

元智大學
資訊管理學系碩士班
碩士論文

文本特徵於新聞立場偵測之效能評估

Investigating the Effectiveness of Content Features
on News Stance Detection

研 究 生：劉芷妤

指導教授：楊錦生 博士

曾淑芬 博士

中華民國一一〇年一月

文本特徵於新聞立場偵測之效能評估

Investigating the Effectiveness of Content Features
on News Stance Detection

研 究 生：劉芷妤

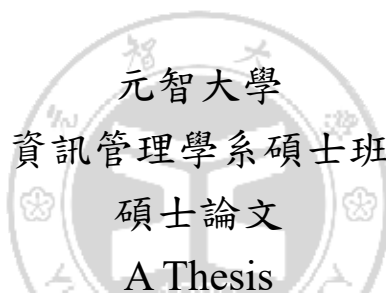
Student：Chih-Yu Liu

指導教授：楊錦生 博士

Advisors：Dr. Chin-Sheng Yang

曾淑芬 博士

Dr. Shu-Fen Tseng



Submitted to Department of Information Management
College of Informatics
Yuan Ze University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Science
in
Information Management

Jan 2021

Chungli, Taiwan, Republic of China

中華民國一一〇年一月

文本特徵於新聞立場偵測之效能評估

學生：劉芷好

指導教授：楊錦生 博士

曾淑芬 博士

元智大學資訊管理學系碩士班

摘要

隨著網際網路與社群媒體的快速發展，我們更容易取得各式各樣的資訊，社群媒體的使用量也不停的增長。然而，大量的訊息在社群媒體之間傳播，其中也包括了假新聞、謠言與錯誤訊息，這些訊息對個人、組織甚至整個國家都有可能造成負面的影響。而在虛假訊息偵測中，立場偵測為其中一個重要的議題，其透過社交媒體、論壇等網站取得的訊息，判斷文本的立場，以收集有用的資訊，並且可以應用於趨勢分析、市場分析、產品用戶評論、意見調查、自動媒體監控、假新聞偵測等，甚至可以過濾未經確認的內容以提供更好的用戶體驗。本研究針對 Fake News Challenge 所提供的 FNC-1 資料集，以基於特徵的機器學習方式進行新聞立場偵測。首先從文本中萃取文本特徵、語義特徵與相似度特徵，並做不同的特徵組合，接著使用天真貝氏、決策樹、隨機森林與梯度提升四種分類演算法，最後進行不同特徵於新聞立場偵測的效能之評估與呈現研究結果。

關鍵字：立場偵測、文本特徵、機器學習、社群媒體

Investigating the Effectiveness of Content Features on News Stance Detection

Student : Chih-Yu Liu

Advisors : Dr. Chin-Sheng Yang

Dr. Shu-Fen Tseng

Department of Information Management

College of Informatics

Yuan Ze University

ABSTRACT

With the rapid development of the internet and social media, we have an easy access to all sorts of information. However, a large amount of information is spread among social media, including fake news, rumors, and misinformation. These messages may affect individuals, organizations, and even entire countries. In the field of natural language processing, stance detection is one of the important issue. It uses information obtained through social media, forums and other websites to determine the stance of the text to collect useful information. It can be applied to trend analysis, market analysis, product user reviews, opinion surveys, automatic media monitoring, fake news detection, etc. Even unconfirmed content can be filtered to provide a better user experience. Based on the FNC-1 dataset provided by Fake News Challenge, this research uses feature-based machine learning to stance detection. First, we extract text features, semantic features and similarity features from the text, and make different feature selection. Then, we use Naive Bayes, Decision Tree, Random Forest and Gradient Boosting classification algorithms. Finally, we evaluate the effectiveness of different features on news stance detection and present the results.

Keywords: Stance detection, Content-based features, Machine Learning, Social Media

誌謝

碩士生涯要結束了，首先要感謝的是我的指導教授楊錦生老師，從剛進來的第一年，就不斷地閱讀了許多論文，增強自己在相關領域的知識，並且透過每次的報告增加自己的口條，過程中遇到問題老師也都會不厭其煩的教導，讓我更能理解每篇論文的內容。接著開始寫程式與論文，每週與老師的開會總是能指引我們往對的方向與目標前進，提供豐富的經驗讓我們遇到困難時也都能迎刃而解。同時要感謝指導教授曾淑芬老師，在寫論文的過程中，也給予了我豐富的建議，讓我能從不同角度去思考，最終順利的完成碩士論文。

另外還要感謝實驗室的學長姊佳萱與志榮，給予我們許多過來人的經驗分享，還會找我們一起吃飯，特別感謝學姊之前有耐心地教我報帳，畢業後也時常關心我們。接著要感謝懿庭與柏欣，在這過程中一起打拼、互相扶持，在程式上曾給了我不少幫助，有你們讓我的碩士生涯順利了許多，還有明靜、楊銘與庭瑜，大家常常一起吃飯、出去走走，為我的碩士生活增添了不少樂趣，很幸運能遇見你們每一個人。最後感謝我的家人，支持我每個決定，讓我能心無旁騖、順利的讀完碩士。

劉芷妤 謹誌

中華民國一一〇年一月

目錄

書名頁.....	i
論文口試委員審定書.....	ii
中文摘要.....	iii
英文摘要.....	iv
誌謝.....	v
目錄.....	vi
表目錄.....	viii
圖目錄.....	xi
第一章 緒論.....	1
1.1 研究背景.....	1
1.2 研究目的.....	2
1.3 研究架構.....	3
第二章 文獻探討.....	4
第三章 研究方法.....	11
3.1 資料來源 (Data Source).....	12
3.2 資料前處理 (Data pre-processing).....	12
3.3 特徵萃取 (Feature Extraction).....	12
3.3.1 文本特徵 (Content Features).....	12
3.3.2 語義特徵 (Semantic Features).....	14
3.3.3 相似度特徵 (Similarity Features).....	15
3.3.4 特徵組合 (Feature selection).....	17
3.4 分類器建構 (Classifiers).....	17
3.5 預測與評估 (Prediction & Evaluation).....	19
第四章 實驗結果與評估.....	20
4.1 資料集.....	20
4.2 評估標準.....	23
4.3 官方基準實驗 (Baseline).....	26
4.4 實驗結果.....	26
4.4.1 分類指標結果.....	26
4.4.2 不同特徵組合對 FNC 分數的影響.....	28

4.4.3 不同分類器對 FNC 分數的影響	32
4.4.4 FNC 競賽結果比較	37
第五章 結論與未來展望	39
5.1 結論	39
5.2 未來展望	40
參考文獻	41
附錄 A GI 辭典類別說明	43
附錄 B LIWC 辭典類別說明	44
附錄 C 分類指標結果—天真貝氏	48
附錄 D 分類指標結果—決策樹	52
附錄 E 分類指標結果—隨機森林	56
附錄 F 分類指標結果—梯度提升	60
附錄 G 所有特徵組合之 FNC 分數展示	64



表目錄

表 1 參考文獻列表.....	8
表 2 Stance Detection dataset for FNC-1 訓練集類別統計數據.....	23
表 3 Stance Detection dataset for FNC-1 測試集類別統計數據.....	23
表 4 新聞立場偵測分類結果之混淆矩陣－範例.....	25
表 5 TF-IDF + TF-IDF 相似度特徵於梯度提升的分類結果.....	27
表 6 TF-IDF + TF-IDF 相似度特徵於隨機森林的分類結果.....	27
表 7 TF-IDF + 所有特徵於梯度提升的分類結果.....	28
表 8 特徵組合 1 之 FNC 分數展示.....	29
表 9 特徵組合 2 之 FNC 分數展示.....	29
表 10 特徵組合 3 之 FNC 分數展示.....	30
表 11 特徵組合 4 之 FNC 分數展示.....	31
表 12 特徵組合 5 之 FNC 分數展示.....	31
表 13 本研究與其他團隊之 FNC 分數結果比較.....	38
表 A 1 GI 辭典類別說明.....	43
表 B 1 LIWC 辭典類別說明.....	44
表 C 1 天真貝氏於 unigram 之分類指標結果.....	48
表 C 2 天真貝氏於 TF-IDF 之分類指標結果.....	48
表 C 3 天真貝氏於 TF-IDF + GI 之分類指標結果.....	49
表 C 4 天真貝氏於 TF-IDF + LIWC 之分類指標結果.....	49
表 C 5 天真貝氏於 TF-IDF + GI + LIWC 之分類指標結果.....	49
表 C 6 天真貝氏於 TF-IDF + TF-IDF Sim 之分類指標結果.....	50
表 C 7 天真貝氏於 TF-IDF + LDA Sim 之分類指標結果.....	50
表 C 8 天真貝氏於 TF-IDF + LSA Sim 之分類指標結果.....	50

表 C 9 天真貝氏於 Sim 之分類指標結果.....	51
表 C 10 天真貝氏於 TF-IDF + Sim 之分類指標結果.....	51
表 C 11 天真貝氏於 TF-IDF + All 之分類指標結果.....	51
表 D 1 決策樹於 unigram 之分類指標結果	52
表 D 2 決策樹於 TF-IDF 之分類指標結果	52
表 D 3 決策樹於 TF-IDF + GI 之分類指標結果.....	53
表 D 4 決策樹於 TF-IDF + LIWC 之分類指標結果.....	53
表 D 5 決策樹於 TF-IDF + GI + LIWC 之分類指標結果	53
表 D 6 決策樹於 TF-IDF + TF-IDF Sim 之分類指標結果	54
表 D 7 決策樹於 TF-IDF + LDA Sim 之分類指標結果.....	54
表 D 8 決策樹於 TF-IDF + LSA Sim 之分類指標結果	54
表 D 9 決策樹於 Sim 之分類指標結果	55
表 D 10 決策樹於 TF-IDF + Sim 之分類指標結果.....	55
表 D 11 決策樹於 TF-IDF +All 之分類指標結果	55
表 E 1 隨機森林於 unigram 之分類指標結果.....	56
表 E 2 隨機森林於 TF-IDF 之分類指標結果.....	56
表 E 3 隨機森林於 TF-IDF + GI 之分類指標結果	57
表 E 4 隨機森林於 TF-IDF + LIWC 之分類指標結果	57
表 E 5 隨機森林於 TF-IDF + GI + LIWC 之分類指標結果.....	57
表 E 6 隨機森林於 TF-IDF + TF-IDF Sim 之分類指標結果.....	58
表 E 7 隨機森林於 TF-IDF + LDA Sim 之分類指標結果	58
表 E 8 隨機森林於 TF-IDF + LSA Sim 之分類指標結果.....	58
表 E 9 隨機森林於 Sim 之分類指標結果.....	59
表 E 10 隨機森林於 TF-IDF + Sim 之分類指標結果	59
表 E 11 隨機森林於 TF-IDF + All 之分類指標結果.....	59

表 F 1 梯度提升於 unigram 之分類指標結果	60
表 F 2 梯度提升於 TF-IDF 之分類指標結果	60
表 F 3 梯度提升於 TF-IDF + GI 之分類指標結果	61
表 F 4 梯度提升於 TF-IDF + LIWC 之分類指標結果	61
表 F 5 梯度提升於 TF-IDF + GI + LIWC 之分類指標結果	61
表 F 6 梯度提升於 TF-IDF + TF-IDF Sim 之分類指標結果	62
表 F 7 梯度提升於 TF-IDF + LDA Sim 之分類指標結果	62
表 F 8 梯度提升於 TF-IDF + LSA Sim 之分類指標結果	62
表 F 9 梯度提升於 Sim 之分類指標結果	63
表 F 10 梯度提升於 TF-IDF + Sim 之分類指標結果	63
表 F 11 梯度提升於 TF-IDF + Sim 之分類指標結果	63
表 G 1 所有特徵組合之 FNC 分數展示	64



圖目錄

圖 1 新聞立場偵測流程圖.....	11
圖 2 標題之原始資料範例.....	20
圖 3 內文之原始資料範例.....	21
圖 4 被標註為同意之文本範例.....	22
圖 5 被標註為不同意之文本範例.....	22
圖 6 效能評估準則流程圖.....	24
圖 7 天真貝氏於新聞立場偵測之效能.....	32
圖 8 決策樹於新聞立場偵測之效能.....	33
圖 9 隨機森林於新聞立場偵測之效能.....	34
圖 10 梯度提升於新聞立場偵測之效能.....	35



第一章 緒論

1.1 研究背景

隨著網際網路與社群媒體的快速發展，人與人之間的互動不再受時間與地點所限制，資訊的取得也不再受限於傳統管道（如：報章雜誌、電視與廣播等），且社群媒體的即時性讓人們能隨時隨地取得最新消息，並且使用者之間也能夠達到雙向的互動性。

根據 DataReportal¹於 2020 年十月最新的調查顯示，世界總人口已達 78.1 億人，全球使用網際網路的用戶為 46.6 億，表示地球上將近 60% 的人在使用網際網路，而其中社群媒體用戶達到 41.4 億，約為總人口的 53%。最新數據表明最新一季的社群媒體使用量大幅增長，從 2020 四月到十月，社群媒體用戶就從 38.1 億提升到 41.4 億，他們認為受 COVID-19 影響，許多人在疫情期間開始接觸與使用社群媒體。然而，大量的訊息在社群媒體之間傳播，其中也包括了假新聞、謠言與錯誤訊息(Ghanem, Rosso, & Rangel, 2018)，這些訊息讓社群媒體中的用戶對於取得的訊息有品質上的疑慮，小則影響到個人（例如：健康或名譽等），嚴重可能甚至會影響到整個國家（例如：國家秩序、安全等）。

有鑑於社交媒體對於現代社會的影響力，在自然語言處理（Natural Language Research, NLP）領域中，從文本中自動萃取有用資訊為重要的研究議題之一，其相關研究包括了觀點方面的情感分析、情緒識別、論點挖掘、真實性與謠言偵測、假新聞偵測等(Küçük & Can, 2020)。在虛假訊息偵測中，立場偵測為其中一個重要的領域，其透過特定主題的貼文內容，偵測社群媒體中文章的作者對於特

¹ <https://datareportal.com/>

定主題、目標或社會議題等的立場（例如：贊同、反對或中立）；或者透過給定的標題與內文文本，偵測兩者之間的立場（例如：同意、不同意、討論或無關）。立場偵測透過社交媒體、論壇等網站取得的訊息，收集可能有助於個人、組織、甚至政府層級做決策的資訊(Al-Ghadir, Azmi, & Hussain, 2020)，例如：趨勢分析、市場調查、產品用戶評論、針對性廣告、選舉預測、自動媒體監控等，甚至可以過濾未經確認的內容以提供更好的用戶體驗(Küçük & Can, 2020)。

根據先前的相關研究顯示，立場偵測的方法包括：基本的機器學習方法(Lai et al., 2016; Al-Ghadir et al., 2020)，例如：從文本中萃取特徵並使用機器學習中的分類演算法像是天真貝氏、決策樹等；使用深度學習的方法(Lai et al., 2016; Riedel et al., 2017; Ghanem et al., 2018)，例如：多層感知機（Multilayer Perceptron，MLP）、神經網路等；集成學習的方法(Liu et al., 2016; Thorne et al., 2017; Mourad et al., 2018)，例如：Voting、Bagging 與 Boosting 等，詳細內容將會在第二章進行說明。

1.2 研究目的

本研究針對 Fake News Challenge²團隊所提供的 FNC-1 資料集做立場偵測，使用監督式學習的方式，透過給定的新聞標題與內文文本，將其分類為「同意」、「不同意」、「討論」或「無關」。先前許多研究皆從文本中萃取特徵，例如：文本特徵（如：n-gram、TF-IDF、目標人物提及、政黨人物提及等）、情感特徵（如：National Research Council，NRC 情感辭典、AFINN 情感辭典等）、結構特徵（如：用戶提及次數、hashtag 次數等）、語言學特徵（如：POS 詞性標註）、語義特徵（如：Linguistic Inquiry and Word Count，LIWC 辭典、General Inquirer，GI 辭典等）、主題模型特徵（如：Latent Dirichlet Allocation，LDA、

