

國立臺灣大學管理學院商學研究所



碩士論文

Graduate Institute of Business Administration

College of Management

National Taiwan University

Master Thesis

群眾募資專案之動態預測：

台灣群眾募資平台 flyingV 之實證研究

The Dynamic Prediction of Crowdfunding Projects:

An Empirical Study of a Taiwanese Crowdfunding Platform, flyingV

王柏偉

Pai-Wei Wang

指導教授：黃俊堯 博士

Dr. Chun-Yao Huang

中華民國 106 年 6 月

June, 2017

誌謝



當開始提筆寫誌謝時，這不僅是論文完稿最後一次的編輯，也是為商研所兩年生活寫下尾聲；作為工作後再回頭考上研究所當學生的短暫插曲，商研所兩年體會了能單純學習是很難得的時刻，也體會了跟一群同樣有熱情、有拚勁的同學一起成長有多麼快樂；然而當我們都成長時，同時也要各奔東西了。

這篇論文的誕生需要感謝黃俊堯教授的指導，在一開始決定論文題目時給予了很大的自由空間，同時也在數次的交流討論下，逐步給予文獻探討與實務運用等建議，使該論文得以在有一定組織架構下完成；同時也十分謝謝口試委員陳彥君教授與陳華寧博士提供的修改建議，讓這篇論文得以完整。

技術上感謝立人與智賢的經常在數據分析上的交流，這篇論文有許多想法及方法的採用皆來自於兩年來的共同耕耘；另一方面也感謝大維的解惑，解決了在研究流程設計上的矛盾想法；而在爬蟲技術上，則是感謝 SAS 實習同事 Maggie 在 XML 語法與網頁格式等議題上的教導與分享。

情感上感謝商研同學們，你們讓我眼界大開，同時又樂於彼此交流與成長，真的是很快樂的兩年日子；同時也謝謝女友慧芳，尊重我的選擇與忙碌，也願意陪窮學生如我這樣節儉度日，接下來一起迎向新的生活；最後感謝我的父母與家人，接受我在職涯上的任性與選擇，不僅跨領域以及辭職回來念書，都是我在闖時貼心的依賴，期待接下來能好好報答你們。

最後，論文作為碩士學生的收尾，期許自己未來會有越來越精彩的表現！

中文摘要



隨著群眾募資的興起，越來越多學者對預測募資結果的研究感到興趣，然而多數的預測文獻為靜態預測，對於募資期間變化進行分析的動態預測卻十分稀少；另一方面過往動態預測文獻所使用的演算法往往需要大量歷程樣本訓練、且其多缺乏易於理解的預測規則，對於沒有相關團隊及資料的提案人而言，過往文獻缺乏立即實務運用的價值。故本研究蒐集 flyingV 募資專案的歷程資料，並以決策樹作為預測演算法，運用其易懂好用的預測規則，提供給台灣募資提案人可直接運用的預測判別準則。為了證明決策樹與過往文獻建議的演算法有相當的預測能力，以及盡可能提供簡單好用的預測規則，本研究設計了三個模型：

- (1) 決策樹模型與 KNN 模型的預測準確率比較
- (2) 以募資第 N 天做為動態預測時間軸的決策樹
- (3) 接續模型二，並僅以募資達成率作為規則用的決策樹

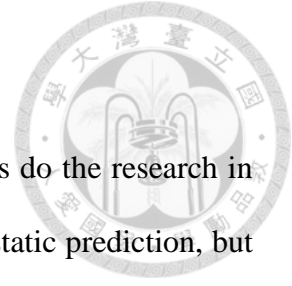
藉此依序證明決策樹有與 KNN 演算法相當的預測能力、使用募資第 N 天作為動態預測時間軸有更好的預測能力，以及僅用募資達成率做為預測規則已有足夠的預測能力。模型實驗結果如下：

- (1) KNN 模型在預測準確率上大致略高於決策樹模型，本研究透過兩相依母體期望值差 T 檢定證明兩者預測準確率上沒有顯著差異。
- (2) 相對過往文獻以募資天數進度百分比，以募資第 N 天作為動態預測時間軸的設計排除單位時間上不同的誤差，而有較佳的預測準確率。
- (3) 僅以募資達成率作為規則的決策樹，演算出只要判斷第一天募資達成率是否超越 4%，就有近 80% 的預測準確率。

本研究證實了可以僅用募資達成率作規則的決策樹能有 80% 以上的預測準確率，並進一步整理出募資前 14 天準確率的變化及對應的募資達成率規則閾值，除了給予提案人簡易現成的預測規則、可作為募資後每日的關鍵績效指標衡量判斷外，本研究也發現前 14 天的募資達成率閾值皆不到 20%，卻仍有相當高的預測準確率，推測多數成功募資案為慢熱型，前期應為口碑推廣的醞釀期。

關鍵字：動態預測、決策樹、群眾募資、flyingV

ABSTRACT

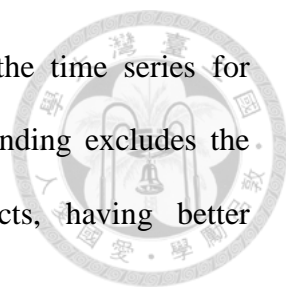


With the rising of the crowdfunding, more and more professors do the research in the prediction of crowdfunding. However, most of their studies are static prediction, but not the dynamic prediction with the changes during the funding period. On the other hand, the algorithm the past dynamic prediction researches use needs large training sample and its predictive rules are not easy-to-use, providing less practical values for funders who may not have relating experts as well as a bunch of samples. This study uses the decision tree, reputed for its easy-understanding predictive rules, as algorithm with the daily data from flyingV to offer the predictive criteria for funders in Taiwan. To prove that the ability of the decision tree is comparable with the algorithm the past researches recommended and to provide simple and useful predictive rules, this study designed the three models:

- (1) The comparison for accuracy between decision tree model and KNN model
- (2) The decision tree model using daily funding as the time series for dynamic prediction
- (3) Same as model 2, but only using the percentage of fund-raising as input variable

Through these models, we can prove sequentially that decision tree is comparable with KNN algorithm, using daily funding as the time series has better predictive performance, and that only using the percentage of fund-raising as predictive rules is enough to predict results. The results are following:

- (1) Though the performance of KNN is slightly higher than decision tree's, we proved there is not significant difference between them with paired samples t-tests.

- 
- (2) Compared using the percentage of funding duration as the time series for dynamic prediction in the past studies, using the daily funding excludes the error within different durations among funding projects, having better performance.
- (3) The decision tree only using the percentage of fund-raising as rules shows that only we check if the percentage of fund-raising at the 1st day is larger than 4%, we have nearly 80% for accuracy.

This study shows that the decision tree only using the percentage of fund-raising as rules can reach above 80% for accuracy, and provides the accuracy chart and paired threshold rules for the percentage of fund-raising for the beginning 14 days. This study provides simple predictive rules as daily KPI for funders, and found the thresholds for the beginning 14 days are under 20% but with relatively high accuracy, implying that most successful crowdfunding projects are slow-to-warm-up, and the early period should be the brewing period for word-of-mouth.

Keywords: dynamic prediction, decision tree, crowdfunding, flyingV

目錄



口試委員會審定書	#
誌謝	i
中文摘要	ii
ABSTRACT	iii
目錄	v
圖目錄	vii
表目錄	viii
第一章 緒論.....	1
第一節 研究背景.....	1
第二節 研究動機.....	2
第三節 研究目的.....	3
第四節 研究流程與章節架構.....	4
第二章 文獻探討.....	6
第一節 群眾募資預測的靜態影響因素.....	6
第二節 群眾募資動態預測.....	8
第三章 研究方法.....	12
第一節 分析流程.....	12
第二節 研究對象.....	16
第三節 研究樣本.....	16
第四節 研究變數定義.....	17
一、 應變數：最終募資狀態.....	17



二、	自變數.....	18
三、	控制變數.....	22
第五節	統計分析方法.....	22
一、	留一交叉驗證法(leave-one-out cross validation, LOOCV)	22
二、	決策樹演算法(decision tree algorithm)	23
第四章	資料分析與結果	26
第一節	敘述性統計.....	27
第二節	募資天數百分比的決策樹分析.....	29
第三節	募資第 N 天的決策樹分析.....	31
一、	模型二 使用所有自變數做訓練.....	32
二、	模型三 僅使用募資金額達成率變數做訓練.....	35
第五章	結論與建議	37
第一節	研究發現與實務意涵.....	37
第二節	研究限制與未來研究建議.....	38
	參考文獻.....	40

圖目錄



圖 1	研究流程圖	5
圖 2	Etter et al.(2013)使用 KNN 演算法之準確率變化	9
圖 3	募資天數百分比模型的分析流程	15
圖 4	募資第 N 天模型的分析流程	15
圖 5	LOOCV 法操作示意圖	23
圖 6	模型一 動態分析的平均預測準確率(募資天數百分比).....	29
圖 7	模型二 動態分析的平均預測準確率(募資第 N 天)，採所有自變數 ..	32
圖 8	模型二 動態分析的最佳決策樹長度	33
圖 9	模型三 動態分析的平均預測準確率(僅採募資金額達成率做自變數)， 並與模型二進行比較。	35
圖 10	模型三 募資金額達成率之閾值與平均預測準確率的動態分析	36
圖 11	模型三動態分析之預測募資成功案件數占比	36

表目錄



表格 1	二元分類預測結果的混淆矩陣	26
表格 2	成功與失敗募資專案的敘述性統計表	28
表格 3	模型一之兩相依母體期望值差 T 檢定表	30
表格 4	模型一之決策樹預測募資成功規則	31
表格 5	模型二之決策樹預測募資成功規則	34

第一章 緒論



第一節 研究背景

在過往，創業或進行產品開發往往需要大量的資金挹注，傳統上若缺乏投資人入股或融資活動來滿足資金需求，擁有再好的人才或技術也會面臨巧婦難為無米之炊。所幸 2006 年起，由 Michael Sullivan 在網路上創辦 fundavlog 後，群眾募資 (crowdfunding) 逐漸成為公司業者融資的新選擇。群眾募資大多由三個角色所組成：提案人、贊助人以及募資平台，由提案人在募資平台發表募資專案，向潛在贊助人提出專案目標、募資金額、募資方式及執行規劃等，若潛在贊助人認同提案人的專案內容，則可透過募資平台進行贊助。

其中 Mollick (2014) 將群眾募資平台依據募資方式的不同，分為四種類別：借貸型 (lending model)、捐贈型 (patronage model)、回饋型 (reward-based crowdfunding)、股權型 (equity crowdfunding)。借貸型的贊助人享有債權，提案人需在協議的時間內歸還本金及利息；捐贈型的贊助人為純資金捐贈，不需給予任何形式上的回饋；相反的回饋型則是由提案人於募資專案中提供數個不同的贊助方案，一般會因贊助方案的金額不同、而提供不同價值的商品或服務給贊助人作為回饋；最後則是股權型，贊助人可以透過贊助而獲得該提案業者的股權。而全球最大的數個群眾募資平台如 Kickstarter 與 Indiegogo 以及國內較大的募資平台如 flyingV 與嘖嘖等皆屬於回饋型平台。

2015 年群眾募資報告指出，全球的群眾募資總金額已達到 344 億美金的水準，其中亞洲地區的總贊助金額就佔了 105.4 億美金，僅次於美洲地區的 172.5 億美金，同時也是年成長率最高的地區¹。而國內群眾募資熱潮也不遑多讓，2013 年起《太白粉路跑》及《看見台灣》的活動類型專案讓群眾募資平台開始嶄露頭角；2014 年適逢多起社會運動，《白色力量》、《反核遊行》、《反服貿紐約時報廣告》及《太陽花學運》等募資專案搭上議題，讓台灣民眾幾乎都知道群眾募資平台的存在；若以台灣最主要的募資平台 flyingV 為例，截至 2015 年的短短三年多時光就

¹Massolution, 2015, <2015CF Crowdfunding Industry Report>

已經累積了 684 件²成功募資案件。而隨著群眾募資越發普遍，其募資案是否能在結案前就提前預測其成功與否之探討，就越來越是提案人及平台業者關注的議題焦點。




第二節 研究動機

絕大多數回饋型的群眾募資專案是採取” all-or-nothing” 模式，其模式代表若最終累積募資金額並未達到或超過當初提案人的募資目標時，提案人便無法獲得所有已累積的募資金額，其募資金額將全數退回所有的贊助人帳戶中；另一方面，群眾募資專案的募資成效可作為市場回應熱度的指標，並且觀察各贊助方案的贊助熱烈程度來進一步了解市場的偏好，以做為後續專案執行的修正參考依據。然而多數的群眾募資專案天數為 30-60 天，若提案人能夠在專案進行前或進行期間的前期就能判定募資是否成功，提案人就能及早調整籌資計畫或資源配置，如更改贊助方案數或減少專案行銷預算等，因此預測成功與否將會是提案人及募資平台相當在乎的議題。

