Websites in Data Driven Generation

Navigation, Searching, Recommendation

其實 ...

我主要是想分享推薦系統建制心得 ...

聽說 ...

推薦系統已經變成網站必備物品了？

在座有人 ...

正在幫自己的網站建制推薦系統的嗎？

照慣例 ...

先來玩個小遊戲 (現場統計) 吧 ...

請大家

兩兩一組

請推薦你的同伴

一本書, 一首音樂, 一種食物, 一種飲料
(給大家 2 min 完成四項推薦)

微信号: 4ilwh

大家發現了嗎？

其實, 每個人都是推薦系統 .....

但是 ...

聽完夥伴介紹，會接受推薦的 ... ？

剛剛 ...

是推薦「自己」最喜歡的？

剛剛 ...

是推薦「最熱門的」？

剛剛 ...

是推薦「最新的」？

剛剛 ...

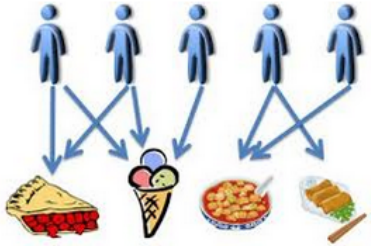
推薦的東西自己完全沒用過的 ？

回到主題

什模是一個推薦系統呢？

基本上，就是一個媒合系統 ...

基本上，就是一個媒合系統 ...



一個系統

有 Users

有 Items

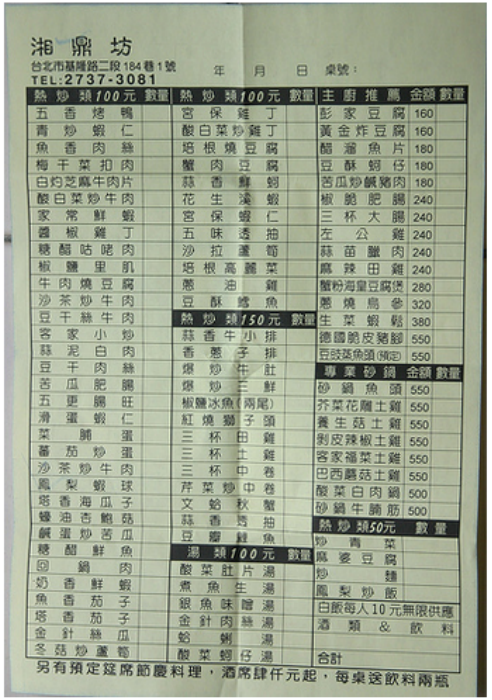
還有 Ranking

而媒合意味著 ...

User 找 Item !

而 Item 也在找 User !

For examples: Item-Listing System



湘鼎坊

台北市基隆路二段184巷1號

TEL:2737-3081

年　月　日　桌號：

熱炒類100元	數量
五香烤鴨	
青炒蝦仁	
魚香肉絲	
梅干菜扣肉	
白灼芝麻牛肉片	
酸白菜炒牛肉	
家常鮮蝦	
醬椒雞丁	
糖醋咕咾肉	
椒鹽里肌	
牛肉燒豆腐	
沙茶炒牛肉	
豆干絲牛肉	
客家小炒	
蒜泥白肉	
豆干肉絲	
苦瓜肥腸	
五更腸旺	
滑蛋蝦仁	
菜脯蛋	
蕃茄炒蛋	
沙茶炒牛肉	
鳳梨蝦球	
塔香海瓜子	
蠔油杏鮑菇	
鹹蛋炒苦瓜	
糖醋鮮魚	
回鍋肉	
奶香鮮蝦	
魚香茄子	
塔香茄子	
金針絲瓜	
冬菇炒蘆筍	

熱炒類100元	數量
宮保雞丁	
酸白菜炒雞丁	
塔根燒豆腐	
蟹肉豆腐	
蒜香鮮蚵	
花生燒蝦	
宮保蝦仁	
五味透抽	
沙拉蘆筍	
塔根高麗菜	
蔥油雞	
豆酥鱈魚	

熱炒類150元	數量
蒜香牛小排	
香蔥子排	
爆炒牛肚	
爆炒三鮮	
椒鹽冰魚(兩尾)	
紅燒獅子頭	
三杯田雞	
三杯土雞	
三杯中卷	
芹菜炒中卷	
文蛤秋蟹	
蒜香透抽	
豆瓣鯉魚	

湯類100元	數量
酸菜肚片湯	
煮魚生湯	
銀魚味噌湯	
金針肉絲湯	
蛤蜊湯	
酸菜蚵仔湯	

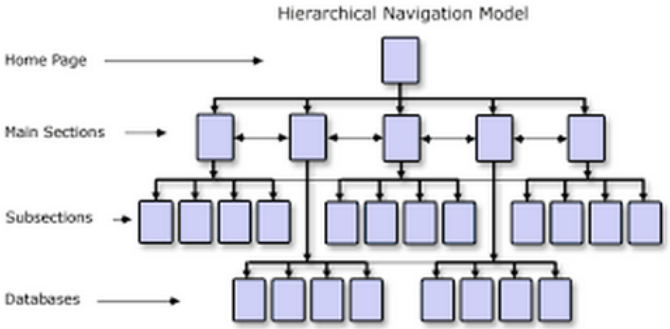
主廚推薦	金額	數量
彭家豆腐	160	
黃金炸豆腐	160	
醋溜魚片	180	
豆酥蚵仔	180	
苦瓜炒鹹豬肉	180	
椒脆肥腸	240	
三杯大腸	240	
左公雞	240	
蒜苗臘肉	240	
麻辣田雞	240	
蟹粉海皇豆腐煲	280	
蔥燒鳥參	320	
生菜蝦鬆	380	
德國脆皮豬腳	550	
豆豉蒸魚頭(預定)	550	

專業砂鍋	金額	數量
砂鍋魚頭	550	
芥菜花雕土雞	550	
養生菇土雞	550	
剝皮辣椒土雞	550	
客家福菜土雞	550	
巴西蘑菇土雞	550	
酸菜白肉鍋	500	
砂鍋牛腩筋	500	

熱炒類50元	數量
炒青菜	
麻婆豆腐	
炒麵	
鳳梨炒飯	
白飯每人10元無限供應	
酒類&飲料	
合計	

另有預定筵席節慶料理，酒席肆仟元起，每桌送飲料兩瓶

有沒有很像 ...


Hierarchical Navigation Model
Home Page
Main Sections
Subsections
Databases

也有一些會走精美奪目路線 ...

Top class steak
Chef Fricksteak

許多類似的事情，也發生在網站進化史上！

因為，網站本身就是一種用來「媒合」Content的環境！

話說 … 在那個沒有「推薦系統」這個口號的年代 …

What's New | Check Email | YAHOO! | Personalize | Help

Yahoo! Mail
free email for life

Know when friends are online!
Click to download Yahoo! Messenger

Yahoo! Auctions
coins, cards, stamps

Search | advanced search

Shopping - **Auctions** - Yellow Pages - People Search - Maps - Travel - Classifieds - Personals - Games - Chat - **Clubs**
Mail - Calendar - Messenger - **Companion** - My Yahoo! - News - Sports - Weather - TV - Stock Quotes - more...

Yahoo! Shopping - Thousands of stores. Millions of products.

Departments		Stores	Products
· Apparel	· Flowers	· Sports Authority	· Digital cameras
· Bath/Beauty	· Food/Drink	· Gap	· Pokemon
· Computers	· Music	· Eddie Bauer	· MP3 players
· Electronics	· Video/DVD	· Macy's	· DVD players

Arts & Humanities
Literature, Photography...

News & Media
Full Coverage, Newspapers, TV...

Business & Economy
Companies, Finance, Jobs...

Recreation & Sports
Sports, Travel, Autos, Outdoors...

Computers & Internet
Internet, WWW, Software, Games...

Reference
Libraries, Dictionaries, Quotations...

Education
College and University, K-12...

Regional
Countries, Regions, US States...

Entertainment
Cool Links, Movies, Humor, Music...

Science
Animals, Astronomy, Engineering...

Government
Elections, Military, Law, Taxes...

Social Science
Archaeology, Economics, Languages...

Health
Medicine, Diseases, Drugs, Fitness...

Society & Culture
People, Environment, Religion...

In the News
- Scores killed in Nigerian riots
- Austria's Haider resigns as party leader
- Floods trap thousands in Mozambique

more...

Marketplace
- Y! Auctions - Peanuts, Pokemon, computers
- Free 56K Internet Access
- Yahoo! Bill Pay - free 3-month trial

more...

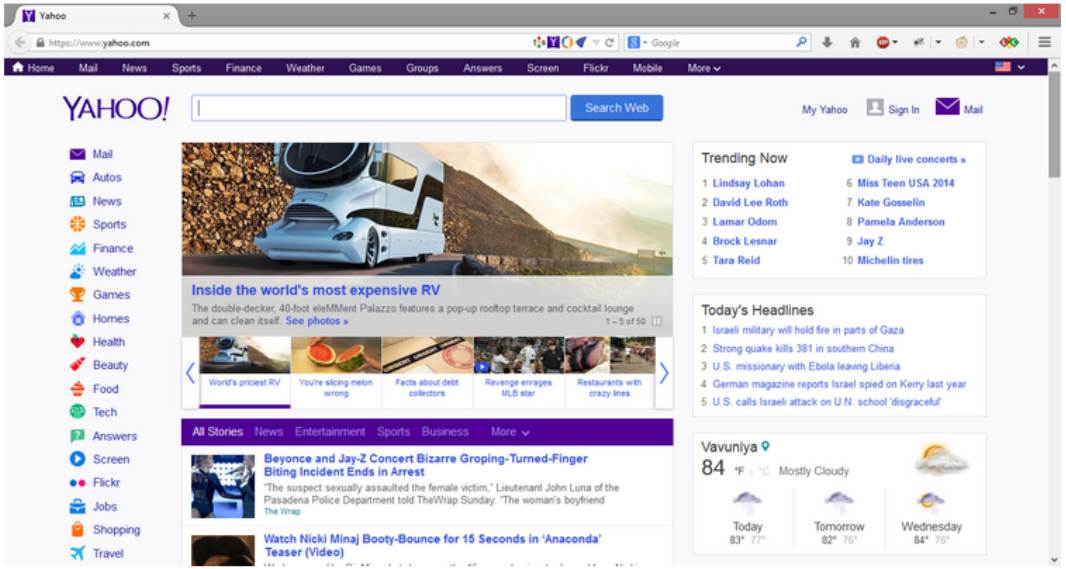
Inside Yahoo!
- Yahoo! GeoCities - build your free home page
- Play free Fantasy Soccer
- Yahoo! Clubs - create your own
- Y! Greetings - free greeting cards

more...