² <http://www.fakenewschallenge.org/>

Latent semantic analysis, LSA 等)、文本相似度特徵(如: TF-IDF 餘弦相似度)等,本研究目的為探討哪些文本特徵對於新聞立場偵測較有幫助。此外,根據競賽提供的結果顯示,前三名團隊皆使用到深度學習的方法,但考慮到深度學習的訓練速度需花費較多時間、需要消耗巨大計算資源和硬體需求高等,本研究在不使用深度學習的前提下,利用文本特徵(unigram、TF-IDF)、語義特徵(GI 辭典、LIWC 辭典)與相似度特徵(TF-IDF 相似度、LDA 相似度、LSA 相似度)做不同的特徵組合,並使用天真貝氏、決策樹、隨機森林與梯度提升四種分類演算法對文本進行立場偵測,最終進行比較與評估結果。

1.3 研究架構

本研究共有五個章節,第一章介紹本論文的研究背景、目的與研究架構。第二章進行了文獻探討,將立場偵測相關的研究分為三類,分別為:基本的機器學習方法、深度學習方法與集成學習的方法。第三章為研究方法,介紹了資料來源、資料前處理與特徵萃取,其中特徵萃取包括了文本特徵、語義特徵與相似度特徵三大類,再做不同的特徵組合。接著為模型的建構,使用了天真貝氏、決策樹、隨機森林與梯度提升四種分類演算法。第四章為實驗結果與效果評估,詳細介紹了資料集內容、評估標準與官方基準實驗,並以不同角度評估效果。最後第五章為結論與未來展望。

第二章 文獻探討

立場偵測透過特定主題的貼文內容，偵測社群媒體中文章的作者對於特定主題、特定目標或社會議題等的立場，可能為「贊同」、「反對」或「中立（無；皆否）」(Al-Ghadir et al., 2020; Lai, Farías, Patti, & Rosso, 2016; Liu et al., 2016; Mourad, Shawky, Fayed, & Badawi, 2018; Vijayaraghavan, Sysoev, Vosoughi, & Roy, 2016)；或者透過給定的新聞標題與內文文本，偵測兩者之間的立場，可能為「同意」、「不同意」、「討論」或「無關」(Ghanem et al., 2018; Riedel, Augenstein, Spithourakis, & Riedel, 2017; Thorne et al., 2017)。

近幾年立場偵測備受重視，許多人做了相關研究，甚至有些組織舉辦了相關競賽，以下將介紹其中兩個立場偵測的相關競賽。SemEval-2016³ (Semantic Evaluation, 語義評估) 為 SIGLEX (Special Interest Group on the Lexicon, 計算語言學協會辭典的特殊興趣小組) 於 2016 年所舉辦的國際知名競賽，Task 6 的任務 A 為其中一項任務，針對 Twitter 文章做立場偵測，涵蓋了語義分析與情感分析，使用 SemEval-2016 task 6 資料集，其中資料集包括了五個特定主題：無神論、氣候變遷、女權主義、希拉蕊與墮胎合法化，分類結果為「贊同 (Favor)」、「反對 (Against)」或「皆否 (Neither)」。

此外，Rao 和 Pomerleau⁴於 2017 年舉辦了針對假新聞的立場偵測競賽，目標為透過競賽促進開發幫助事實查核人員使用機器學習、自然語言處理與人工智慧，在新聞報導中識別惡作劇和虛假訊息的工具，從而解決假新聞的問題。其使用 Fake News Challenge (FNC-1) 資料集，對於新聞標題和內文做立場偵測，分類結果為「同意 (Agree)」、「不同意 (Disagree)」、「討論 (Discuss)」或「無關 (Unrelated)」。

³ <https://alt.qcri.org/semeval2016/>

⁴ <http://www.fakenewschallenge.org/>

在機器學習領域中，立場偵測為一種多分類的問題，通常會先將文本做資料前處理，將文本處理成結構性的資料，再進行特徵萃取，讓分類模型學習與訓練，最終對未標記樣本進行預測分類。

立場偵測常見的特徵包括：詞彙特徵（如：詞袋、n-gram、hashtag、立場指示詞、同義字、貼文長度等）、使用者貼文互動特徵（如：轉貼數、回覆數、按讚數等）、情感或情緒詞、詞向量（如：word2vec、GloVe）、TF-IDF、主題模型特徵（如：LDA、LSA 等）、POS 詞性標註與命名實體等(Küçük & Can, 2020)。

本章節將立場偵測分為三類，分別為：基本的機器學習方法（Feature-based Machine Learning）、深度學習方法（Deep Learning）與集成學習方法（Ensemble Learning）。立場偵測常見的方法，機器學習例如：SVM（Support Vector Machine，支持向量機）、Logistic Regression（邏輯迴歸）、Naïve Bayes（天真貝氏）與 Decision Tree（決策樹）等；深度學習例如：MLP（Multilayer Perceptron，多層感知機）、LSTM（Long Short-Term Memory，長短期記憶）、RNN（Recurrent Neural Network，遞迴神經網路）與 CNN（Convolutional Neural Networks，卷積神經網路）等；集成學習例如：Voting、Boosting 與 Bagging 等(Küçük & Can, 2020)，以下將介紹立場偵測相關之研究。

基本的機器學習方法：

(Lai et al., 2016) 使用 SemEval-2016 資料集針對 2016 美國總統大選進行立場偵測，使用高斯天真貝氏分類器，特徵包括：情感特徵（LIWC、DAL 辭典等）、結構特徵（hashtags 次數、用戶提及次數、標點符號等）、文本特徵（目標人物的提及、政黨提及等），本篇強調文本特徵對於立場偵測的重要，且認為立場偵測高度依賴於手中目標領域的知識。

(Al-Ghadir et al., 2020) 使用 SemEval-2016 task 6 資料集，使用 SVM、KNN (K Nearest Neighbor, K 近鄰演算法) 與四種不同 KNN 的延伸版本 (區別度、權重、類別與範例)，特徵使用：1. 以詞頻篩選 top-k 個單詞與詞幹清單，計算 TF-IDF 作為特徵；2. NRC 八個情緒特徵 (憤怒、恐懼、期待、信任、驚喜、悲傷、喜悅、厭惡) 與兩個情感特徵 (正面、負面)，最終結果為基於權重的 KNN (Weighted-based K-NN) 表現最佳且優於現有研究之技術。

深度學習方法：

(Vijayaraghavan et al., 2016) 使用 SemEval-2016 task 6 資料集，提出了字元級 (character-level) 和單詞級 (word-level) 的卷積神經網路，使用 one-hot encoding 將推文以二進制矩陣表示，結果為字元級效果較佳，Macro F1 score 為 63.5%，但認為需要數量更大且更多樣性的數據，才能達到更好的效果。

(Riedel et al., 2017) 使用 FNC-1 資料集，將標題與內文文本的詞頻 (TF) 和標題與內文文本的 TF-IDF 相似度連接作為特徵，輸入至具有一個隱藏層的 MLP 中，獲得 FNC 競賽第三名。

(Ghanem et al., 2018) 使用 FNC-1 資料集，特徵為：1. Cue words 清單，類別包括信仰、否認、懷疑、報告、知識、否定和假的，並使用 Google 新聞 word2vec 擴展，對每個單詞取得 5 個最相似單詞；2. 對於每個標題內文計算詞嵌入的餘弦相似度；3. FNC-1 Baseline 特徵，包括 n-gram、單詞重疊 (overlap)、極性 (polarity) 與反駁詞 (refuting) 等。其測試了 SVM、梯度提升、隨機森林、天真貝氏與神經網路五種方法，最終為包含兩個隱藏層的神經網路表現最佳，Macro F1 score 為 59.6%，僅次於 FNC 第二名。

集成學習方法：

(Liu et al., 2016) 使用 SemEval-2016 task 6 資料集，提出的方法為：梯度提升

決策樹 (Gradient Boosting Decision Tree, GBDT) 使用 GloVe 詞向量特徵、隨機森林與 SVM 使用 unigram 特徵並用 Information Gain 做特徵選擇，最終將所有包含的分類器合併到一個集成系統中，分別加上或不加上全域特徵 (此研究定義之情感分數公式)。結果顯示集成方法僅在墮胎與女權議題表現較佳，對於其他主題則是隨機森林與梯度提升決策樹的效果較佳。

(Thorne et al., 2017) 使用 FNC-1 資料集，提出的方法由五種獨立分類器的堆疊集成 (Stacked Ensemble) 分別為：1. 詞向量、標題與內文 TF-IDF 餘弦相似度、反駁詞使用 MLP；2. 詞嵌入向量排除停用詞與標點符號、單詞重疊、反駁詞使用 MLP；3. unigram 與 bigram TF-IDF 使用邏輯迴歸；4. 詞嵌入向量使用 MLP；5. FNC baseline 方法，包括 n-gram、單詞重疊 (overlap)、極性

(polarity) 與反駁詞 (refuting) 等特徵，並使用梯度提升分類器。最終使用梯度提升決策樹與上述五種預測結果堆疊集成。此研究獲得 FNC 競賽第十一名。

(Mourad et al., 2018) 使用 SemEval-2016 task 6 資料集，因不同分類器分別於不同主題表現較佳，此篇使用由隨機森林、SVM 與高斯天真貝氏，三個分類器組成的投票方法決定最終分類結果。特徵為語言學 (如：n-gram、POS、TF-IDF 等)、主題特徵 (LDA)、推文特徵 (如：推文長度、hashtag 等)、標記特徵 (如：情感、意見)、詞嵌入、相似度、文本特徵、情感辭典 (如：GI、AFINN 等)。其中使用 ReliefF 挑選特徵，最重要的特徵包括：標記特徵、n-gram、POS、hashtag。結果顯示該模型於「希拉蕊」與「無神論」兩主題皆優於其他研究，總體平均也只較最佳方法相差 0.01%，強調基於特徵的模型不依賴於神經網路也能具有競爭力的評估結果。

表 1 為上述所有文獻探討之列表整理，依照年份排序，表格欄位依序為參考文獻、資料集、方法或模型與摘要。

表 1 參考文獻列表

參考文獻	資料集	方法／模型	摘要
Liu et al. (2016)	SemEval-2016 task 6	1. Gradient Boosting Decision Tree (GBDT) 2. Random Forest (RF) 3. SVM 4. Ensemble Classifier	<ul style="list-style-type: none"> ✧ GBDT 使用 GloVe 詞向量特徵；RF 和 SVM 使用 unigram + Information Gain 特徵選擇；集成方法使用上述三種分類器，分別加或不加全域特徵（本篇定義情感分數公式） ✧ 集成方法僅在墮胎與女權議題表現較佳，對於其他主題則是 RF 與 GBDT 的效能更好
Vijayaraghavan, Sysoev, Vosoughi, & Roy (2016)	SemEval-2016 task 6	character and word level Convolutional Neural Networks (CNN)	<ul style="list-style-type: none"> ✧ 使用 character-level 和 word-level 的 CNN ✧ character-level 表現較佳，但認為需更大更多樣的數據，才能達到更好的效果 ✧ Macro F1 score: 63.5%
Lai, Farías, Patti, & Rosso (2016)	SemEval-2016 (2016 U.S. presidential elections)	Gaussian Naive Bayes	<ul style="list-style-type: none"> ✧ 針對 2016 美國總統大選做立場偵測 ✧ 使用情感特徵 (LIWC、DAL 等)、結構特徵 (hashtags 頻率、標點符號等)、文本特徵 (目標人物提及、政黨提及等) ✧ 強調文本特徵對於立場偵測的重要，且立場偵測高度依賴於手中目標領域的知識。

Thorne et al. (2017)	Fake News Challenge (FNC-1)	1. Logistic Regression 2. Multi-Layer Perceptron (MLP) 3. Gradient Boosted Decision Tree	✧ 五個獨立分類器的堆疊集成 (stacked ensemble) ✧ 特徵：1. 詞向量、餘弦相似度、反駁詞 2. 詞嵌入向量排除停用詞與標點符號、單詞重疊、反駁詞 3. unigram 與 bigram TF-IDF 4. 詞嵌入向量 5. FNC baseline ✧ 獲得 FNC 競賽第十一名
Riedel et al. (2017)	Fake News Challenge (FNC-1)	MLP with one hidden layer	✧ 將標題與內文文本的 TF 和標題與內文文本的 TF-IDF 相似度連接作為特徵，輸入至具有一個隱藏層的 MLP 中 ✧ 為 FNC 競賽結果第三名
Ghanem, Rosso, & Rangel (2018)	Fake News Challenge (FNC-1).	1. SVM 2. Gradient Boost 3. Random Forest 4. Naive Bayes 5. Neural Network	✧ Cue words 清單，類別包括：信仰、否認、懷疑、報告、知識、否定和假的，並使用 Google 新聞 word2vec 擴展，對每個單詞取得 5 個最相似單詞 ✧ 對於每個標題內文計算詞嵌入的餘弦相似度 ✧ FNC-1 features ✧ Macro F1 score: 59.6%，僅次於 FNC 第二名
Mourad et al. (2018)	SemEval-2016 task 6	1. Random Forest 2. linear SVM 3. Gaussian Naïve Bayes 最終使用 Vote 決定分類結果	✧ 因不同分類器分別於不同主題表現較佳，本篇使用投票決定最終分類結果 ✧ 特徵：語言學 (n-gram、POS、TF-IDF 等)、主題特徵 (LDA)、推文特徵 (推文長度、hashtag 等)、標記特徵 (情感、意見)、詞嵌

			<p>入、相似度、文本特徵、情感辭典（GI、AFINN 等）。</p> <p>✧ 使用 ReliefF 挑選特徵，重要特徵包括：標記特徵、n-gram、POS、hashtag</p>
Al-Ghadir, Azmi, & Hussain (2020)	SemEval-2016 task 6	<p>1. SVM</p> <p>2. KNN</p> <p>3. DKNN (Discernibility-based)</p> <p>4. WKNN (Weighted-based)</p> <p>5. CKNN (Class-based)</p> <p>6. EKNN (Exemplar-based)</p>	<p>✧ 以詞頻篩選 top-k 個單詞與詞幹清單，計算 TF-IDF 作為特徵</p> <p>✧ NRC 八個情緒特徵（憤怒、恐懼、期待、信任、驚喜、悲傷、喜悅、厭惡）與兩個情感特徵（正面、負面）</p> <p>✧ Weighted-based K-NN 基於權重的 KNN 表現最佳且優於現有研究之技術</p>
Küçük & Can (2020)	線上辯論、Twitter、新聞等立場偵測相關資料集	<p>立場偵測常見方法：</p> <p>1. 機器學習：SVM、Logistic Regression、Naïve Bayes、Decision Tree</p> <p>2. 深度學習：MLP、LSTM、RNN、CNN</p> <p>3. 集成學習方法：Voting、Boosting、Bagging</p>	<p>✧ 對於立場偵測做了調查，說明了概念、定義、通用架構、相關資料集、指標及應用等</p> <p>✧ 常見特徵：詞彙特徵（詞袋、n-gram、hashtag 等）、使用者貼文互動特徵（轉貼、回覆等）、情感情緒詞、詞向量、TF-IDF、主題模型（LDA、LSA）、POS 標記、命名實體等</p>

第三章 研究方法

本章節將詳細介紹本研究所使用的新聞立場偵測技術之架構與實驗設計，包括資料集的資料來源、資料前處理與特徵萃取。特徵將分為三大類，分別為：文本特徵、語義特徵與相似度特徵。最後為模型的建構，使用了四種常見的分類器做訓練，下一章節將會展示預測結果與效能評估。圖 1 為本研究新聞立場偵測（News Stance Detection）之流程圖：