儘管群眾募資仍是相當新穎的一種募資方式，學者依舊提出了不少可能影響專案成功率的關鍵因素，如不少學者都指出專案類型、目標金額大小、募資天數等專案本身的設定都與募資結果有關；而群眾募資的本質就是群眾，因此學者也探討提案人本身的社會資本(social capital)狀態，指出提案人的社群網站(如 facebook、twitter 等)粉絲及朋友人數等也會影響募資成功率(Giudici et al., 2014; Mollick, 2014; Zheng et al., 2014)；而募資平台分享到社群網站的次數及早期支持行為也可作為社會資本的影響依據(Colombo et al., 2015; Liao et al., 2015)，並與成功募資案有顯著相關；甚至是專案的介紹影像或影片(Greenberg et al., 2013; Mollick, 2014)以及文字描述(Zhou et al, 2015)也都指出與募資結果有關。

然而絕大多數的預測研究都是以靜態資料來進行因素的影響力探討，若欲加入時間構面、討論募資期間的動態資訊如何影響募資結果，該類動態預測(dynamic

² flyingV 官網 - 2015 年年度回顧。 <https://report.flyingv.cc/annual/2015>



prediction)的相關研究往往因資料蒐集費時且不易而相對稀少。多數的學者主要探討募資開始後3-7天的早期募資情形做預測因素之一使用(Chen et al., 2015; Li et al., 2016); 或者將募資期間分為前中後三階段, 探討各自不同的專案更新類型如何影響成功率(Xu et al., 2014); 惟少數學者(Etter et al., 2013; Rao et al., 2014; Zhang et al., 2015)分別對 Kickstarter 及中國追夢籌平台的募資專案做完整動態預測, 並專注在演算法模型的修正與比較; 雖有不錯的預測準確率及演算法建議提供給平台團隊運用, 但對一般提案人而言, 其往往缺乏相關知識及足夠的訓練資料來預測專案成功率, 而使該研究顯得缺乏實務價值; 另一方面針對台灣群眾募資平台的動態預測研究則幾近稀少, 因此本研究將嘗試對台灣的群眾募資平台, 使用規則易懂好用的決策樹演算法進行動態預測研究分析, 並以站在協助提案人預測的立場, 提供實用且簡單的預測規則。

第三節 研究目的

根據上述的研究背景及研究動機, 本研究目的為針對台灣主要的群眾募資平台進行動態預測分析, 並站在提案人的立場, 使用決策樹模型以建立簡單易懂的預測規則; 而為了達成該目的, 本研究首先探討決策樹與過往文獻建議的演算法, 比較兩者的動態預測準確率; 接著將動態預測的時間軸, 改以募資第 1 天、第 2 天……等不同天(下文起以「募資第 N 天」表示)的資訊來進行動態預測, 而非過往文獻將募資天數(duration)平均切割成數個時間點(下文起以「募資天數百分比」表示)來逐一進行動態預測, 利於提案人直覺判斷; 最後本研究嘗試以較少的自變數來進行決策樹預測, 讓提案人得以專注少數變數的動態變化即可, 同時本研究也會與一般決策樹進行比較, 觀察預測準確率的差異。一言以蔽之, 本研究主要的三個研究問題如下:

- 一、 決策樹模型是否在預測準確率上與過往文獻的演算法表現更加優異?
- 二、 改以募資第 N 天做為動態預測的時間軸, 是否比以募資天數百分比為時間軸的動態預測來得更加精準?
- 三、 若以較少的關鍵自變數進行決策樹預測, 是否仍擁有優異的預測能力?



第四節 研究流程與章節架構

本論文的研究架構如圖 1 所示，個別對應到本研究的五個章節。第一章緒論，主要說明研究背景、研究動機、研究目的及說明研究流程與章節架構。

第二章為文獻探討，此章會先針對影響募資成效之靜態因素作逐一探討，接著檢索群眾募資專案的動態預測文獻，了解一般演算法模型建置流程及演算法缺點，並以文獻建議的演算法作為本研究決策樹模型的比較基準(benchmark)。

第三章為研究設計與方法。說明分析流程、研究對象、研究樣本、研究變數定義、統計分析方法，本研究將針對台灣最大的群眾募資平台 flyingV 進行連續 5 個月的資料蒐集與整理，來做為此次動態預測分析的主要樣本；並使用決策樹演算法進行動態預測，考量避免模型過度配適(overfitting)及有限樣本，本研究使用留一交叉驗證法(leave-one out cross validation, LOOCV)來進行模型準確率驗證。

第四章為資料分析與討論。此章節除了呈現蒐集的專案中各項變數的敘述性統計外，也會設計三個操作模型，分別進行募資天數百分比與募資第 N 天的動態預測分析，並嘗試找出少量變數即可有效預測，並指出其預測準確率的變化。

第五章為結論與建議。主要會依據第四章的分析結果，歸納出本研究的結論，指出研究發現、實務運用、研究限制以及未來的研究發展方向。

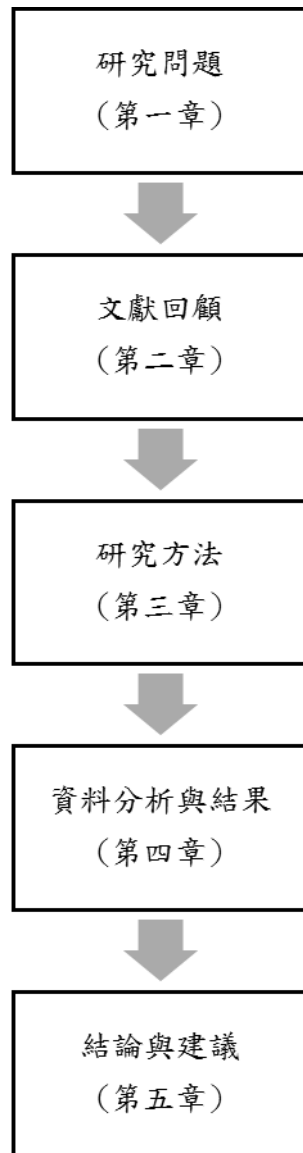


圖 1 研究流程圖

第二章 文獻探討



儘管本研究將著重在群眾募資專案的動態預測分析上，一些非動態的專案訊息如募資天數等仍有可能扮演募資成功與否的重要角色；此外，本研究著重使用的決策樹模型，其相對於常見的迴歸分析較可避免自變數過多而造成的解釋力不顯著狀態，因此越多可能的解釋自變數，對本研究決策樹找出關鍵變數越有幫助。本章節將分為群眾募資的靜態預測文獻回顧以及動態預測文獻回顧，前者將注重在影響群眾募資成功可能性的靜態因素，而後者著重在動態預測文獻上慣用的研究架構及可能相關的動態資訊上。藉由本章文獻探討，提供本研究在資料採集與分析流程設計上的參考依據。

第一節 群眾募資預測的靜態影響因素

事實上，從提案人開始在募資平台上發起群眾募資專案時，一些基礎設定就已經影響了募資成功與否的關鍵。Muller et al.(2013)指出，若將募資成功與募資失敗的專案相比，可以發現募資成功的專案普遍來說有較小的募資金額目標，推測與需求的募資金額較小有關，因此只需要少數贊助人支持就能輕鬆達標；而文獻也發現，募資天數長短與募資成功是呈現負相關(Mollick, 2013)，可能因較長的募資天數會導致提案人態度鬆懈，使得募資成效較差。而 Mollick(2013)其研究指出專案種類如科技、藝術等並沒有對募資成功與否有顯著影響，但其文獻中其他模型仍將專案種類視為控制變數的一種。

對於回饋型的群眾募資專案而言，部分贊助人會願意投資的原因之一就是為了募資成功後的回饋獎勵，因此贊助方案的內容設計也將影響潛在贊助人的贊助意願及其金額。Kuppuswamy & Bayus(2013), de Witt(2012), Mod(2010) & Steinberg(2012)都認為也證實，贊助方案的數量與募資是否成功呈現正相關，Steinberg(2012)解釋越多的贊助方案種類，越能滿足較多不同贊助人的潛在需求，因此能提高募資成效。而 Mod(2010)的研究認為提案人應該避免提供金額過低的贊助方案，他認為原擁有較高預算的潛在贊助者可能會選較低的贊助方案取代之，而非提高更多預算選擇較高金額的贊助方案，林亭佑(2015)的研究也證實贊助方案

金額中位數與募資成功率有正向關係，而衡量贊助方案金額分佈的四分位數則有負向關係，支持了 Mod(2010)的論點。

然而 Mollick(2013)也指出，群眾募資專案屬於一種資訊不透明的募資型態，贊助人僅能透過專案上的敘述來了解提案人的需求及募資原因，因此專案品質(Project quality)就扮演訊號(signal)傳送的重要角色。而能透露出專案品質的關鍵，不外乎影片上的精緻程度、專案文字的描述、專案進度更新的頻率、提案人的提案次數、提案人的社會資本(social capital)等。事實上在 Kickstarter 給予募資者的建議³上第一守則，就是建議提案人做一個影片來描述其專案，其聲稱很少有其他事會比這件事來得更重要，缺少影片將會嚴重影響其募資成果，而 Mollick(2013)及 Kuppuswamy & Bayus(2013)的研究也證實了專案有無影片顯著地與募資成功與否有所關聯。而 Kickstarter 也提及，在專案開始募資後盡速進行進度更新，因為很少專案能夠在初期就募資達標成功，因此迅速進行更新也對潛在贊助人透露出其提案人是已做足準備的訊號，同時 Mollick(2013)也證明了三天內更新進度與募資成功存在顯著關聯性。Xu et al.(2014)更進一步明確指出，在募資初期應著重社群推廣(呼籲贊助者主動分享專案)上的更新，募資中期則著重在進度報告(如研發進度完成 30%)上，募資後期則是新的贊助方案發佈上的更新，都能有效提高募資成功率。

而用以描述專案的文案，多篇文獻(Greenberg et al. 2013, Xu et al.,2014 & Zhou et al., 2015) 也都做了相關的研究，將文字字數、句子數、標題字數都列入預測模型的變數中，而 Xu et al(2014)指出文字字數與募資成功呈現負相關；Greenberg et al.(2013)與 Zhou et al.(2015)更進一步使用文字採礦技術，萃取出情感判斷作為模型訓練的變數，惟可惜未指出其影響程度。

群眾募資相對其他募資方法最大的特色就是社會資本的重要性，因此群眾募資專案與社群上的結合推廣，可以接觸更多潛在的贊助者以及提高贊助人數的機會。不過由於社會資本的具體描述不易，Gudici et al.(2014), Mollick(2014) & Zheng et al.(2014)的研究多以外部社群網站如 Facebook、LinkedIn 等的朋友及粉絲人數作

³ Kickstarter School。 <https://www.kickstarter.com/help/handbook>

為社會資本的衡量，而其研究結果也都證實了社群對募資成功有顯著的正向結果。另一方面專案本身在社群上的話題性也是衡量社會資本的方式之一，Colombo et al.(2015)及 Liao et al.(2015)也曾以募資平台分享到社群的次數作為社會資本的依據，且也證實與募資成功的正向相關性。

另外，Gerber et al.(2012)也指出除了透過社會資本向贊助人發出訊號，提案人也可以透過擁有較多的成功專案經驗，吸引潛在的投資者參與群眾募資，Kickstarter 統計(2015)數據也證實，第二次提案成功機率是 73%，比第一次提案就成功的機率 38%高出近一倍。總的來說，不論是社會資本、提案經驗、影片與文案的描述以及專案的更新頻率，越能協助贊助人消除資訊不對等的困擾，就越能提升其贊助募資專案的可能性，進而提高募資成功的機會。

第二節 群眾募資動態預測

儘管在第一節描述了眾多學者對於群眾募資的靜態預測研究，企圖找出影響群眾募資成功的關鍵因素，以利於在提出募資專案前設計出合適的細節設定，然而關於募資期間的動態預測研究，就相對少了許多；除了前一節提及的專案更新數如何影響成功率(Xu et al., 2014)，並進一步提出募資中不同時期建議的更新類型外，主要以 Etter et al.(2013)為首，建立起較完整的動態預測分析研究。

Etter et al. (2013) 的動態預測研究主要透過募資結束以前的累計募資金額作為資訊，隨著募資進度進展使資訊量逐漸增加，並個別以 KNN 等演算法進行動態預測專案最終的募資結果，呈現其預測準確率。其研究對象為 Kickstarter 的 16042 件專案，從每件專案開始募資起，以平均每 15 分鐘紀錄一次每件專案的累計募資金額及贊助人數，直到該專案募資結束為止。資料處理部分則依照各專案的募資天數(duration)來平均切割成 1000 個時間點，將其對應的動態資訊抽出作為樣本，也就是每個專案都會有 1000 個不同時間點的累計募資金額之動態資訊；接著使用 KNN 演算法進行 1000 次的預測，每一次的訓練資料會逐時增加，舉例而言當進行第 2 次預測時，會使用專案開始募資後的前 2 個時間點資訊做為訓練資料，依此類推；而 KNN 演算法的運作說明如圖 2，演算法會將累計到募資天數百分比 70% 的訓練樣本之各專案曲線繪出，再將待預測樣本的累計募資金額曲線(虛線)放入，