一路走來 ...

始終如一

YAHOO!
Search:
Web Search
Sneaky credit killers
Think you're a word whiz
Popular Actor Photo Galleries


Yahoo
https://www.yahoo.com
Google
Home Mail News Sports Finance Weather Games Groups Answers Screen Flickr Mobile More
YAHOO!
Search Web
My Yahoo
Sign In
Mail
Mail
Autos
News
Sports
Finance
Weather
Games
Homes
Health
Beauty
Food
Tech
Answers
Screen
Flickr
Jobs
Shopping
Travel
Inside the world's most expensive RV
The double-decker, 40-foot eleMMent Palazzo features a pop-up rooftop terrace and cocktail lounge and can clean itself. See photos »
1 – 5 of 50
World's priciest RV
You're slicing melon wrong
Facts about debt collectors
Revenge enrages MLB star
Restaurants with crazy lines
All Stories News Entertainment Sports Business More
Beyonce and Jay-Z Concert Bizarre Groping-Turned-Finger Biting Incident Ends in Arrest
"The suspect sexually assaulted the female victim," Lieutenant John Luna of the Pasadena Police Department told TheWrap Sunday. "The woman's boyfriend
The Wrap
Watch Nicki Minaj Booty-Bounce for 15 Seconds in 'Anaconda' Teaser (Video)
Trending Now
Daily live concerts »
1 Lindsay Lohan
2 David Lee Roth
3 Lamar Odom
4 Brock Lesnar
5 Tara Reid
6 Miss Teen USA 2014
7 Kate Gosselin
8 Pamela Anderson
9 Jay Z
10 Michelin tires
Today's Headlines
1 Israeli military will hold fire in parts of Gaza
2 Strong quake kills 381 in southern China
3 U.S. missionary with Ebola leaving Liberia
4 German magazine reports Israel spied on Kerry last year
5 U.S. calls Israeli attack on U.N. school 'disgraceful'
Vavuniya
84 °F | °C Mostly Cloudy
Today 83° 77°
Tomorrow 82° 76°
Wednesday 84° 76°

象徵著 Hierarchical Navigator Style (Item Listing) 的網站世代

Google!

Search the web using Google!

10 results | Google Search | I'm feeling lucky

Index contains ~25 million pages (soon to be much bigger)

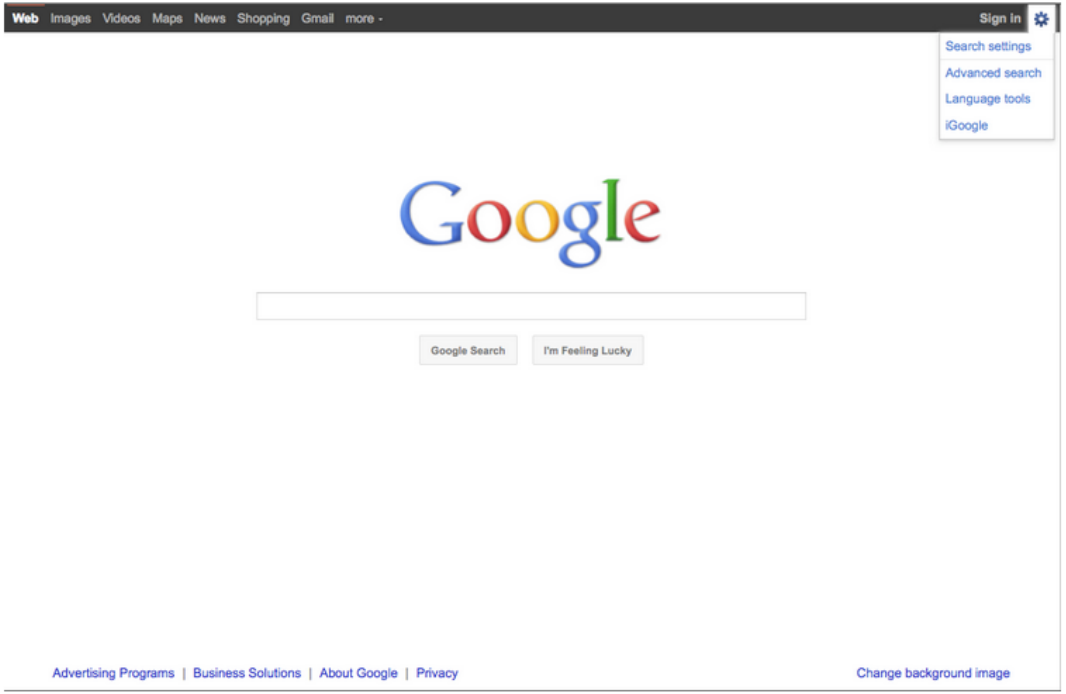
About Google!

Stanford Search Linux Search

Get Google! updates monthly!

your e-mail | Subscribe | Archive

Copyright ©1997-8 Stanford University


Web
Images
Videos
Maps
News
Shopping
Gmail
more
Sign in
Search settings
Advanced search
Language tools
iGoogle
Google
Google Search
I'm Feeling Lucky
Advertising Programs | Business Solutions | About Google | Privacy
Change background image

象徵著 Searching & Tagging Navigator Style 的網站世代

Type to search for people, places and things

Home

"I don't wait in lines, that's why I don't understand Ted's class." Could be the quote of the quarter

Like · Comment · 9 hours ago near Seattle, WA ·

12 people like this.

Write a comment...

The Great Gatsby

"There I was, way off my ambitions, getting deeper in love every minute."

Like · Comment · Share · 47,078 642 4,478 · Sponsored

It doesn't matter if the glass is half empty or half full. Clearly there's room for more beer!!

Like · Comment · Share · 6 hours ago via Facebook for BlackBerry 10 ·

5 people like this.

Write a comment...

Your Ads Create an Ad

Engage 121 Today Monthly

12	7	8
Ads	Campaigns	Page Likes

Joshua

a new photo.

the New York Post

exhibit he visited at

Coll...

Adkins are

May 8

others invited you

Sponsored Create Ad

CableTV.com asked: What is the best vampire TV show or movie of all time? (Fire off in the comments and tell is why)

The Vampire Diaries

True Blood

98 More...

Like This Page

Knicks on MSG Network

The Knicks look to finish the series TONIGHT! Tipoff is at 7PM on MSG Network.

Like · 182,002 people like MSG Networks.

象徵著 Social Networking 的網站世代

雖然各有千秋 ...

各領風騷數十年 ...

不變的是 ...

從一家到每家的趨勢 ...

看著網路世代的前浪後浪 ...

想到了一些小啟示 ...

Tree Search << Graph Search << Human Search

商業的 Domain Know How << 使用者 Know How

機器不動 << 機器被動 << 機器主動

使用者 Know How 真的還蠻重要的 ...

就好像我常去吃的 ... (釋道素食)

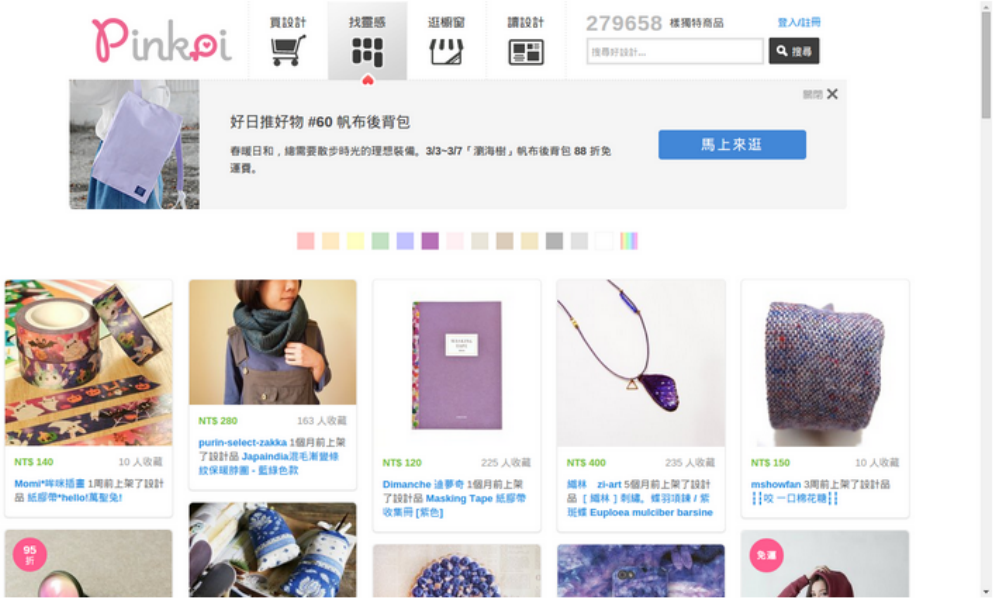
我每次去…老闆娘總會「主動」問我…

「一樣嗎？」

可是基本上菜單上是沒有「一樣」這到菜的XD

不過 Domain Know How 也不是完全沒用

強大的 Domain Know How Navigators ....



所以 ...