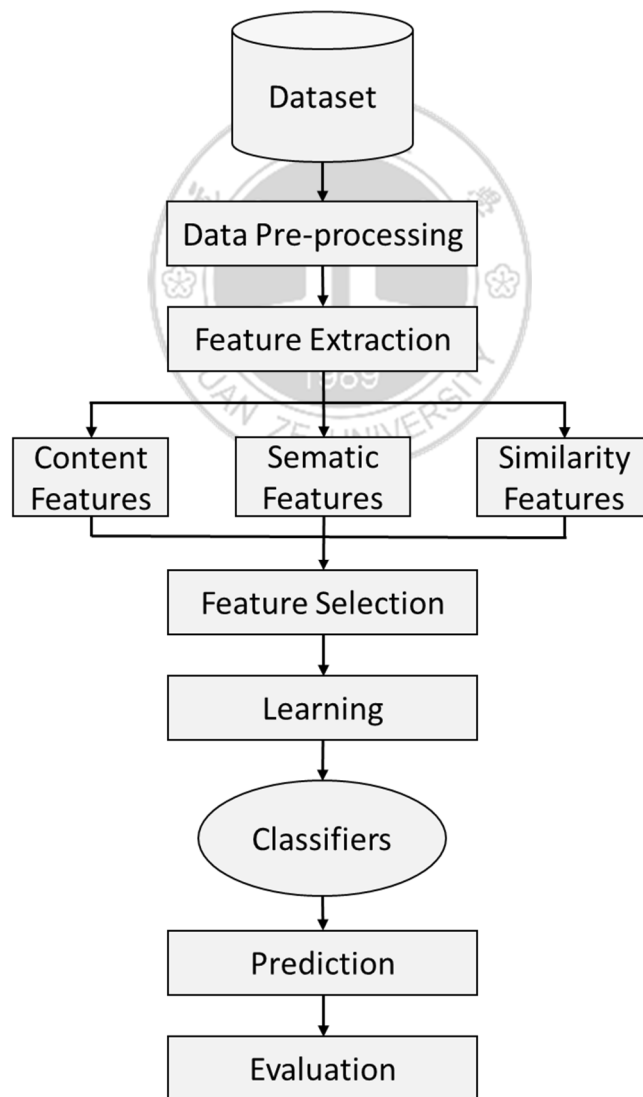


圖 1 新聞立場偵測流程圖

3.1 資料來源 (Data Source)

本研究使用 Fake News Challenge⁵ (FNC) 所提供的資料集，此數據來自於 Craig Silverman 創建的 Emergent (Ferreira & Vlachos, 2016) 資料集，由記者們從各種來源像是謠言網站(例如：snopes.com)、Twitter 帳戶(例如：@Hoaxalizer)蒐集而來，其中主題包括全世界和美國的新聞與技術報導等。資料集包含了訓練集與測試集，資料主要由新聞標題(Headline)與內文(Body)所組成，第四章將會詳細說明資料集內容。

3.2 資料前處理 (Data pre-processing)

為了使資料能有效輸入至機器學習的模型訓練，必須先將資料做一些前處理。首先將所有文本轉為小寫英文字母，並使用 Python 自然語言處理工具箱 NLTK (Natural Language Tool Kit) 的 word_tokenize 模組，對資料集中每個文本包括標題與內文，進行斷詞處理(Tokenization)，將文本的句子處理成分開的單詞，以利在下一步特徵萃取時做單詞的計算。停用詞(Stop words)為出現頻率高卻沒有統計意義的單詞，本研究利用 NLTK 的停用詞列表，將文本中的停用詞去除。

3.3 特徵萃取 (Feature Extraction)

本研究探討的是文本特徵於新聞立場偵測的效能，因此選擇與新聞內容相關的特徵，且將其分為三大類，分別是文本特徵、語義特徵與相似度特徵，接下來會依序做詳細的介紹。

3.3.1 文本特徵 (Content Features)

本研究利用計算新聞文本中單詞的數量與 TFIDF 計算，找出新聞內容中較具有影響力的單詞作為特徵，以下將分別介紹本研究使用到的 unigram 與 TF-IDF 兩種文本特徵。

⁵ <http://www.fakenewschallenge.org/>

➤ **unigram**

Unigram 計算單詞在單一文本中出現的次數。本研究將標題文本做 unigram 得出的 4,477 個單詞，加上內文文本做 unigram 得出的 29,264 個單詞，共 33,741 個單詞作為 unigram 特徵。

➤ **TF-IDF**

TF-IDF (Salton, Buckley, & management, 1988) (Term Frequency - Inverse Document Frequency，詞頻-反向文件頻率)，是用於資訊檢索與文字探勘常見的加權技術，評估單詞對於一個文本集或是單詞對於文本集中其中一份文本的重要程度。單詞的重要性與它在單一文本中出現的次數成正比；與它在整個文本集中出現的次數成反比。

詞頻 (Term Frequency, TF) 為單詞出現在單一文本的次數；反向文件頻率 (Inverse Document Frequency, IDF) 為單詞普遍重要性的度量。若單詞在單一文本中出現很多次，也就是 TF 很高，表示 w 在單一文本中有一定的重要性，但同時考慮到整個文本集時，若 w 在所有文本中出現次數愈多，表示參考價值愈低，意即 IDF 小；反之若 w 在單一文本出現次數高，但在所有文本中出現次數低，則表示 w 具有獨特性，參考價值也就高，意即 IDF 大。將 TF 與 IDF 相乘即可得到 TF-IDF。

為了挑選出參考價值較高的單詞，本研究首先計算在整個文本集 D_j 中的單詞 t_i 之 TF-IDF 值，公式如下：

$$tf_{i,j} = n_{i,j} \quad (n_{i,j} : \text{單詞 } t_i \text{ 在文本集 } D_j \text{ 中出現的次數})$$

$$idf_i = \log \frac{|D|}{|\{j : t_i \in d_j\}|}$$

$$(|D| : \text{文本集中文本總數} ; |\{j : t_i \in d_j\}| : \text{包含單詞 } t_i \text{ 的文本數})$$

$$tf \times idf_{i,j} = tf_{i,j} \times idf_i \quad (\text{式 1})$$

透過 TF-IDF 公式(式 1)，計算整個文本集中所有單詞的 TF-IDF，分別從標題文本與內文文本中，篩選出 TF-IDF 前 3,000 高的單詞，將它們聯集後共得出 4,282 個單詞，將其作為特徵。

接著針對單一文本，計算在單一文本 d_j 中的單詞 t_i 之 TF-IDF 值，公式如下：

$$tf_{ij} = n_{ij} \quad (n_{ij} : \text{單詞 } t_i \text{ 在單一文本 } d_j \text{ 中出現的次數})$$

$$idf_i = \log \frac{|D|}{|\{j : t_i \in d_j\}|}$$

($|D|$ ：文本集中文本總數； $|\{j : t_i \in d_j\}|$ ：包含單詞 t_i 的文本數)

$$tf \times idf_{ij} = tf_{ij} \times idf_i \quad (\text{式 2})$$

透過 TF-IDF 公式 (式 2) 計算每筆標題文本與內文文本對於前面篩選出 4,282 個單詞的 TF-IDF 當作特徵值，再將其連接起來。

3.3.2 語義特徵 (Semantic Features)

考慮到語義特徵可能對新聞文本的立場偵測有幫助，本研究選擇了兩個知名的語義辭典，分別為 GI 辭典與 LIWC 辭典，透過計算新聞內文中針對不同語義類別的單詞數作為特徵。

➤ GI 辭典

哈佛大學建置的 General Inquirer (Stone, Dunphy, & Smith, 1966)，簡稱 GI 辭典，合併了哈佛字典 (Harvard IV-4 Dictionary) 和拉斯維爾字典 (Lasswell's Dictionary)，共有 182 個標籤類別，11,788 個單詞。因 GI 辭典類別數與各類別中的單詞數眾多，本研究挑出我們認為和新聞立場偵測較相關的類別，分別是：Positiv、Negativ、Affil、Hostile、Strong、Power、Weak、Submit、Active、Passive、Virtue、Vice、Yes、

No、Negate 與 Intrj，上述 16 個類別之詳細說明請見附錄 A，接著計算每筆內文文本對於該類別所包含的單詞個數作為特徵。我們挑選的類別像是正面詞、負面詞、強烈的、消極的、是與否…等，而其他像是特定機構語言、宗教、位置路線、物體等較無關的類別則未列入當作特徵。

➤ LIWC 辭典

Linguistic Inquiry and Word Count (Pennebaker, Francis, & Booth, 2001) (LIWC，語文探索與字詞計算)，包含 4,486 個單詞，64 個類別，例如：語文特性類別（如：代名詞、冠詞等）、副語言學類別（如：應和詞、停頓詞等）、心理特性類別（如，情感詞、認知詞等）、個人化類別（如工作、休閒活動等）等。因 LIWC 辭典之單詞數與類別數量適中，本研究採用全部共 64 個類別，計算出每筆內文文本對於該類別所包含的單詞數作為特徵，詳細的類別說明請見附錄 B。

3.3.3 相似度特徵 (Similarity Features)

為了檢測新聞的立場為相關（包括同意、不同意、討論）或無關，本研究將計算新聞標題與內容的相似度，包括直接計算標題與內文文本的 TF-IDF 餘弦相似度和使用兩種主題模型去計算餘弦相似度，其中在主題特徵 LDA 與 LSA 方面，我們執行過主題數分別為 5、10 與 15 的實驗，但結果非常相似，因此本研究最終皆採用主題數為 10 作為特徵，以下將作詳細的介紹。

➤ 文本相似度 (TF-IDF Similarity)

餘弦相似度 (Cosine Similarity⁶)，藉由測量兩個向量夾角的餘弦值來度量它們之間的相似性，與向量大小無關，僅與向量的方向有關。向

⁶ https://en.wikipedia.org/wiki/Cosine_similarity

量之間的夾角愈小，表示兩向量的方向愈接近，意即相似度愈高。當兩向量夾角為 0° 時，餘弦值為 1；當兩向量夾角為 90° ，餘弦值為 0；當兩向量夾角為 180° ，餘弦值為 -1。餘弦值愈接近 1 表示兩者愈相似。本研究以標題文本與內文文本的 TF-IDF 向量，計算每筆標題文本與內文文本之間的餘弦相似度，作為一個特徵。

餘弦相似度公式如下：

$$\text{similarity} = \cos(\theta) = \frac{\vec{h} \cdot \vec{b}}{\|\vec{h}\| \|\vec{b}\|} = \frac{\sum_{i=1}^n h_i \times b_i}{\sqrt{\sum_{i=1}^n (h_i)^2} \times \sqrt{\sum_{i=1}^n (b_i)^2}} \quad (\text{式 3})$$

(\vec{h} ：標題文本的 TF-IDF 向量； \vec{b} ：內文文本的 TF-IDF 向量)

➤ LDA 相似度

Latent Dirichlet Allocation (Blei, Ng, & Jordan, 2003) (LDA，隱含狄利克雷分布)，是一種主題模型，主題模型在機器學習和自然語言處理等領域中，是用來在一系列文檔中找出抽象主題的一種統計模型，它會自動分析每個文檔，統計文檔中的詞語，根據統計的訊息來斷定當前文檔含有哪些主題，以及每個主題所占的比例。LDA 將文檔集中每篇文檔的主題以概率的方式呈現，為無監督式學習的算法，意即訓練時不需要手工標註的訓練集，只需要給定文檔集並指定主題的數量。本研究先利用訓練集的標題文本與內文文本訓練 LDA 模型，並將主題數設為 10，接著再透過餘弦相似度公式（式 3），計算每筆標題文本的 LDA 向量與內文文本的 LDA 向量之間的餘弦相似度。

➤ LSA 相似度

Latent Semantic Analysis (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990) (LSA，潛在語義分析) 探討的是隱藏在字詞背後的某

種關係，這種關係不是以詞典上的定義為基礎，而是以字詞的使用環境作為最基本的參考。它以向量模型為基底，常以單詞-文檔 (term-document) 矩陣表示單詞與文檔之間的關聯。本研究使用 sklearn 中的 TruncatedSVD (aka LSA) 將訓練集的標題文本與內文文本降維至 10 個主題的向量空間，接著透過餘弦相似度的公式 (式 3)，計算每筆標題文本的 LSA 向量與內文文本的 LSA 向量之間的餘弦相似度。

3.3.4 特徵組合 (Feature selection)

將上述 7 種特徵做不同的組合，並分別丟入分類器做訓練，使用的分類器將在下一小節說明。本研究提出了五種特徵組合分別如下，透過不同的特徵組合，評估文本特徵於新聞立場偵測的效能：

- 特徵組合 1. 文本特徵
- 特徵組合 2. 文本特徵 + 語義特徵
- 特徵組合 3. 文本特徵 + 相似度特徵
- 特徵組合 4. 相似度特徵
- 特徵組合 5. 文本特徵 + 語義特徵 + 相似度特徵

3.4 分類器建構 (Classifiers)

Python 的 Scikit-learn (SKlearn) 為針對 Python 語言的機器學習庫，具有各種分類、迴歸和分群演算法。本研究針對前一小節所提的五大特徵組合，使用 SKlearn 套件下的天真貝氏 (Naive Bayes)、決策樹 (Decision Tree)、隨機森林 (Random Forest) 與梯度提升 (Gradient Boosting) 四種分類演算法做分類器建構。以下依序介紹本研究所使用的四種分類演算法：

➤ **天真貝氏 (Naive Bayes)**

天真貝氏(Agrawal, Bayardo, & Srikant, 2000)是一種簡單並以機率為基礎的分類器，它假定在判別特徵為某一類別時的機率為獨立的，也就是樣本的每個特徵之間都是不相關的。考慮到本研究的相似度特徵為連續值，這裡使用 SKlearn 中的高斯天真貝氏 (GaussianNB) 作為分類演算法。

➤ **決策樹 (Decision Tree)**

決策樹(Quinlan, 2014)由節點和有向邊組成。其中每個內部節點皆表示一個特徵，葉節點表示為某一類別，樣本從根節點通過內部節點各個特徵的判斷，最終達到屬於它的葉節點，也就是它的所屬分類。決策樹易於理解和實現，能在相對短的時間內對大型數據做出可行且效果良好的結果。這裡使用 SKlearn 中的決策樹分類器 (DecisionTreeClassifier) 作為分類演算法。

➤ **隨機森林 (Random Forest)**

隨機森林(Breiman, 1996)是一個包含多個決策樹的分類器，分類結果由多個決策樹決定，哪個類別多就分為哪一類。其概念就是結合多個弱學習器來建構一個更強的學習器模型，這種方法又稱為集成方法 (Ensemble method)。隨機森林的學習過程快速，可以處理大量的輸入變數，產生高準確度的分類器。這裡使用 SKlearn 中的隨機森林分類器 (RandomForestClassifier) 作為分類演算法。

➤ **梯度提升 (Gradient Boosting)**

梯度提升(Friedman, 2001)也是一種集成方法，在每一次的預測中，都保留了先前的學習函數，並多加入一個新的函數再學習，期望得出最

好的模型。本研究使用 SKlearn 中的梯度提升分類器
(GradientBoostingClassifier) 作為分類演算法。

3.5 預測與評估 (Prediction & Evaluation)

將前一小節建構好的天真貝氏、決策樹、隨機森林與梯度提升共四種分類器，針對不同的特徵組合對測試集做新聞立場預測，而預測結果與效能的評估與比較，將在第四章做進一步的介紹。



