

# 推薦系統 API 技術簡介 (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：

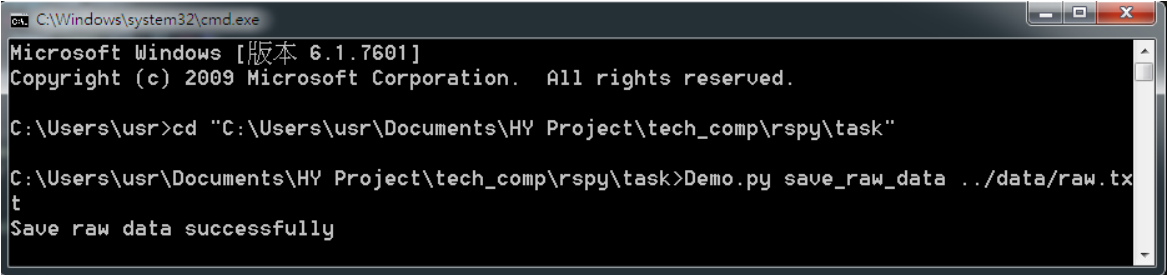
使用方便簡單，足以滿足最基本推薦系統的需求。輕量化設計，相較於其他開放源程式碼的推薦系統套件(e.g., Drupal)更加簡單易懂，且支援動態更新資料。此外，此套件雛型能夠輕易轉型成其他推薦形式，如將商品評分紀錄改成商品點擊紀錄等，便能計算顧客對該商品的未來點擊數以預測其喜好程度。

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。



```
C:\Windows\system32\cmd.exe
Microsoft Windows [版本 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

C:\Users\usr>cd "C:\Users\usr\Documents\HY Project\tech_comp\rspy\task"
C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>Demo.py save_raw_data ../data/raw_data.txt
Save raw data successfully
```

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。

```
C:\Windows\system32\cmd.exe

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>Demo.py train 8 50000
Step: 0
Step: 10000
Step: 20000
Step: 30000
Step: 40000
Time taken: 7.25999999046 seconds.
Ein
1382.13803178

Train data successfully
```

3. 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。

```
C:\Windows\system32\cmd.exe

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>Demo.py rec_for_a_cus 43 3
(id: 12, name: 模仿遊戲 The Imitation Game, rate: 9.37)
(id: 25, name: 曼哈頓戀習曲 Begin Again, rate: 9.28)
(id: 17, name: 求婚大作戰(日語劇), rate: 8.95)

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>_
```

4. 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。

```
C:\Windows\system32\cmd.exe

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>Demo.py add_a_good 鴻揚超人首部曲
8
29
Add a new good successfully

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>
```

5. 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。

```
C:\Windows\system32\cmd.exe

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>Demo.py add_a_cus ../data/new_cus.txt
Add a new customer successfully

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>
```

6. 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。

```
C:\Windows\system32\cmd.exe

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>Demo.py update_a_cus 44 ../data/new_cus.txt
Add a new customer successfully

C:\Users\usr\Documents\HY Project\tech_comp\rspy\task>
```