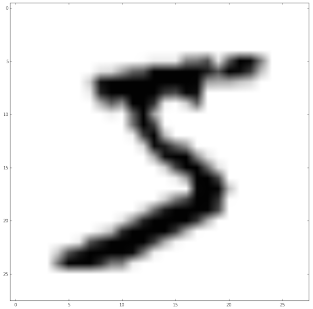
Navigation, Searching, Recommendation 本一家

每個成功的推薦系統背後

都一定有一個成功的 Tracking System！

因為 Machine
從來就不會自己 Learning

他就像小孩子一樣
都是靠 Training Data 教它的 …
偶爾也會學壞！



一般來說，收好 Data 之後 Model 幾乎就自動出現了

但多數廠商都沒把 Data 都沒收好

Math Behind

Recommendation System

幾種常見的推薦手法：

Newest

Hottest

Content-Based Similarity

Tagging & Searching

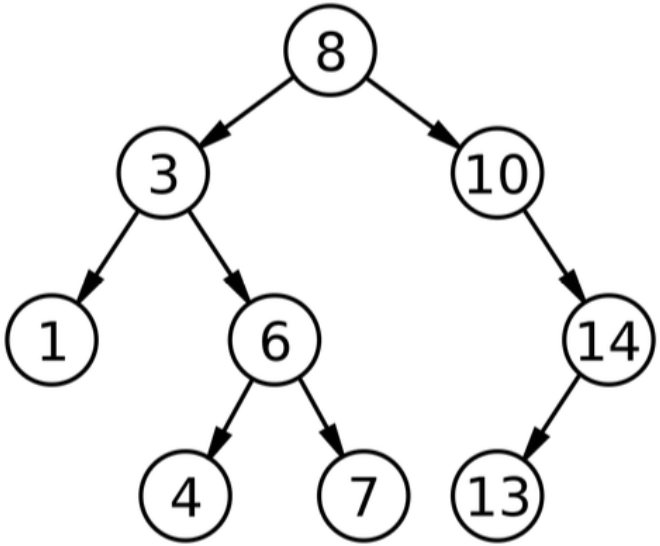
Collaborative Filtering

讓我們來看看這些方法背後的數學吧！

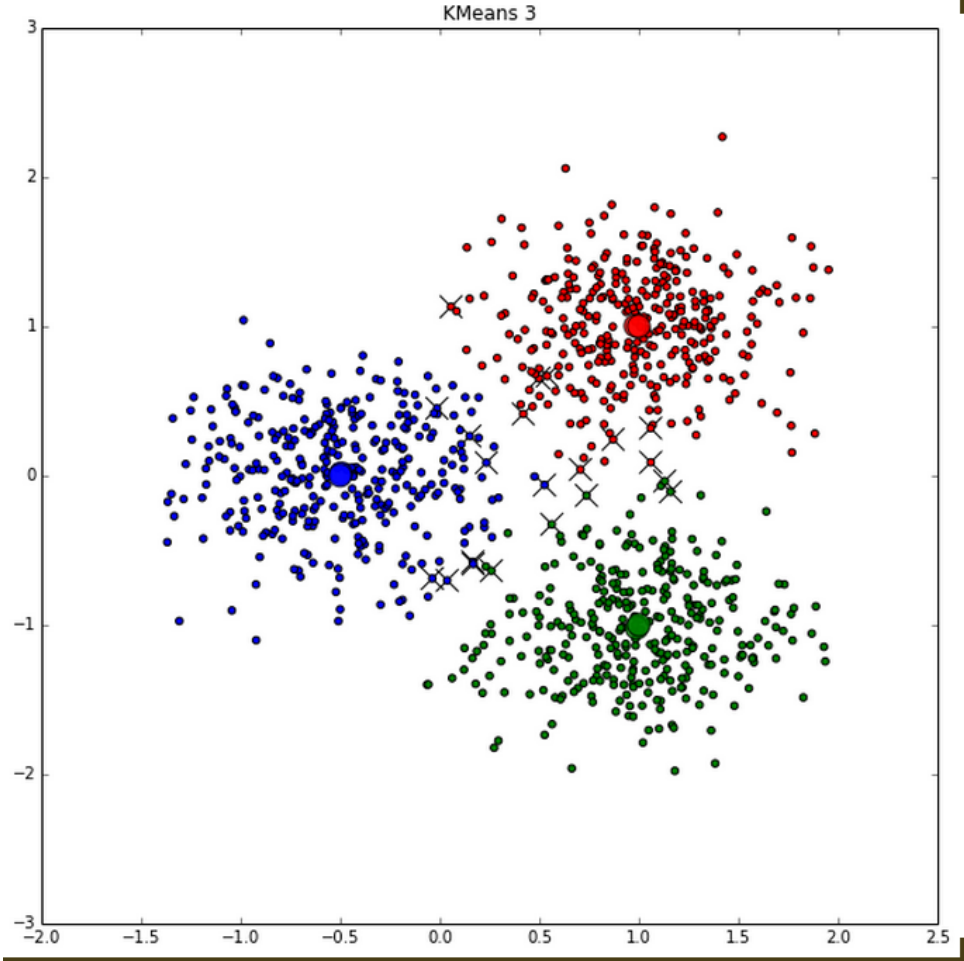
讓我們來看看這些方法背後的數學吧！

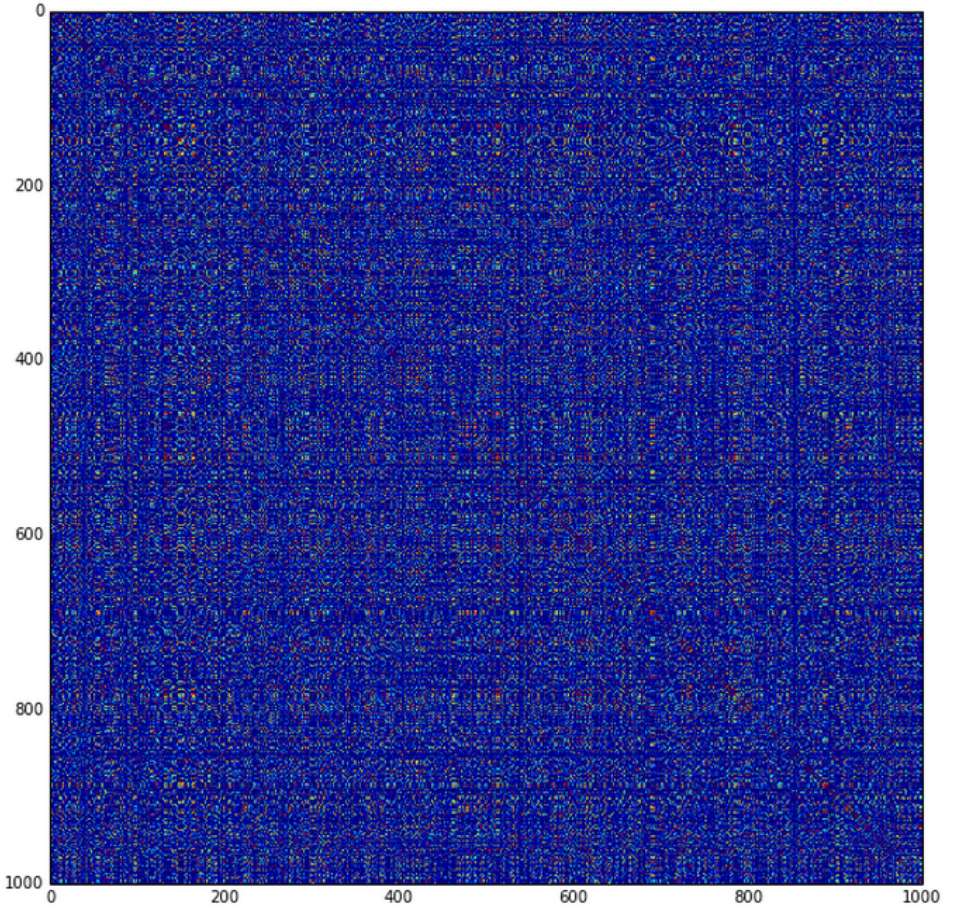
Newest & Hottest --- Sorting !

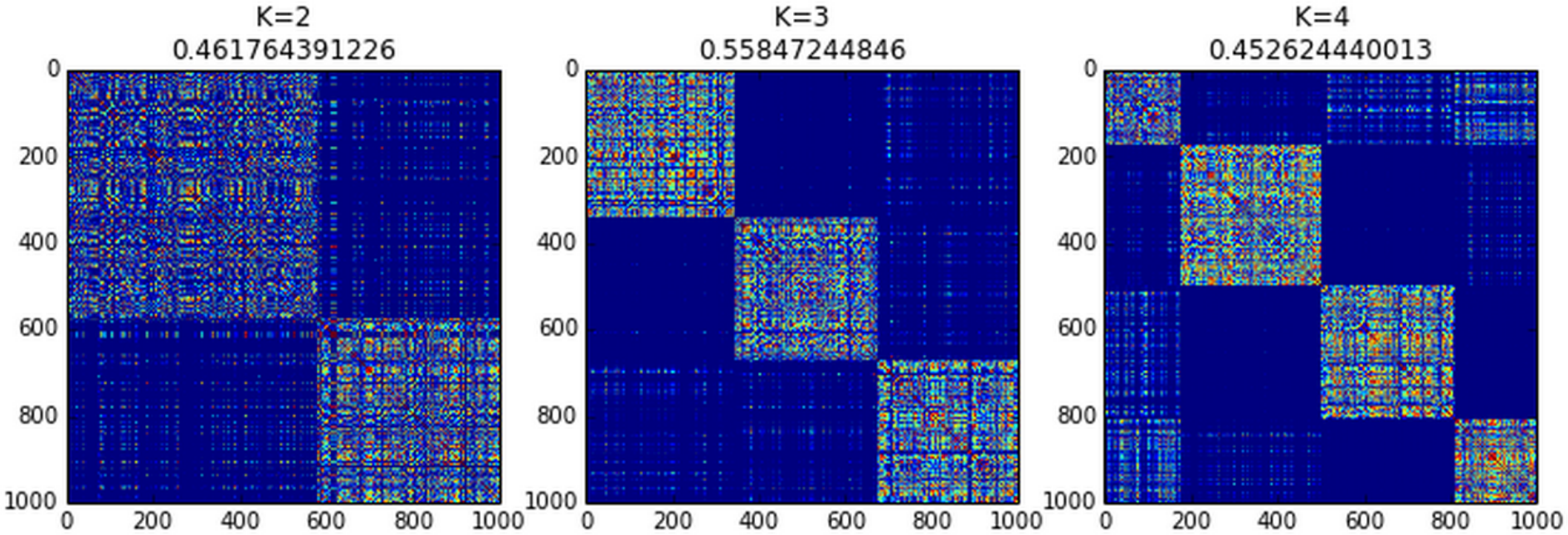
Download from
Dreamstime.com
This watermarked comp image is for previewing purposes only.
ID 231188
Sparkia | Dreamstime.com