第四章 實驗結果與評估

以 Fake News Challenge 團隊所提出的評估方法為標準，評估本研究所提出文本特徵於新聞立場偵測的方法，本章節將詳細介紹使用的資料集、評估標準，以及最終實驗結果與效能之評估。

4.1 資料集

本研究使用的資料集為 Fake News Challenge 團隊提供的新聞立場偵測資料集 Stance Detection dataset for FNC-1⁷，此資料集包含了 49,972 筆訓練集與 25,418 筆測試集。其中訓練集包含了標題（Headline）與內文（Body）兩份檔案，標題有 49,972 筆，資料欄位為：[Headline, Body ID, Stance]，原始資料範例如圖 2 所示；內文有 2,533 筆，資料欄位為：[Body ID, articleBody]，原始資料範例如圖 3 所示。

1	Headline	Body ID	Stance
2	Police find mass graves with at least '15 bodies' near Mexico town where 43 students disappeared after police clash	712	unrelated
3	Hundreds of Palestinians flee floods in Gaza as Israel opens dams	158	agree
4	Christian Bale passes on role of Steve Jobs, actor reportedly felt he wasn't right for part	137	unrelated
5	HBO and Apple in Talks for \$15/Month Apple TV Streaming Service Launching in April	1034	unrelated
6	Spider burrowed through tourist's stomach and up into his chest	1923	disagree
7	'Nasa Confirms Earth Will Experience 6 Days of Total Darkness in December' Fake News Story Goes Viral	154	agree
8	Accused Boston Marathon Bomber Severely Injured In Prison, May Never Walk Or Talk Again	962	unrelated
9	Identity of ISIS terrorist known as 'Jihadi John' reportedly revealed	2033	unrelated
10	Banksy 'Arrested & Real Identity Revealed' Is The Same Hoax From Last Year	1739	agree
11	British Aid Worker Confirmed Murdered By ISIS	882	unrelated
12	Gateway Pundit	2327	discuss
13	Woman detained in Lebanon is not al-Baghdadi's wife, Iraq says	1468	agree
14	Kidnapped Nigerian schoolgirls: Government claims ceasefire deal with Boko Haram that will bring missing girls home	1003	unrelated

圖 2 標題之原始資料範例

⁷ <https://github.com/FakeNewsChallenge/fnc-1>

1	Body ID	articleBody
		<p>A small meteorite crashed into a wooded area in Nicaragua's capital of Managua overnight, the government said Sunday. Residents reported hearing a mysterious boom that left a 16-foot deep crater near the city's airport, the Associated Press reports.</p> <p>Government spokeswoman Rosario Murillo said a committee formed by the government to study the event determined it was a "relatively small" meteorite that "appears to have come off an asteroid that was passing close to Earth." House-sized asteroid 2014 RC, which measured 60 feet in diameter, skimmed the Earth this weekend, ABC News reports.</p> <p>0 Murillo said Nicaragua will ask international experts to help local scientists in understanding what happened.</p> <p>The crater left by the meteorite had a radius of 39 feet and a depth of 16 feet, said Humberto Saballos, a volcanologist with the Nicaraguan Institute of Territorial Studies who was on the committee. He said it is still not clear if the meteorite disintegrated or was buried.</p> <p>Humberto Garcia, of the Astronomy Center at the National Autonomous University of Nicaragua, said the meteorite could be related to an asteroid that was forecast to pass by the planet Saturday night.</p>
2		<p>Last week we hinted at what was to come as Ebola fears spread across America. Today, we get confirmation. As The Daily Caller reports, one passenger at Dulles International Airport outside Washington, D.C. is apparently not taking any chances. A female passenger dressed in a hazmat suit - complete with a full body gown, mask and gloves - was spotted Wednesday waiting for a flight at the airport.</p>

圖 3 內文之原始資料範例

實驗時，我們會先根據標題的 Body ID 去對應該 Body ID 所指向其所屬的內文，換句話說 Body ID 介於 0 至 2532 之間，在 49,972 筆不同的標題中，他們有可能對應到相同的內文，但其被標註的分類結果不盡相同。我們將標題與其所屬內文合併為同一個文本，合併後欄位為 [Headline, Body ID, articleBody, Stance]，共 49,972 筆。其中 Stance 為新聞內文相對於標題的立場類別，其包含「同意 (Agree)」、「不同意 (Disagree)」、「討論 (Discuss)」、與「無關 (Unrelated)」。

文本類別由蒐集資料的記者們根據立場之定義，人工標註而成(Ferreira & Vlachos, 2016)，其中新聞內文相對於標題的立場之定義為：若內文與標題一致，被標註為「同意」；若內文不同意標題，被標註為「不同意」；若內文討論與標題相同的主題，但並未表明立場，被標註為「討論」；若內文討論的主題與標題不同，則被標註為「無關」。

舉例來說，圖 4 為被標註為「同意 (Agree)」的文本（因範圍過大只顯示片段），其標題為「以色列開放水壩時，數百名巴勒斯坦人逃離加沙的洪水」，而新聞內文一開始就提到「由於最近冬季嚴峻的暴風雨，以色列當局在邊界附近開了許多水壩，淹沒了加沙谷地，星期天早晨，數百名巴勒斯坦人被撤離家園...」，因此本篇內文為完全同意標題的立場，被標記為「同意 (Agree)」。

index	Headline	Body ID	Stance	articleBody
39	Hundreds of Palestinians flee floods in Gaza as Israel opens dams	158	agree	"Hundreds of Palestinians were evacuated from their homes Sunday morning after Israeli authorities opened a number of dams near the border, flooding the Gaza Valley in the wake of a recent severe winter storm. The Gaza Ministry of Interior said in a statement that civil defense services and teams from the Ministry of Public Works had evacuated more than 80 families from both sides of the Gaza Valley (Wadi Gaza) after their homes flooded as water levels reached more than three meters. Gaza has experienced flooding in recent days amid a major storm that saw temperatures drop and frigid rain pour down. The storm displaced dozens and caused hardship for tens of thousands, including many of the approximately 110,000 Palestinians left homeless by Israel's assault over summer. The suffering is compounded by the fact that Israel has maintained a complete siege over Gaza for the last eight years, severely limiting electricity and the availability of fuel for generators. It has also prevented the displaced from rebuilding their homes, as construction materials are largely banned from entering."

圖 4 被標註為同意之文本範例

接著為被標註為「不同意 (Disagree)」之文本範例，（因範圍過大只顯示片段），其標題為「蜘蛛從遊客的胃中鑽入並伸入他的胸部」，但內文中提到「...我們可能都被錯誤訊息所誤導...」，並且在後續內文中證實了此為虛構的故事，因此此文本被標註為「不同意 (Disagree)」。

index	Headline	Body ID	Stance	articleBody
251	Spider burrowed through tourist's stomach and up into his chest	1923	disagree	"Fear not arachnophobes, the story of Bunbury's ""spiderman"" might not be all it seemed. Perth scientists have cast doubt over claims that a spider burrowed into a man's body during his first trip to Bali. The story went global on Thursday, generating hundreds of stories online. Earlier this month, Dylan Thomas headed to the holiday island and sought medical help after experiencing ""a really burning sensation like a searing feeling"" in his abdomen. Dylan Thomas says he had a spider crawl underneath his skin. Thomas said a specialist dermatologist was called in and later used tweezers to remove what was believed to be a ""tropical spider"". But it seems we may have all been caught in a web... of <u>misinformation</u> ."

圖 5 被標註為不同意之文本範例

接著介紹資料集的類別分佈，表 2 為訓練集中，樣本在四個分類分別的筆數與比率，「無關」的文本數超過七成；「討論」與「同意」分別佔了 17.82% 和 7.36%；「不同意」的文本數最少，僅佔了 1.68%：

表 2 Stance Detection dataset for FNC-1 訓練集類別統計數據

	無關	討論	同意	不同意	總計
筆數	36,545	8,909	3,678	840	49,972
百分比	73.13	17.82	7.36	1.68	100

表 3 為測試集中，樣本分別在四個類別中的筆數與比率，和訓練集一樣的是「無關」文本數也超過七成；「不同意」文本數比例較訓練集高，佔了 17.57%；「討論」為 7.49%；文本數最少的為「同意」類別，僅佔了 2.74%：

表 3 Stance Detection dataset for FNC-1 測試集類別統計數據

	無關	討論	同意	不同意	總計
筆數	18,349	1,903	697	4,464	25,413
百分比	72.20	7.49	2.74	17.57	100

4.2 評估標準

分類問題常使用分類指標（例如：準確率、精準率、召回率...等）來評估訓練出模型的好壞，此外，本研究主要使用了 Fake News Challenge 團隊提供的 FNC 分數標準來評估本實驗之結果，並與其他參賽團隊做比較，下方將依序說明兩種評估標準。

分類指標：

本研究使用分類器對新聞立場做預測時，計算了分類問題常用的指標，包括：混淆矩陣（Confusion matrix）與其衍伸出的準確率（Accuracy）、精準率

(Precision)、召回率 (Recall) 與調和平均數 (F1-score)，其中混淆矩陣將在下一個段落進行說明，以下簡單介紹四個指標：

- 準確率：預測正確的結果占總樣本的百分比。
- 精準率：所有被預測為正的樣本中，實際為正樣本的比率。
- 召回率：所有實際為正的樣本中，被預測為正的比率。
- 調和平均數： $(2 * \text{精準率} * \text{召回率}) / (\text{精準率} + \text{召回率})$

FNC 分數：

Fake News Challenge 團隊提供了一套評估準則，在此稱之為 FNC 分數⁸，本研究將使用 FNC 分數評估分類模型效能，評分包含下列兩階段：

第一階段：將標題與內文分為相關或無關，佔 25% 權重分數。

第二階段：將相關分類為同意、不同意或討論，佔 75% 權重分數。

圖 6 為 Fake News Challenge 團隊提出的效能評估準則流程圖⁹：

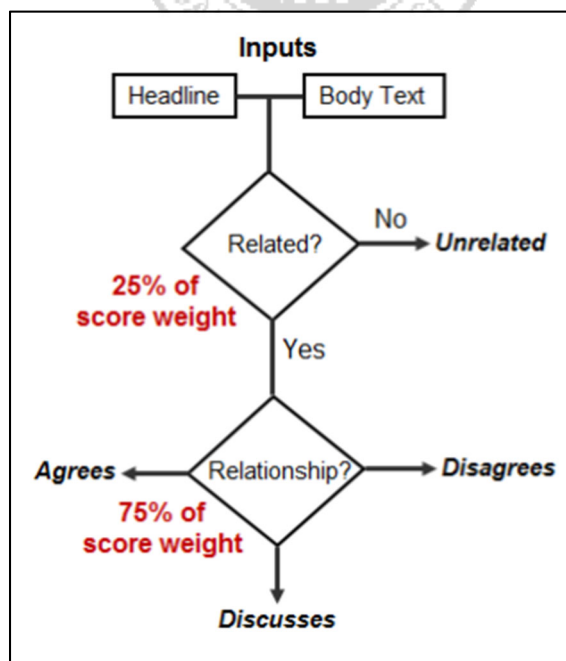


圖 6 效能評估準則流程圖

⁸ <https://github.com/FakeNewsChallenge/fnc-1/blob/29d473af2d15278f0464d5e41e4cbe7eb58231f2/scorer.py>

⁹ <http://www.fakenewschallenge.org/>

舉個例子，表 4 為新聞立場偵測某分類結果之混淆矩陣：

表 4 新聞立場偵測分類結果之混淆矩陣－範例

預測 \ 真實	同意	不同意	討論	無關
同意	606	4	1,172	121
不同意	199	3	386	109
討論	979	11	3,206	268
無關	89	2	361	17,897

計算方法：

「獲得分數」：

- 正確將相關分類為相關（包括同意、不同意、討論）者，皆獲得 0.25 分
- 正確將無關分類為無關者，獲得 0.25 分
- 正確將相關分類為同意、不同意、討論者，獲得 0.75 分

「最高總分」：

- 四項類別文本總數*0.25 + 相關類別（同意、不同意、討論）文本數*0.75
- 上述「獲得分數總和」除以「最高總分」，即為結果評估的分數。

計算過程如下：

「獲得分數」：

- $(606 + 4 + 1,172 + 199 + 3 + 386 + 979 + 11 + 3,206) * 0.25 = 6,566 * 0.25 = 1,641.5$
- $17,897 * 0.25 = 4,474.25$
- $(606 + 3 + 3,206) * 0.75 = 3815 * 0.75 = 2,861.25$

「最高總分」：（同意：1,903，不同意：697，討論：4,464，無關：18,349）

- $(1,903 + 697 + 4,464 + 18,349) * 0.25 + (1,903 + 697 + 4,464) * 0.75$
 $= 6,353.25 + 5,298 = 11,651.25$

- $(1,641.5 + 4,474.25 + 2,861.25) / 11,651.25 = 8,977 / 11,651.25 = 77.05\%$

由上述計算結果得知 77.05% 即為評估結果的 FNC 分數。

4.3 官方基準實驗 (Baseline)

Fake News Challenge 團隊所提出的 Baseline 實驗，首先對文本做了轉小寫、刪除非字母數字、移除停用詞等資料前處理，接著使用的特徵包括 n-gram、單詞重疊 (overlap)、極性 (polarity) 與反駁詞 (refuting) 等，再使用梯度提升 (Gradient Boosting) 分類器進行新聞立場偵測，最終 FNC 分數達到 75.20%。

4.4 實驗結果

首先展示其中幾個表現較佳的特徵組合在使用特定分類器條件下的分類指標結果 (準確率、精準率、召回率、調和平均數與混淆矩陣)；接著第二小節以 3.3.4 章節所提出五大特徵組合的角度，展示天真貝氏、決策樹與隨機森林三種分類器在不同特徵組合的情況下，於新聞立場偵測的實驗結果 (FNC 分數) 與效能比較；第三小節則分別從每個分類器的角度做分析；最後再將本研究較佳的結果與 Fake News Challenge 競賽基準實驗和其他表現優異的團隊比較。

4.4.1 分類指標結果

以下將展示幾個表現較佳的分類結果與其各個分類指標，首先表 5 為 TF-IDF 加上 TF-IDF 相似度特徵在使用梯度提升分類器的情況下之分類結果，其準確度達 88.19%。對於不同類別的精準率、召回率、調和平均數也都不同，透過表 5 可以發現在「無關」類別中的表現較最佳，也是樣本數最多的類別，三個指標皆達到 96%-98%；接著為樣本數第二多的「討論」類別，調和平均數為 74.72%；樣本數第三的「同意」和樣本數最少的「不同意」類別之精準率只有大約 50%，召回率則是非常低。此結果顯示，TF-IDF 加上 TF-IDF 相似度特徵使用梯度提升分類器時，其精準率、召回率與調和平均數皆與樣本數成正比。

表 5 TF-IDF + TF-IDF 相似度特徵於梯度提升的分類結果

Gradient Boosting	TFIDF + TFIDF Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8819	0.5318	0.2680	0.3564
不同意		0.5000	0.0201	0.0386
討論		0.6686	0.8468	0.7472
無關		0.9646	0.9868	0.9756
混淆矩陣	[[510 5 1227 161] [122 14 421 140] [311 9 3780 364] [16 0 226 18107]]			
FNC 分數	9354.5 / 11651.25 = <u>80.29%</u>			

表 6 為 TF-IDF 加上 TF-IDF 相似度特徵於隨機森林的分類結果，其中準確率達 86.99%，在「無關」這個最大類別中的表現較佳，三個指標皆達到 96%；接著依序為第二、第三類別的「討論」與「同意」；最小類別的「不同意」在隨機森林分類器中，則未將預測樣本分類到此類別。此結果顯示，在使用 TF-IDF 加上 TF-IDF 相似度特徵的情況下，隨機森林的精準率、召回率與調和平均數也皆與樣本數成正比。

