

商品推薦系統

─推薦系統概念

Lecture 1

講師：周光宇博士

2018.08

講師簡介

周光宇博士 Dr. Kuang-Yu Chow



新加坡 KYC Global Pte. Ltd. 執行長
台灣東吳大學 巨量資料管理學院 兼任副教授
新加坡國立大學 大數據人工智能專題 講座
Frost & Sullivan 亞太地區 ICT 高級顧問



經歷

- 32年AT&T工作經驗
- AT&T系統集成 亞太地區 執行董事
- AT&T批發 亞太地區 銷售副總裁
- Concert (AT&T / BT合資公司) 亞太地區 營運副總裁
- AT&T中國公司 營運副總裁
- AT&T貝爾實驗室 技術經理

學歷

美國加州大學洛杉磯分校 (UCLA) 工程碩士和博士
台灣清華大學數學理學士

課程簡介

- 本課程將重點介紹推薦系統及其應用。隨著大數據和人工智慧引起企業和產業的革命性變化，個性化營銷已經成為有數據和分析能力企業的關鍵競爭優勢，而推薦系統在其中起著至關重要的作用。例如，亞馬遜35%的收入和70%的Netflix電影都來自推薦系統提供的個性化推薦。然而，推薦系統的應用不限於零售和娛樂。他們也用於網路廣告、房地產選擇、政治選舉等等。
- 在這個課程中，學生將學習各種基本類型的推薦系統、算法及其應用。完成課程後，學生能夠在本次比賽中應用一些學到的概念和算法，並能夠設計和實施至少一種或兩種基本的推薦系統。

參考書籍和報告

- ❑ **Recommender Systems – The Text Book** by Charu C. Aggarwal
- ❑ **Recommender Systems – An Introduction** by Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich, Cambridge University Press
- ❑ **Recommender Systems Handbook** by Francesco Ricci, Lior Rokach, BrachaShapira, Paul B. Kantor
- ❑ **A Gentle Introduction to Recommender Systems with Implicit Feedback** by Jesse Steinweg-Woods
- ❑ **Collaborative Filtering Recommender Systems** by Michael D. Ekstrand, John T. Riedl and Joseph A. Konstan, Foundations and Trends in Human–Computer Interaction
- ❑ **Movies Recommender System:** <https://www.kaggle.com/rounakbanik/movie-recommender-systems>
- ❑ **Recommender Systems in Python 101:** <https://www.kaggle.com/gspmoreira/recommender-systems-in-python-101>
- ❑ **Multi-criteria Recommender System for Life Insurance Plans based on Utility Theory** by Akshay Hinduja and Manju Pandey

課程大綱

時間	內容	備註
10:00-10:50	Lecture 1 推薦系統概念	包含3題搶答
10:50-11:00	休息	
11:00-12:00	Lecture 2 協同過濾推薦	包含1題搶答
12:00-13:00	中午休息	
13:00-13:50	Lecture 3 基於內容推薦	包含1題搶答
13:50-14:00	休息	
14:00-14:50	Lecture 4 基於知識的推薦系統(I+II)	包含1題搶答
14:50-15:00	休息	
15:00-15:50	Lecture 5 評估推薦系統	包含4題搶答

更多選擇更好?



