全方位客戶關係視圖(comprehensive customer relationship view)

~~行內單一平台查閱客戶360角度關係圖~~

~~大數據結合NLP分析 AI半監督式客戶關係尋找~~

Please give a brief description of the technology project(250) 你們做了什麼

本專案以信用卡交易為起點，利用知識圖譜儲存信用卡交易資料，納入了客戶以及特店作為知識實體並建立連結；引入外部非結構化數據，配合NLP技術進行知識抽取、共指消解；知識推理的過程並導入AI半監督式的機器學習模型，增進可用性；最後透過內部可視化平台讓內部客戶使用。

據此我們掌握了過去不曾瞭解的客戶喜好、活動範圍、生活習慣；同時掌握卡戶喜愛的特店，能更針對性地去與需要的特店簽約；最後這些特店清單也提供企金作為客戶清單，可以拜訪目前正在擴張的中小企業戶。

Please list the key dates for the stages of the project i.e. when it was started, when implementation began, when it went live: 專案規劃是?

|  |  |
| --- | --- |
| Date | Stage |
| 2018/07~2018/09 | 知識實體建立 |
| 2018/09~2018/10 | 外部特店資訊收集 |
| 2018/10~2018/11 | 知識融合 |
| 2018/11~2019/01 | 建立知識實體連結 |
| 2019/02~2019/03 | 建立半監督式的關係挖掘模型 |
| 2019/03~2019/04 | 評估連結質量並建立回饋系統 |
| 2019/04~2019/05 | 投入行內應用、開放接口 |

**How was your project innovative?** What was the thinking and rationale behind the project? How does it differ from other products or solutions that already exist? (max 500 words)

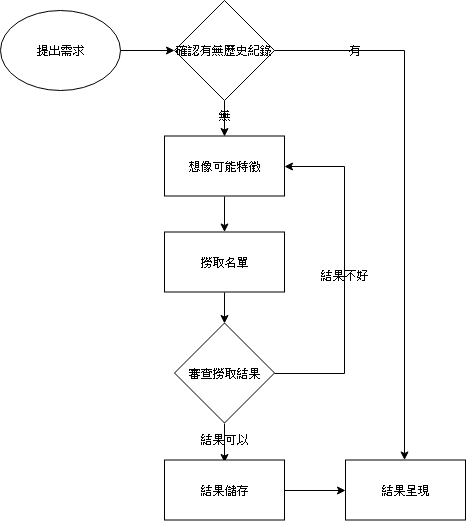
哪裡創新?為什麼要創新?

銀行擁有大量的客戶資料，但目前大多僅僅是作為一種紀錄而已。隨著對資料採礦與分析的需求越來越多，我們認為過去by任務、純人工處理、每個任務獨立、缺乏統籌、缺乏紀錄、結果呈現不友善的情況需要改變。市場上有很多呈現客戶關係的方案，但僅僅是呈現面的改善，我們認為儲存、分析、管理方面也有很多可以改善的地方，因此我們的專案創新了以下幾點:

1. 使用NLP技術進行知識抽取、知識融合等，減少大量人工標籤工程
2. 採用知識圖譜的結構儲存資料間的交互關係，而非過去在需要的時候從DB join資料重現，方便儲存分析結果
3. 採用知識圖譜的流程，建立、清洗、連結、加工、推理我們所擁有的資料，除了保證資料的質量外，也提供高可擴充性的架構
4. 將純人工寫規則撈取資料改成人機交互的AI方式，半監督的機器學習模型，以保證足夠準確性的同時有良好的彈性。同時也讓機器學習人類業務經驗
5. 建立內部可視化平台，將平台作為分析單位交付/業務單位查看的統一平台，在增加圖形視覺化能力的同時，減少兩方的溝通成本
6. 統一的資料分析結果庫、使用簡便的查詢平台、單一視圖即可查詢所有結果，使得管理與整合成本大幅降低。

**How was the project useful?** What goals did the project have? What need does it fulfil? What problem was it trying to solve? How was the project successful in meeting its goals? How has the project benefitted its users i.e. the end-user, your organisation, customers or client? (max 500 words) 解決了什麼問題?做的結果怎麼樣?有用嗎?

如前段所述，這個專案是為了改善過去方法的問題，一般的工作流程如下:



而我們改善的部分如下

|  |  |  |  |
| --- | --- | --- | --- |
| stage | past | now | Benefit |
| 確認有無歷史紀錄 | 向同事詢問 | 紀錄在知識圖譜 | 去除遺忘的風險 |
| 撈取名單 | 人工撰寫規則 | 給AI a few例子 | 省去編寫規則的時間，並導入機器學習算法增進效果 |
| 結果儲存 | code | Edge on knowledge graph | 紀錄可以圖形化呈現，而且有單一查詢入口 |
| 結果呈現 | 表格 | 網路圖 | 更直觀可見 |

在我們試行一段時間後，統計的資料顯示:統一的資料結構減少了30%的資料整合時間；而AI半監督式方法，減少了50%的程式開發時間，20% more business opportunities from missing data sets。對業務單位而言，有了一個可視化的內部平台方便查看，簡單的客戶分析因為不用編寫程式，甚至自己完成。

而對組織整體而言，良好的資料結構與系統規劃能減少資料整合的難度，更方便併入複雜多樣的資料來源；而內部可視化平台作為分析單位與業務單位的橋樑，減少兩者的溝通成本；這樣的結構可以很容易的結合更多來源資料。

**How was the project transformational?** Does this project have the potential to transform the market? Is it likely to be replicated? How could the solution be applied to a wider market? (max 500 words) 你認為這件事可以改變市場嗎?

我們相信這件事不只會改變金融業，也會改變整個市場，因為這是每個data driven公司會經歷的過程。

這個專案的範圍雖然只有信用卡資料，但其實是我們對未來趨勢的對策，我們看到三個未來的問題，1.更多更雜的資料2.指數成長的分析結果3.缺少已標註的數據。

1.我們並不是指資料大小的問題，那已經有許多不錯的解決方案；而是指越來越多的資料源。不同來源的資料如何整合?如何展示他們的關係?如何表示它們的交互關係?我們認為透過語意網路，也就是知識圖譜的本質，是一個不錯解決方案，在知識融合階段，能夠有效的整合不同來源資料，而在知識融合、知識加工階段，則能一探資料交互作用所帶來的可能性，且為搜尋資料提供一個統一的簡便入口。

2.資料分析帶來很大的價值，相應而來的是指數成長的分析結果，我們可以預期有很多好的分析結果，但卻沒看到有太多人關注如何儲存這些結果?如何展現這些結果?使用者有多少時間看完這些結果?而我們認為利用節點、邊、權重以及相應的說明文字，用一張關係圖來呈現這些結果，雖然簡化了不少分析結果，但不失為一種對使用者友善的方式。使用者無需關注背後的運算，開發者遵循單一且簡單的呈現方式，對兩方的溝通能起到良好結果。

3.AI演算法的發展雖然一日千里，但事實上存在著缺少標註資料的問題。大部分的AI都是採用監督式算法，這無疑有較好的準確性，但卻必須要有相應的標註資料來訓練，而現實中最缺少的就是已標註的資料集。所以半監督式模型也是現在的AI發展重點，我們建立一套可以半監督學習的模型，同時提供友善的人機交互介面，解決已標註資料缺少的問題。而隨著時間演進，這些半學習出來的資料，都會成為行內最珍貴的已標註資料，為未來Bank 4.0貢獻。