表 6 TF-IDF + TF-IDF 相似度特徵於隨機森林的分類結果

Random Forest	TFIDF + TFIDF Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8699	0.5079	0.4404	0.4717
不同意		0.0000	0.0000	0.0000
討論		0.6632	0.7787	0.7163
無關		0.9606	0.9696	0.9651
混淆矩陣	[[838 0 870 195] [198 0 355 144] [597 0 3476 391] [17 0 540 17792]]			
FNC 分數	9267 / 11651.25 = <u>79.54%</u>			

表 7 為 TF-IDF 加上所有特徵並使用梯度提升分類器之結果，準確率達 86.97%，「無關」類別的三個指標皆達 96-98%，其次為「討論」類別，調和平均數達 68.70%；「同意」類別的調和平均數僅有 48.55%；「不同意」類別則表現最差，調和平均數不到 1%。

表 7 TF-IDF + 所有特徵於梯度提升的分類結果

Gradient Boosting	TFIDF + ALL			
	準確率	精準率	召回率	調和平均數
同意	0.8697	0.4221	0.5712	0.4855
不同意		0.2692	0.0301	0.0542
討論		0.7229	0.6546	0.6870
無關		0.9654	0.9848	0.9750
混淆矩陣	[[1087 17 645 154] [315 21 229 132] [1142 39 2922 361] [31 1 246 18071]]			
FNC 分數	9144.5 / 11651.25 = <u>78.49%</u>			

本研究最終發現，所有實驗結果皆與上述三個例子相似，分類結果之精準率、召回率與調和平均數皆與樣本數成正比，樣本數越多的類別的預測效果皆越高，甚至有些情況分類器並未將樣本分類到最小類別「不同意」（例如：隨機森林），其他結果請參考附錄 C 至附錄 F。

4.4.2 不同特徵組合對 FNC 分數的影響

接著透過前一小節所得出的混淆矩陣，套用 4.2 章節所提 FNC 分數的算法，計算出 FNC 分數作為評估指標，再分別從不同特徵組合和不同分類器的角度，進行評估與比較，所有結果之表格整合請見附錄 E。

特徵組合 1（文本特徵）：

首先特徵組合 1 針對兩個文本特徵：unigram 與 TF-IDF，選擇效果較佳

的特徵作為後續建構分類器的基礎。偵測結果如表 8 所示，TF-IDF 於四種分類器的 FNC 分數皆達 41% 以上。本研究考慮到兩個因素：一、unigram 特徵數眾多，影響分類演算法執行效率；二、表 8 中 TF-IDF 的 FNC 分數表現較穩定，且於天真貝氏與決策樹分類器的效果較 unigram 佳，因此本研究最終選擇 TF-IDF 作為文本特徵，建構後續分類器。

表 8 特徵組合 1 之 FNC 分數展示

FNC score	Naïve Bayes	DecisionTree	Random Forest	Gradient Boosting
unigram	26.67%	35.97%	46.92%	43.54%
TFIDF	42.56%	47.89%	44.64%	41.60%

特徵組合 2 (文本特徵 + 語義特徵):

特徵組合 2 為文本特徵加上語義特徵，其中包括 GI 辭典與 LIWC 辭典，表格中兩種以上特徵組合的 FNC 分數高於單一文本特徵 TF-IDF 時，以粗體表示。如表 9 所示加了語義特徵的 FNC 分數效果幾乎差異不大甚至有些變差，提高的部分最多也不到 1%。此結果顯示在我們的做法中，TF-IDF 加上 GI 或 LIWC 辭典特徵時，對於新聞立場偵測並沒有太大的幫助。

表 9 特徵組合 2 之 FNC 分數展示

FNC score	Naïve Bayes	Decision Tree	Random Forest	Gradient Boosting
TFIDF	42.56%	47.89%	44.64%	41.60%
TFIDF + GI	43.01%	48.87%	43.88%	41.97%
TFIDF + LIWC	43.00%	45.52%	43.75%	42.25%
TFIDF + GI + LIWC	43.01%	47.04%	39.75%	42.16%

特徵組合 3 (文本特徵 + 相似度特徵):

特徵組合 3 為文本特徵加上相似度特徵，其中包括 TF-IDF 相似度 (TF-IDF Sim)、LDA 相似度 (LDA Sim) 與 LSA 相似度 (LSA Sim)。表格中兩種以上特徵組合的 FNC 分數高於單一文本特徵 TF-IDF 時以粗體表示，若分數高於 FNC Baseline (如 4.3 節所示為 75.20%)，則以雙底線表示。

透過表 10 可以看到加了相似度特徵時，分數皆大於或等於單一文本特徵 TF-IDF 的分數，其中天真貝氏提升的效果較不明顯，但在決策樹、隨機森林與梯度提升分類器皆有較明顯的效果提升。TF-IDF 加上 TF-IDF 相似度時，決策樹提升了 24.99%，到達 72.88%；隨機森林提升了 34.9%，達到 79.54%且高於 FNC Baseline；梯度提升更是提高了 38.69%，達到 80.29%且高於 FNC Baseline。而 TF-IDF 同時加上三種相似度特徵 (TF-IDF 相似度+LDA 相似度+LSA 相似度) 時，決策樹、隨機森林與梯度提升分別提高了 28.33%、31.49%與 36.54%，達到 76.22%、76.13%與 78.14%，皆高於 FNC Baseline。此結果說明相似度特徵對於新聞立場偵測，在使用決策樹、隨機森林與梯度提升的情況下，帶來的正面效果頗為顯著。

表 10 特徵組合 3 之 FNC 分數展示

FNC score	Naïve Bayes	Decision Tree	Random Forest	Gradient Boosting
TFIDF	42.56%	47.89%	44.64%	41.60%
TFIDF + TFIDF Sim	42.59%	72.88%	<u>79.54%</u>	<u>80.29%</u>
TFIDF + LDA Sim	42.56%	53.98%	54.40%	54.67%
TFIDF + LSA Sim	42.56%	52.37%	51.38%	53.99%
TFIDF + TFIDF Sim + LDA Sim + LSA Sim	42.60%	<u>76.22%</u>	<u>76.13%</u>	<u>78.14%</u>

特徵組合 4 (相似度特徵):

本研究在特徵組合 4 測試了在不考慮文本特徵 TF-IDF 的情況下，單純只有三種相似度特徵 (TF-IDF 相似度+ LDA 相似度+ LSA 相似度)，對於新聞立場偵測的效果。結果如表 11 所示，在四種分類演算法的情況下，FNC 分數皆高於 70%，皆較單一文本特徵 TF-IDF 高，其中天真貝氏之 FNC 分數最高，到達 77.19% 且高於 FNC Baseline。

表 11 特徵組合 4 之 FNC 分數展示

FNC score	Naïve Bayes	Decision Tree	Random Forest	Gradient Boosting
TFIDF	42.56%	47.89%	44.64%	41.60%
TFIDF Sim + LDA Sim + LSA Sim	<u>77.19%</u>	70.24%	73.87%	74.16%

特徵組合 5 (文本特徵 + 語義特徵 + 相似度特徵):

特徵組合 5 合併了所有特徵，結果如表 12 所示，天真貝氏的 FNC 分數較單一文本特徵 TF-IDF 只提升不到 1%；決策樹與隨機森林分別提升了 25.6%與 29.06%，高於單一文本特徵 TF-IDF，但皆未超越 FNC Baseline；梯度提升分類器則較單一文本特徵 TF-IDF 提升了 36.52%，達到 78.49%且高於 FNC Baseline。

表 12 特徵組合 5 之 FNC 分數展示

FNC score	Naïve Bayes	Decision Tree	Random Forest	Gradient Boosting
TFIDF	42.56%	47.89%	44.64%	41.60%
TFIDF + TFIDF Sim + LDA Sim + LSA Sim+ GI + LIWC	43.07%	73.49%	73.70%	<u>78.49%</u>

4.4.3 不同分類器對 FNC 分數的影響

這裡將以四個分類器的角度，依序以圖表方式展現各個特徵組合的 FNC 分數，並且與 FNC Baseline 做比較。圖表中的 Y 軸表示 FNC 分數；X 軸表示不同特徵組合，其中三個相似度特徵（TF-IDF 相似度+ LDA 相似度+ LSA 相似度）將以「Sim」表示，兩個語義特徵（GI 辭典 + LIWC 辭典）加上三個相似度特徵將以「All」表示。

天真貝氏：

首先為天真貝氏分類器，如圖 7 所示，TF-IDF 的 FNC 分數（42.56%）高於 unigram（26.67%）；此外，只有在三個相似度特徵（Sim）時，FNC 分數到達 77.19%且高於 FNC Baseline（75.20%），其他特徵組合的效果皆在 43%上下，沒有較明顯的效果提升。

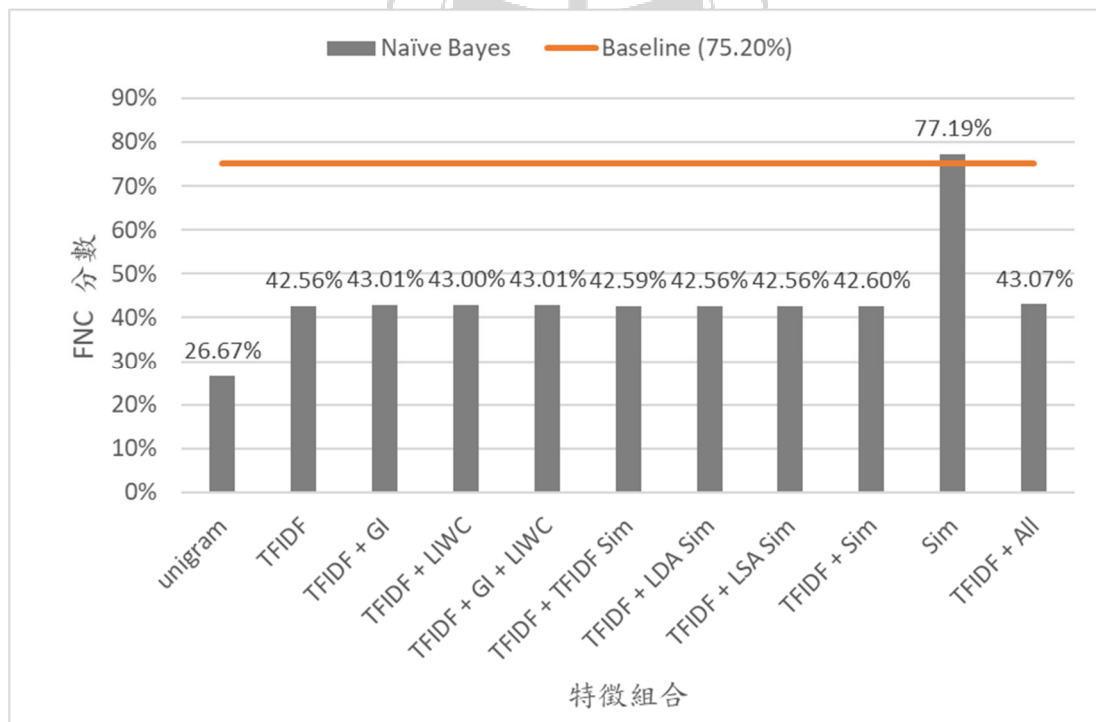


圖 7 天真貝氏於新聞立場偵測之效能

決策樹：

接著為決策樹分類器，如圖 8 可見 TF-IDF 加上三個相似度特徵（TF-IDF + Sim）時，效果高於 FNC Baseline（75.20%），到達 76.22%；TF-IDF 加上 TF-IDF 相似度特徵（TF-IDF + TF-IDF Sim）時，效果為 72.88%、TF-IDF 加上所有特徵（TF-IDF + All）時效果為 73.49%，兩者皆接近 FNC Baseline。其餘 TF-IDF 加上語義特徵（GI、LIWC）時，無明顯的提升效果；TF-IDF 加上主題模型 LDA 與 LSA 時，平均只提升了 5%。

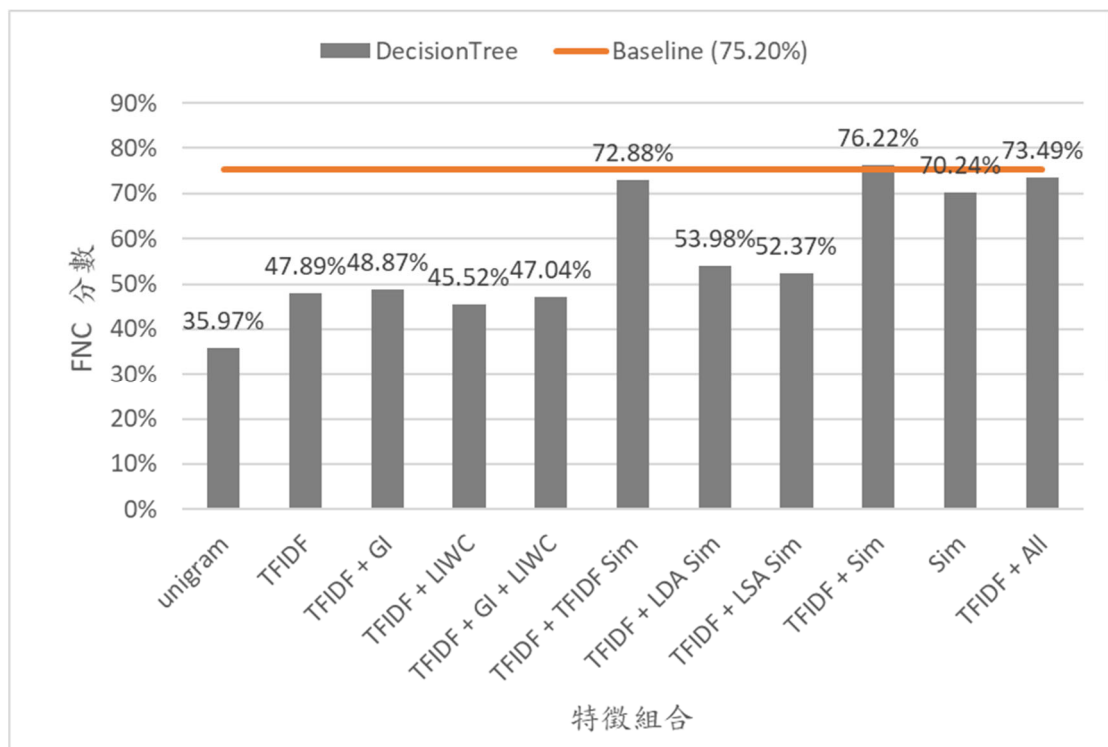


圖 8 決策樹於新聞立場偵測之效能

隨機森林：

隨機森林分類器如圖 9 所示，TF-IDF 加上 TF-IDF 相似度（TF-IDF + TF-IDF Sim）時，FNC 分數最高，到達 79.54%且高於 FNC Baseline（75.20%）；TF-IDF 加上三個相似度特徵（TF-IDF + Sim）時，到達了 76.13%，高於 FNC Baseline；除此之外，只有三種相似度特徵（Sim）和 TF-IDF 加上所有特徵（TF-IDF + All）兩種情況下，效果分別到達 73.87%和 73.70%，接近 FNC Baseline。