24種口味



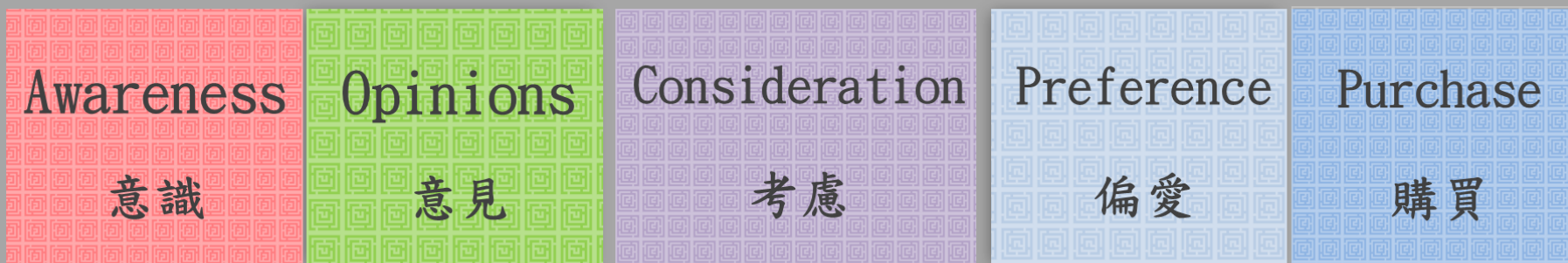
6種口味

哪種情況賣的更好?

典型購買流程

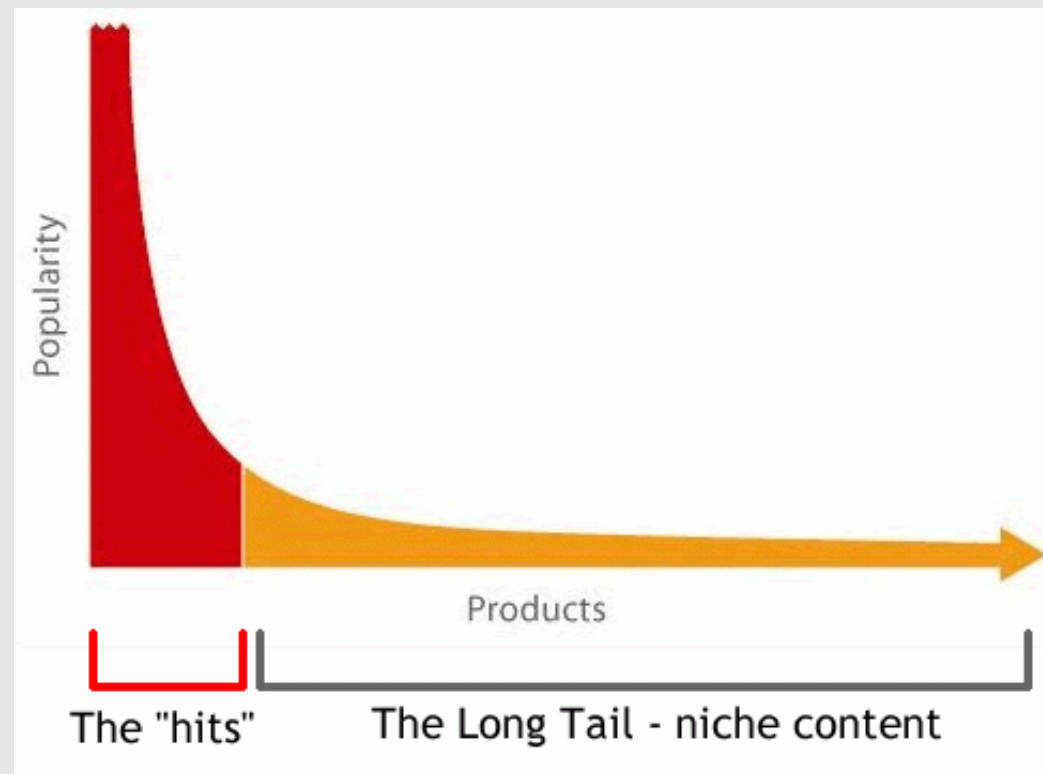
AIDA: **A**ware, **I**nterest, **D**esire, **A**ction