8
3
10
1
6
14
4
7
13

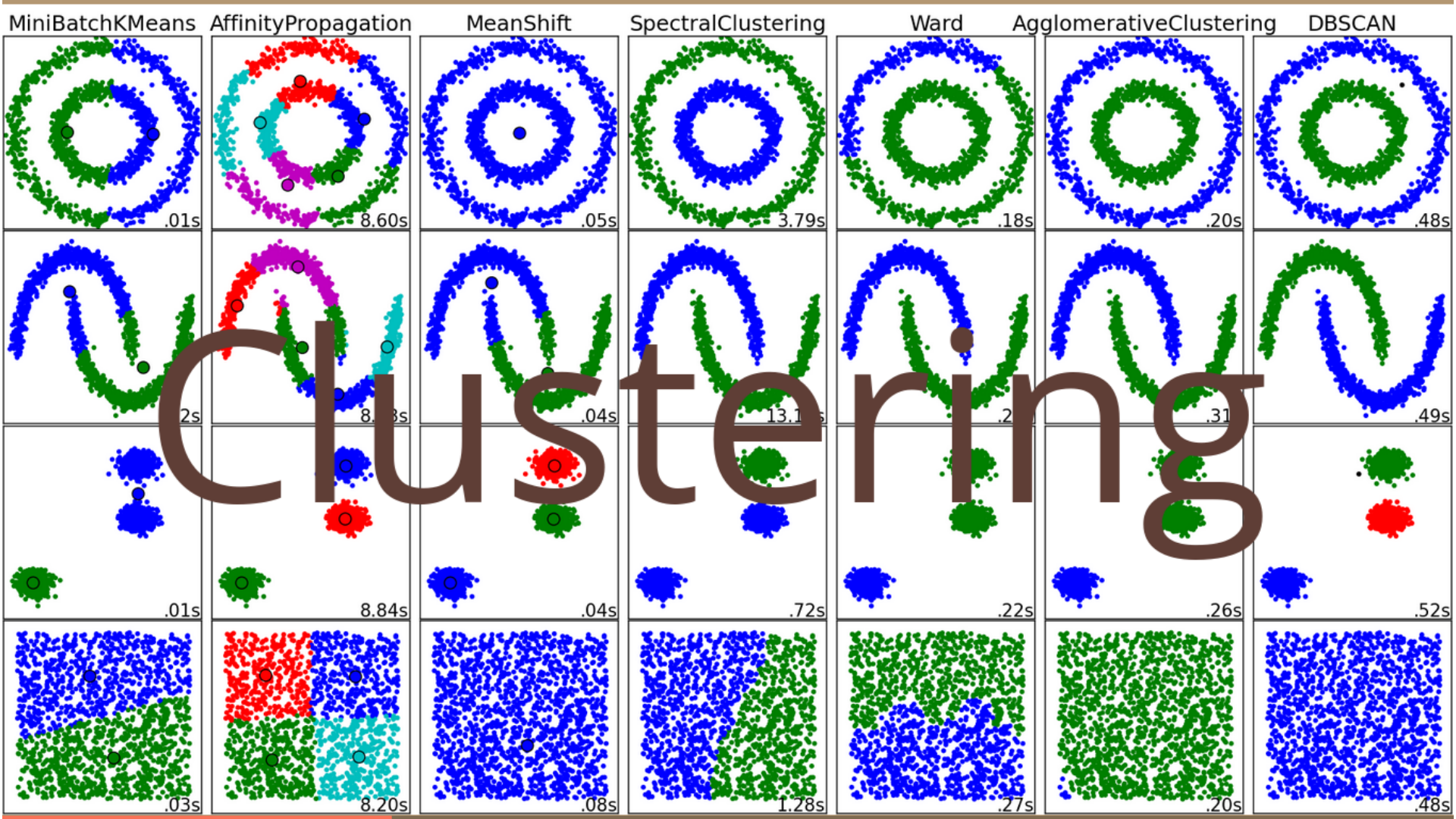
Content-Based Similarity --- Clustering Methods

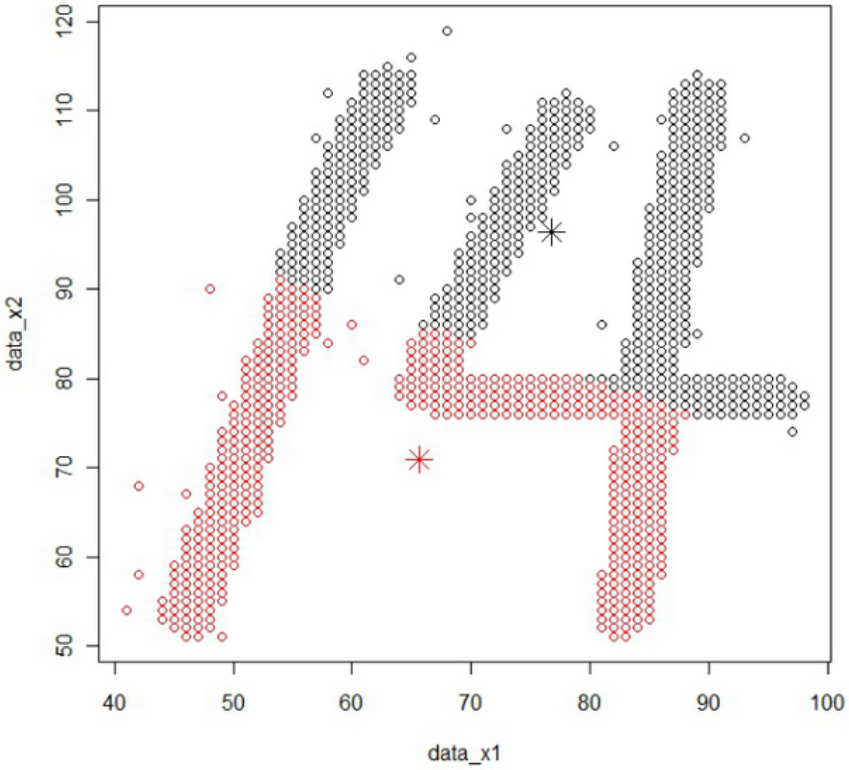

KMeans 3
3
2
1
0
−1
−2
−3
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5


0
200
400
600
800
1000
0
200
400
600
800
1000

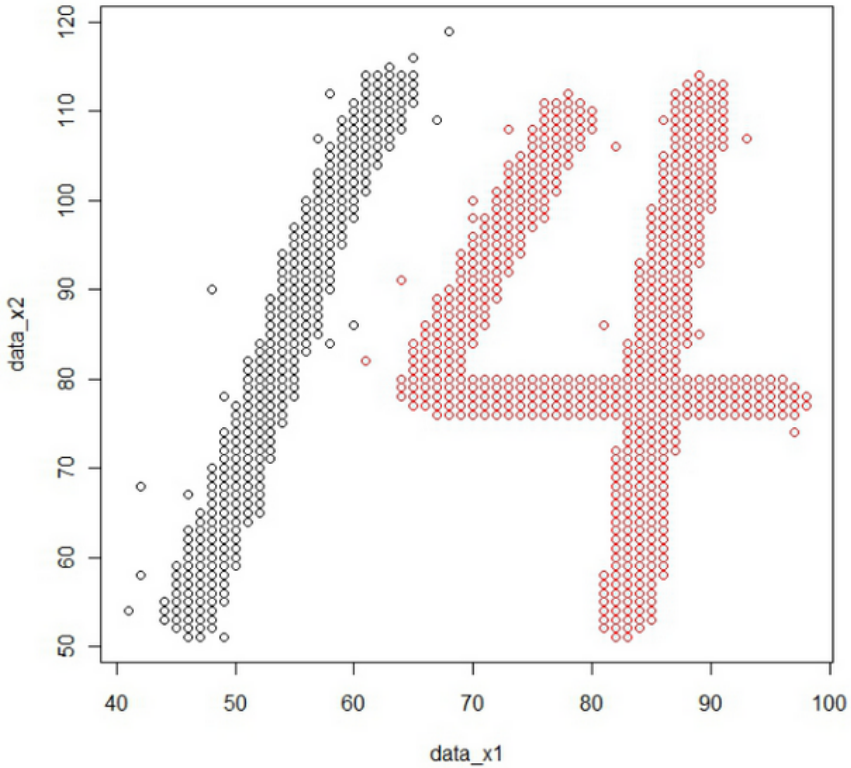

K=2
0.461764391226
K=3
0.55847244846
K=4
0.452624440013
0
200
400
600
800
1000

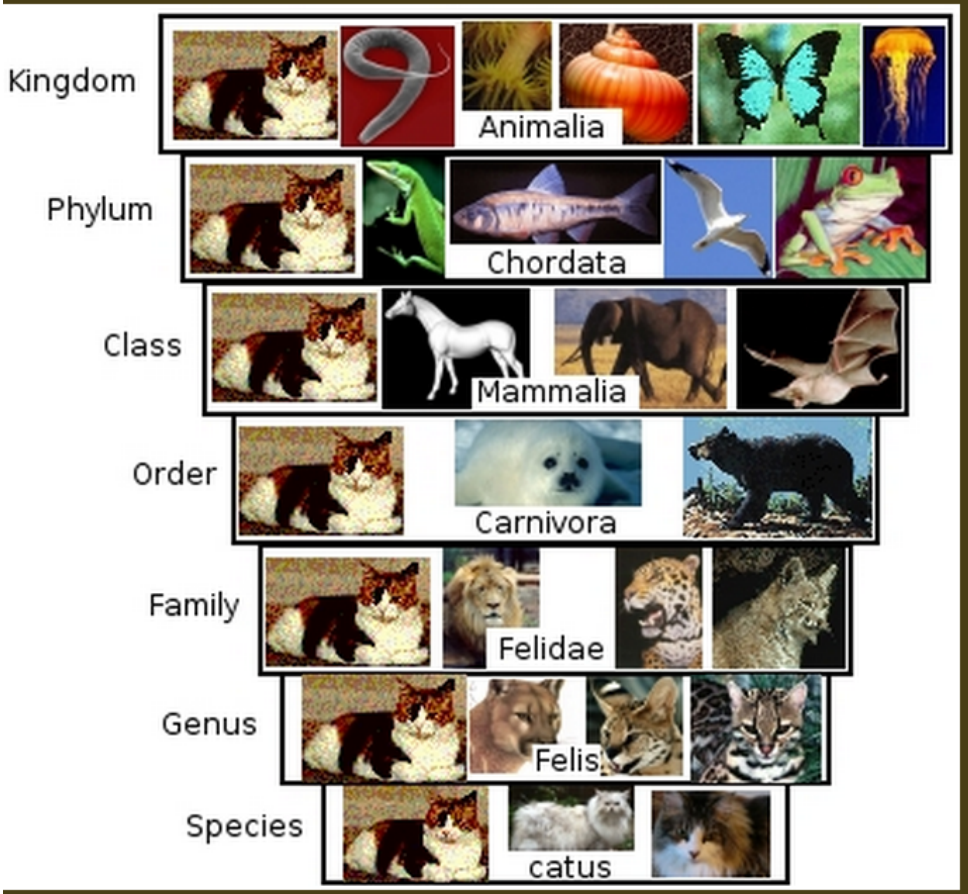
Clustering


MiniBatchKMeans
AffinityPropagation
MeanShift
SpectralClustering
Ward
AgglomerativeClustering
DBSCAN
.01s
8.60s
.05s
3.79s
.18s
.20s
.48s
.04s
.49s
.01s
8.84s
.04s
.72s
.22s
.26s
.52s
.03s
8.20s
.08s
1.28s
.27s
.20s
.48s


data_x2
120
110
100
90
80
70
60
50
40
50
60
70
80
90
100
data_x1

複習一下





Imagine If you meet …

Clustering Methods 的千古難題：

要分幾群？

要如何分？

無法「直接強迫」控制！

jupyter news_similarity (autosaved)

