

Task 2 : Debiasing

進度報告

組別： 第六組

組員： M10815113 賴啓明

M10815103 陳子揚

B 10532037 王韋翔

B 10532030 楊博惟

指導老師： 李漢銘 教授

目錄

Part 1

開發環境

問題分析&方法

Part 2

Part 3

目前進度

未來目標

Part 4

開發環境

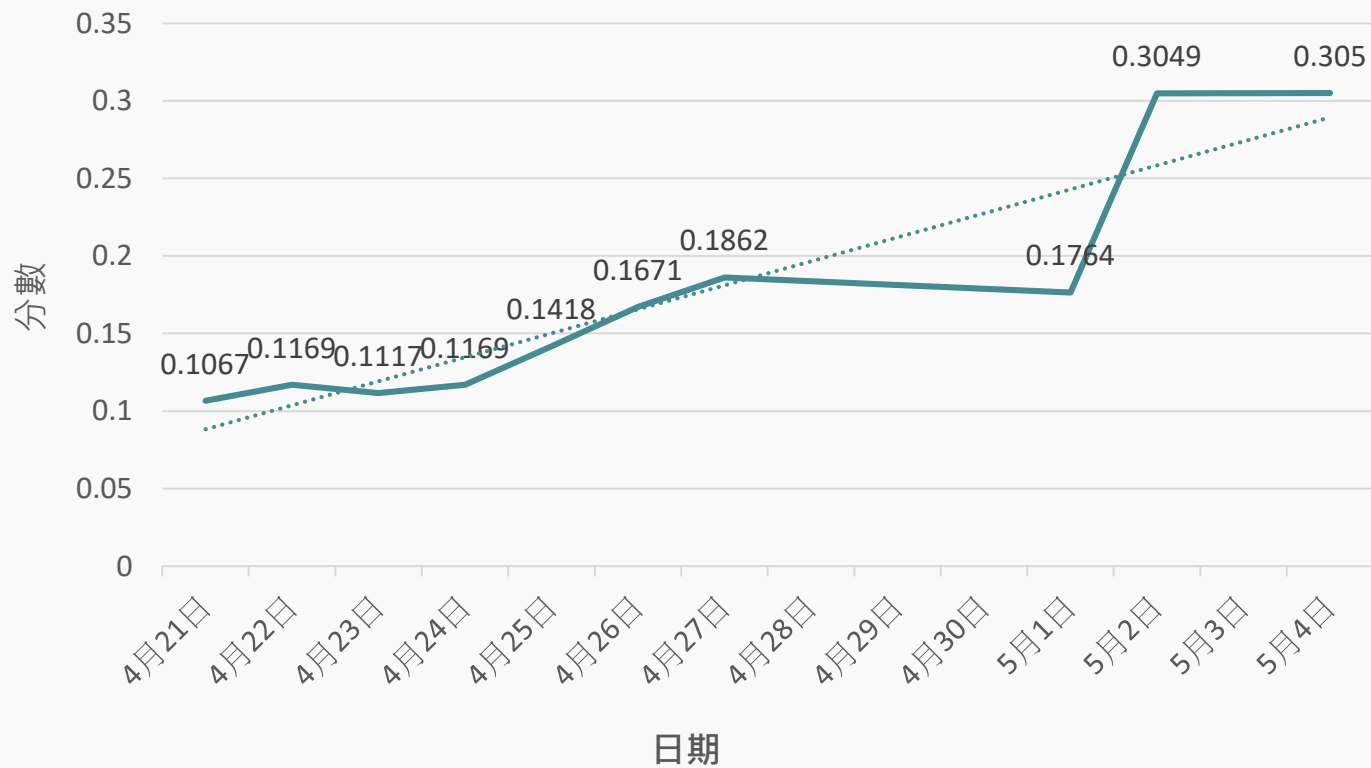
- Anaconda
- +
- Jupyter Notebook

- 根據用戶過往的點擊紀錄，推薦較少曝光的商品
- 分析資料後，使用者不重覆點擊相同商品
- 資料集的部分商品沒有對應的**Vector**表示

- 對每個使用者點擊的商品，產生item-item similarity matrix M
- $M[i][j] = \frac{M[i][j]}{C_i * C_j}$, C_i is item i click counts
- 利用使用者點擊的歷史和 M ，推薦出前50名商品

目前進度

成績



目前進度

97	waterrrmelon	第二幼儿园	0.309	0.684	0.309	0.391	0
98	8 64	Others-1	0.308	0.662	0.308	0.441	0
99	Waterrrmelon	华中科技大学	0.308	0.660	0.308	0.387	0
100	8 大忽悠工程院	大忽悠工程院	0.307	0.661	0.307	0.433	0

Date: 2020/05/06 20:20

目前進度

分數



Track 1: 98 /0.308/0.662/0.308/0.441/0.202

Track 1

目前排名



○ **Date:** 2020-05-05 20:23:54 **Ranking:** None

score: 0.3082

hitrate_50_full: 0.6616

ndcg_50_full: 0.3082

hitrate_50_half: 0.4414

ndcg_50_half: 0.2016

Date: 2020/05/06 20:20

未 來 目 標

- 嘗試使用DSSM等model提升recall
- 考慮時間因素
- 對user或item做word embedding來計算相似性