

經濟部 107 年度
《人工智慧產業關鍵技術拔尖計畫》
合作研究計畫

《信評資料與特徵萃取》
建議書徵求文件

財團法人資訊工業策進會

中華民國 107 年 03 月 16 日

107 年度合作研究計畫建議書徵求文件

一、 簡介

「信評」是金融產業甚至商業環境中極為重要的基礎建設，一方面現代的各種商業行為有大部分的比例不是以現金交易而是以信用交易為主，而不論是借貸，授信，租賃，還是賒帳等商業行為，據以決策的依據主要就是分析與判斷交易對手未來是否可能違約？違約的可能性如何？與如何盡量控管好未來對方願意付帳或還款的風險？而信用風險評估更是銀行等金融機構或是企業授信決策的重要決策過程，其中交易對手的「信評」，就是最寶貴的參考依據了。

傳統的信評或徵信，靠著所謂的五 C 或五 P 的原則來進行，五 C 指的是品格 (Character)，能力 (Capacity)，資本 (Capital)，擔保品 (Collateral) 與經濟情況 (Condition)。而五 P 則是指被授信人 (People) 資金用途 (Purpose) 還款財源 (Payment) 保障 (Protection) 與展望 (Perspective)。

不論是五 C 還是五 P，雖然一直以來都是實務界與教科書中重視的面向，但畢竟只是觀念，不易量化成為知識經濟將經驗加以承傳累積。一直到紐約大學的 Altman 所發展的 Z score 模型有了革命性的改變，他是第一個將企業的信評應用財務會計數字計算出一個分數，而且推動其在實務上應用甚廣的學者，金融機構因此多了一個工具來協助它們提高授信成功的機率與減少呆帳的損失，或是提供授信人員的某種的「保險」。

許多的研究人員後來發展出各式各樣的信評模型，包括了以市場資料與複雜數學推倒應用的 KMV 模型，而一些國際知名的標準普爾與穆迪等信評公司，也參考了許多量化模型來作為其提供企業信評的最後參考。

除了對於大公司的企業金融評等以外，對於個人消費金融信用或是中小企業與小微金融信用，國際最知名的算是 FICO 的評分卡了，中小企業與大企業信評方式大的不同，是前者在負責人的個人信評重要性比重極大。

近年來隨著大數據，AI 與金融科技的快速發展，國際上的信評分析又有一個新的突破，信評由過去之內容 (content) 分析轉變為脈絡場景 (context) 分析。這也是本計畫的主要目的，進行信評資料與特徵萃取技術研究，除了傳統的會計資料與市場資料以外，利用各種社群，交易平台所取得的表外資料，一方面能與表內資料進行驗證，一方面也可以補足表內資料所欠缺的資訊。

二、計畫目標

本計畫的主要目標為應用近年來應用甚廣的風險因子擷取技術與輪廓描繪技術，蒐集各類大數據包括新聞，社群，部落格和網頁，交易平台等資訊，來歸納出各種具有決策有用性的表外風險因子，建立適當的模型，以期能對於國內金融機構放貸給中小企業時，更能提升其準確評估的能力，也希望能因此提高國內優質中小企業融通資金的機會。

三、計畫範圍

本計畫以台灣某產業(例如餐飲業)的信評為研究對象。蒐集與其信用風險相關的各類大數據資料，分析出適當的信評風險因子，並與目前現有的基準資料(benchmark)相互比對驗證，發展出適合之信評模型。

工作項目包括：

1. 蒐集目前各國相關之實務應用文線，提出適合國內之假設，應用各類資料探勘與爬蟲技術，從社群網絡，新聞媒體與交易平台等內容中蒐集過濾有信評關聯性的風險因子。
2. 應用風險因子從事國內適當之中小企業之信評預測，並與目前現存之基準相互比對，測試是否有增額之預測效果。
3. 整合傳統之信評模型與本研究發現之風險因子，提出整合之信評模型建議，分析與建議適當權數，並進行適當之驗證(Validation)，以確定其效能。