File Edit View Insert Cell Kernel Help

Code Cell Toolbar: None

In [224]:
```
article_df[labels == 4]["title"]
```

Out[224]:
```
56              大嬸還魂！  趙薇您真的回來了
1434          娘娘孫儷挺肚出巡  5侍衛隨侍
3017        笑死人不償命  《星星》東北話噴飯版
3408        蛤~是金秀賢陰謀  李敏鎬被偷拍賭博
3593       中國嫩模雙人組       爆乳寫真撩人
3973      Chanel救了她！楊冪靠這咖拯救孕味
4370            外星人現身台北  大家快逃啊
5249         美人胚子！ 短髮若曦  靈氣不滅
6715     冰上佳人！  17歲花式花冰小將賈萌博眼球
6756          影帝爆輸千萬被追債  賭場照曝光
7087            馮小剛：中國導演戴腳鐐跳舞
7630          楊丞琳嗆愛心便當  驚露飯桶本性
7957            王麟祥回鍋  當華視新總經理
Name: title, dtype: object
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [12]:
```
# from sklearn.cluster import DBSCAN
```

jupyter news_similarity (autosaved)

File Edit View Insert Cell Kernel Help

Code Cell Toolbar: None

In [30]:
```
for topic_idx, topic in enumerate(nmf.components_):
    print("Topic #%d:" % topic_idx)
    print(" ".join([temp_ws["word"][i]
                    for i in topic.argsort()[:-n_top_words - 1:-1]]))
    print()
```

```
Topic #0:
方法 一部分 驗證 Facebook 816 國產 典範 反觀 官邸 進帳 美聯社 改為 財長 不再 所得 結論 中將 Yanukovich 公路 國
()
Topic #1:
黃岩 抑制 進帳 200 大家 銷售 美元 圓仔 一路 最低 近年來 想法 創業 感冒 高山 避險 多元化 國外 台電 差距
()
Topic #2:
輻射 生涯 放大 搭配 首度 劇情 初期 疼痛 行為 LED 稅制 歷經 稍早 截至 女人 PADDING 失業率 多人 改組 少女
()
Topic #3:
方法 驗證 失分 證件 選情 最低 一方面 目的 氣象 台灣 氛圍 全壘打 質疑 近年來 高山 北京 買進 而是 黃金 截至
()
Topic #4:
目的 最低 離職 國外 理性 圓仔 200 農藥 穿著 公開賽 做愛 AppleEnews calibri 感動 政務委員 思考 經典 臺灣 做出 太
()
Topic #5:
個股 致死 幸好 感動 抑制 進帳 離職 中將 圓仔 cell 最低 決策 巴黎 公開賽 內線交易 名將 近年來 實驗 國產 翻轉
()
Topic #6:
TPP 恐龍 致死 進帳 有利於 阿富汗 rdquo 內線交易 做愛 縣政府 近年來 單日 差不多 降低 離職 年將 周日 國產 理性 原住
()
Topic #7:
```

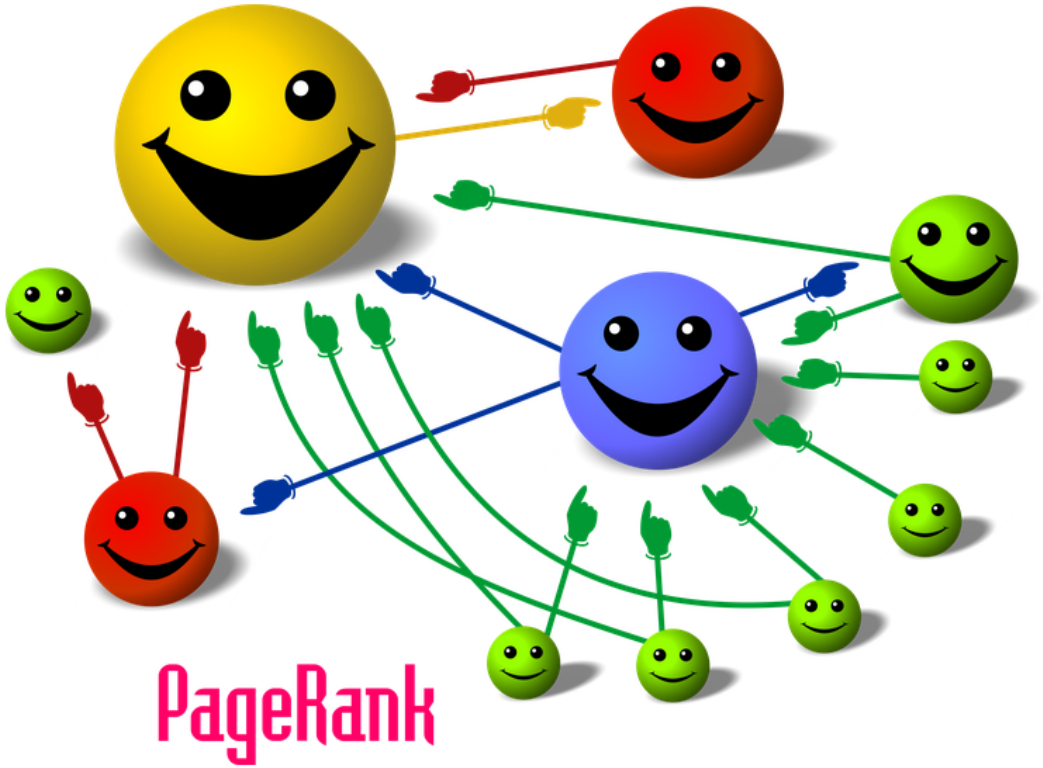
jupyter news_similarity (autosaved)

File Edit View Insert Cell Kernel Help

Code Cell Toolbar: None