提出推薦可以加速購買過程



長尾理論

- 實體世界和網路世界的區別被稱為長尾現象。
- 長尾現象迫使網路商家向個人用戶推薦商品。



Amazon推薦系統 - 銷售額的35%

KC

Hi, Kuang-Yu

Your Orders

0 recent orders

Amazon Prime

Try Prime >

Gift Card Balance

Reload \$100, Get \$5 >


AMAZON


PILOTseason

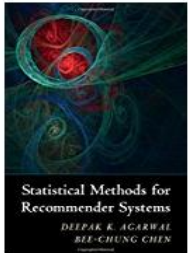
presented by
GEICO

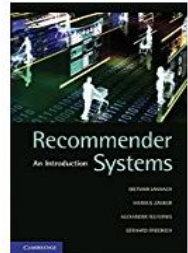
Related to items you've viewed


See more

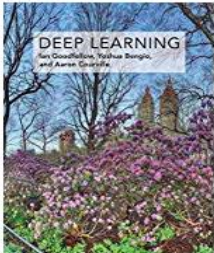


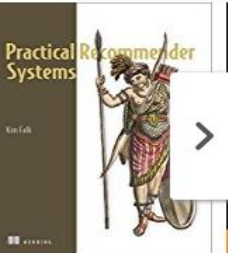










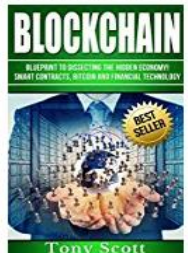



Get your taxes done right with TurboTax® 2016

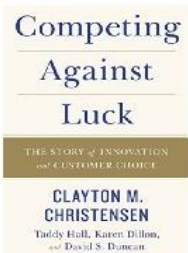
Shop now


Ad feedback

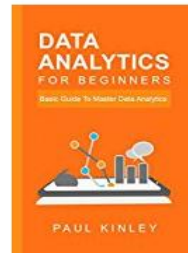
Recommended for You in Kindle Books









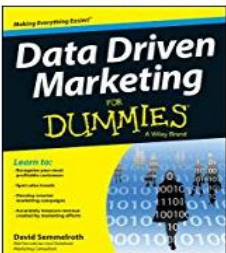


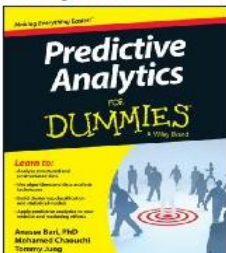


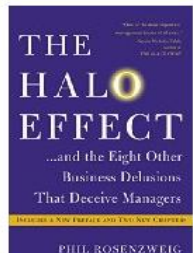


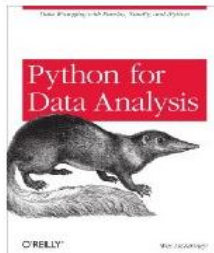


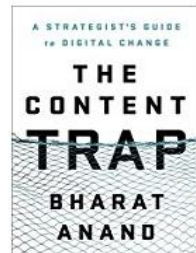
Recommendations for you in Kindle Store

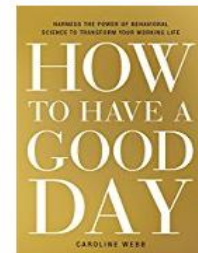