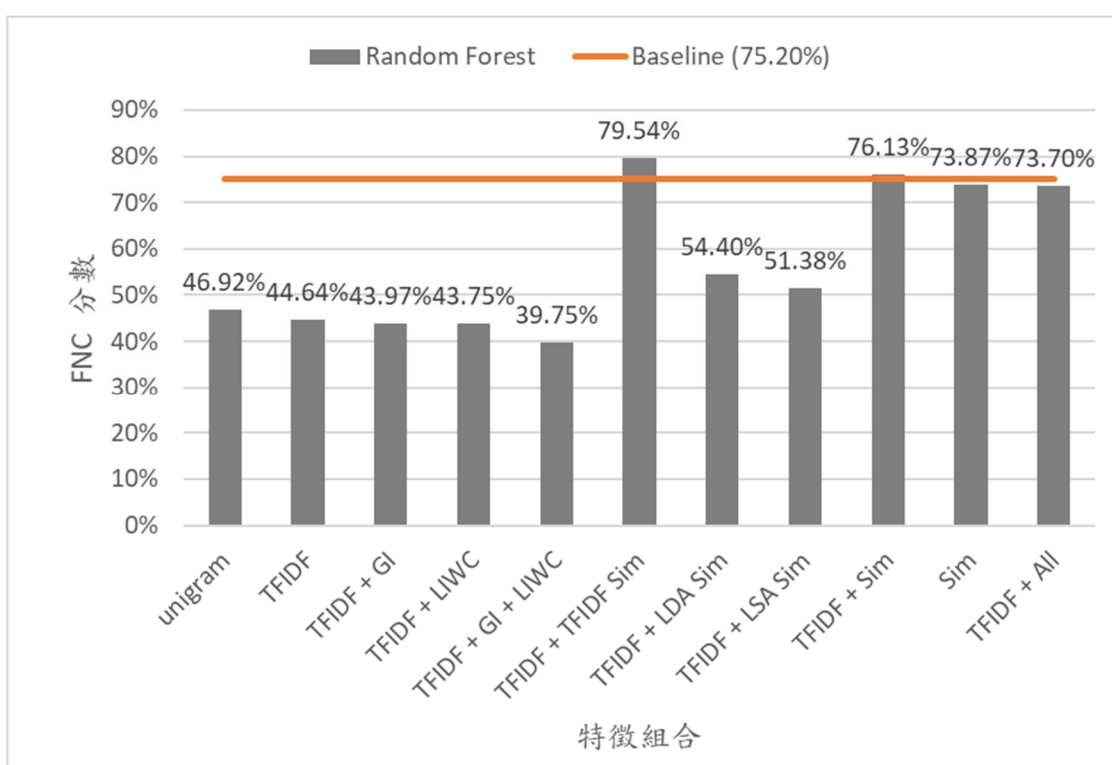


圖 9 隨機森林於新聞立場偵測之效能

梯度提升：

如圖 10 所示，梯度提升分類器在 TF-IDF 加上 TF-IDF 相似度（TF-IDF + TF-IDF Sim）時，FNC 分數最高，到達 80.29%且高於 FNC Baseline（75.20%）；TF-IDF 加上三個相似度特徵（TF-IDF + Sim）和 TF-IDF 加上所有特徵（TF-IDF + All）時，FNC 分數分別達到 78.14%和 78.49%，皆高於

FNC Baseline (75.20%)；單純只有三個相似度特徵 (Sim) 時的效果達到 74.16%，接近 FNC Baseline。

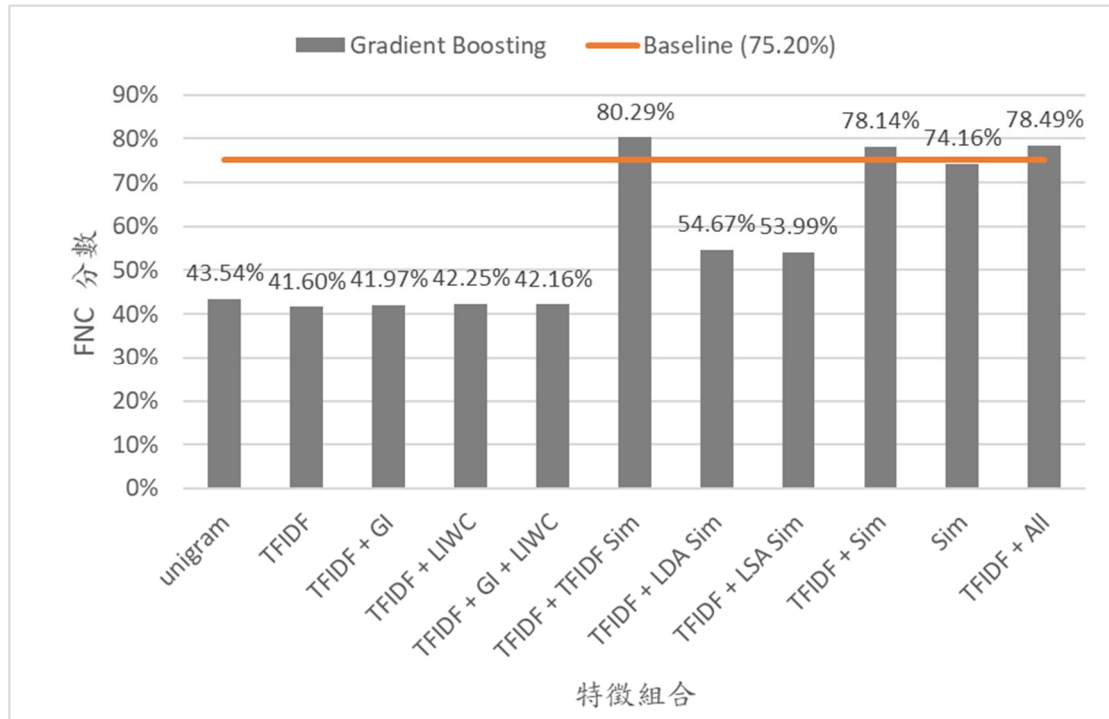


圖 10 梯度提升於新聞立場偵測之效能

綜觀以上，梯度提升分類器之分數普遍高於其他三者，其次為隨機森林。其中 FNC 分數最高者為 TF-IDF 加上 TF-IDF 相似度特徵 (TF-IDF + TF-IDF Sim)，在使用梯度提升分類器時，效果達到 80.29%；第二高同樣為此特徵但在使用隨機森林分類器的情況下，效果為 79.54%；接著為梯度提升分類器在 TF-IDF 加上三種相似度特徵 (TF-IDF + Sim) 時與 TF-IDF 加上所有特徵 (TF-IDF + All) 時，FNC 分數分別達到 78.14%與 78.49%。天真貝氏的分數普遍較低，只有在三個相似度特徵 (Sim) 時最高，達到 77.19%。另外 TF-IDF 加上三個相似度特徵 (TF-IDF + Sim) 時，決策樹與隨機森林兩種分類器的效果分別到達 76.22%與 76.13%。上述所有情況下，本研究之新聞立場偵測的效能皆高於 FNC Baseline (75.20%)，且皆和相似度特徵有關，特別是 TF-IDF 相似度特徵，由此可見 TF-IDF 相似度對於新聞立場偵測有顯著

的重要性。

另外從圖 8 到圖 10 可以發現 TF-IDF 分別加上兩種主題模型 LDA 與 LSA 在使用決策樹、隨機森林與梯度提升三種分類器的情況下，較單一文本特徵 TF-IDF 提升了 10% 左右；而語義特徵包括 GI 辭典與 LIWC 辭典，使用四種分類器的效果皆無明顯提升，可見在我們的方法中，語義特徵對於新聞立場偵測沒有太大的幫助。

總結來說，在 TF-IDF 加上 TF-IDF 相似度特徵 (TF-IDF + TF-IDF Sim)、TF-IDF 加上三個相似度特徵 (TF-IDF + Sim) 與 TF-IDF 加上所有特徵 (TF-IDF + All) 時，皆為使用梯度提升分類器的效果最佳。



4.4.4 FNC 競賽結果比較

Fake News Challenge 公布了競賽前三名團隊的資訊與前 50 名團隊分別獲得的 FNC 分數，前三名團隊皆使用了多層感知器（Multilayer Perceptron，MLP），為一種神經網路，基本結構由輸入層、隱藏層與輸出層所組成，其中的隱藏層做資料的特徵擷取（降維），這個過程由資料去學習而來，最後的輸出層即為分類。以下簡單介紹前三名團隊分別所使用的方法：

1. SOLAT in the SWEN¹⁰:

來自思科（Cisco）系統的 Talos Intelligence 團隊，提出以 50/50 加權平均套用在兩個模型：1. 使用兩個一維卷積神經網路（Convolutional Neural Networks，CNN），將標題與內文文本的詞嵌入送到具有三個隱藏層的 MLP 中；2. 計數（Count）、TF-IDF、奇異值分解（Singular Value Decomposition，SVD）、詞向量（Word2Vec）與情感（Sentiment）共五大特徵集並使用梯度提升決策樹（Gradient Boosted Decision Trees）。最終所得 FNC 分數為 82.02%，獲得 FNC 競賽第一名。

2. Athene (UKP Lab) (Hanselowski, Avinesh, Schiller, & Caspelherr, 2017)

來自德國達姆施塔特工業大學（Technische Universität Darmstadt）的團隊，使用五個單獨的 MLP 集合，其中每個都有七個隱藏層，並且提出七個特徵集：FNC Baseline 特徵、詞袋一元語法（Bag-of-words unigram）、非負矩陣分解餘弦距離、非負矩陣分解連接、潛在語意學（Latent Semantic Indexing，LSI）與單詞相似度，分類結果由五個單獨的 MLP 投票而成。最終所得 FNC 分數為 81.97%，獲得競賽第二名。

¹⁰ <https://github.com/Cisco-Talos/fnc-1>

3. UCL Machine Reading (Riedel et al., 2017)

來自英國倫敦大學學院（University College London，UCL）的 Machine Reading 團隊，將標題與內文文本的 TF 和標題與內文文本的 TF-IDF 相似度連接作為特徵，輸入至具有一個隱藏層的 MLP 中。最終結果 FNC 分數為 81.72%，為此競賽的第三名。

本研究挑選三種特徵組合下的最佳結果，和 FNC 競賽前五名團隊與官方 FNC Baseline 比較，結果如表 13 所示，TF-IDF + TF-IDF 相似度（TFIDF + TFIDF Sim）特徵在使用梯度提升分類器時的效果最好，FNC 分數到達 80.29%，高於 FNC 競賽團隊中的第四名；此外，TF-IDF 加上三個相似度特徵（TF-IDF + Sim）和 TF-IDF 加上所有特徵（TF-IDF + All）時，也是使用梯度提升分類器的效果最佳，結果雖未超越競賽前五名之團隊，但皆高於 FNC Baseline。

表 13 本研究與其他團隊之 FNC 分數結果比較

團隊 / 本研究方法	所得分數	FNC 分數
1. SOLAT in the SWEN	9556.500	82.02%
2. Athene (UKP Lab)	9550.750	81.97%
3. UCL Machine Reading	9521.500	81.72%
本研究：TFIDF + TFIDF Sim，梯度提升	9354.500	80.29%
4. Chips Ahoy!	9345.500	80.21%
5. CLUlings	9285.000	79.69%
本研究：TF-IDF + Sim，梯度提升	9103.750	78.14%
本研究：TF-IDF + All，梯度提升	9144.500	78.49%
FNC Baseline	8761.74	75.20%

第五章 結論與未來展望

5.1 結論

Fake News Challenge 為著名的立場偵測競賽，競賽中前三名皆使用深度學習的方法，本研究的目的是在於使用簡單的機器學習方法，讓新聞立場偵測在不使用深度學習方法的情況下，依然能得到好的分類效果。本研究使用 Fake News Challenge 提供的資料集，對新聞的標題與內文文本做新聞立場偵測。利用文本特徵、語義特徵與相似度特徵做不同的特徵組合，並使用天真貝氏、決策樹、隨機森林與梯度提升四種分類器，最終透過計算 FNC 分數比較各個特徵組合於不同分類器之效能。

實驗結果顯示，以分類器的角度來看，天真貝氏普遍表現較差；集成方法的分類器（隨機森林與梯度提升）表現較佳，其中又為梯度提升表現最好。

在特徵方面，文本特徵中 TF-IDF 的效果較 unigram 佳；TF-IDF 相似度為最有效的特徵，其對於新聞立場偵測存在著重要的影響力；而 TF-IDF 單獨加上 LDA 與 LSA 主題模型時，對分類效果有些提升，但皆無法超越 FNC Baseline；TF-IDF 加上語義特徵（LIWC 與 GI 辭典）則皆無明顯的提升效果，因此語義特徵於本研究的作法中，對新聞立場偵測沒有太大的幫助。

總結來說，我們使用 TFIDF + TFIDF 相似度特徵，在使用梯度提升分類器時，FNC 分數到達 80.29%，接近競賽前三名的 81.72%，這說明了本研究在不使用深度學習的情況下，依然具有不錯的分類結果。

5.2 未來展望

以下以列點方式描述本研究之未來展望：

分類方法之改善：

- (1) 本研究在語義特徵的效果較不明顯，我們是單純計算標題與內文相對於語義類別所出現的單詞數作為特徵，未來可以針對計算標題與內文之間的語義相似度作為特徵，也許能提高效果；另外也可以嘗試其他特徵像是詞嵌入（word embedding）。
- (2) 本研究使用了天真貝氏、決策樹、隨機森林與梯度提升，共四種單獨的分類演算法，未來可以進行不同方法的嘗試，例如：集成學習中的 Stacking。
- (3) 目前採取分類的方法為直接分四類（同意、不同意、討論與無關），未來期望將分類過程分成兩階段：一、先使用二元分類器將樣本分類為「相關」與「無關」兩類；二、將第一階段被分為「相關」的樣本抽出，再使用多分類的分類器將樣本分類為「同意」、「不同意」或「討論」。第一階段為針對標題與內文是無相關進行分類，而第二階段較主要是判斷標題與內文之間的立場，兩階段分別所適用效果較佳的特徵與分類器應該有所差異，期許能透過此方式提升預測效果。

於不同資料集之應用：

- (1) 因本研究僅針對 FNC-1 資料集進行文本特徵的新聞立場偵測，但因不同資料集會有不同的特性，若能將我們的方法應用在其他包含標題與內文的資料集中，就能檢測我們的方法是否於不同資料集也能表現一致的效果。

參考文獻

- Agrawal, R., Bayardo, R., & Srikant, R. (2000). *Athena: Mining-based interactive management of text databases*. Paper presented at the International Conference on Extending Database Technology.
- Al-Ghadir, A. I., Azmi, A. M., & Hussain, A. J. I. F. (2020). A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments. *67*, 29-40.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. J. J. o. m. L. r. (2003). Latent dirichlet allocation. *3*(Jan), 993-1022.
- Breiman, L. J. M. l. (1996). Bagging predictors. *24*(2), 123-140.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. J. J. o. t. A. s. f. i. s. (1990). Indexing by latent semantic analysis. *41*(6), 391-407.
- Ferreira, W., & Vlachos, A. (2016). *Emergent: a novel data-set for stance classification*. Paper presented at the Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies.
- Friedman, J. H. J. A. o. s. (2001). Greedy function approximation: a gradient boosting machine. 1189-1232.
- Ghanem, B., Rosso, P., & Rangel, F. (2018). *Stance detection in fake news a combined feature representation*. Paper presented at the Proceedings of the First Workshop on Fact Extraction and VERification (FEVER).
- Hanselowski, A., Avinesh, P., Schiller, B., & Caspelherr, F. J. F. N. C. (2017). Description of the system developed by team athene in the fnc-1.
- Küçük, D., & Can, F. J. A. C. S. (2020). Stance detection: A survey. *53*(1), 1-37.
- Lai, M., Farias, D. I. H., Patti, V., & Rosso, P. (2016). *Friends and enemies of clinton and trump: using context for detecting stance in political tweets*. Paper presented at the Mexican International Conference on Artificial Intelligence.
- Liu, C., Li, W., Demarest, B., Chen, Y., Couture, S., Dakota, D., . . . Pancholi, M. (2016). *Iucl at semeval-2016 task 6: An ensemble model for stance detection in twitter*. Paper presented at the Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016).
- Mourad, S. S., Shawky, D. M., Fayed, H. A., & Badawi, A. H. (2018). *Stance detection in tweets using a majority vote classifier*. Paper presented at the International Conference on Advanced Machine Learning Technologies and Applications.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. J. M. L. E. A. (2001). Linguistic inquiry and word count: LIWC 2001. *71*(2001), 2001.