```
Topic #7:
報價 廢鋼 致死 圓仔 感動 內線交易 離職 翻轉 單日 施政 最低 原住民 國外 相對 cell 巴黎 國防 有個 銷售 保持
()
Topic #8:
0cm 意思 200 銷售 地檢署 理性 抑制 大家 感冒 物料 效應 進帳 感動 最低 決策 臺灣 一路 北京 近年來 處分
()
Topic #9:
方法 一部分 驗證 結婚 816 總計 調閱 大多 監督 典範 巴黎 課稅 大聲 對戰 平安 女將 而已 安倍 考慮 在於
()
Topic #10:
一級 抑制 決策 高山 銷售 穿著 大家 驚訝 理性 中將 610 感動 200 氣象 有個 calibri 做愛 寬頻 中央委員 安倍
()
Topic #11:
唯一 抑制 穿著 200 816 近年來 罷黜 典範 將為 雞蛋 決策 銷售 排名第 臺灣 一路 農曆 iPad 中將 理性 國外
()
Topic #12:
帳號 日後 心理 中將 大雅 國產 天母 Yanukovich 名將 保全 同仁 三重 清理 原住民 出包 做愛 得標 接待 阿富汗 Firefox
()
Topic #13:
失望 決策 抑制 寬頻 610 國外 200 離職 大家 美元 結婚 高山 公開賽 創業 圓仔 感動 市政 rdquo 品質 越來越
()
Topic #14:
油價 進帳 近年來 圓仔 感動 610 190 隨後 張盛 200 所得 街頭 龍頭 還要 寬頻 最大 高山 予以 決策 根據
()
Topic #15:
概念股 施政 離職 決策 圓仔 進帳 節省 感動 TPP 推特 銷售 保全 波段 2884 人均 抑制 Yanukovich 二月 葉克膜 小鳥
()
Topic #16:
受傷 引來 決策 銷售 200 美元 恐龍 大家 抑制 感動 公路 多元化 190 離職 公開賽 向上 施政 國外 影響力 援助
()
```

Tagging & Searching --- Page Rank !


PageRank

$$PageRank\ of\ site = \sum \frac{PageRank\ of\ inbound\ link}{Number\ of\ links\ on\ that\ page}$$

OR

$$PR(u) = (1 - d) + d \times \sum \frac{PR(v)}{N(v)}$$

Exploration of the Google PageRank Algorithm

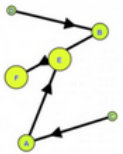
Circles correspond to web pages, links correspond to hyperlinks

Questions/Comments
A Few Notes About This

Damping Factor d:
0.85

Generate New

- *Drag nodes to rearrange*
- *Click on node, then click on another to create link*
- *Double-click to add node (max 9)*
- ***Del+mouseover** deletes node*
- *Click on link to delete*
- *Mouseover link to see location in matrix*



Sorted PageRank Vector
(10 iterations)

E	0.24
F	0.24
A	0.17
B	0.17
C	0.09
D	0.09

The PageRank vector **x** satisfies **Gx** = **x**, where

Google Matrix **G** = d * [(Hyperlink Matrix **H**) + (Dangling Nodes Matrix **A**)] + ((1-d)/N) * (NxN Matrix **U** of all 1's)

$$G = 0.85 * \left[\text{Hyperlink Matrix } H + \text{Dangling Nodes Matrix } A \right] + \frac{0.15}{6} * \text{Matrix } U \text{ with all 1's}$$

Hyperlink Matrix H

	A	B	C	D	E	F
A			1			
B				1		