四、預期成果

1. 交付成果：

- 1份表外資料風險因子篩選初步報告(2018/05/31)
- 1份專利構想(2018/05/31)
- 1份表外資料風險因子篩選進階報告(2018/09/30)
- 1份期末報告(2018/11/30)

查核點	時間	產出物	報告架構/大綱
M1	2018/05/31	表外資料 風險因子 初步篩選 與 專利構想	<ul style="list-style-type: none">●文獻探討●蒐集資料(資料模擬)與分析●表外資料風險因子初步篩選結果●可行之專利構想說明
M2	2018/09/30	表外資料 風險因子 進階篩選	<ul style="list-style-type: none">●文獻探討●蒐集資料(資料模擬)與分析●表外資料風險因子進階篩選結果
M3	2018/11/30	期末報告	<ul style="list-style-type: none">●研究方法設計之說明●表外資料風險因子篩選成果與說明●與現行風險因子之比較

2. 預期效益：

- 提供專家對於表外資料風險因子定義及重要性之評估技術研發參考。

- 解讀及分析資料之技術，可作為專利構思。
- 產出可鑑定系統之可信度、多元性及通用性。

※前述成果如有專利構想或專利申請產出時，需注意專利申請之新穎性(novelty)。因凡經公開發表之研發成果，如擬申請專利，須於公開發表後6個月內完成，前述成果如是以論文方式公開發表，將無法取得大陸與歐盟等國之專利。

五、執行方式

以下條列說明合作研究內容：

- 合作計畫執行單位需定期舉行計畫控管會議
- 合作計畫執行單位需按時繳交表外資料風險因子初步與進階篩選報告各1份
- 合作計畫執行單位需按時繳交期末報告1份
- 本單位要求合作計畫執行單位完成1件專利構想交付
- 合作計畫執行單位提供教育訓練2小時，訓練內容為提供萃取因子原由說明

六、計畫期程及預估計畫總經費

計畫執行區間：107年01月01日至107年12月15日

總經費：600,000元

七、驗收標準

完成表外資料風險因子初步與進階篩選報告各1份、提供1件專利構想、提供教育訓練2小時、完成期末報告1份

八、技術能力需求

- 具備對於文字探勘應用於風險因子擷取與輪廓描繪技術之研究人員
- 對於大數據分析應用有所經驗之研究人員
- 對於信評分析與模型驗證有經驗之研究人員
- 對於財務與會計資料分析有經驗，並了解其限制之研究人員

參考資料

1. Al-Najjar, D., & Al-Najjar, B. (2014). Developing a multi stage predicting system for corporate credit rating in emerging markets. *Journal of Enterprise Information Management*, 27(4), 475-487. Retrieved from <https://search.proquest.com/docview/1660759527?accountid=13952>
2. Chen, Yehning, Rachel J. Huang, John Tsai, and Larry Y. Tzeng, "Soft Information and Small Business Lending," *Journal of Financial Services Research*, February 2015, Volume 47, Issue 1, pp 115-133
3. Crosman, P. (2017). Can AI improve credit scores? *ISO & Agent*, 6(2), 42. Retrieved from <https://search.proquest.com/docview/1883350000?accountid=13952>
4. The Economic Times (Online) Artificial intelligence revolution in lending: Hype or reality? [technology]. (2016, Nov 08). Retrieved from <https://search.proquest.com/docview/1836754289?accountid=13952>
5. Khashei, M., & Mirahmadi, A. (2015). A soft intelligent risk evaluation model for credit scoring classification. *International Journal of Financial Studies*, 3(3), 411-422. <http://dx.doi.org/10.3390/ijfs3030411> Retrieved from <https://search.proquest.com/docview/1721904082?accountid=13952>

附件 1：契約書格式

1-1：計畫書格式

1-2：經費動支報表

1-3：成果報告撰寫須知

1-4：報告格式

1-5：論文格式

1-6：保密聲明書

1-7：委託匯款同意書