尋找出 K 個最靠近該虛線的訓練樣本，並計算其募資成功比例，進而推估待預測樣本的募資成功機率。而文獻最終得到動態預測的準確率表現如圖 3。



募資天數百分比=70%

募資金額

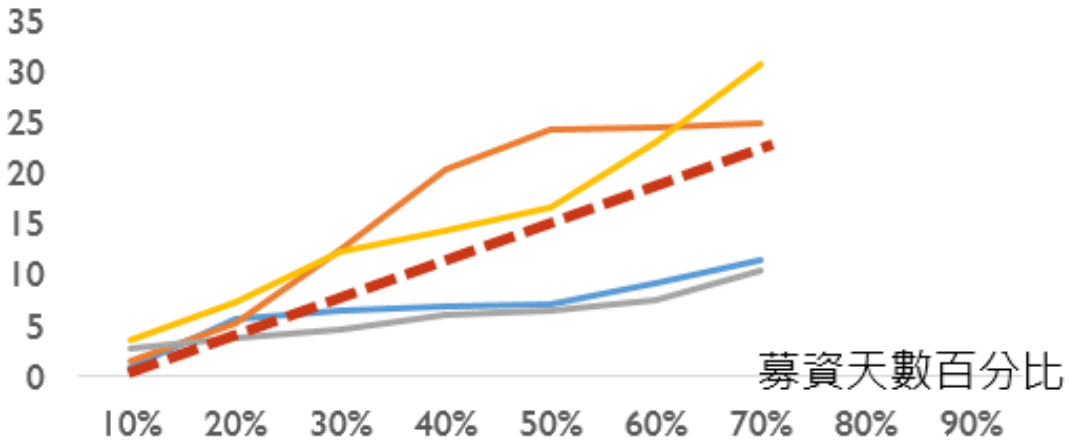


圖 2 KNN 演算法示意圖

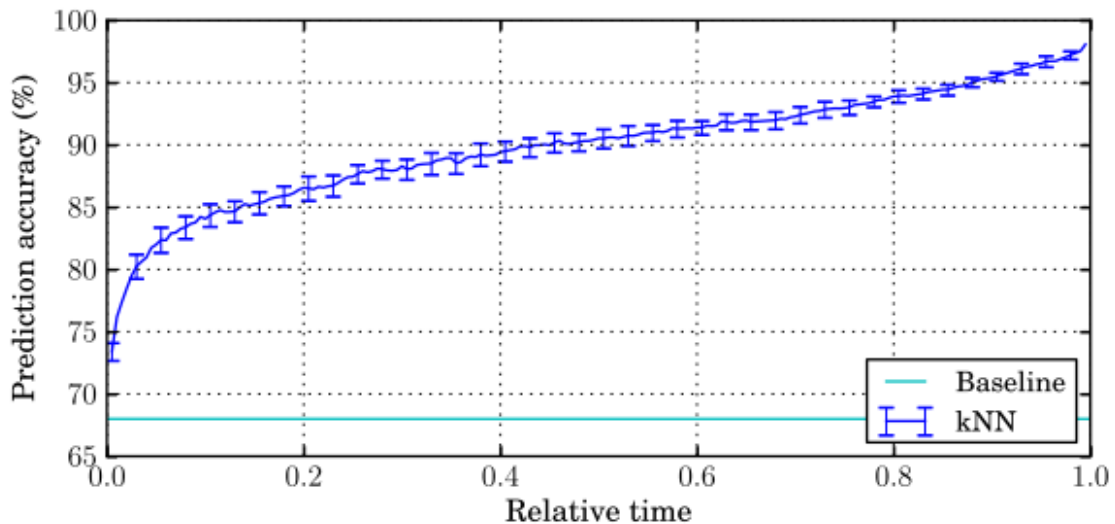


圖 3 Etter et al.(2013)使用 KNN 演算法之準確率變化

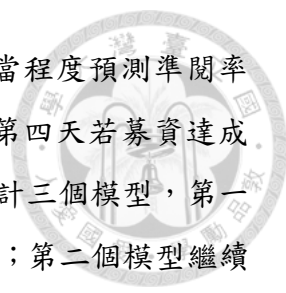
由圖 3 可以發現，其預測準確率平均在過了 15%的募資天數後，即可達到 85%，相對於圖中淺藍色線顯示的 68%，其為 Greenberg et al.(2013)使用靜態資料的預測精準度，顯示動態預測確實有其貢獻。然而 Etter et al.(2013)也提及，KNN 模型本身形同黑盒子，並不能解釋專案為何成功或失敗，且其模型需要大量的動

態資訊來做為訓練資料，對於一般提案人而言不容易取得，因此較難被實務所用。儘管 Zhang et al.(2015)也針對中國追夢籌群眾募資平台的 2092 件專案進行動態預測分析研究，並以反向傳播類神經(backpropagation neural network, BPNN)演算法預測募資結果，也仍是不能解決模型黑盒子與難為實務所用的困境。

Rao et al.(2014)也嘗試研究群眾募資的動態預測分析，不過其採用了決策樹演算法進行預測，該研究提及使用決策樹演算法原因為簡單易懂，也適用於動態分析，並且絕大多數軟體都擁有決策樹演算法，因此其實驗結果也容易重現。詳細的決策樹演算法將於第三章第五節介紹。其研究流程前半段與 Etter et al.(2013)大致相同，紀錄 8529 件 Kickstarter 募資專案的動態過程資訊，取募資天數介於 28 至 31 天者，並取出 20 個時間點的對應動態資訊以進行動態預測；其主要使用募資金額增加量作為自變數，並允許使用歷史訊息，舉例進行第 3 次預測時，其自變數包含了第 1、2、3 個時間點下的募資金額增加量，提供給決策樹進行預測；而預測結果在第一個時間點(5%)即可達到 82%的預測準確率，最高可達 93%準確率。唯一可惜的是，該研究並未進一步與其他演算法進行預測準確率上的比較，以及列出對應時間下的分類規則。

綜合該節動態預測相關文獻所述，雖已有研究以演算法進行動態預測，但多屬黑盒子且不易操作，同時也缺少台灣募資專案的相關動態預測研究；站在台灣群眾募資提案人的立場下，本研究認為可嘗試以 Rao et al.(2014)所使用較簡單易懂的決策樹演算法來進行台灣群眾募資的動態預測，並與 Etter et al.(2013)的 KNN 演算法準確率進行比較；而為了提供提案人較直覺的分類規則，本研究將對過往進行動態預測常見的流程做下述更動：

- 一、 進行每一次動態預測時，僅使用當期的動態資訊，而不使用歷史資訊。(舉例如進行第 2 次動態預測時，只使用第 2 個時間點的變數資料，而不使用第 1 個時間點的變數資料)
- 二、 另外建立以天數為時間點(募資第 N 天)的動態預測，而非僅以平均分成數個時間點(募資天數百分比)的時間軸方式來進行動態預測
- 三、 盡可能減少決策樹產生的分支規則及使用到的自變數，儘管可能減少些微的預測準確率



而本研究進行這三點預測流程調整，主要目的為在維持相當程度預測準確率下，盡可能給予提案人直觀易操作的預測規則；舉例如在募資第四天若募資達成率超過 19%，則有 92%的機率募資成功。因此本研究將對應設計三個模型，第一個模型確認決策樹相對 KNN 演算法有相當或更優秀的預測能力；第二個模型繼續使用決策樹模型，但使用募資第 N 天作為動態預測的時間軸，觀察其預測能力變化；第三個模型修改第二個模型，但使用較少的自變數來進行預測，同時比較預測率的差異程度。

第三章 研究方法



第一節 分析流程

由於動態預測的分析流程略為複雜，本節將逐一講解如何建構分析流程。一般動態預測基本的分析流程依序為：資料蒐集、資料清洗與整理、依時間點不同(即動態)將資料分組、依組別各自進行演算法預測，經過此流程，即可得到最基本的動態預測結果，計算出各時間點的預測準確率。如圖 4 所示。

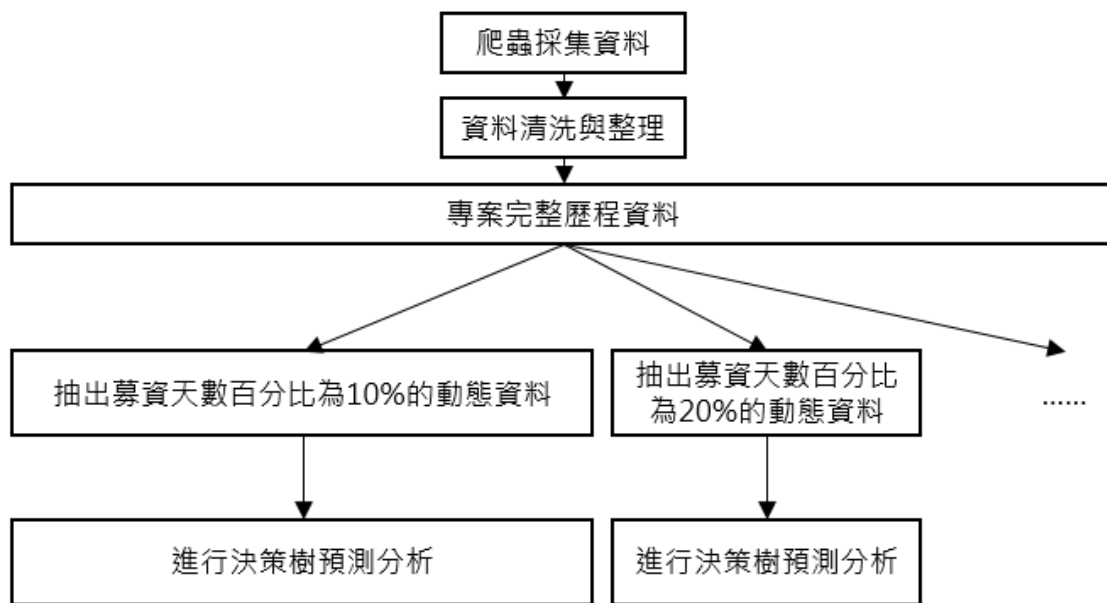


圖 4 動態預測分析流程說明步驟一

另外在機器學習中，為了避免高估預測準確率，通常會將資料分為訓練組與驗證組，訓練組用以訓練演算法模型參數，再以訓練好的模型來預測驗證組的資料，藉此衡量預測準確率，此方法即是交叉驗證法。然而這樣的預測準確率仍存在分組誤差，因此學者也發展出將樣本分割成數組，取其中一組為驗證組，其他組為訓練組來訓練模型，以計算預測準確率，接著每一組輪流做驗證組各別計算預測準確率，再將所有預測準確率進行平均即可，此即是著名的 K 次交叉驗證法 (K-folds cross validation)。本研究因研究樣本較少，故採留一交叉驗證法 (Leave-one-out cross validation)，詳細介紹將於第五節講解之。因此加入交叉驗證法的分析流程將如圖 5 所示。

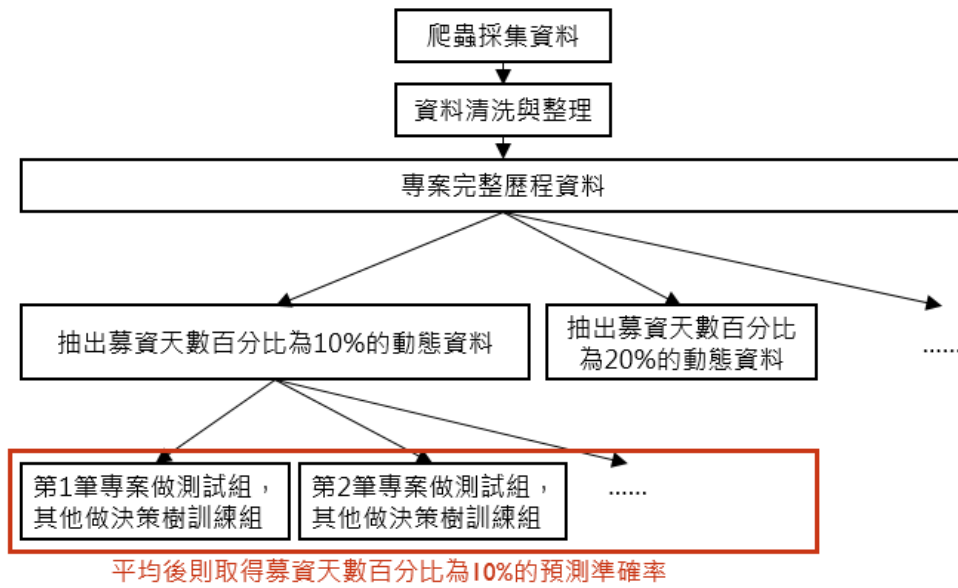


圖 5 動態預測分析流程說明步驟二

而決策樹演算法可以用修剪長度來進行模型優化，然而依預測情境不同等狀況適合的決策樹長度可能不同，導致在預測準確率上也会有所差異，因此在機器學習中也會進行不同決策樹長度設定來找出最佳預測準確率組合。故綜上所述，進階的動態預測分析流程將依序為：資料蒐集、資料清洗與整理、依時間點不同(即動態)將資料分組、依組別各自進行演算法預測(同時進行交叉驗證與參數優化)，而後者多以撰寫程式多重迴圈方式來取得最佳平均預測準確率。如圖 6 所示。