C						
D						
E	1					
F		1				

Dangling Nodes Matrix A

	A	B	C	D	E	F
A					1/6	1/6
B					1/6	1/6
C					1/6	1/6
D					1/6	1/6
E					1/6	1/6
F					1/6	1/6

Matrix U with all 1's

	A	B	C	D	E	F
A	1	1	1	1	1	1
B	1	1	1	1	1	1
C	1	1	1	1	1	1
D	1	1	1	1	1	1
E	1	1	1	1	1	1
F	1	1	1	1	1	1

$$G = \begin{bmatrix} 0.03 & 0.03 & 0.88 & 0.03 & 0.17 & 0.17 \\ 0.03 & 0.03 & 0.03 & 0.88 & 0.17 & 0.17 \\ 0.03 & 0.03 & 0.03 & 0.03 & 0.17 & 0.17 \\ 0.03 & 0.03 & 0.03 & 0.03 & 0.17 & 0.17 \\ 0.88 & 0.03 & 0.03 & 0.03 & 0.17 & 0.17 \\ 0.03 & 0.88 & 0.03 & 0.03 & 0.17 & 0.17 \end{bmatrix}$$

Collaborative Filtering

Matrix Completion

理論上是這樣 ...

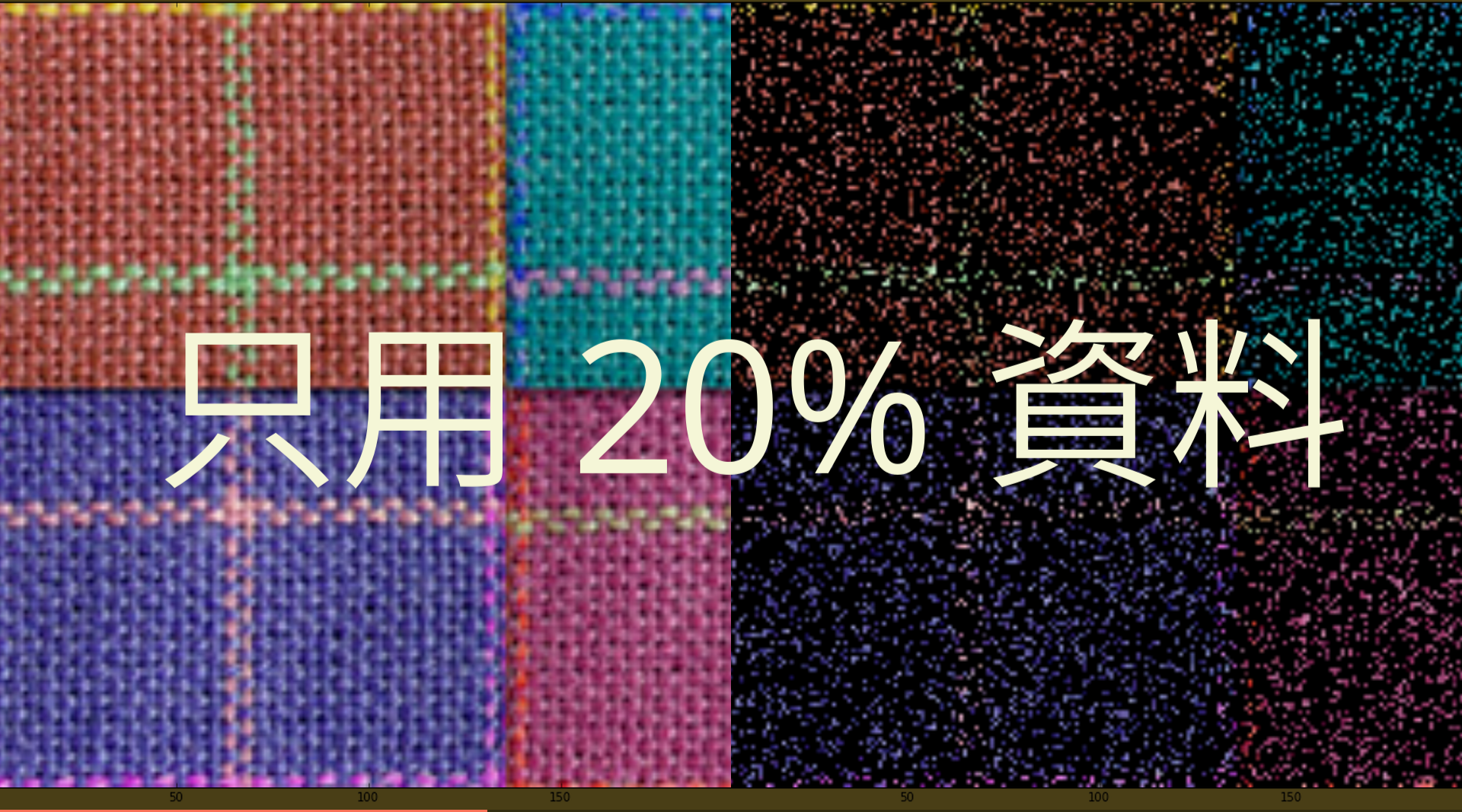
但是調參數很重要！

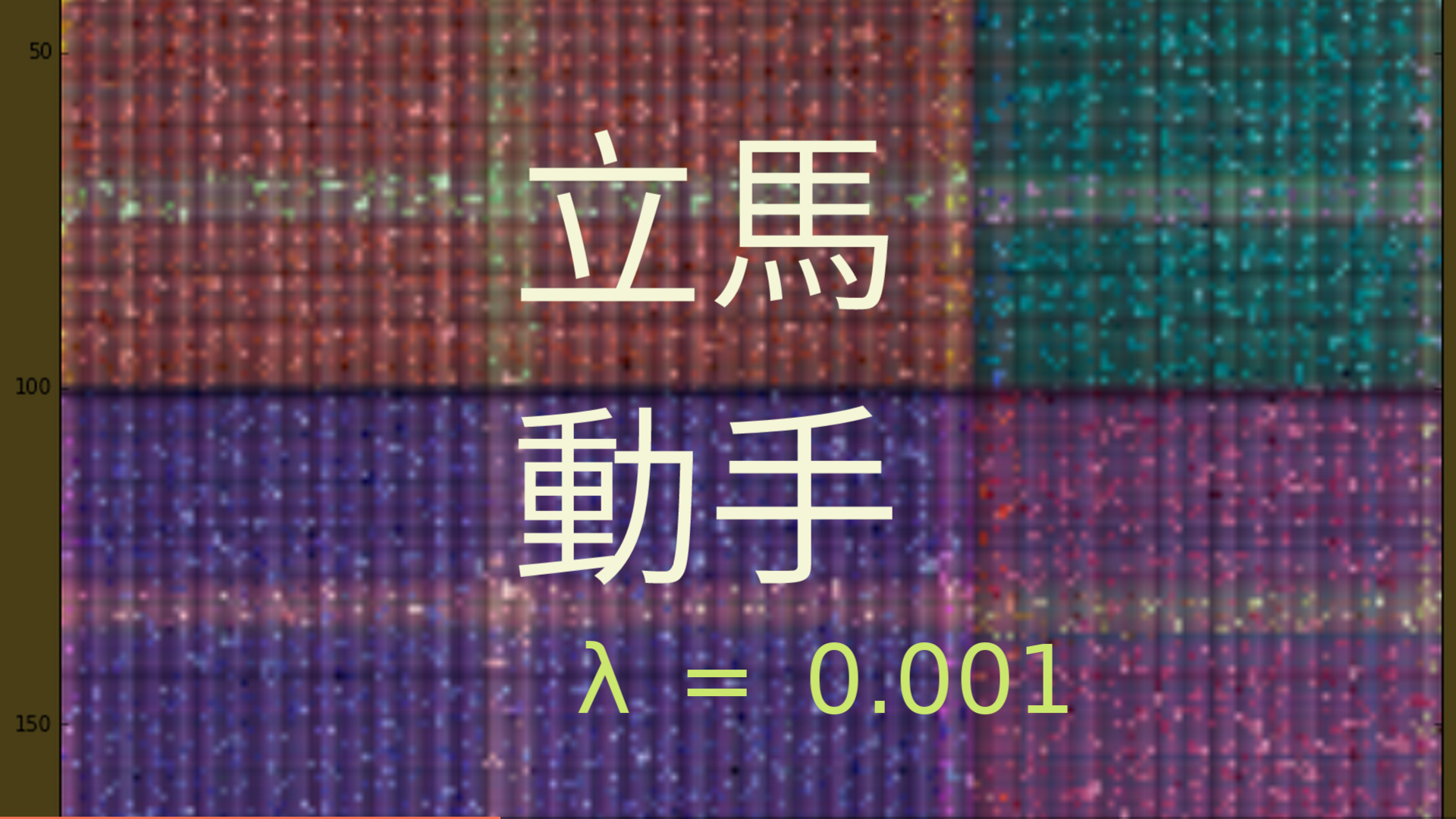
Matrix Completion

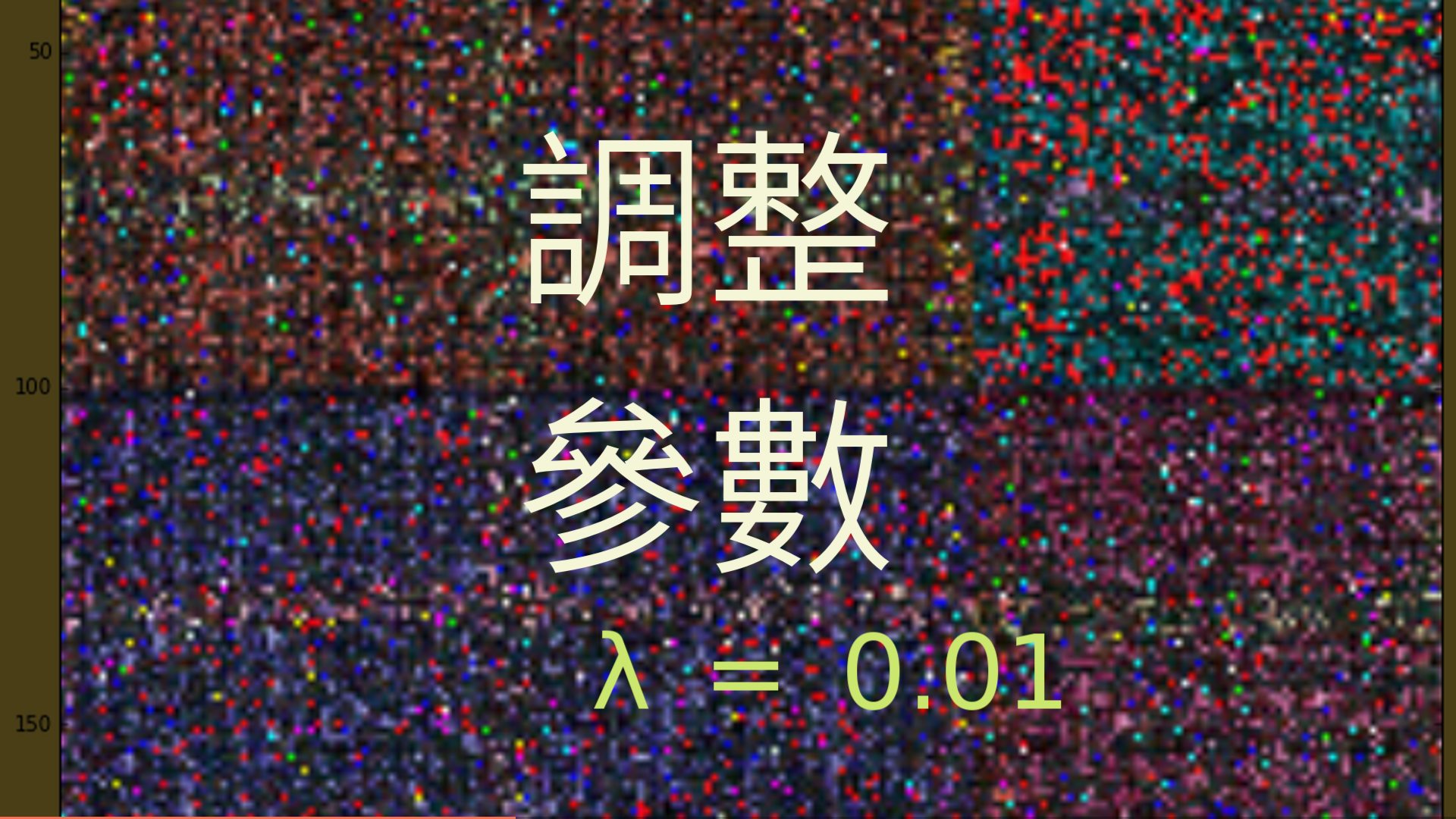
B : data matrix, Ω : indice of observed pixels

$$\text{minimize}_{X \in R^{m \times n}} \quad \frac{1}{2}\left\| P_\Omega(X - B) \right\|_2^2 + \lambda \left\| X \right\|_*,$$

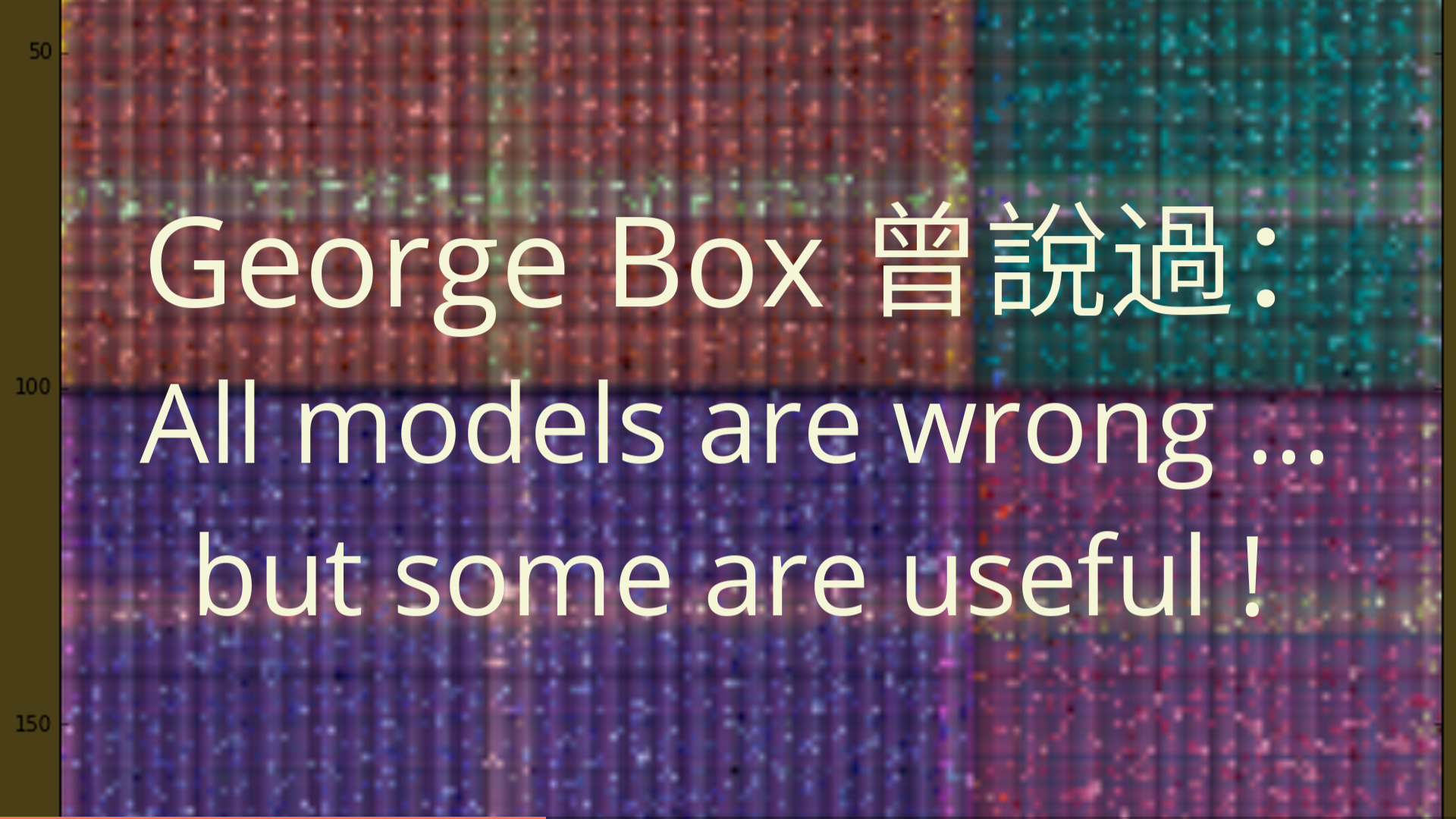
$$\begin{pmatrix} 1 & ? & 9 & ? \\ ? & ? & ? & 14 \\ 3 & ? & ? & ? \\ ? & 8 & 12 & 16 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \\ 4 & 8 & 12 & 16 \end{pmatrix}$$


只用 20% 資料
50
100
150
50
100
150


50
100
150
立馬
動手
λ = 0.001

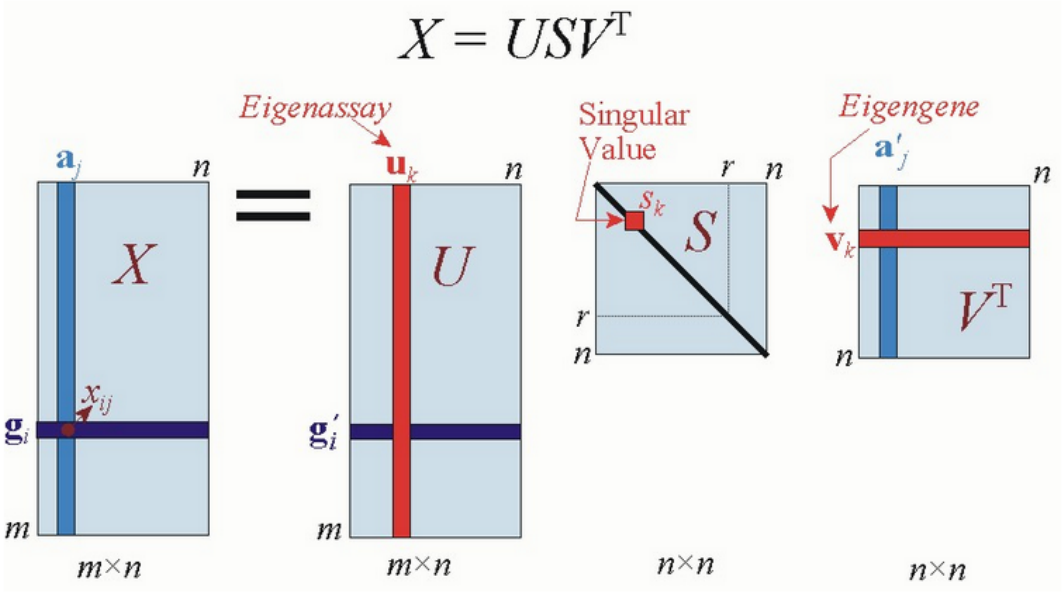

50
100
150
調整
參數
λ = 0.01

50
100
150
調整
參數
λ = 0.0001


50
100
150
George Box 曾說過:
All models are wrong ...
but some are useful !

Collaborative Filtering

Matrix Factorization


$X = USV^{T}$
Eigenassay
Singular Value
Eigengene
$\mathbf{a}_j$
n
X
x_{ij}
$\mathbf{g}_i$
m
$m \times n$
$\mathbf{u}_k$
U
$\mathbf{g}'_i$
$m \times n$
r
s_k
S
$n \times n$
$\mathbf{a}'_j$
$\mathbf{v}_k$
V^{T}
$n \times n$

它真正想解的是 Reconstruction Error 最小的東西 ...

$$\min_{||\vec{w}||=1} \sum_i ||\vec{x}_i - < \vec{x}_i, \vec{w} > \vec{w}||$$

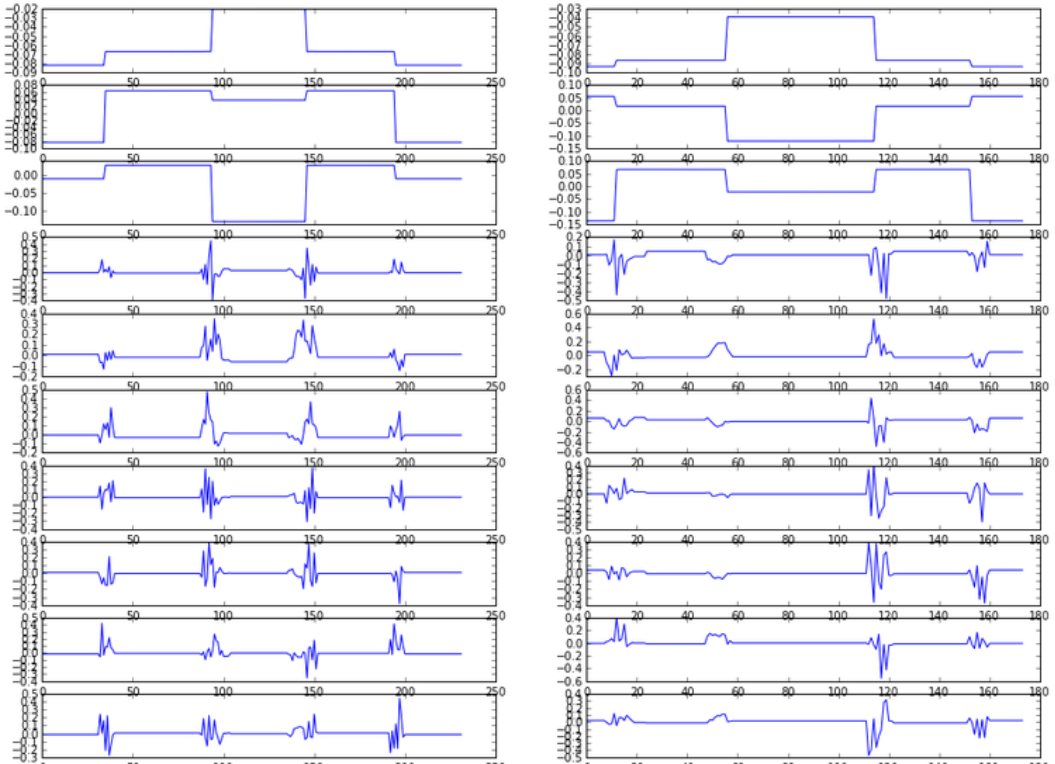
只是剛好 2-norm 的時候，解剛好是 SVD 而已 ...

看圖猜猜看

猜出它的 PCA 有幾個 PCs？

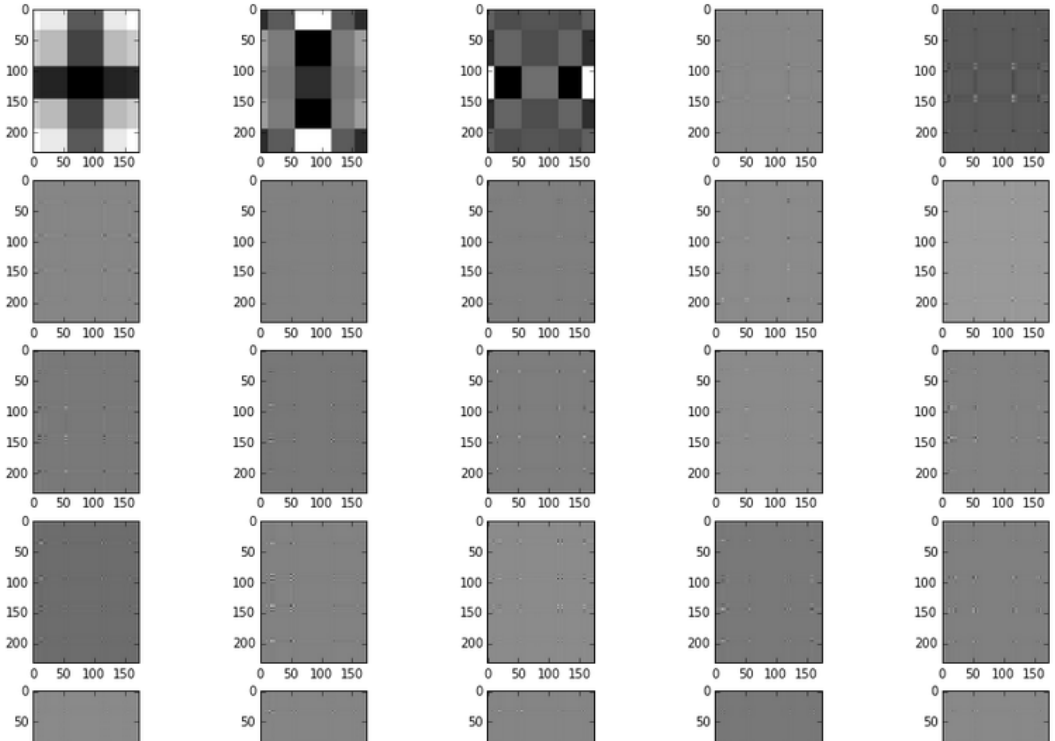
In [6]:

