語句結構」特徵與「句法結構」特徵，將有助於本系統建立的意見單元抽取模式品質提升。
2. 在「智慧型手機」產品評價文章中，語句結構在意見單元抽取較句法結構具影響。
3. 與基於字詞距離的意見單元抽取方法比較中，本系統取得的 F-Measure 值較佳，且準確率大幅提升約 24%，顯示在「智慧型手機」產品的意見單元識別中，本研究提出的方法有較好的效果。
4. 與基於句法路徑的意見單元抽取方法比較中，本系統取得的 F-Measure 值較佳，且準確率大幅提升約 23%，顯示在「智慧型手機」產品的意見單元識別中，本研究提出的方法有較好的效果。

本研究之研究貢獻整理如下：

1. 本研究採用系統發展研究方法，建置一套實務的雛型系統，此系統實作了建立意見單元抽取模式的流程，並使用資料探勘技術中的類神經網路對資料進行訓練與測試，進一步歸納出意見單元的抽取規則，以建立意見單元抽取模式。
2. 本研究有別於過去意見單元抽取的相關研究，提出基於語句的資料探勘分析方法，在特徵屬性的選取，結合了「語句結構」特徵與「句法結構」特徵，並證明「語句結構」特徵與「句法結構」特徵，對於意見單元抽取具影響性。
3. 本研究所定義的「語句結構」與「句法結構」特徵屬性，提供研究者若欲使用資料探勘技術進行相關研究時，對於特徵屬性的選擇，可以參考本研究的建議，進行相關實驗之探討與應用。
4. 從應用的角度，本研究建立一套基於語句層級上，應用於中文「智慧型手機」產品之意見單元抽取方法，提供相關領域評價文章之意見單元抽取。

本研究之研究限制說明如下：

1. 本研究所建立的雛型系統目標範圍是以中文「智慧型手機」產品評論文章為例，由於時間及人力的限制，未能驗證本研究方法是否通用於其他領域的評論文章，未來使用更多領域的評論文章對本研究之方法進行驗證。
2. 本研究在建立的雛型系統與預測模式時，為求準確性在實作時需要以人工的方式執行，受限於時間與人力因此較無法做到大量處理，使得本系統在透過類神經網路的資料探勘分類技術建立意見單元抽取模式時，所提供的

訓練資料數量較為有限，若能夠增加訓練資料的數量，可以讓訓練資料的組合更為全面，以利再提升意見單元抽取模式的有效性。

3. 本研究建置的雛型系統僅以中文「智慧型手機」產品評價文章做為評估來源，建立相關產品領域的意見單元抽取模式，未來能透過網路爬蟲擷取其他領域評價文章，例如：電腦、汽車、美食…等，以建立跨領域的意見單元抽取模式，並加以探討。
4. 本研究在意見單元識別的部分是以類神經網路的資料探勘分類技術為基礎來進行推論和實作，但資料探勘分類技術有很多，例如：支援向量機（Support Vector Machine; SVM）、決策樹（Decision Tree）、K-th Nearest Neighbor（K-NN）、Bayesian Statistics 等其它方法，未來能使用各種監督式學習方法並將推論結果進行比較，以找出更適合的方法。
5. 本研究在建立雛型系統的各環節需要人力的參與，雖然人工的參與能夠提升各環節工作的準確性和有效性，但也提高了人力與時間的成本，未來能夠在不降低原有人工參與的準確性和有效性下，設計替代方案以解決人力與時間成本的耗費。

參考文獻

中央研究院詞庫小組（1998），中研院平衡語料庫詞類標記集，http://ckipsvr.iis.sinica.edu.tw/papers/category_list.doc（存取日期 2013/01/20）。

王正豪、李啟菁（2010），『中文部落格文章之意見分析』，未出版碩士論文，國立臺北科技大學資訊工程研究所，臺北市。

台灣大學自然語言處理實驗室（2007），台大意見詞典（NTUSD），http://nlg18.csie.ntu.edu.tw:8080/opinion/pub1.html（存取日期 2012/12/25）。

李林琳（2008），『基於特定領域的漢語句子意見挖掘』，未出版碩士論文，上海交通大學電子資訊與電氣工程學院電腦系，上海市。

邱皓政（2010），*量化研究與統計分析*，五南圖書出版公司，臺北市。

唐都鈺（2012），『領域自我調整的中文情感分析詞典構建研究』，未出版碩士論文，哈爾濱工業大學電腦科學技術研究所，哈爾濱市。

楊盛帆（2009），『以整合式規則來做網路論壇上的 3C 產品口碑分析』，未出版碩士論文，元智大學資訊管理研究所，桃園縣。

葉怡成（2001），*應用類神經網路*，儒林圖書有限公司，臺北市。

簡之文（2012），『部落格文章情感分析之研究』，未出版碩士論文，淡江大學資訊管理研究所，新北市。

ACNielsen（2009），AC 尼爾森研究：口碑行銷極具廣告說服力，

http://tw.nielsen.com/site/news20090716.shtml（存取日期 2013/01/15）。

Aleksander, I., Morton, H.B. and Myers, C.E. (1990), 'HCI: a cognitive neural net prospects', *Proceedings of the IEEE Colloquium*, South America, Argentina, Brazil, Chile, Uruguay, August 31-September 15, pp. 1-4.