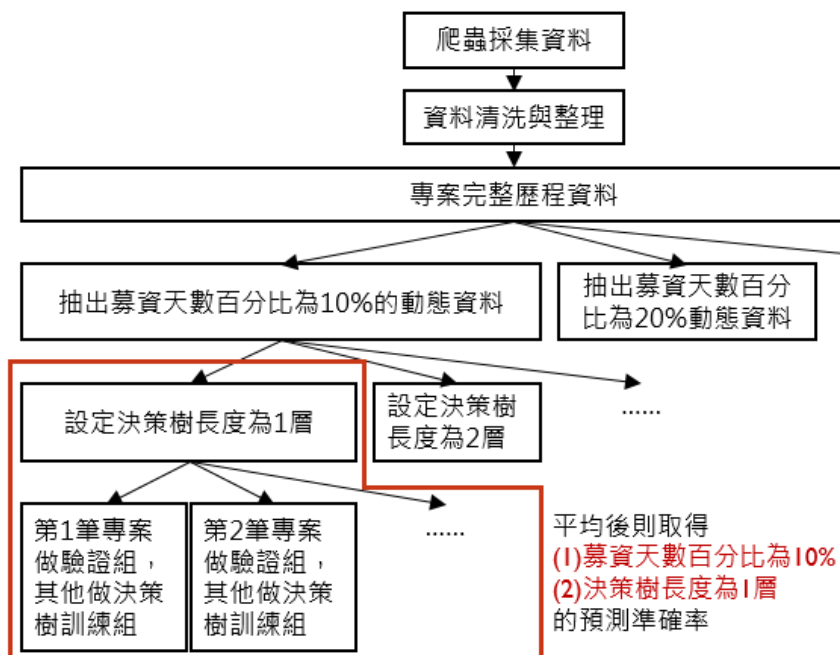


圖 6 動態預測分析流程說明步驟三

事實上透過圖 6 流程設計後，本研究將會挑出驗證組中平均預測準確率最高且其決策樹最短者，作為動態預測的預測準確率。如圖 7 所示，本研究將挑出平均預測準確率最高的 86% 作為募資天數百分比為 10% 的動態預測表現，而最佳的決策樹長度即平均預測準確率下最短的決策樹長度，挑選長度 4 作為代表。

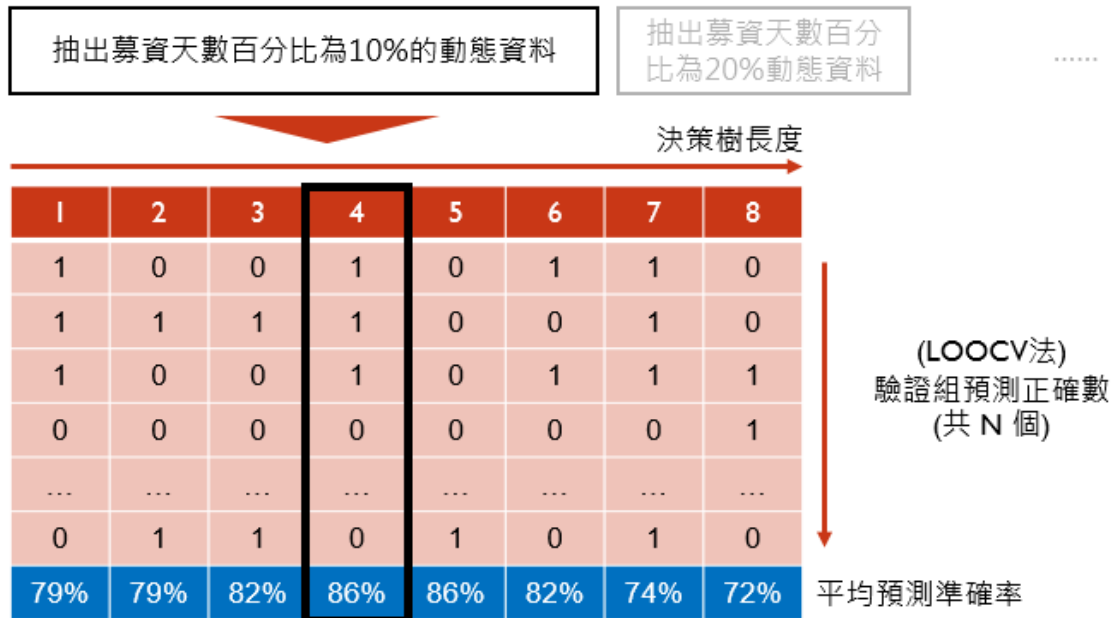


圖 7 決策樹長度優化及預測準確率挑選示意

最後若依照動態預測時間軸上的不同，本研究的分析流程可分為兩種類型，如圖 8 及圖 9 圖 9 所示，待爬蟲採集完資料並整理成專案完整歷程資料後將會進行三個迴圈的運算，第一個迴圈依模型不同而拆分成依募資天數百分比及募資第 N 天分別進行決策樹模型優化及預測率驗證；第二個迴圈為修剪決策樹長度優化之，優化方式為重複進行決策樹長度 1 至 8 層的機器學習訓練，找出平均預測準確率最高的長度作為其優化設定；第三個迴圈則是交叉驗證法，以輪流將樣本做驗證組來確認其模型預測準確率的平均值。

因此在 1 號最內層的虛線範圍裡，可得知給定募資天數百分比/募資第 N 天及給定決策樹長度下的平均預測準確率；在 2 號中間層的虛線範圍裡，則可得知給定募資天數百分比/募資第 N 天的狀況下，最合適的決策樹長度、平均預測準確率及關鍵變數與閾值為何；而在 3 號最外層的虛線範圍內，則可進一步描繪出隨著募資天數百分比/募資第 N 天的增長下，平均預測準確率將會如何變化，並進一步與過往動態預測分析文獻做比較。

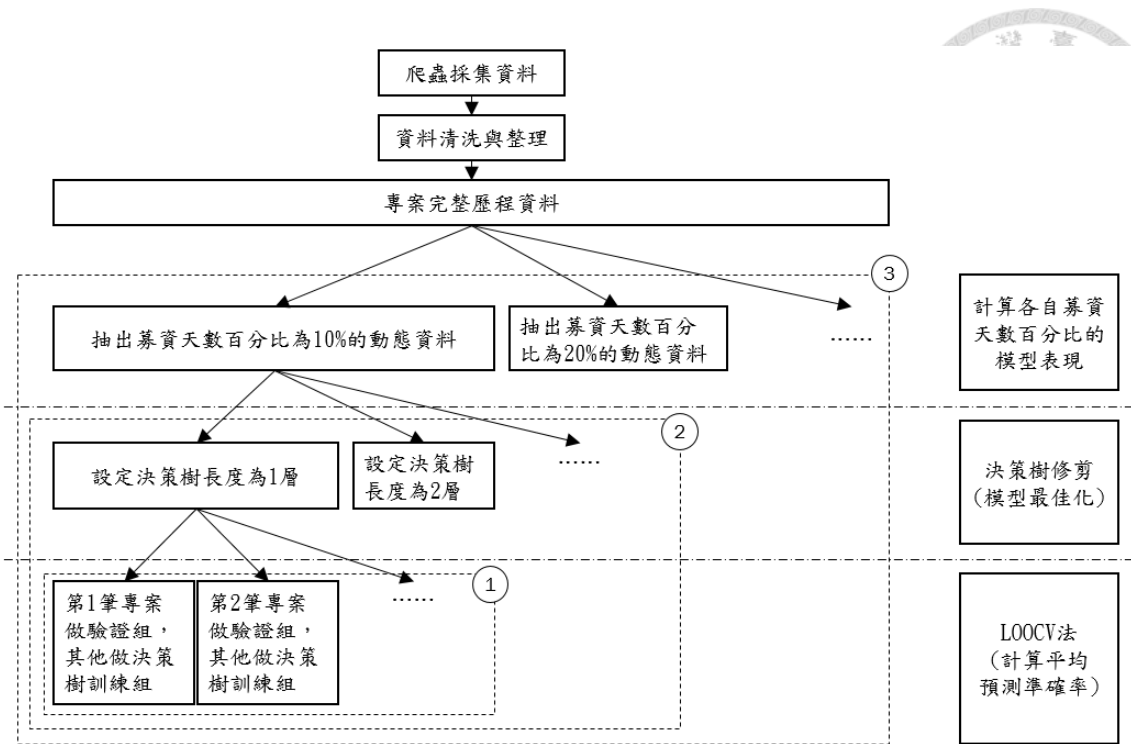


圖 8 募資天數百分比模型的分析流程

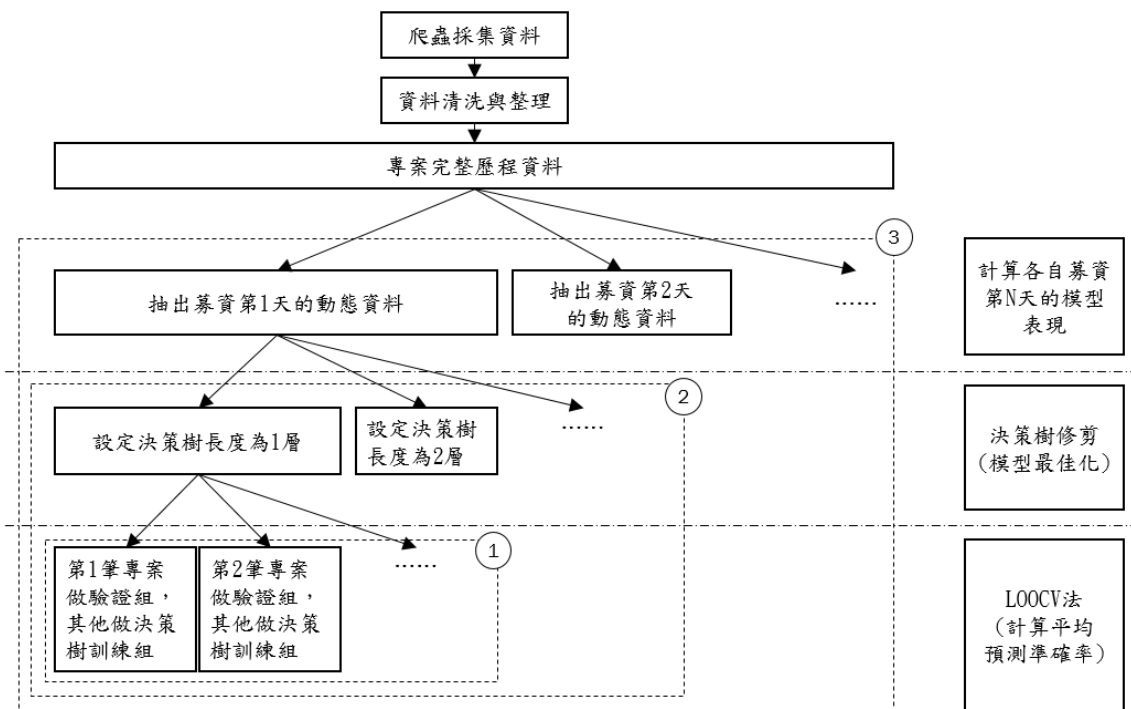


圖 9 募資第 N 天模型的分析流程



第二節 研究對象

本研究的研究對象為台灣知名的群眾募資平台 flyingV，其成立於 2012 年 4 月，截至 2015 年為止，已成功募資了共 3 億台幣，累積了 684 件成功募資案，單單 2016 年就有 237 件成功專案，為台灣最大的群眾募資平台。平台採用 all-or-nothing 模式，募資專案若成功，flyingV 將收取 8% 募資金額作為手續費，相較於過往文獻所探討的國際群眾募資平台如 Kickstarter，flyingV 網站的專案主要針對國內的用戶，故本研究較能適用於台灣群眾募資專案的預測關鍵因素。

第三節 研究樣本

相較於傳統群眾募資預測研究，其大多僅針對募資結束後的專案靜態資料作為研究樣本，本研究所蒐集的樣本則來自於 flyingV 網站每日仍在募資中的專案，希望能從每一份募資專案開始募資的第一天到募資結束期間內，追蹤並記錄每一天的動態資料，因此本研究以 R 語言撰寫的爬蟲程式進行抓取資料，蒐集時間為 2016 年 10 月 5 日至 2017 年 3 月 2 日為止近 5 個月，於每日晚間 11 點進行專案資料蒐集，合計蒐集了 8047 筆資料，並排除未能完整記錄起始的專案歷程資料 3848 筆後，成功蒐集了共 90 份募資專案起始的每日動態資料，以專案網址及記錄日期作為不重複值計共 4199 筆資料。

本研究在每日所採集的原始資料欄位，包含了專案網址、專案名稱、專案類型、紀錄日期、倒數天數、預計募資結束日期、募資現狀(已成功募資或尚未達成)、募資目標金額、累計募資金額、贊助人數、進度更新數、互動數、互動回應數、影片瀏覽人次、影片喜歡人數、影片不喜歡人數、影片頻道訂閱數、臉書響應總次數(為讚、評論、分享次數加總)、提案人歷史募資次數、完整文案、贊助方案種類、贊助方案價格，以及各贊助方案贊助人數。

除了蒐集完上述的原始資料欄位外，另會新增計算資料欄位，包含了募資總天數、募資第 N 天、募資金額達成率(累計募資金額/募資目標金額)、募資天數百分比(募資第 N 天/募資總天數)以及該專案最終募資狀態(募資成功或募資失敗)，其

中少數募資專案會突然延後或提前結束時間一至數天不等，故上述募資總天數將以募資專案正式起訖日計算得之。上述所有資料將會根據本章第三節研究變數定義，以最終募資狀態作為應變數，募資第 N 天與募資天數百分比分別做控制變數，進行決策樹預測分析。