- Quinlan, J. R. (2014). *C4. 5: programs for machine learning*: Elsevier.
- Rapp, A., Beitelspacher, L. S., Grewal, D., & Hughes, D. E. J. J. o. t. A. o. M. S. (2013). Understanding social media effects across seller, retailer, and consumer interactions. *41*(5), 547-566.
- Riedel, B., Augenstein, I., Spithourakis, G. P., & Riedel, S. J. a. p. a. (2017). A simple but tough-to-beat baseline for the Fake News Challenge stance detection task.
- Salton, G., Buckley, C. J. I. p., & management. (1988). Term-weighting approaches in automatic text retrieval. *24*(5), 513-523.
- Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). The general inquirer: A computer approach to content analysis.
- Thorne, J., Chen, M., Myrianthous, G., Pu, J., Wang, X., & Vlachos, A. (2017). *Fake news stance detection using stacked ensemble of classifiers*. Paper presented at the Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism.
- Vijayaraghavan, P., Sysoev, I., Vosoughi, S., & Roy, D. J. a. p. a. (2016). Deepstance at semeval-2016 task 6: Detecting stance in tweets using character and word-level cnns.



附錄 A GI 辭典類別說明

表 A 1 GI 辭典類別說明

類別名稱	單詞個數	類別說明	單詞舉例
Positiv	1915	正面詞	
Negativ	2291	負面詞	
Affil	557	從屬或支持	abide 遵守, agree 同意, admire 欣賞, approval 認可
Hostile	833	敵意或侵略性的態度	attack 攻擊, avenge 報仇, annoy 惱人的, anger 憤怒
Strong	1902	較強烈、積極的	aggressive 積極的, absolute 絕對, achieve 實現
Power	689	權利、控制或權威的	abolish 廢除, administer 管理, authorize 授權
Weak	755	較消極的	abandon 放棄, absence 缺席, amateur 業餘
Submit	284	服從、脆弱或是對他人依賴；提交	Accept 接受, admit 承認, apology 道歉, apply 申請
Active	2045	主動、活躍的	Abolish 廢除, accede 加入, accentuate 強調
Passive	911	被動的	abdicate 退位, accustomed 習慣, affliction 痛苦
Virtue	719	美德：對道德認同或好運的	Able, absolute 絕對的, acceptable 可接受的, admirable 令人欽佩的
Vice	685	惡：對道德不滿或不幸的	Abnormal 異常, abominable 可惡, abuse 濫用, acrimony 刻薄
Yes	20	直接表示同意 "of course", "all right".	Absolute 絕對的, agree, course 當然, definitely 肯定的
No	7	直接表明不同意，可以分別標識出是否存在或否定	Disagree, no, nope, wrong, mean#5 (by no means)絕不
Negate	217	20 個 "dis", 40 個 "in", 100 個 "un",和本身有"no"; 通常表示不利的一面	ain't 不是, aversion 厭惡, can't, disadvantage
Intrj	42	包括感嘆號，隨意詞和俚語	Alas 唉, alright, blah 等等, damn, goddamn

附錄 B LIWC 辭典類別說明

表 B 1 LIWC 辭典類別說明

編號	類別名稱	中文	單詞舉例
1	funct	功能詞	A, about
2	pronoun	代名詞	I, them, itself
3	ppron	特定人稱代名詞	I, them, her
4	i	第一人稱單數代名詞	I, me, mine
5	we	第一人稱複數代名詞	We, us, our
6	you	第二人稱代名詞	You, your
7	Shehe	第三人稱單數代名詞	She, her, him
8	they	第三人稱複數代名詞	They, their
9	ipron	非特定人稱代名詞	It, it's, those
10	article	冠詞	A, an, the
11	Verb	動詞	Walk, went, see
12	Auxverb	助動詞	Am, will, have
13	past	過去式	Went, ran, had
14	Present	現在式	Is, does, hear
15	Future	未來式	Will, gonna
16	Adverb	副詞	Very, really, quickly

17	preps	介系詞	To, with, above
18	conj	連接詞	And, but, whereas
19	Negate	否定詞	No, not, never
20	quant	概數詞	Total, unique, various
21	number	數字	Second, thousand
22	swear	髒話	Damm, piss, fuck
121	social	社會歷程詞	Gossip, group, grownup
122	family	家族詞	Hubby, husband
123	Friend	朋友詞	Lover, mate, neighbor
124	Humans	人類詞	Participant, partner
125	affect	情感歷程詞	Peace, passion
126	posemo	正向情緒詞	Love, nice, sweet
127	negemo	負向情緒詞	Hurt, ugly, nasty
128	Anx	焦慮詞	Worried, fearful, nervous
129	Anger	生氣詞	Hate, kill, annoyed
130	Sad	悲傷詞	Crying, grief, sad
131	Cogmech	認知歷程詞	Abandom, absolutely, accept
132	Insight	洞察詞	Accept, acknowledge, admit
133	Cause	因果詞	Allow, affect, because

134	discrep	差距詞	Could, expect, hope
135	Tentat	暫定詞	Almost, alot, any
136	Certain	確切詞	Absolute, all, always
137	Inhib	限制詞	Avert, avoid, ban
138	Incl	包含詞	And, both, each
139	Excl	排除詞	Just, not, or
140	percept	感知歷程詞	Acid, aroma, bitter, black
141	See	視覺詞	View, saw, seen
142	Hear	聽覺詞	Listen, hearing
143	Feel	感覺詞	Cold, cool, dry, feel
146	Bio	生理歷程詞	Abortion, ache, alive
147	Body	身體詞	Cheek, hands, spit
148	Health	健康詞	Clinic, flu, pill
149	Sexual	性詞	Horny, love, incest
150	ingest	攝食詞	Dish, eat, pizza
250	Relative	相對詞	Area, bend, exit, stop
251	Motion	移動詞	Arrive, car, go
252	Space	空間詞	Down, in, thin
253	Time	時間詞	End, until, season

354	Work	工作詞	Job, mojobs, xerox
355	Achieve	成就詞	Earn, hero, win
356	Leisure	休閒詞	Cook, chat, movie
357	Home	家庭詞	Family
358	Money	金錢詞	Aduit, cash, owe
359	Relig	宗教詞	Altar, church, mosque
360	Death	死亡詞	Bury, coffin, kill
462	Assent	應和詞	Agree, ah, alright
463	Nonfl	停頓贅詞	Er, hm, sigh, uh
464	filler	填充贅詞	Blah, ohwell, orwhatever



附錄 C 分類指標結果—天真貝氏

表 C 1 天真貝氏於 unigram 之分類指標結果

天真貝氏	unigram			
	準確率	精準率	召回率	調和平均數
同意	0.1179	0.0797	0.2186	0.1168
不同意		0.0343	0.7360	0.0655
討論		0.1688	0.1190	0.1396
無關		0.7369	0.0837	0.1503
混淆矩陣	[[416 1256 157 74] [132 513 24 28] [863 2624 531 446] [3808 10573 2433 1535]]			
FNC 分數	3107.75 / 11651.25 = 26.67%			

表 C 2 天真貝氏於 TF-IDF 之分類指標結果

天真貝氏	TF-IDF			
	準確率	精準率	召回率	調和平均數
同意	0.3286	0.1038	0.3831	0.1633
不同意		0.1235	0.0301	0.0484
討論		0.2053	0.5166	0.2938
無關		0.7577	0.2885	0.4179
混淆矩陣	[[729 22 533 619] [256 21 224 196] [1243 37 2306 878] [4795 90 8170 5294]]			
FNC 分數	4958.25 / 11651.25 = 42.56%			

表 C 3 天真貝氏於 TF-IDF + GI 之分類指標結果

天真貝氏	TF-IDF + GI			
	準確率	精準率	召回率	調和平均數
同意	0.3398	0.1054	0.3720	0.1642
不同意		0.1235	0.0301	0.0484
討論		0.2075	0.5199	0.2967
無關		0.7609	0.3044	0.4348
混淆矩陣	[[708 22 535 638] [239 21 233 204] [1193 37 2321 913] [4580 90 8094 5585]]			
FNC 分數	5011 / 11651.25 = 43.01%			

表 C 4 天真貝氏於 TF-IDF + LIWC 之分類指標結果

天真貝氏	TF-IDF + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.3393	0.1052	0.3720	0.1640
不同意		0.1235	0.0301	0.0484
討論		0.2075	0.5202	0.2967
無關		0.7609	0.3037	0.4341
混淆矩陣	[[708 22 535 638] [240 21 233 203] [1195 37 2322 910] [4589 90 8098 5572]]			
FNC 分數	5009.5 / 11651.25 = 43.00%			

表 C 5 天真貝氏於 TF-IDF + GI + LIWC 之分類指標結果

天真貝氏	TF-IDF + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.3397	0.1053	0.3720	0.1642
不同意		0.1235	0.0301	0.0484
討論		0.2075	0.5199	0.2967
無關		0.7610	0.3043	0.4348
混淆矩陣	[[708 22 535 638] [240 21 233 203] [1193 37 2321 913] [4581 90 8094 5584]]			
FNC 分數	5011 / 11651.25 = 43.01%			

表 C 6 天真貝氏於 TF-IDF + TF-IDF Sim 之分類指標結果

天真貝氏	TF-IDF + TF-IDF Sim			
	準確率	精準率	召回率	調和平均數
同意	0.3287	0.1038	0.3831	0.1634
不同意		0.1235	0.0301	0.0484
討論		0.2055	0.5172	0.2941
無關		0.7583	0.2886	0.4180
混淆矩陣	[[729 22 535 617] [256 21 224 196] [1243 37 2309 875] [4794 90 8170 5295]]			
FNC 分數	4962 / 11651.25 = 42.59%			

表 C 7 天真貝氏於 TF-IDF + LDA Sim 之分類指標結果

天真貝氏	TF-IDF + LDA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.3287	0.1038	0.3831	0.1634
不同意		0.1235	0.0301	0.0484
討論		0.2053	0.5166	0.2938
無關		0.7578	0.2886	0.4180
混淆矩陣	[[729 22 533 619] [256 21 224 196] [1243 37 2306 878] [4794 90 8169 5296]]			
FNC 分數	4958.75 / 11651.25 = 42.56%			

表 C 8 天真貝氏於 TF-IDF + LSA Sim 之分類指標結果

天真貝氏	TF-IDF + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.3287	0.1038	0.3831	0.1633
不同意		0.1235	0.0301	0.0484
討論		0.2054	0.5166	0.2939
無關		0.7578	0.2887	0.4181
混淆矩陣	[[729 22 533 619] [256 21 224 196] [1243 37 2306 878] [4796 90 8166 5297]]			
FNC 分數	4958.75 / 11651.25 = 42.56%			

表 C 9 天真貝氏於 Sim 之分類指標結果

天真貝氏	TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8494	0.3591	0.3631	0.3611
不同意		0.0000	0.0000	0.0000
討論		0.6104	0.7245	0.6626
無關		0.9710	0.9626	0.9667
混淆矩陣	[[691 0 1073 139] [220 0 361 116] [956 1 3234 273] [57 0 630 17662]]			
FNC 分數	8993.25 / 11651.25 = <u>77.19%</u>			

表 C 10 天真貝氏於 TF-IDF + Sim 之分類指標結果

天真貝氏	TF-IDF + TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.3289	0.1038	0.3831	0.1633
不同意		0.1235	0.0301	0.0484
討論		0.2056	0.5172	0.2942
無關		0.7584	0.2888	0.4183
混淆矩陣	[[729 22 535 617] [256 21 224 196] [1243 37 2309 875] [4796 90 8164 5299]]			
FNC 分數	4963 / 11651.25 = 42.60%			

表 C 11 天真貝氏於 TF-IDF + All 之分類指標結果

天真貝氏	TFIDF + TFIDF Sim + LDA Sim + LSA Sim + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.3401	0.1053	0.3720	0.1642
不同意		0.1235	0.0301	0.0484
討論		0.2080	0.5213	0.2973
無關		0.7621	0.3045	0.4351
混淆矩陣	[[708 22 538 635] [240 21 233 203] [1194 37 2327 906] [4581 90 8091 5587]]			
FNC 分數	5018.75 / 11651.25 = 43.07%			

附錄 D 分類指標結果－決策樹

表 D 1 決策樹於 unigram 之分類指標結果

決策樹	unigram			
	準確率	精準率	召回率	調和平均數
同意	0.1496	0.0909	0.6774	0.1604
不同意		0.0633	0.0445	0.0522
討論		0.1751	0.3846	0.2406
無關		0.8123	0.0417	0.0794
混淆矩陣	[[1289 40 523 51] [473 31 164 29] [2550 100 1717 97] [9861 319 7403 766]]			
FNC 分數	4191 / 11651.25 = 35.97%			

表 D 2 決策樹於 TF-IDF 之分類指標結果

決策樹	TF-IDF			
	準確率	精準率	召回率	調和平均數
同意	0.6321	0.1249	0.2275	0.1612
不同意		0.0652	0.0043	0.0081
討論		0.4077	0.2979	0.3443
無關		0.7671	0.7792	0.7731
混淆矩陣	[[433 11 294 1165] [115 3 49 530] [485 4 1330 2645] [2435 28 1589 14297]]			
FNC 分數	5579.75 / 11651.25 = 47.89%			

表 D 3 決策樹於 TF-IDF + GI 之分類指標結果

決策樹	TF-IDF + GI			
	準確率	精準率	召回率	調和平均數
同意	0.6411	0.1429	0.2181	0.1727
不同意		0.1961	0.0143	0.0267
討論		0.3820	0.3197	0.3480
無關		0.7713	0.7870	0.7790
混淆矩陣	[[415 15 276 1197] [122 10 90 475] [424 3 1427 2610] [1943 23 1943 14440]]			
FNC 分數	5694.5 / 11651.25 = 48.87%			

表 D 4 決策樹於 TF-IDF + LIWC 之分類指標結果

決策樹	TF-IDF + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.6312	0.1114	0.1597	0.1312
不同意		0.0866	0.0158	0.0267
討論		0.3266	0.2377	0.2751
無關		0.7595	0.7992	0.7788
混淆矩陣	[[304 12 327 1260] [103 89 11 84 499] [503 16 1061 2884] [1820 88 1777 14664]]			
FNC 分數	5303.25 / 11651.25 = 45.52%			

表 D 5 決策樹於 TF-IDF + GI + LIWC 之分類指標結果

決策樹	TF-IDF + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.6315	0.1443	0.2160	0.1730
不同意		0.0764	0.0172	0.0281
討論		0.3531	0.2773	0.3107
無關		0.7612	0.7841	0.7725
混淆矩陣	[[411 19 264 1209] [97 12 84 504] [414 11 1238 2801] [1926 115 1920 14388]]			
FNC 分數	5480.25 / 11651.25 = 47.04%			

表 D 6 決策樹於 TF-IDF + TF-IDF Sim 之分類指標結果

決策樹	TF-IDF + TF-IDF Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8151	0.3007	0.4235	0.3517
不同意		0.0815	0.0215	0.0341
討論		0.5718	0.5871	0.5794
無關		0.9614	0.9413	0.9512
混淆矩陣	[[806 65 858 174] [275 15 287 120] [1347 96 2621 400] [252 8 818 17271]]			
FNC 分數	8491.75 / 11651.25 = 72.88%			