Bloom, K., Garg, N. and Argamon, S. (2007), 'Extracting appraisal expressions', *Proceedings of NAACL HLT*, Rochester, New York, USA, April 22-27, pp. 308-315.

Fish, K.E., Barnes, J.H. and Aiken, M.W. (1995), 'Artificial neural networks: a new methodology for industrial market segmentation', *Industrial Marketing Management*, Vol. 24, No. 5, pp. 431-438.

Hu, M. and Liu, B. (2004), 'Mining opinion features in customer reviews', *Proceedings of the 19th National Conference on Artificial Intelligence (AAAI 2004)*, San Jose, USA, July 25-29, pp. 755-760.

Hu, M. and Liu, B. (2004), 'Mining and summarizing customer reviews', *Proceedings of the 10th ACM International Conference on Knowledge Discovery and Data Mining (KDD 2004)*, Seattle, Washington, USA, August 22-25, pp. 168-177.

Huang, Y.H., Pu, X.J., Yuan, C.F. and Wu, G.S. (2011), 'Appraisal expression extraction based on parse tree structure', *Application Research of Computers*, Vol. 28, No. 9, pp. 3229-3234.

Kim, S.M. and Hovy, E. (2004), 'Determining the sentiment of opinions', *Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics (COLING 2004)*, Geneva, Switzerland, August 23-27, pp. 1367-1374.

Kobayashi, N., Inui, K. and Matsumoto, Y. (2007), 'Opinion mining from web documents: Extraction and structurization', *Transactions of the Japanese Society for Artificial Intelligence*, Vol. 22, No. 2, pp. 227-238.

Kobayashi, N., Inui, K., Matsumoto, Y., Tateishi, K. and Fukushima, T. (2004), 'Collecting evaluative expressions for opinion extraction', *Proceedings of the International Joint Conference on Natural Language Processing (IJCNLP 2004)*, New York, USA, March 22-24, pp. 596-605.

Landis, J.R. and Koch, G.G. (1977), 'The measurement of observer agreement for categorical data', *Biometrics*, Vol. 33, No. 1, pp. 159-174.

Larsen, B. and Aone, C. (1999), 'Fast and effective text mining using linear-time document clustering', *Proceedings of the 5th ACM SIGKDD*, San Diego, CA, USA, August 15-18, pp. 16-22.

Liu, B. (2010), 'Sentiment analysis and subjectivity', in Nitin, I. and Fred, J.D. (Eds.), *Handbook of Natural Language Processing*, CRC Press, Taylor and Francis Group, Boca Raton, FL, pp. 627-666.

Liu, B., Hu, M. and Cheng, J. (2005), 'Opinion observer: analyzing and comparing opinions on the web', *Proceedings of the 14th international Conference on World Wide Web (WWW 2005)*, Chiba, Japan, May 10-14, pp. 342-351.

Morinaga, S., Yamanishi, K., Tateishi, K. and Fukushima, T. (2002), 'Mining product reputations on the Web', *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, Edmonton, Alberta, Canada, July 23-25, pp. 341-349.

Nunamaker, J.F.Jr, Chen, M. and Purdin, T.D.M. (1990), 'Systems Development in Information Systems Research', *Journal of Management Information Systems*, Vol. 7, No. 3, pp. 89-106.

Popescu, A.M. and Etzioni, O. (2005), 'Extracting product features and opinions from reviews', *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing*, Vancouver, B.C., Canada, October 6-8, pp. 339-346.

Qiu, G., Liu, B., Bu, J. and Chen, C. (2011), 'Opinion word expansion and target extraction through double propagation', *Journal of Computational Linguistics*, Vol. 37, No. 1, pp. 9-27.

Qu, L., Toprak, C., Jakob, N. and Gurevych, I. (2008), 'Sentence Level Subjectivity and Sentiment Analysis Experiments in NTCIR-7 MOAT Challenge', *Proceedings of NTCIR-7 Workshop Meeting*, Tokyo, Japan, December 16-19, pp. 210-217.

Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H. and Jin, C. (2007), 'Red opal: product-feature scoring from reviews', *Proceedings of the 8th ACM conference on Electronic commerce*, San Diego, CA, USA, June 11-15, pp. 182-191.

Turban, E., Sharda, R. and Delen, D. (2011), *Decision support and business intelligence systems*, Pearson Education, USA.

Wu, Y., Zhang, Q., Huang, X. and Wu, L. (2009), 'Phrase dependency parsing for opinion mining', *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP 2009)*, Singapore, August 6-7, pp. 1533-1541.

Zhao, Y.Y., Qin, B., Che, W.X. and Liu, T. (2011), 'Appraisal expression recognition with syntactic path for sentence sentiment classification', *International Journal of Computer Processing of Languages*, Vol. 23, No. 1, pp. 21-37.