第四節 研究變數定義

由於本研究所採用的預測模型為決策樹演算法，決策樹演算法主要是透過在眾多自變數下逐步挑出最能區辨結果的自變數，故沒有如迴歸分析中需要小心挑選研究變數以避免共線性等狀況，反而越多越有可能區辨的自變數越好。因此除了應變數及控制變數外，本研究也將在第三節花費較大篇幅逐一介紹所挑選的自變數。

一、 應變數：最終募資狀態

本研究以該募資專案最終的募資金額作為主要依據判斷。計算專案結束時的募資金額達成率是否超過 100%，若超過 100%則視為成功，反之則失敗。



二、 自變數

(一) 募資天數

多篇文獻中如 Etter et al.(2013)比較募資成功與失敗專案的募資天數平均值，都可以普遍發現到成功群眾募資的專案擁有平均較短的募資天數，可解釋為過長的募資天數，容易使提案人過於鬆懈，以及潛在贊助者缺乏立即讚住的動機。

(二) 募資目標金額

如自變數募資天數段落所述，Etter et al.(2013)也曾比較成功募資專案與失敗專案的目標金額平均值，普遍而言成功的群眾募資專案擁有平均較低的募資目標金額，可合理推斷較低的募資目標金額會使潛在贊助者認為該提案可能被實踐，也間接使其專案較容易達成。

(三) 累計募資金額

累計募資金額為一動態自變數，紀錄募資專案第 N 天時已累計的募資金額。本研究認為若該專案在募資期間擁有足夠大的累計募資金額時，可能會驅使贊助者有跟風的現象，使得募資成功率加速提升。

(四) 募資金額達成率

相對於前者的累計募資金額，募資金額達成率則是累計募資金額除以募資目標金額所計算而得。本研究認為前者代表的是金額的絕對大小，若考量相對大小，贊助者對於有一定比例募資金額達成率的專案，可能會有較高的意願進行贊助，進而提升募資成功率。

(五) 贊助人數

概念類似累計贊助金額，本研究認為若該專案有較多的贊助人數，除了可以營造潮流趨勢感外，也會給予其他贊助者較高的信賴感，並且較願意進行贊助之行為，故可能提升募資成功率。

(六) 進度更新數

Xu et al. (2014) 指出，若專案在籌資期間擁有較密切的進度更新，會讓潛在贊助者感受到提案人的重視及誠意，也代表提案人經常與贊助者進行溝通，故



本研究也將其變數納入自變數中。

(七) 互動數

此為潛在贊助人或贊助人於平台上向提案人進行提問，Xu et al. (2014) 研究也指出問答類型的更新也與募資專案成功與否有相關，本研究認為贊助人等願意主動與提案人進行問題交流，代表該專案有較高的吸引力，相對容易使募資專案成功。

(八) 互動回應數

同互動數概念，此變數代表提案人回答潛在贊助人或贊助人的問題，除了可視為提案人對專案贊助人的重視程度外，也可讓潛在贊助人感受到募資誠意，故可能是募資成功的重要指標。

(九) 有無影片

Kuppuswamy & Bayus (2013)的研究指出，擁有影片的募資專案有較高的募資成功率，高達八成的成功專案皆附有影片介紹，故本研究將該變數計算之。

(十) 影片平台類型

在 flyingV 平台上所使用的外部影片平台多為 YouTube 及 Vimeo 平台，一般而言 YouTube 平台擁有較高的流量，故本研究也將影片平台類型視為潛在影響指標。若該募資提案沒有影片，則以” none” 計。

(十一) 影片瀏覽人次

在網路盛行的時代，影片若擁有較高的瀏覽人次也暗示有較高的曝光機會，故將影片瀏覽人次納入自變數中。若該募資提案沒有影片，則以遺漏值計。

(十二) 影片喜歡人數

YouTube 及 vimeo 平台提供了使用者評比影片的機制，本研究認為若網友願意給予影片正評，有較高機率給予該專案贊助。若該募資提案的影片平台為 vimeo 或沒有影片，則以遺漏值計。

(十三) 影片不喜歡人數

YouTube 平台除了提供給網友點選喜歡該影片的機制外，也提供點選不喜歡的機制，因此本研究也將其納入自變數中。若該募資提案的影片平台為 vimeo



或沒有影片，則以遺漏值計。

(十四) 影片頻道訂閱人數

相對於僅討論募資案的影片，若提案人的 YouTube 頻道擁有較多的訂閱人數，也隱含了提案人擁有較多的網路影響力，呼應了眾多文獻中(Giudici et al., 2014; Mollick, 2014; Zheng et al., 2014) 提及群眾募資與社會資本的重要相關性，故本研究認為其有潛在可能影響募資成功率。若該募資提案的影片平台為 vimeo 或沒有影片，則以遺漏值計。

(十五) 臉書響應總次數

為該募資專案連結在 FB 平台上被按讚、評論與分享的加總次數，此為 FB 提供的查詢功能，只需提供募資專案連結即可了解該募資專案在社群上的影響力如何。其呼應了文獻中(Colombo, Franzoni, & Rossi Lamastra, 2015; Liao et al., 2015) 所指出的社群分享次數與募資成功率的相關性。

(十六) 專案類型

Harms(2007)及 Lambert & Schwienbacher(2010)的研究指出，如果是有形的回饋品，較容易吸引潛在贊助者。故本研究將不同的專案類型視為潛在的影響募資成功因素之一。

(十七) 專案簡介字數

此為在 flyingV 平台上一次瀏覽眾多募資專案時，所出現的簡介。雖然字數較多不見得有較高的募資成功率，但本研究也認為較少的簡介字數可能較難吸引潛在贊助者點入連結進一步了解，故也將其放入自變數名單中。

(十八) 專案文案字數

此為 flyingV 平台上已進入該募資專案，募資專案所提供的專案介紹文案。雖然字數較多不見得有較高的募資成功率，但本研究也認為較少的文案可能會讓潛在贊助者因缺乏了解，而減少贊助的可能性。

(十九) 提案人歷史募資次數

根據 Kickstarter 統計(2015)指出，提案人平均第一次提出專案就成功的比率是 38%，然而第二次提案成功的比率卻提升至 73%。因此本研究也將該提案人曾在 flyingV 提案的次數視為潛在的影響募資成功率因素之一。



(二十) 贊助方案種類數

Kuppuswamy & Bayus(2013)發現贊助方案的種類數與募資是否成功有所關聯，未達成目標的專案平均有 7.57 個回饋方案，而超越目標金額的專案則平均有 8.16 個回饋方案。故本研究將其列入自變數中。

(二十一) 贊助方案金額中位數

林亭佑(2015)的研究指出，贊助方案金額的中位數與募資成功率有正向關係，因為中位數可代表整體贊助方案的價格帶，越高的價格帶因較多的贊助金額而較容易達成募資目標金額。

(二十二) 贊助方案金額四分位差

林亭佑(2015)的研究指出，贊助方案金額的四分位差與募資成功率有反向關係，因為四分位差可代表整體贊助方案的價格帶分佈情形，而越大的四分位差代表贊助方案價格落差越大，呼應了 Mods(2010)認為選擇中間方案的贊助人，比起增加金額贊助，更有可能減少贊助金額來選擇基本的回饋方案。

(二十三) 贊助方案最低金額回饋方案

Mods(2010)的研究認為，提案人應該避免提供 25 美元以下的回饋方案。而 flyingV 平台的募資方案普遍提供 100 元的純粹贊助選項，本研究將人工排除純粹贊助的方案，並找出擁有最低金額的有形物回饋方案的價格，作為可能影響募資成功率的自變數之一。



三、 控制變數

(一) 募資天數百分比

其為募資第 N 天除以募資天數而得。多數文獻如 Colomo et al(2015)指出早期的資金挹注對募資成功率有其相關性，另一方面探討群眾募資動態預測的 Etter et al(2013)等皆以募資天數的百分比做為預測率變化的時間軸，為了以決策樹的預測準確率與文獻進行比較，故本研究也採募資天數百分比進行實驗操作。其中本研究與 Zhang et al(2015)相同，以 10% 為分組依據，作為時間軸上如 10%、20% 依序做預測率上的比較。

(二) 募資第 N 天

雖然大多數進行群眾募資動態預測的文獻都以募資天數百分比作為預測率變化的時間軸，但本研究認為除非到了接近募資的尾聲，否則募資天數百分比的單位對於潛在贊助者的影響力根本不客觀，只能作為專案間標準化的比較基準。

因此本研究決定嘗試採募資前 N 天來做為預測率變化的時間軸，一來在募資初期時募資天數百分比的影響力不大外，二來也能共同以相同的一天作為比較單位，比較不同專案上在第 N 天的動態變化，進而找出合適的預測指標及預測率逐日的變化。

第五節 統計分析方法

一、 留一交叉驗證法(leave-one-out cross validation, LOOCV)

為機器學習進行交叉驗證的方法之一。由於許多機器學習方法容易產生過度配適(overfitting)的問題，因此在操作上常常需要將資料母體進行事先分成訓練組及驗證組，使用機器學習演算法對訓練組資料進行模型訓練，再將驗證組資料丟入訓練好的模型查看預測準確度，來確保預測準確度的可信程度。

LOOCV 法代表只使用原樣本的一個樣本做為驗證資料，而剩下的樣本皆作為訓練資料；而這驗證法會進行 N 次(N 為樣本數)，讓每一個樣本都當過一次的驗證

資料，最後將 N 次的預測結果做平均，即是該模型的平均預測準確率。其運作流程如圖 10 所示。

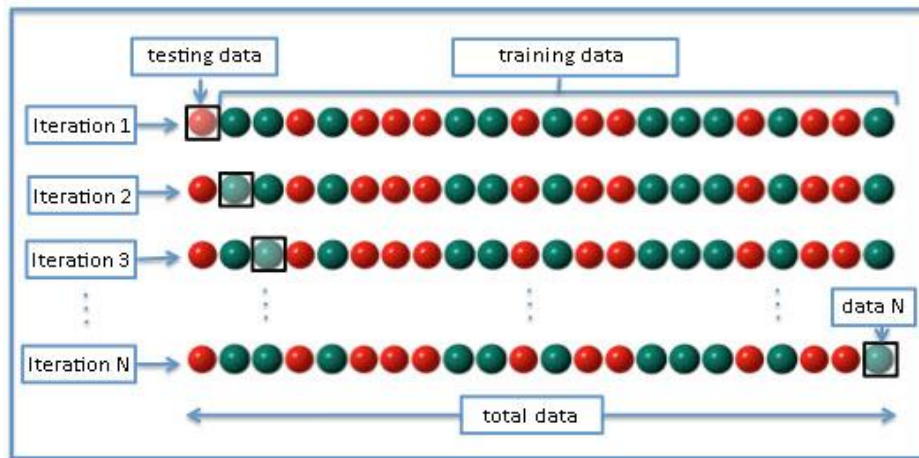


圖 10 LOOCV 法操作示意圖⁴

LOOCV 法所計算的預測準確率一般而言較為準確，但在計算效率上較為費時，因此 LOOCV 法適用於樣本數較少的資料母體。而本研究的專案樣本僅 90 件有完整歷程記錄，故本研究採用 LOOCV 法來驗證決策樹模型的預測準確率表現，而非一般動態分析文獻上常用的 K 次交叉驗證法(K-folds cross validation)。

二、 決策樹演算法(decision tree algorithm)

決策樹是機器學習的一種，與一般資料結構中的樹結構一樣，擁有節點、樹葉等結構，最早的演算法可追溯至 1960 年代。其為監督式學習的模型，可針對已給定的分類目標變數找出合適的變數分類準則，進而產生新的節點與分類準確率，這也是決策樹在學業界廣受歡迎的主因，其建構出來的分類原則簡單易懂，且容易運用在商業決策上，適合作為機器學習入門的演算法之一。

其原理為以資料母群體做為根節點，將所有自變數逐一找出最適合的分類閾值，將母群體分為二群以上，其分群後的應變數比例有較高的同質性，再比較所有自變數的分類閾值分類後，分類最乾淨的作為分支的原則，即分類原則，將母

⁴ 修改 wiki 圖片，原圖作者為 Joan.domenech91
<https://commons.wikimedia.org/wiki/File:Leave-one-out.jpg>

群體進行向下分群，接著重複上述動作；隨著決策樹不斷分支變長，直到找不到合適的分類閾值為止。圖 11 就是典型的決策樹模型，右圖顯示分類規則，左圖則顯示規則在散佈圖上的意涵即是找出分界界線，藉此將不同目標值分類出來。