表 D 7 決策樹於 TF-IDF + LDA Sim 之分類指標結果

決策樹	TF-IDF + LDA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.6888	0.1689	0.2239	0.1925
不同意		0.0579	0.0215	0.0314
討論		0.4468	0.3649	0.4017
無關		0.8129	0.8411	0.8268
混淆矩陣	[[426 56 436 985] [156 15 145 381] [608 41 1629 2186] [1332 147 1436 15434]]			
FNC 分數	6289 / 11651.25 = 53.98%			

表 D 8 決策樹於 TF-IDF + LSA Sim 之分類指標結果

決策樹	TF-IDF + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.6781	0.1637	0.1456	0.1541
不同意		0.0322	0.0344	0.0333
討論		0.4305	0.3741	0.4003
無關		0.7992	0.8318	0.8152
混淆矩陣	[[277 64 378 1184] [68 24 137 468] [486 126 1670 2182] [861 532 1694 15262]]			
FNC 分數	6101.25 / 11651.25 = 52.37%			

表 D 9 決策樹於 Sim 之分類指標結果

決策樹	TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8115	0.2807	0.4020	0.3306
不同意		0.0870	0.0846	0.0858
討論		0.5979	0.4704	0.5266
無關		0.9568	0.9646	0.9607
混淆矩陣	[[765 203 754 181] [233 59 255 150] [1526 370 2100 468] [201 46 403 17699]]			
FNC 分數	8184 / 11651.25 = 70.24%			

表 D 10 決策樹於 TF-IDF + Sim 之分類指標結果

決策樹	TF-IDF + TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8456	0.3815	0.5260	0.4422
不同意		0.1687	0.0588	0.0872
討論		0.6602	0.6223	0.6407
無關		0.9635	0.9629	0.9632
混淆矩陣	[[1001 53 675 174] [305 41 224 127] [1191 127 2778 368] [127 22 531 17669]]			
FNC 分數	8881 / 11651.25 = 76.22%			

表 D 11 決策樹於 TF-IDF +All 之分類指標結果

決策樹	TF-IDF + TF-IDF Sim + LDA Sim + LSA Sim + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.8271	0.3069	0.4046	0.3490
不同意		0.0965	0.0473	0.0635
討論		0.6161	0.5901	0.6028
無關		0.9614	0.9582	0.9598
混淆矩陣	[[770 81 881 171] [237 33 291 136] [1222 210 2634 398] [280 18 469 17582]]			
FNC 分數	8563 / 11651.25 = 73.49%			

附錄 E 分類指標結果－隨機森林

表 E 1 隨機森林於 unigram 之分類指標結果

隨機森林	unigram			
	準確率	精準率	召回率	調和平均數
同意	0.6622	0.1608	0.0384	0.0619
不同意		0.0000	0.0000	0.0000
討論		0.3198	0.3275	0.3236
無關		0.7501	0.8335	0.7896
混淆矩陣	$\begin{bmatrix} 73 & 0 & 245 & 1585 \\ 38 & 0 & 65 & 594 \\ 87 & 0 & 1462 & 2915 \\ 256 & 0 & 2799 & 15294 \end{bmatrix}$			
FNC 分數	5467.25 / 11651.25 = 46.92%			

表 E 2 隨機森林於 TF-IDF 之分類指標結果

隨機森林	TF-IDF			
	準確率	精準率	召回率	調和平均數
同意	0.7187	0.7692	0.0053	0.0104
不同意		0.0000	0.0000	0.0000
討論		0.4310	0.1756	0.2496
無關		0.7409	0.9521	0.8333
混淆矩陣	$\begin{bmatrix} 10 & 0 & 132 & 1761 \\ 0 & 0 & 24 & 673 \\ 3 & 0 & 784 & 3677 \\ 0 & 0 & 879 & 17470 \end{bmatrix}$			
FNC 分數	5201.25 / 11651.25 = 44.64%			

表 E 3 隨機森林於 TF-IDF + GI 之分類指標結果

隨機森林	TF-IDF + GI			
	準確率	精準率	召回率	調和平均數
同意	0.7147	0.8000	0.0063	0.0125
不同意		0.0000	0.0000	0.0000
討論		0.4049	0.1606	0.2300
無關		0.7378	0.9501	0.8306
混淆矩陣	$\begin{bmatrix} 12 & 0 & 125 & 1766 \\ 0 & 0 & 14 & 683 \\ 2 & 0 & 717 & 3745 \\ 1 & 0 & 915 & 17433 \end{bmatrix}$			
FNC 分數	5112.5 / 11651.25 = 43.88%			

表 E 4 隨機森林於 TF-IDF + LIWC 之分類指標結果

隨機森林	TF-IDF + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.7202	1.0000	0.0016	0.0031
不同意		0.0000	0.0000	0.0000
討論		0.4314	0.1436	0.2155
無關		0.7381	0.9623	0.8354
混淆矩陣	$\begin{bmatrix} 3 & 0 & 127 & 1773 \\ 0 & 0 & 27 & 670 \\ 0 & 0 & 641 & 3823 \\ 0 & 0 & 691 & 17658 \end{bmatrix}$			
FNC 分數	5097 / 11651.25 = 43.75%			

表 E 5 隨機森林於 TF-IDF + GI + LIWC 之分類指標結果

隨機森林	TF-IDF + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.7236	1.0000	0.0011	0.0021
不同意		0.0000	0.0000	0.0000
討論		0.4630	0.1360	0.2102
無關		0.7378	0.9690	0.8377
混淆矩陣	$\begin{bmatrix} 2 & 0 & 108 & 1793 \\ 0 & 0 & 27 & 670 \\ 0 & 0 & 607 & 3857 \\ 0 & 0 & 569 & 17780 \end{bmatrix}$			
FNC 分數	4631 / 11651.25 = 39.75%			

表 E 6 隨機森林於 TF-IDF + TF-IDF Sim 之分類指標結果

隨機森林	TF-IDF + TF-IDF Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8699	0.5079	0.4404	0.4717
不同意		0.0000	0.0000	0.0000
討論		0.6632	0.7787	0.7163
無關		0.9606	0.9696	0.9651
混淆矩陣	[[838 0 870 195] [198 0 355 144] [597 0 3476 391] [17 0 540 17792]]			
FNC 分數	9267 / 11651.25 = 79.54%			

表 E 7 隨機森林於 TF-IDF + LDA Sim 之分類指標結果

隨機森林	TF-IDF + LDA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.7556	0.3167	0.0300	0.0547
不同意		0.0000	0.0000	0.0000
討論		0.5651	0.4093	0.4747
無關		0.7872	0.9439	0.8585
混淆矩陣	[[57 0 335 1511] [179 0 83 597] [64 0 1827 2573] [42 0 988 17319]]			
FNC 分數	6338.5 / 11651.25 = 54.40%			

表 E 8 隨機森林於 TF-IDF + LSA Sim 之分類指標結果

隨機森林	TF-IDF + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.7442	0.5250	0.0110	0.0216
不同意		0.0000	0.0000	0.0000
討論		0.5490	0.3486	0.4264
無關		0.7691	0.9447	0.8479
混淆矩陣	[[21 0 230 1652] [1 0 45 651] [7 0 1556 2901] [11 0 1003 17335]]			
FNC 分數	5981.5 / 11651.25 = 51.34%			

表 E 9 隨機森林於 Sim 之分類指標結果

隨機森林	TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8413	0.3146	0.4777	0.3794
不同意		0.0957	0.0129	0.0228
討論		0.6548	0.5421	0.5931
無關		0.9631	0.9833	0.9731
混淆矩陣	[[909 36 792 166] [319 9 232 137] [1613 42 2420 389] [48 7 252 18042]]			
FNC 分數	8607 / 11651.25 = 73.87%			

表 E 10 隨機森林於 TF-IDF + Sim 之分類指標結果

隨機森林	TF-IDF + TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8638	0.5134	0.4230	0.4638
不同意		0.0000	0.0000	0.0000
討論		0.7131	0.7043	0.7087
無關		0.9262	0.9811	0.9529
混淆矩陣	[[805 0 672 426] [165 0 262 270] [582 0 3144 738] [16 0 331 18002]]			
FNC 分數	8869.75 / 11651.25 = 76.13%			

表 E 11 隨機森林於 TF-IDF + All 之分類指標結果

隨機森林	TF-IDF + TF-IDF Sim + LDA Sim + LSA Sim + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.8553	0.5282	0.3442	0.4168
不同意		0.0000	0.0000	0.0000
討論		0.7091	0.6815	0.6950
無關		0.9072	0.9831	0.9436
混淆矩陣	[[655 0 697 551] [113 0 256 328] [456 0 3042 966] [16 0 295 18038]]			
FNC 分數	8587 / 11651.25 = 73.70%			

附錄 F 分類指標結果－梯度提升

表 F 1 梯度提升於 unigram 之分類指標結果

梯度提升	unigram			
	準確率	精準率	召回率	調和平均數
同意	0.1736	0.1180	0.5460	0.1941
不同意		0.2353	0.0057	0.0112
討論		0.2029	0.7543	0.3198
無關		0.3333	0.0001	0.0001
混淆矩陣	$\begin{bmatrix} 1039 & 4 & 859 & 1 \\ 313 & 4 & 380 & 0 \\ 1090 & 6 & 3367 & 1 \\ 6360 & 3 & 11985 & 1 \end{bmatrix}$			
FNC 分數	5073.25 / 11651.25 = 43.54%			

表 F 2 梯度提升於 TF-IDF 之分類指標結果

梯度提升	TF-IDF			
	準確率	精準率	召回率	調和平均數
同意	0.7146	0.1658	0.0173	0.0314
不同意		0.0000	0.0000	0.0000
討論		0.4096	0.0766	0.1291
無關		0.7297	0.9693	0.8326
混淆矩陣	$\begin{bmatrix} 33 & 0 & 78 & 1792 \\ 0 & 0 & 11 & 686 \\ 11 & 0 & 342 & 4111 \\ 155 & 4 & 404 & 17786 \end{bmatrix}$			
FNC 分數	4846.5 / 11651.25 = 41.60%			

表 F 3 梯度提升於 TF-IDF + GI 之分類指標結果

梯度提升	TF-IDF + GI			
	準確率	精準率	召回率	調和平均數
同意	0.7196	0.8571	0.0032	0.0063
不同意		0.2857	0.0029	0.0057
討論		0.4161	0.0867	0.1435
無關		0.7312	0.9750	0.8357
混淆矩陣	[[6 5 74 1818] [0 2 12 683] [0 0 387 4077] [1 0 457 17891]]			
FNC 分數	4890.5 / 11651.25 = 41.97%			

表 F 4 梯度提升於 TF-IDF + LIWC 之分類指標結果

梯度提升	TF-IDF + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.7209	0.5938	0.0100	0.0196
不同意		0.0000	0.0000	0.0000
討論		0.4360	0.0916	0.1514
無關		0.7321	0.9751	0.8363
混淆矩陣	[[19 0 70 1814] [1 0 9 687] [7 0 409 4048] [5 1 450 17893]]			
FNC 分數	4923 / 11651.25 = 42.25%			

表 F 5 梯度提升於 TF-IDF + GI + LIWC 之分類指標結果

梯度提升	TF-IDF + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.7210	0.3733	0.0147	0.0283
不同意		0.0571	0.0029	0.0055
討論		0.4486	0.0840	0.1415
無關		0.7324	0.9766	0.8370
混淆矩陣	[[28 7 77 1791] [1 2 11 683] [8 7 375 4074] [38 19 373 17919]]			
FNC 分數	4912.5 / 11651.25 = 42.16%			

表 F 6 梯度提升於 TF-IDF + TF-IDF Sim 之分類指標結果

梯度提升	TF-IDF + TF-IDF Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8819	0.5318	0.2680	0.3564
不同意		0.5000	0.0201	0.0386
討論		0.6686	0.8468	0.7472
無關		0.9646	0.9868	0.9756
混淆矩陣	[[510 5 1227 161] [122 14 421 140] [311 9 3780 364] [16 0 226 18107]]			
FNC 分數	9354.5 / 11651.25 = 80.29%			

表 F 7 梯度提升於 TF-IDF + LDA Sim 之分類指標結果

梯度提升	TF-IDF + LDA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.7511	0.2847	0.0841	0.1298
不同意		0.0000	0.0000	0.0000
討論		0.5276	0.3815	0.4428
無關		0.7966	0.9387	0.8619
混淆矩陣	[[160 0 484 1259] [50 0 113 534] [156 0 1703 2605] [196 0 928 17225]]			
FNC 分數	6370 / 11651.25 = 54.67%			

表 F 8 梯度提升於 TF-IDF + LSA Sim 之分類指標結果

梯度提升	TF-IDF + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.7576	0.3221	0.0552	0.0942
不同意		0.0000	0.0000	0.0000
討論		0.5901	0.3799	0.4623
無關		0.7858	0.9512	0.8606
混淆矩陣	[[105 0 322 1476] [14 0 99 584] [70 0 1696 2698] [137 2 757 17453]]			
FNC 分數	6290.5 / 11651.25 = 53.99%			

表 F 9 梯度提升於 Sim 之分類指標結果

梯度提升	TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8415	0.3261	0.5292	0.4035
不同意		0.0417	0.0014	0.0028
討論		0.6498	0.5311	0.5845
無關		0.9654	0.9814	0.9733
混淆矩陣	[[1007 6 728 162] [326 1 236 134] [1733 11 2371 349] [22 6 314 18007]]			
FNC 分數	8640.75 / 11651.25 = 74.16%			

表 F 10 梯度提升於 TF-IDF + Sim 之分類指標結果

梯度提升	TF-IDF + TF-IDF Sim + LDA Sim + LSA Sim			
	準確率	精準率	召回率	調和平均數
同意	0.8676	0.4131	0.5749	0.4808
不同意		0.3000	0.0258	0.0476
討論		0.7236	0.6438	0.6814
無關		0.9642	0.9844	0.9742
混淆矩陣	[[1094 11 634 164] [330 18 216 133] [1196 21 2874 373] [28 10 248 18063]]			
FNC 分數	9103.75 / 11651.25 = <u>78.14%</u>			

表 F 11 梯度提升於 TF-IDF + Sim 之分類指標結果

梯度提升	TF-IDF + TF-IDF Sim + LDA Sim + LSA Sim + GI + LIWC			
	準確率	精準率	召回率	調和平均數
同意	0.8697	0.4221	0.5712	0.4855
不同意		0.2692	0.0301	0.0542
討論		0.7229	0.6546	0.6870
無關		0.9654	0.9848	0.9750
混淆矩陣	[[1087 17 645 154] [315 21 229 132] [1142 39 2922 361] [31 1 246 18071]]			
FNC 分數	9144.5 / 11651.25 = <u>78.49%</u>			

附錄 G 所有特徵組合之 FNC 分數展示

表 G 1 所有特徵組合之 FNC 分數展示

特徵 \ 分類器	天真貝氏	決策樹	隨機森林	梯度提升
unigram	26.67%	35.97%	46.92%	43.54%
TFIDF	42.56%	47.89%	44.64%	41.60%
TFIDF + GI	43.01%	48.87%	43.88%	41.97%
TFIDF + LIWC	43.00%	45.52%	43.75%	42.25%
TFIDF + GI + LIWC	43.01%	47.04%	39.75%	42.16%
TFIDF + TFIDF Sim	42.59%	72.88%	<u>79.54%</u>	<u>80.29%</u>
TFIDF + LDA Sim	42.56%	53.98%	54.40%	54.67%
TFIDF + LSA Sim	42.56%	52.37%	51.38%	53.99%
TFIDF + Sim	42.60%	<u>76.22%</u>	<u>76.13%</u>	<u>78.14%</u>
Sim	<u>77.19%</u>	70.24%	73.87%	74.16%
TFIDF + All	43.07%	73.49%	73.70%	<u>78.49%</u>