

# Hands-on 2 - Recommendation System in SVD

---

王嘉澤 (Chiatse Wang)

資料科學計算, Spring, 2018

# Outline

- Recommendation system
- Homework 1
- SVD Demo

# Origin

- Netflix prize

A year into the competition, the Korbell team won the first Progress Prize with an 8.43% improvement. They reported more than 2000 hours of work in order to come up with the final combination of 107 algorithms that gave them this prize. And, they gave us the source code. We looked at the two underlying algorithms with the best performance in the ensemble: *Matrix Factorization* (which the community generally called SVD, *Singular Value Decomposition*) and *Restricted Boltzmann Machines* (RBM). SVD by itself provided a 0.8914 RMSE, while RBM alone provided a competitive but slightly worse 0.8990 RMSE. A linear blend of these two reduced the error to 0.88. To put these algorithms to use, we had to work to overcome some limitations, for instance that they were built to handle 100 million ratings, instead of the more than 5 billion that we have, and that they were not built to adapt as members added more ratings. But once we overcame those challenges, we put the two algorithms into production, where they are still used as part of our recommendation engine.


How to recommend?

## Most Popular Items




# Recommendation System

- Content based algorithms
- Collaborative filtering algorithms

My Courses | Course Catalog

## Recommendations for you.


We combed our catalog and found courses and Specializations that we think match your interests. Browse our recommendations below, and start learning something new today!



### Applied AI with DeepLearning, IBM Watson IoT Data Science Certificate

IBM


Starts Mar 12, 2018



### Robotics

University of Pennsylvania | 6 Course Specialization


Starts Mar 12, 2018



### IBM Blockchain Foundation for Developers

IBM


Starts Mar 12, 2018



### Machine Learning

University of Washington | 4 Course Specialization

Starts Mar 12, 2018



### Genomic Data Science

Johns Hopkins University | 8 Course Specialization

Starts Mar 12, 2018

# Content Based

開全站分類

博客來 售票網 企業採購 福利平台 海外專館

0 結帳 登入 加入會員 購物金 會員專區 電子書權

博客來 > 中文書 > 心理勵志 > 勵志故事/散文 > 真實人生故事 > 商品介紹



主題活動

試閱

### 廢文

可購買版本(2):

電子書 優惠價224元	平裝 優惠價288元
----------------	---------------

作者: 蔡嘉佳 [追蹤作者](#)

出版社: 時報出版 [訂閱出版社新書快訊](#)

出版日期: 2017/09/19

語言: 繁體中文

定價: 320元

優惠價: 9折 288元

本商品單次購買10本85折 272元

抵用購物金最多再省\$66 詳情

運送方式:

- 可配送點: 台灣、蘭嶼、綠島、澎湖、金門、馬祖、全球
- 可取貨點: 台灣、蘭嶼、綠島、澎湖、金門、馬祖、香港、澳門、新加坡

中午前訂 可明天 拿

庫存 > 10

放入購物車

直接結帳

加入下次再買清單

買就送 (贈品)『成為怪胎之必要』布條  
[> more](#)  
剩餘數量: 314

我要寫評鑑

分享

讚 106

買了此商品的人，也買了...



親愛的我Oh! Dear Me: 250天憂鬱症紀實

原價: 350  
優惠價: 9折315元



我只擔心你會不會一直下到明天早上

原價: 380  
優惠價: 79折300元



與自己和好如初(精美雙書封設計)

原價: 320  
優惠價: 79折253元



那些最靠近你的

原價: 350  
優惠價: 79折277元



下雨的人(隨書附贈下雨的人空白手抄本)