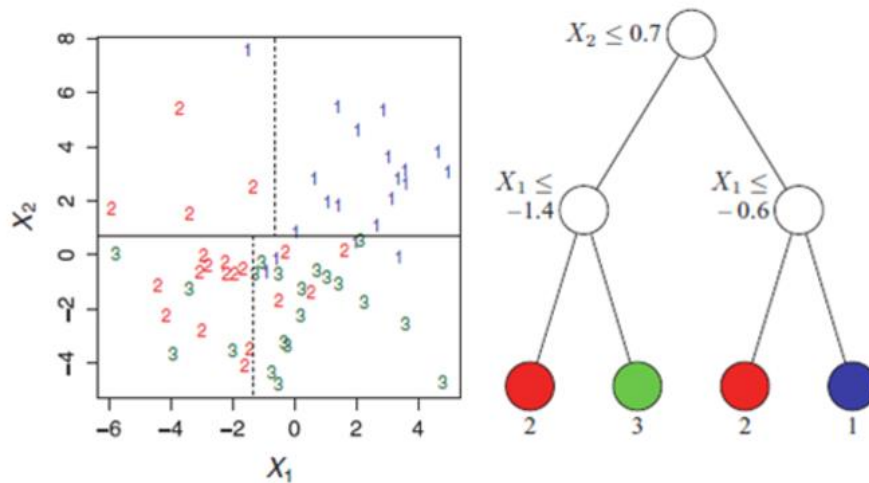


圖 11 典型的決策樹示意⁵

然而當決策樹過長時，會將分界界線繪製的過於細膩，若以此模型預測新的資料時反而容易預測失準；圖 12 則顯示了決策樹長度越長，訓練組的預測準確率(黑色實線)會越高，然而驗證組(黑色虛線)卻是先升高後在逐漸降低，即說明了模型過度配適的影響，因此一般會進行決策樹修剪，慣例上是找出驗證組預測準確率最高下，決策樹長度最短者作為最合適的決策樹長度。

⁵ 決策樹示意圖。 http://img.fanli7.net/2012-12-10/1355054667_7826.png

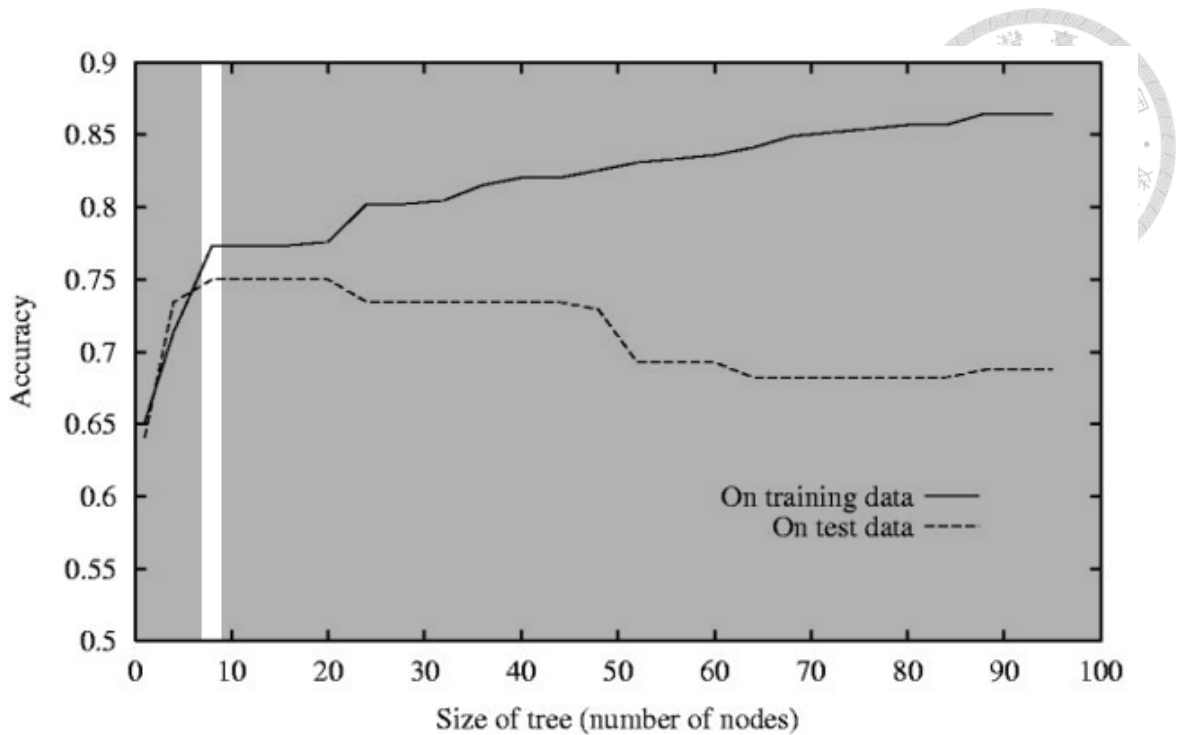


圖 12 不同決策樹長度下，訓練組與驗證組的預測準確率表現

相對於 Etter et al.(2013)與 Zhang et al(2015)的群眾募資專案動態分析研究中，分別採 KNN、馬可夫鏈(Markov Chain)及反向傳播類神經(backpropagation neural network, BPNN)演算法進行預測，其演算法缺點皆為預測準則難理解、不好運用；而 Rao et al.(2014)同樣採用決策樹演算法進行預測，卻未公布其所挑選出變數及其分類閾值，只顯示有不錯的預測準確率。故本研究將續採決策樹演算法，並嘗試找出其分類準則，提供給提案人有較明確的預測參考基準。

第四章 資料分析與結果



本章節將會進行資料分析，第一節將進行專案歷程資料的敘述性統計；第二節為模型一，以募資天數百分比為時間軸的決策樹預測分析，藉此與過往文獻建議的演算法做預測模型上的優劣比較；第三節為模型二與模型三，以募資第 N 天為時間軸的決策樹預測分析，其中模型二為將全部自變數都做為變數選擇的決策樹，模型三為僅用募資金額達成率作為決策樹的輸入變數，並比較模型二、三之間的預測率差異。其模型差異整理如下：

模型一：以募資天數百分比作為時間軸，將全自變數做為決策樹輸入變數。

模型二：以募資第 N 天作為時間軸，將全部自變數做為決策樹輸入變數。

模型三：以募資第 N 天作為時間軸，僅募資達成率做為決策樹輸入變數。

模型的表現將由預測準確率(accuracy, ACC)來判定，表格 1 為描述二元分類所有預測狀況的混淆矩陣(confusion matrix)，而本研究的預測準確率即為(真陽性+真陰性)/(總樣本數)而得之。

		真實募資結果	
		募資成功	募資失敗
預測募資結果	募資成功	真陽性(TP)	偽陽性(FP)
	募資失敗	偽陰性(FN)	真陰性(TN)

表格 1 二元分類預測結果的混淆矩陣

其中在進行募資天數百分比或募資第 N 天挑選(如挑出募資第 13 天的紀錄)時，本研究會先將已提前達成目標專案(如第 3 天時就已經達標)或已超過募資總天數(如總共僅募資 9 天)的專案排除掉，方進行決策樹模型訓練，以避免過度樂觀估計演算法的預測表現。而決策樹的演算法則使用 R 語言的 tree 函數操作之。




第一節 敘述性統計

本研究蒐集了共 90 份專案完整歷程紀錄，合計 4199 筆資料，表格 2 則顯示了以最後一天紀錄各變數的敘述性統計。專案結果顯示平均有 50% 的募資成功率，比過往文獻平均 40% 左右高出 10%，推測應是群眾募資風潮已起，已有較多潛在贊助人關注。成功募資專案的平均目標金額僅約 11 萬元左右，遠少於失敗專案近 23 萬，映證了文獻所說較低的目標金額有助於募資成功。此外也與文獻相同的是成功專案募資天數依然較短，平均約 45 天。值得一提的是，募資成功與募資失敗的平均募資達成率差異相當大，募資失敗的平均達成率僅不到 11%，可以說幾乎不受到潛在贊助人的喜愛，可見其群眾募資專案的達成率往往有 M 型化的趨勢。

在進度更新數、互動提問數、平均募資提案數、影片瀏覽人次、臉書響應總次數上，募資成功專案都有較高的數值，合乎情理也與過往文獻提及相符。有趣的是在文案字數上，募資成功專案反而平均僅有 1462 字，失敗專案的文案字數平均卻高達 2,202 字，儘管未有顯著差異，但其 P 值也已達 0.079，本研究推測可能近期已有趨勢將文案圖像化有關，可以在排版及字型上有較優的美化表現，也是未來可進一步探討議題。

在探討贊助方案上，成功募資專案平均有較多的贊助方案選擇，約有 10 種類型，高於失敗專案的 7.8 種，與文獻相符；然而在贊助方案價格的中位數及四分位數上，成功專案偏低且分布偏短，顯示成功專案的贊助方案普遍為較平價的選擇，與文獻略有出入，在平均贊助方案最低金額回饋方案上也較低；儘管如此，在探討贊助方案上成功專案與失敗專案皆未有顯著差異。

表格 2 整理出成功與失敗專案的敘述統計表，並進行了兩者平均值的 T 檢定，列出其 P 值；本研究發現僅平均募資目標、平均贊助人數、平均募資金額、平均募資達成率及平均進度更新數有顯著差異，大多過往文獻提及的靜態影響因素皆無顯著差異。



	募資成功專案	募資失敗專案	總體	T檢定之P值
專案數	45	45	90	-
專案數占比	50%	50%	100%	-
平均募資目標	\$109,467	\$340,073	\$224,770	0.0086**
平均募資天數(天)	44.98	48.33	46.66	0.3117
平均贊助人數(人)	93.98	12.98	53.48	0***
平均募資金額	\$128,842	\$21,241	\$75,041	0***
平均募資達成率	144.6%	10.29%	77.45%	0***
平均進度更新數	3.289	1.067	2.178	0.0038*
平均互動提問數	0.422	0.222	0.322	0.2724
平均募資提案數	1.22	1.09	1.16	0.3027
平均簡介字數	35.4	32.84	34.12	0.3918
平均文案字數	1,462	2,202	1,832	0.0794
平均影片瀏覽人次	2,261	504.1	1,402	0.2136
平均臉書 響應總次數	1,383	402.8	893	0.0053
平均贊助方案數	10	7.8	8.9	0.0540
平均贊助方案 價格中位數	\$1,253	\$1,670	\$1,461	0.2848
平均贊助方案 四分位數	\$2,619	\$3,595	\$3,107	0.5806
平均贊助方案 最低金額回饋方案	\$498	\$787	\$642.5	0.2276

表格 2 成功與失敗募資專案的敘述性統計表

第二節 募資天數百分比的決策樹分析



在此章節將會以募資天數百分比為時間軸做動態預測分析，即決策樹模型的模型一，也是過往文獻慣用的動態預測時間軸架構，並進一步與文獻建議的演算法做準確率比較。本研究將取出募資天數百分比 10%、20%、30% 等依序至 90% 共 9 個區間來分別進行預測，並繪製出如圖 6 的預測率變化圖。

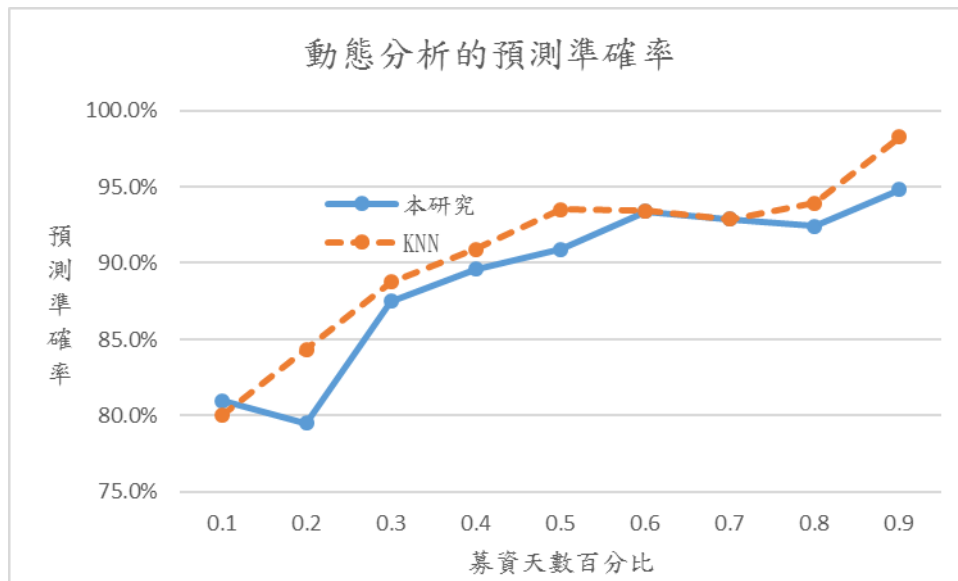


圖 13 模型一 動態分析的平均預測準確率(募資天數百分比)

由圖 13 可以發現，本研究的預測準確率在募資天數 10% 時，就已經有 80% 以上的預測準確率，並且整體預測準確率隨著募資天數百分比增加而上升，最高可達近 95% 的預測準確率。惟在募資天數 20%、70%、80% 的平均預測準確率略為下滑，推測該時間點的資訊本身可能較不適合作為預測的領先指標。

另外圖 13 也與 Etter et al.(2013)研究的動態分析預測率相比較，本研究也採其建議的 KNN 演算法同樣進行預測，發現乍看之下整體而言 KNN 演算法仍有較佳的預測準確率水準；但值得一提的是，兩者間的平均準確率差距不到 2%，因此本研究對兩者結果進行兩相依母體期望值差 T 檢定，各募資天數百分比下的 P 值及樣本數如表格 3，也證實兩者預測準確率近乎沒有顯著差異，故決策樹演算法可以替代 KNN 演算法用於募資結果預測。

