推薦系統API雛形 ─ 技術簡介與展示步驟(Ver. 1.0)

**作者：**陳韋辛 Frank Chen

**題目：**

「以人工神經網路演算法和回歸分析實作個性化推薦系統之API雛形」

**實作動機：**

商品評分與推薦系統日趨普遍，本人也對具有實用性的演算法相當感興趣，因此決定研讀並實作一個即便不懂該演算法也容易使用的推薦系統API，期望能協助拓展公司未來接案的種類，如購物網站或是旅遊網站等具有推薦功能之專案。

**功能：**

預測所有顧客對未知電影或產品之評價，以便未來對該顧客進行商品推薦，此外，此套件可以動態新增或更新顧客之評價或新增新產品，而不會干擾到原先已處理過的數據導致重新計算，相當符合實際使用情形。

**程式碼：**

* <https://github.com/WeiHsinChen/RecommendationAPI>

**使用環境：**

* 以Python語言實作
* 使用SQLite 資料庫
* 須下載SQLAlchemy便於資料庫存取以及numpy套件以便矩陣運算

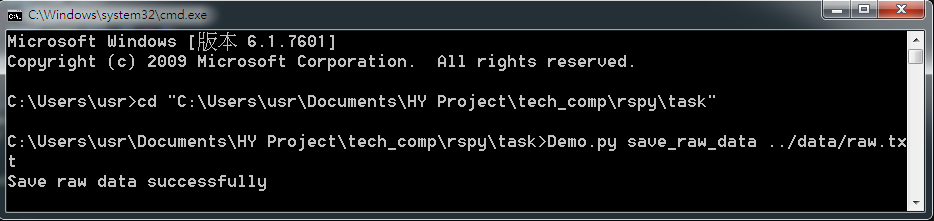
**套件優點：**

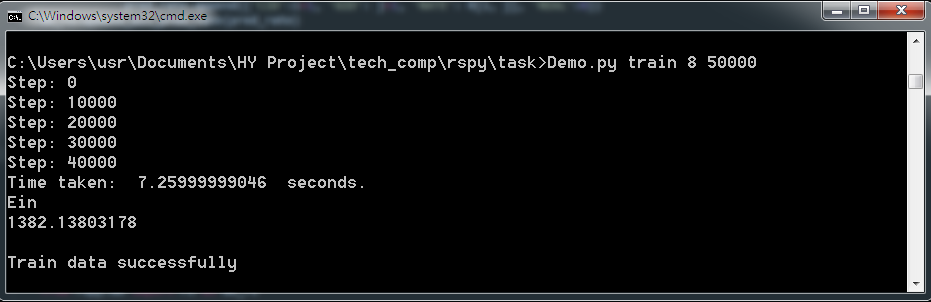
使用方便簡單，足以滿足最基本推薦系統的需求。輕量化設計，相較於其他開放源程式碼的推薦系統套件(e.g., Drupal)更加簡單易懂，且支援動態更新資料。此外，此套件雛型能夠輕易轉型成其他推薦形式，如將商品評分紀錄改成商品點擊紀錄等，便能計算顧客對該商品的未來點擊數以預測其喜好程度。

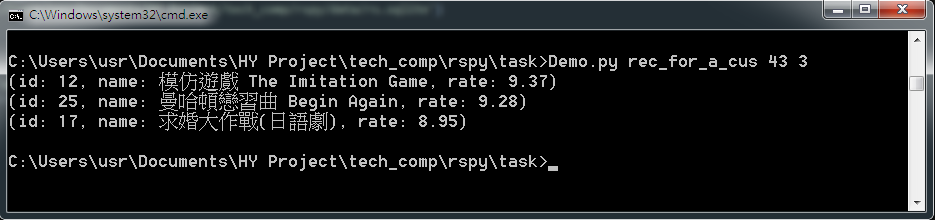
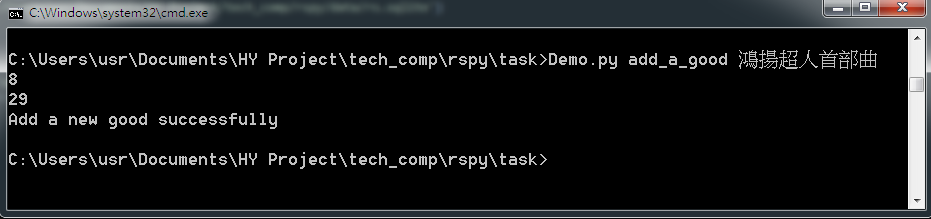
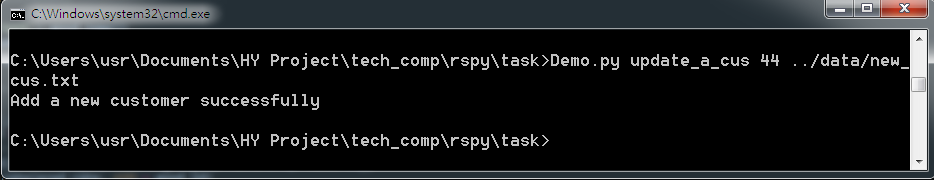
**待改善處：**

目前的輸入參數之格式完全是便於現場示範，未來實際參數僅須改成JSON格式即可。此外，推薦準度以及運算速度有待最佳化，並且還有許多可以擴充之功能，如數據過大時使用Hadoop等Map reduce技巧降低計算之需求或是denoising提升準確性。

**使用方式：**

1. 第一次使用時，必須將全體用戶之評分紀錄文字檔(.txt)匯入資料庫，以便存取與分析，因此使用者須將評分文字檔之「相對路徑」為輸入，呼叫Demo.py中的save\_raw\_data方程式，輸入格式如Git/data/raw\_data.txt。
2. 第一次使用時，必須計算所有用戶之評分資料，以產生矩陣 V 和矩陣 W以便預測全體用戶之行為，因此使用者必須呼叫Demo.py中的train方程式，並輸入參數「神經元數量」以及「計算迴圈數」，若無輸入參數，則預設神經元數量為10，計算迴圈數為10000000。



1. 推薦顧客未評價之商品時，使用者須要呼叫Demo.py中的rec\_for\_a\_cus方程式，並輸入參數「顧客ID」以及「欲推薦數量」，若無輸入「欲推薦數量」，則欲設為3，返回值為該數量之推薦商品字串，實際使用時轉為JSON即可‧
2. 新增商品時，須呼叫Demo.py中的add\_a\_good方程式，並輸入參數「商品名稱」即可。
3. 新增顧客時，須呼叫Demo.py中的add\_a\_cus方程式，並將該用戶之評分文字檔(.txt)作為輸入參數即可，格式如Git/data/new\_cus.txt。
4. 更新顧客評分時，須呼叫Demo.py中的update\_a\_cus方程式，並將該用戶之評分文字檔(.txt)作為輸入參數即可，格式如Git/data/new\_cus.txt。