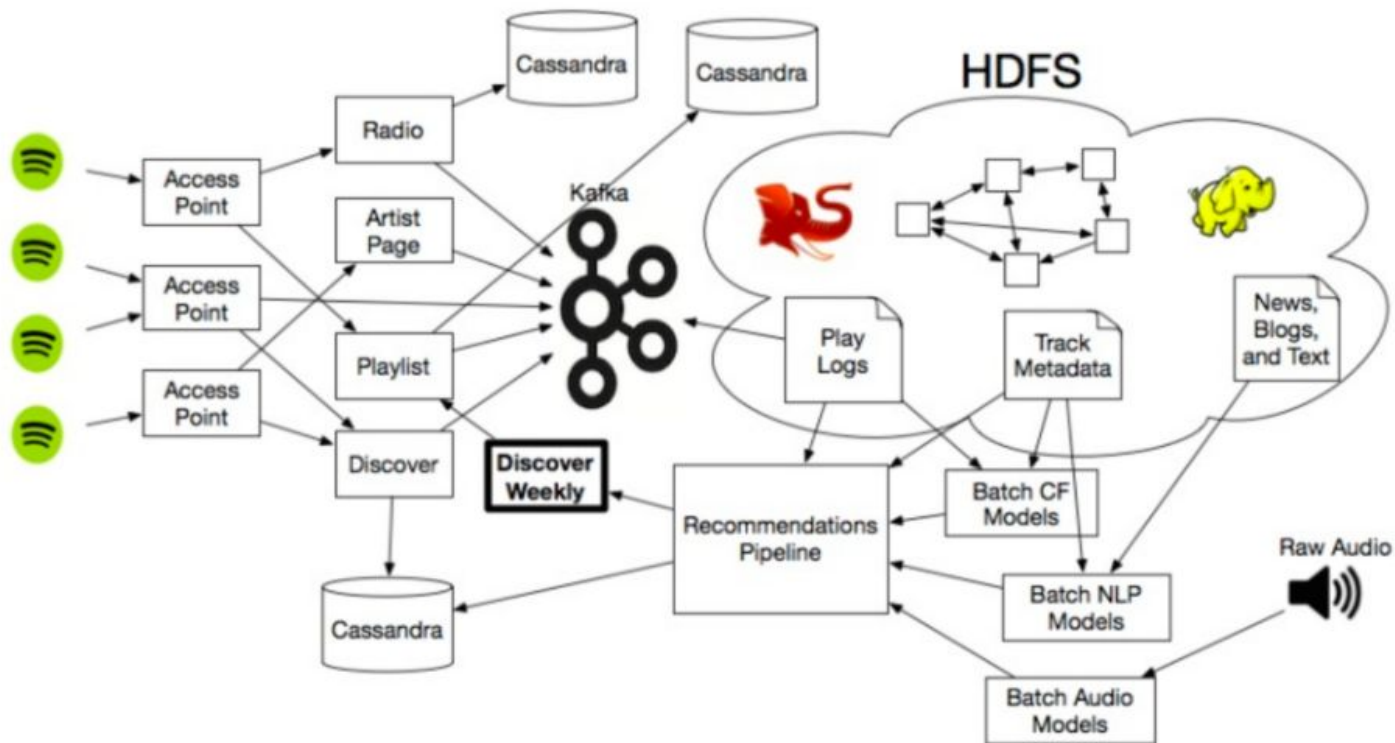
原價: 350  
優惠價: 79折277元

# Collaborative Filtering

- Idea:  
A likes item 1, 2, 3 and B likes item 2, 3, 4  
If A and B have ***similar interest***,  
then A should like item 4 and B should like item 1
- User-user collaborative filtering
- Market basket analysis (Beer and diaper)



# Spotify



Ref 1: <https://hackernoon.com/spotifys-discover-weekly-how-machine-learning-finds-your-new-music-19a41ab76efe>

Ref 2: <https://buzzorange.com/techorange/2018/01/24/how-ai-find-spotify-weekly-discovery/>

# Homework 1

# Content

- Handwriting (3 problems)
- Programming

# Dataset

- MovieLens (<https://grouplens.org/datasets/movielens/>)
- MovieLens 1M Dataset
  - 1 million ratings
  - 6000 users
  - 4000 movies
  - Released 02/2003

# Programming

- In class, we have show a brief workflow for movie recommendation. However, we didn't evaluate the performance. Thus, we should split the data into training set and testing set (You can add a validation set).
- You should
  - Split the data to two sets.
  - Find the best rank  $k$  (latent).
  - Calculate the error by RMSE (Root Mean Square Error) or SSE (Error Sum of Square).
  - Answer the three questions in the notebook.

# SVD Demo

# Singular Value Decomposition

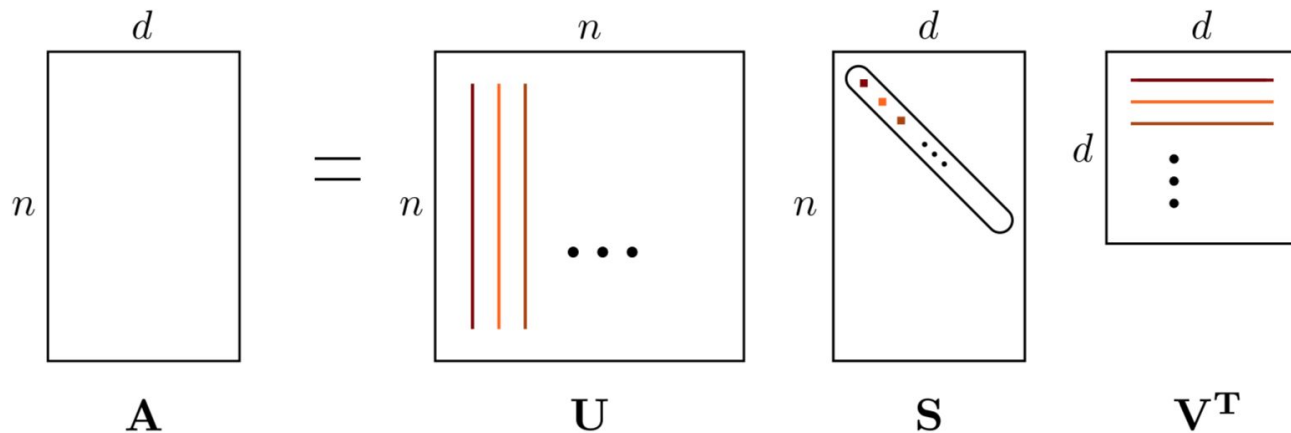


Figure 2: The singular value decomposition (SVD). Each singular value in  $\mathbf{S}$  has an associated left singular vector in  $\mathbf{U}$ , and right singular vector in  $\mathbf{V}$ .

# Rating Matrix

$$R = \begin{pmatrix} 1 & ? & 2 & ? & ? \\ ? & ? & ? & ? & 4 \\ 2 & ? & 4 & 5 & ? \\ ? & ? & 3 & ? & ? \\ ? & 1 & ? & 3 & ? \\ 5 & ? & ? & ? & 2 \end{pmatrix} \begin{matrix} \text{Alice} \\ \text{Bob} \\ \text{Charlie} \\ \text{Daniel} \\ \text{Eric} \\ \text{Frank} \end{matrix}$$