此外 KNN 演算法需要有大量的訓練樣本方能進行預測，對沒有專業分析人員

及其他專案歷程記錄樣本的提案人而言，相較缺乏實務價值；而決策樹提供的分類預測規則，如表格 4 所示，給予提案人現成易懂的預測邏輯，並可直接套用之，對提案人有相當程度的幫助。

募資天數 百分比	10%	20%	30%	40%	50%	60%	70%	80%	90%
P 值	0.159	0.045 [*]	0.320	0.321	0.159	1	1	0.568	0.159
樣本數	85	83	80	77	77	76	70	66	58

表格 3 模型一之兩相依母體期望值差 T 檢定表



募資天數百分比	決策樹預測募資成功分類規則
10%	(1) 募資達成率>9% (2) 募資達成率<9% & 分類屬於休閒、音樂、旅行、設計、遊戲、運動 & 影片有人按 like & 影片瀏覽人次>339 (3) 募資達成率<9% & 分類屬於休閒、音樂、旅行、設計、遊戲、運動 & 影片沒有人按 like & 募資目標金額<27,500
20%	募資達成率>12.4%
30%	募資達成率>15.7%
40%	募資達成率>16.2%
50%	(1) 一定成功：募資達成率>22.5% & 累計募資金額>11,750 (2) 可能成功：募資達成率>22.5% & 累計募資金額<11,750 (3) 可能失敗：募資達成率<22.5% & 募資天數<28 天 (4) 一定失敗：募資達成率<22.5% & 募資天數≥28 天
60%	募資達成率>16.1%
70%	募資達成率>16.1%
80%	募資達成率>22.8%
90%	募資達成率>51.2%

表格 4 模型一之決策樹預測募資成功規則

第三節 募資第 N 天的決策樹分析

相對於絕大多數的文獻皆採用募資天數百分比，以因應不同募資天數的專案能作模型的共同訓練與預測，本研究卻認為其方法會受限在時間單位上的不同而有所落差，舉例而言若有兩個募資專案，其募資天數分別為 10 天與 60 天，若取募資天數百分比之 10% 做訓練與比較，則會取出募資 1 天與募資 6 天的結果，在動態代表性上會有一定缺陷。故本研究接下來將採募資第 N 天方式進行動態預

測，以共同的時間長度(天)來進行模型的共同訓練與預測。

其中由於多數文獻指出募資初期為成功與否的關鍵觀察期，故本研究將取 N 小於等於 14 天(2 個禮拜)的狀況下進行預測，同時為確保 N 仍處於募資初期，故研究樣本將排除募資天數小於 42 天(6 個禮拜)者，仍有 54 件專案完整歷程記錄，共 3948 筆資料。

一、 模型二 使用所有自變數做訓練

由圖 14 的動態分析預測準確率可以發現，其預測準確率也有優秀的表現，在募資第二天(約不到募資天數百分比 4.7%)就能擁有 83%的平均預測準確率，募資第五天(約不到募資天數百分比 11.9%)則達到 85%準確率，募資第 9 天(約不到募資天數百分比 21.4%)即達到 92.5%的準確率；而本研究也依照比例與模型一相比，分別按照募資天數百分比 10%、20%、30%對應募資第 4、8、12 天(募資總天數以 42 天計)，在排除募資總天數長度不同後模型二也擁有較佳的預測準確率表現，而模型二使用募資第 N 天作時間軸呈現，也對提案人更容易理解及判讀。

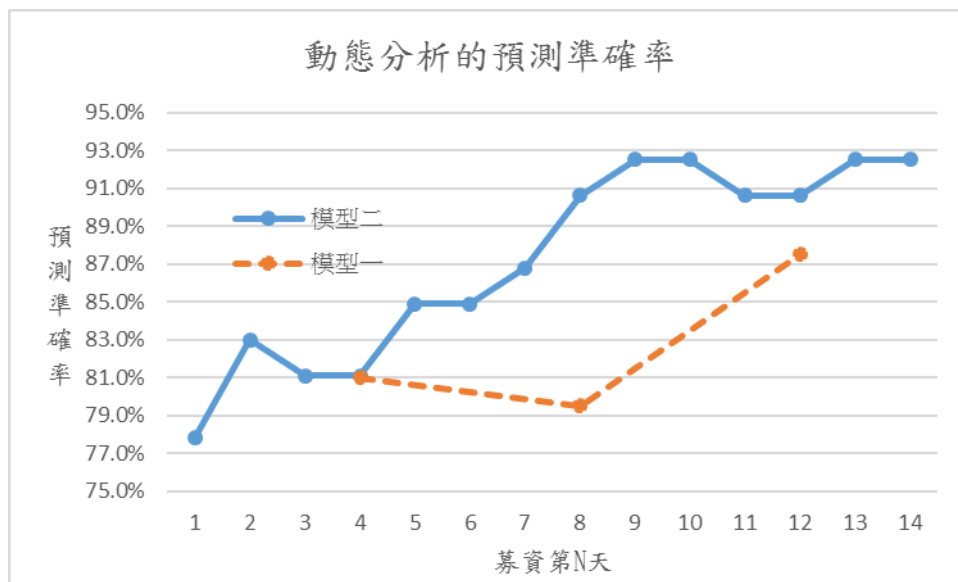


圖 14 模型二 動態分析的平均預測準確率(募資第 N 天)，採所有自變數

有趣的是，本研究觀察模型其每日的最佳決策樹長度如圖 15，發現除了第八天建議的決策樹長度為 2 層外，其餘天數的最佳決策樹長度皆為 1 層；再進一步

觀察模型挑選出的關鍵變數如表格 5，大多以募資金額達成率為主，儘管直覺上很合理、缺乏洞見，但本研究發現其分類閾值皆不到 20%，換句話說僅觀察募資專案在前 14 天內募資金額達成率是否能超過 20%，將有高達 90%的預測準確率判斷該專案是否能募資成功。

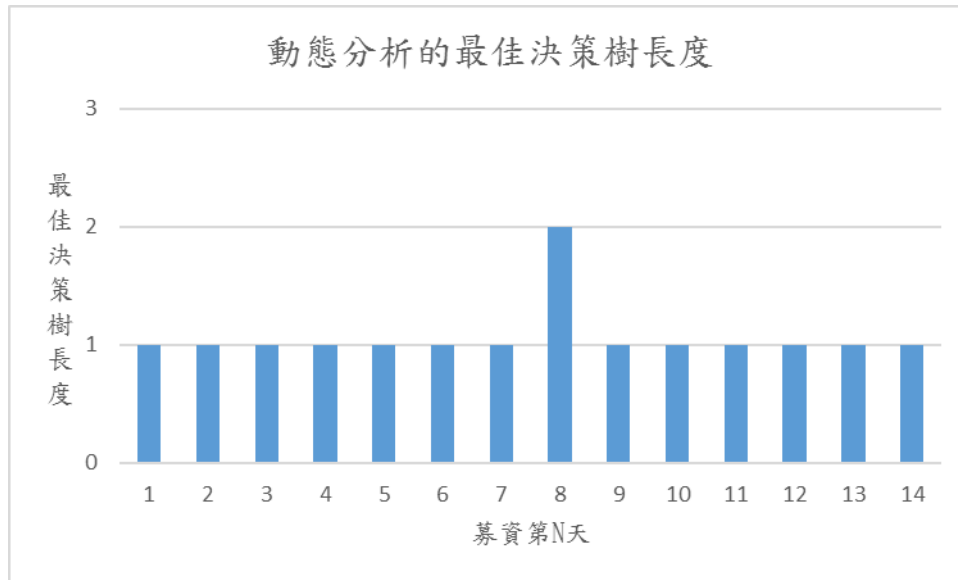


圖 15 模型二 動態分析的最佳決策樹長度



募資第 N 天	決策樹預測募資成功分類規則
第 1 天	募資達成率>3.6%
第 2 天	募資達成率>5.6%
第 3 天	募資達成率>6.7%
第 4 天	募資達成率>7.1%
第 5 天	募資達成率>9.6%
第 6 天	募資達成率>9%
第 7 天	募資達成率>11.8%
第 8 天	(1) 一定成功：募資達成率>8.4% & 影片瀏覽人次>116 人 (2) 可能成功：募資達成率>8.4% & 影片瀏覽人次<116 人 (3) 可能失敗：募資達成率<8.4% & 類型屬於出版、在地 (4) 一定失敗：募資達成率<8.4% & 類型不屬於出版、在地
第 9 天	募資達成率>11.6%
第 10 天	募資達成率>11.6%
第 11 天	募資達成率>10.1%
第 12 天	募資達成率>10.3%
第 13 天	募資達成率>13.8%
第 14 天	募資達成率>15.1%

表格 5 模型二之決策樹預測募資成功規則

透過上述的發現，因此本研究設計模型三，僅將募資金額達成率作為決策樹的自變數，其餘設定皆與模型二相同，其模型表現於下段詳述。



二、 模型三 僅使用募資金額達成率變數做訓練

圖 16 顯示了模型三與模型二的預測準確率比較，可以發現其預測準確率近乎一致，說明了僅採用募資金額達成率做預測分類依據，也能有一定的預測表現。其中部分天數模型三的預測準確率低於模型二，代表該天模型二有其他變數更適合做預測；而部分天數模型三的預測準確率高於模型二，推測應是模型二在進行交叉驗證時，部分樣本因採取不同的自變數使得預測錯誤，進而在計算總體預測準確率時拉低了平均值，本研究認為當研究樣本足夠多時應可避免此一狀況。

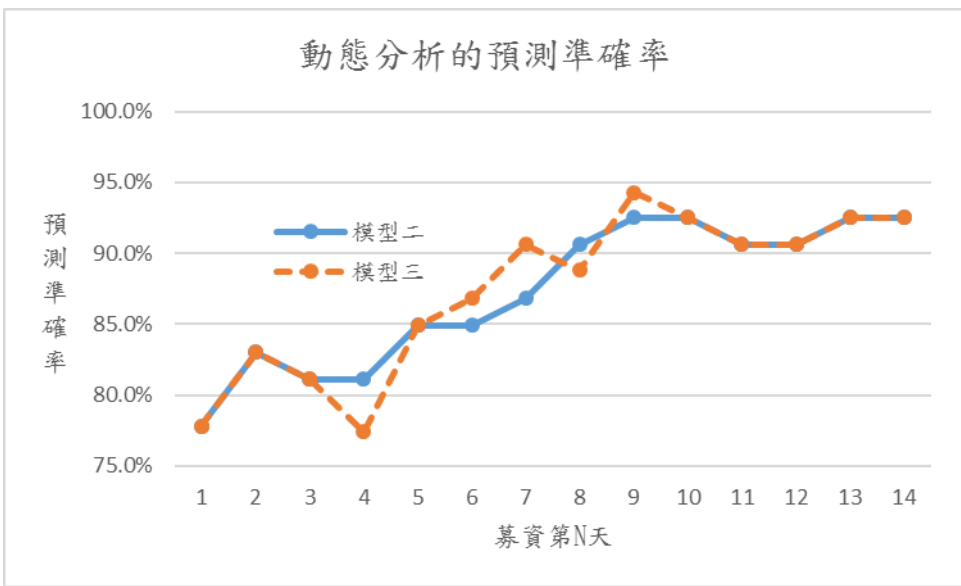


圖 16 模型三 動態分析的平均預測準確率(僅採募資金額達成率做自變數)，並與模型二進行比較。

而圖 17 則進一步整理出每日的預測規則(即募資金額達成率閾值)，並與平均預測準確率做比較。該圖可提供提案人做為參考，在募資期間能以簡單易懂的方式來預測專案是否成功；舉例而言，在募資第二天，僅需要觀察該專案的募資金額達成率是否超過 5.6%，則有高達 83%的平均預測準確率判斷該專案是否最後將募資成功。圖 18 則顯示了募資第 N 天的預測募資成功的占比數，可觀察其真陽率；同樣舉例而言，在募資第二天，若該專案的募資金額達成率超過 5.6%，則有 100%的機率最後將募資成功。