```
show_n = 10

fig, ax = pylab.subplots(show_n, 2)
for i in range(show_n):
    ax[i,0].plot(U[:,i])
    ax[i,1].plot(V[i,:])
```

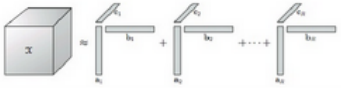


In [7]:

```
show_n_col = 5
show_n_row = 5


fig, ax = pylab.subplots(show_n_row, show_n_col)
for i in range(show_n_row):
    for j in range(show_n_col):
        ax[i,j].imshow(U[:,i*show_n_col+j].reshape(232,1)*V[i*show_n_col+j,:]*S[i*sh
ow_n_col+j],cmap=cm.gray)
```



canonical tensor decomposition



OPTIMIZATION PROBLEM

given tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$, find rank-$R$ canonical tensor $\mathcal{A}_R \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ that minimizes

$$f(\mathcal{A}_R) = \frac{1}{2} \|\mathcal{T} - \mathcal{A}_R\|_F^2.$$

FIRST-ORDER OPTIMALITY EQUATIONS

$$\nabla f(\mathcal{A}_R) = \mathbf{g}(\mathcal{A}_R) = 0.$$

(problem is non-convex, multiple (local) minima, solution may not exist (ill-posed), ... ; but smooth, and we assume there is a local minimum)

(de Silva and Lim, SIMAX, 2009)

UNIVERSITY OF WATERLOO

Boosting Models

Ensemble Learning for Ranking

jupyter NI_Simple_Rcomm_DEV_20140507_3 Last Checkpoint: 10/21/2014 (autosaved)

File Edit View Insert Cell Kernel Help

Code Cell Toolbar: None

In [38]:
```
time_ratio = 0.7
lookup_sorted_df["average_score"] = (1-time_ratio)*lookup_sorted_df["similar_score"] + time_ratio*l
lookup_sorted_df.sort(["average_score"])
```

Out[38]:

	title	dt_pubDate	time_score	similar_score	average_scor
2368750	學運女王遭控假公益真斂財　劉喬安否認但坦言有瑕疵	2014-05-03 02:25:00	0.180716	0.000000	0.126501
2372650	蘋果協力參與慈善 對SF Gives捐款50萬美元	2014-05-06 05:20:19	0.000000	0.962236	0.288671
2367689	太陽花學運女王　遭爆募物資斂財	2014-05-02 07:38:00	0.188985	0.640105	0.324321
2368563	學運女王募款 慈善協會未立案	2014-05-02 22:00:00	0.182798	0.700959	0.338247
2368592	未立案協會募款　學運女王恐違法	2014-05-02 22:00:00	0.182798	0.818920	0.373635
2359496	女王爆氣譙蔡正元　「終生抵制國民黨」	2014-04-27 03:48:00	0.223657	0.813661	0.400658
2368576	蘋果日報5月01日 蘋果日報網站總瀏覽量1871萬次	2014-05-02 22:00:00	0.182798	0.937968	0.409349
2368679	控施明德侵占紅衫軍捐款　黃越綏判賠百萬	2014-05-03 00:08:00	0.181804	0.945284	0.410848
2367896	爆料許聖梅隆乳　主播王瑞玲挨告不起訴	2014-05-02 09:41:00	0.188151	0.946307	0.415598
2359526	「太陽花女王」臉書發怒　罵蔡正元：瞧不起你	2014-04-27 04:25:00	0.223530	0.898349	0.425976
2351919	太陽花女王好風度　為「低級」說向幹哥致歉	2014-04-22 05:57:00	0.243885	0.878903	0.434390
2350846	人美心更美！　太陽花女王神似張柏芝	2014-04-21 12:53:00	0.246318	0.888847	0.439076

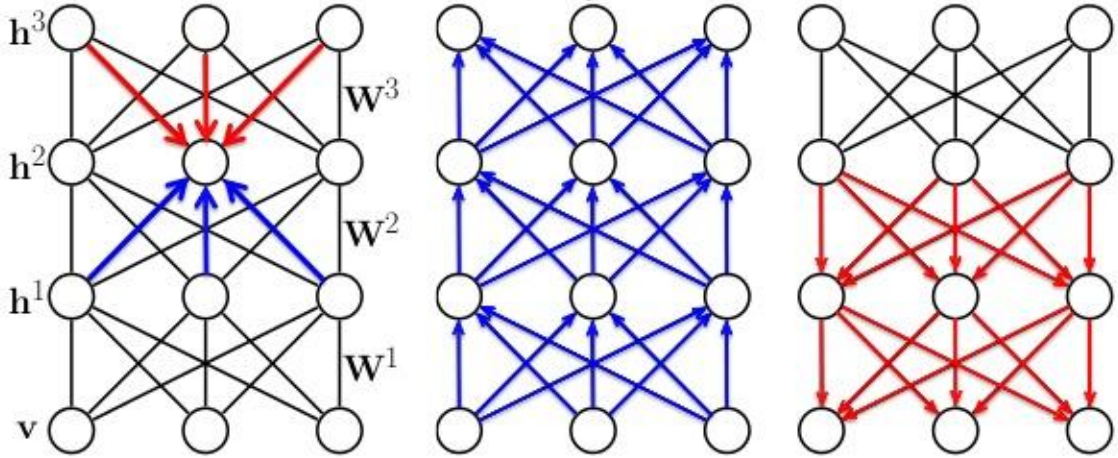
Mathematical Formulation

$$P_\theta(\mathbf{v}) = \frac{P^*(\mathbf{v})}{\mathcal{Z}(\theta)} = \frac{1}{\mathcal{Z}(\theta)} \sum_{\mathbf{h}^1,\mathbf{h}^2,\mathbf{h}^3} \exp\left[\mathbf{v}^\top W^1 \mathbf{h}^1 + \mathbf{h}^{1\top} W^2 \mathbf{h}^2 + \mathbf{h}^{2\top} W^3 \mathbf{h}^3\right]$$

Deep Boltzmann Machine

Neural Network Output

Deep Belief Network



Input

Unlike many existing feed-forward models: ConvNet (LeCun), HMAX (Poggio), Deep Belief Nets (Hinton)

一般來說，都會用

多個 Week Learner 組合 Strong Learner
讓 Strong Learner 有 back-off

複習一下，幾種常見的推薦手法：

Newest

Hottest

Content-Based Similarity

Tagging & Searching

Collaborative Filtering

Last View & After View

Mixture Model

感謝大家！

如有問題，歡迎大家來信

info@agilearning.io