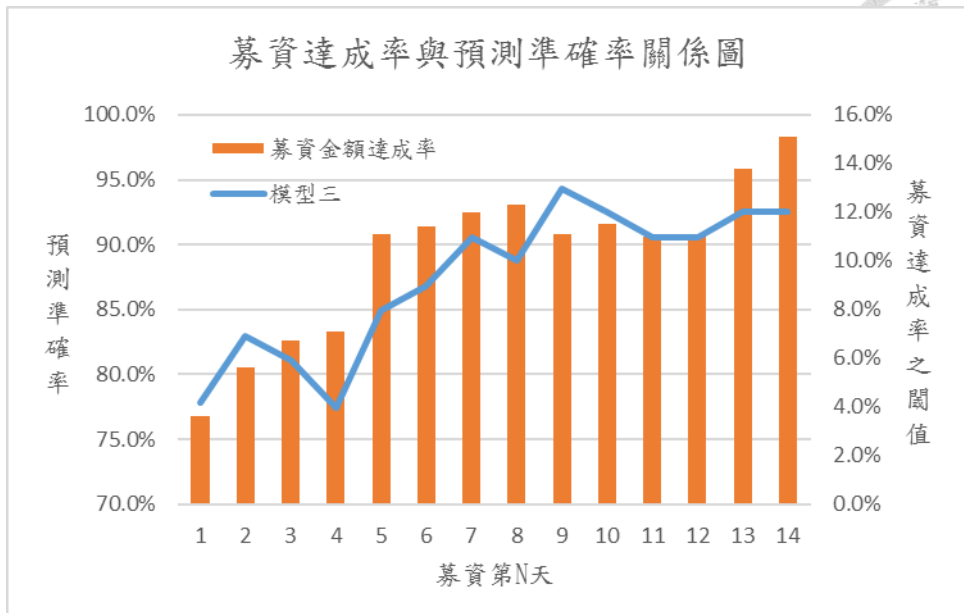


圖 17 模型三 募資金額達成率之閾值與平均預測準確率的動態分析

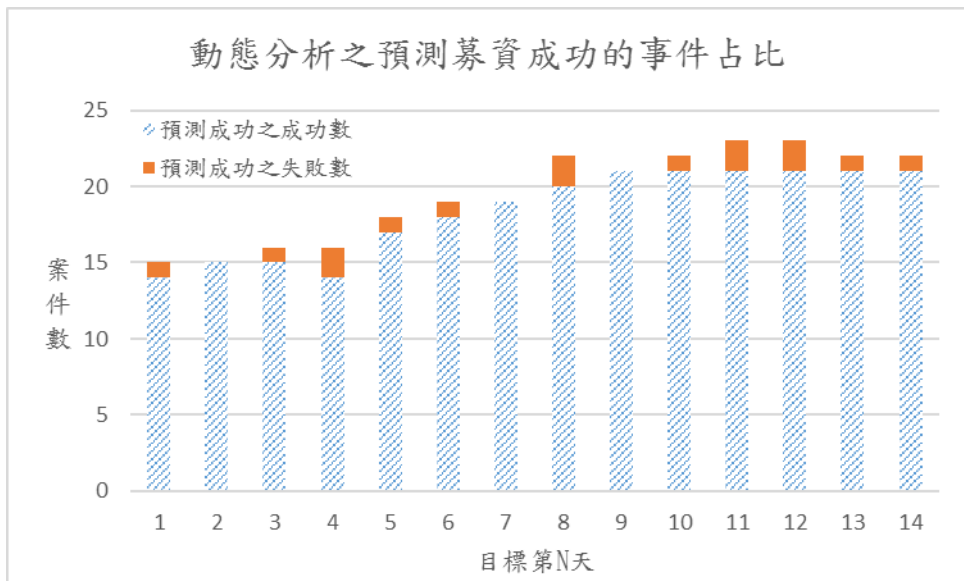


圖 18 模型三動態分析之預測募資成功案件數占比

第五章 結論與建議



本研究目的在於在群眾募資專案中進行動態預測，並以提案人角度找出預測精準且易用的模型演算法，故嘗試使用決策樹操作之。因此從第一章的研究動機衍生出三個研究問題：第一，使用決策樹做預測，其準確率變化是否與文獻建議的演算法相當？第二，募資專案若改以募資第 N 天做動態預測的時間軸，是否有較佳的預測準確率表現？第三、是否能以較少的關鍵變數進行預測，同時預測表現仍可與一般動態預測方法相近？在第二章的文獻探討中，發現不僅較少探討群眾募資專案動態預測的研究，而探討動態預測的研究也多強調於機器學習的預測表現優化，卻不利於一般提案人使用及操作；因此透過第二章文獻探討，本研究於第三章建立分析流程及研究模型設計，嘗試以 flyingV 為研究對象，進行不同設定下的決策樹模型預測。第四章則呈現了三種模型訓練後的預測結果，顯示了決策樹提供相當水準的預測結果，同時以募資第 N 天作為時間軸的動態預測也有更佳的預測表現，最後只以每日募資達成率做預測指標，也指出募資第二天就有突破 80% 的準確率；本研究透過這些模型設定，證明只需要簡單的觀察特定指標，即可協助提案人在早期評估募資專案的最終結果。

第一節 研究發現與實務意涵

本章節將依序對本研究的三個研究問題及實務意涵分別做出探討。

首先，在圖 13 及表格 3 中說明了決策樹演算法確實與 Etter et al(2013)研究提及的 KNN 演算法預測準確率幾乎不分軒輊，但相對於 KNN 演算法需要大量訓練樣本作為模型，決策樹提供現成且明確的規則給提案人做專案操作上的改善，提案人本身不需了解且建立模型；因此若考量略微犧牲預測率與明確預測規則帶來的效益，對提案人而言不可否認地決策樹是實用的機器學習預測方法。

第二，本研究取出募資天數長於 42 天以上的專案，以募資第 N 天作時間軸進行動態預測分析，研究數據也指出募資第二天就 80% 以上的預測準確率水準，第九天就達到 90% 以上的水準。若以天數比例計算與文獻比較，也有比同期更優秀

的預測表現，證明了在排除過往文獻慣用的募資天數百分比而採募資第 N 天做動態分析後，反而有更佳的預測能力。本研究推測應是募資天數百分比法無法排除不同專案在募資天數上的差異，造成同樣 10% 部分專案只有一天的籌資時間，而部分專案卻可能有長達六天的籌資時間，使其預測準確率下降。

第三，本研究延續募資第 N 天為時間軸的動態預測分析，但僅使用每日的募資達成率做決策樹模型的解釋變數，依然有與原模型不相上下的預測能力，本研究數據證明了募資達成率是動態預測中最好的預測指標。有趣的是，初期的募資達成率閾值都相當的低，即使到了募資第 14 天，募資達成率閾值也僅有 15%，但其預測準確率卻高達 92%；本研究推測相對於初期投入大量行銷資源而快速達標的專案，仍存在缺少曝光資源但仍具吸引力的群眾募資專案，因此在早期雖沒有爆炸性的成長，卻也能緩慢卻穩定地籌資，累積口耳相傳的基數後在中後期開始巨幅成長起來，近似於傳統的成長曲線。

本研究在實務貢獻上有三，第一是相對於過往動態預測文獻上，集中於 Kickstarter 及中國的追夢籌上，對於台灣群眾募資平台上的實證研究相當少，因此本研究提供了台灣 flyingV 在動態預測上的論證數據，也證明了決策樹的機器學習演算法也能在台灣有相似的預測能力。第二，就我能力所及，本研究應是少數先採用募資第 N 天的方式進行動態預測的方法，也證明相對於募資天數百分比，本研究有較佳的預測表現。第三，本研究數據證明僅用募資達成率的決策樹模型就有良好的預測能力表現，並整理出前 14 天的募資達成率參考閾值，不僅說明初期沒有爆炸性的募資進度不需過於操心，同時也提供提案人明確的每日關鍵績效指標，作為省察募資專案表現的客觀參考依據。

第二節 研究限制與未來研究建議

雖本研究已蒐集了近五個月的 flyingV 研究樣本，卻仍僅有 90 件專案擁有完整的歷程記錄，相對於國外文獻以 kickstarter 為主的樣本動輒上萬件下，樣本數仍稍嫌不足，使得本研究進行決策樹的動態預測分析時，較難依據專案類型等方式進一步分類比較，否則將碰到樣本數過少代表性不足的問題；因此本研究建議未

來可以往跨平台如嘖嘖募資平台同時進行資料蒐集，並且將蒐集時間拉長至一整年，藉此可比較不同平台及不同季度下的動態預測差異，探索關鍵預測指標是否有所不同。

另一方面，雖本研究已指出前 14 天的募資達成率閾值及對應的預測準確率，提供給提案人與平台業者良好可靠的指標數據，卻可能僅適用於募資天數長於 42 天的群眾募資專案；42 天以下募資專案其對應的募資達成率閾值可能會有所出入，因此本研究也建議未來可依照募資天數短中長分別進行分析，探查其募資達成率閾值的差異以及對應的預測準確率。

此外，群眾募資中社群如何交互影響仍扮演相當重要的角色，儘管台灣不如國外對 twitter 平台來得盛行，使得社群的影響力研究相對難以調查，但仍可嘗試對臉書平台的討論時間軸上的變化，如何與動態預測的因素交互影響觀察，雖然相對挑戰，但仍是一個不錯的未來研究方向。

在本研究的結果上，本研究發現普遍初期的募資達成率閾值都相當的低，前 14 天的募資達成率只要突破 16% 就有近 90% 的預測準確率，雖然本研究推論應是口碑累積的前期、仍在累積口耳相傳的群眾基數，但仍缺乏詳細進一步的研究數據，因此本研究建議未來研究可針對此一特別的募資成長曲線，做進一步的研究探討。

最後，儘管本研究為提案人提供了簡單易懂的每日績效指標，辨別專案的執行成效，然而當提案人的累計募資達成率未達當天的績效指標時，若有進一步的行動建議，相信會能帶給提案人更多改善的空間；本研究的表格 2 雖驗證了各種靜態因素的影響力如募資目標設定等，但在開始募資後只有進度更新數有顯著地指出成敗差異，卻也缺乏因果驗證；因此本研究建議可針對募資期間，探討何種行為可有助於提升募資達成率，協助提案人達到每日目標。

參考文獻



- Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. *Journal of business venturing*, 29(5), 585-609.
- Chen, S. Y., Chen, C. N., Chen, Y. R., Yang, C. W., Lin, W. C., & Wei, C. P. (2015). Will Your Project Get the Green Light? Predicting the Success of Crowdfunding Campaigns. In *PACIS* (p. 79).
- Chuanhui, L. I. A. O., Yunhao, Z. H. U., & Xi, L. I. A. O. (2015). The Role of Internal and External Social Capital in Crowdfunding: Evidence from China. *Revista de Cercetare si Interventie Sociala*, 49, 187.
- Colombo, M. G., Franzoni, C., & Rossi-Lamastra, C. (2015). Internal social capital and the attraction of early contributions in crowdfunding. *Entrepreneurship Theory and Practice*, 39(1), 75-100.
- De Witt, N. (2012). A Kickstarter's Guide to Kickstarter. Retrieved from <https://kickstarterguide.com/files/2012/06/A-Kickstarters-Guide.pdf>
- Etter, V., Grossglauser, M., & Thiran, P. (2013, October). Launch hard or go home!: predicting the success of kickstarter campaigns. In *Proceedings of the first ACM conference on Online social networks* (pp. 177-182). ACM.
- Gerber, E. M., Hui, J. S., & Kuo, P. Y. (2012, February). Crowdfunding: Why people are motivated to post and fund projects on crowdfunding platforms. In *Proceedings of the International Workshop on Design, Influence, and Social Technologies: Techniques, Impacts and Ethics* (Vol. 2, p. 11).
- Giudici, Giancarlo and Guerini, Massimiliano and Rossi Lamastra, Cristina, Why Crowdfunding Projects Can Succeed: The Role of Proponents' Individual and Territorial Social Capital (April 24, 2013). Available at SSRN:

<https://ssrn.com/abstract=2255944> or <http://dx.doi.org/10.2139/ssrn.2255944>

Greenberg, M. D., Pardo, B., Hariharan, K., & Gerber, E. (2013, April). Crowdfunding support tools: predicting success & failure. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1815-1820). ACM.

Greenberg, M. D., Pardo, B., Hariharan, K., & Gerber, E. (2013, April). Crowdfunding support tools: predicting success & failure. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1815-1820). ACM.

Harms, M. (2007). What drives motivation to participate financially in a crowdfunding community?. Available at SSRN 2269242.

Kuppuswamy, Venkat and Bayus, Barry L., Crowdfunding Creative Ideas: The Dynamics of Project Backers in Kickstarter (November 2, 2015). *UNC Kenan-Flagler Research Paper No. 2013-15*.

Li, Y., Rakesh, V., & Reddy, C. K. (2016, February). Project success prediction in crowdfunding environments. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining* (pp. 247-256). ACM.

Mod, C. (2010). KICKSTARTUP–Successful fundraising with Kickstarter. com & (re) making Art Space Tokyo. *craigmod–Journal.[web document].[referenced 18.2. 2014]*. Available: *craigmod.com/-journal/kickstartup*.

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of business venturing*, 29(1), 1-16.

Rao, H., Xu, A., Yang, X., & Fu, W. T. (2014, April). Emerging dynamics in crowdfunding campaigns. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction* (pp. 333-340). Springer International Publishing.

Steinberg, D. (2012). The Kickstarter handbook: Real-life success stories of artists,

inventors, and entrepreneurs. *Quirk Books*.

Xu, A., Yang, X., Rao, H., Fu, W. T., Huang, S. W., & Bailey, B. P. (2014, April). Show me the money!: An analysis of project updates during crowdfunding campaigns. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 591-600). ACM.

Xuan Zhang, and Alan G. Wang. 2015. Money Talks: A Predictive Model on Crowdfunding Success Using Project Description. *Proc. ACIS 2015*, 1–8.

Zhang, B., Liu, C., & Xu, W. (2015) Make the Right Choices: Dynamic Prediction on the Success of Crowdfunding Projects. Retrieved from http://epic.is.cityu.edu.hk/sigbps/BPS15/docs/SIGBPS2015_paper_12.pdf

Zheng, H., Li, D., Wu, J., & Xu, Y. (2014). The role of multidimensional social capital in crowdfunding: A comparative study in China and US. *Information & Management*, 51(4), 488-496.

林亭佑. (2015). 群眾募資專案成效之影響因素：台灣群眾募資平台 flyingV 之實證研究. 臺灣大學商學研究所學位論文, 1-50.