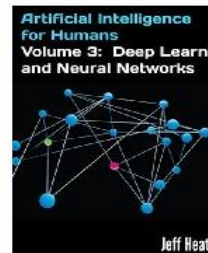














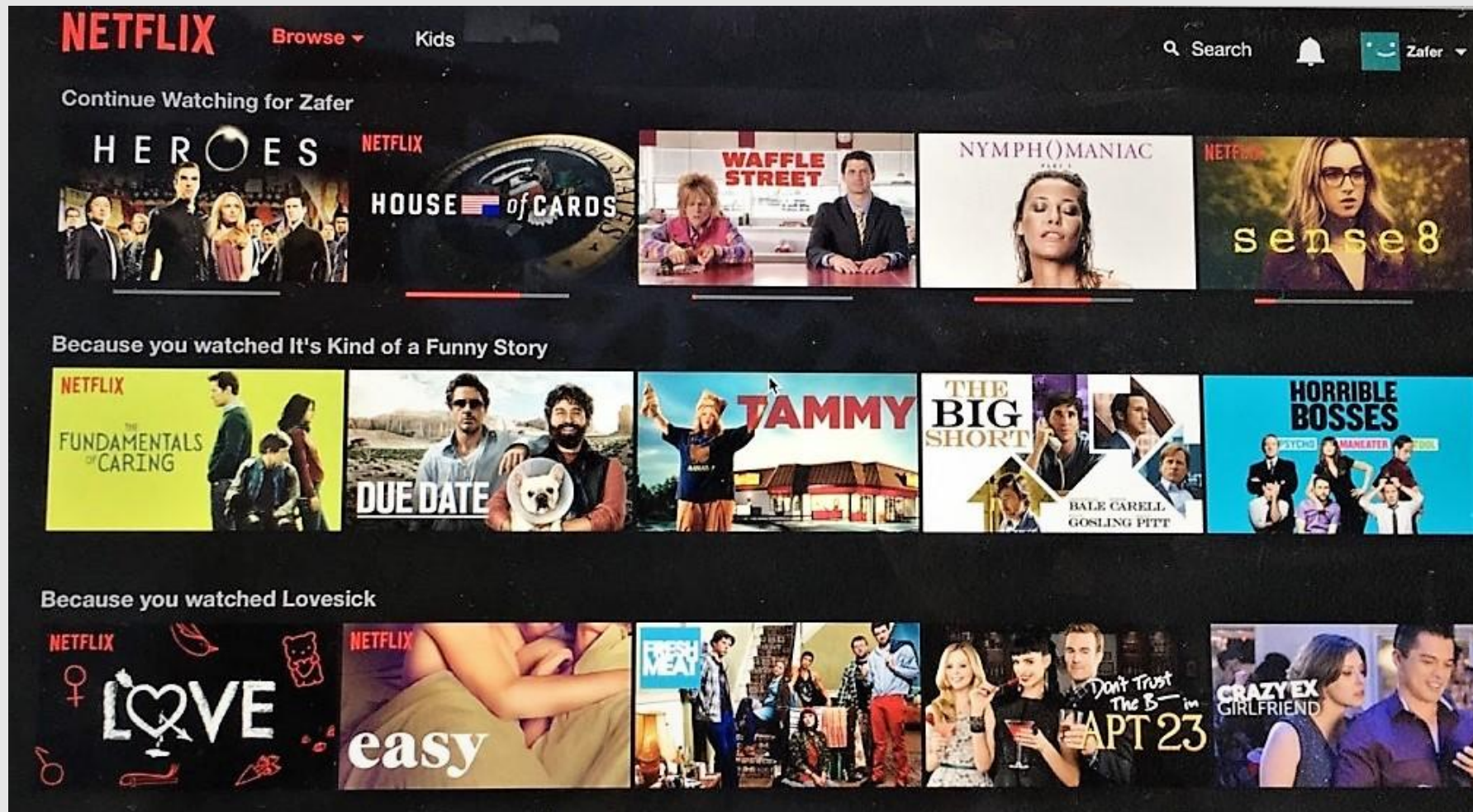
music | New album

DEPECHE MODE






Netflix推薦系統 - 幫助提高忠誠度和降低成本



Netflix: 70%用戶觀看的電影來自推薦

LinkedIn: People You May Know - 你可能認識的人




[Home](#)
[My Network](#)
[Jobs](#)
[Messaging](#)
[Notifications](#)
[Me](#)
[Work](#)
[Try Premium for Free](#)

415

Your connections

[See all](#)









Add personal contacts

[Continue](#)

[More options](#)


We'll import your address book to suggest connections and help you manage your contacts. [Learn more](#)

People you may know

 <p>John Maguire Director of Sales, ASEAN at TIME dotCom Berhad 38 mutual connections</p>	Dismiss	Connect
 <p>Jorge Matar Executive Director Global Data Sales - Wholesale AT&T Partner Solutions 30 mutual connections</p>	Dismiss	Connect
 <p>Hang Sin hbsin@live.com.sg</p>	Dismiss	Invite
 <p>Shaw Irina Director at AT&T Australia 37 mutual connections</p>	Dismiss	Connect
 <p>Sandro Natale Director Business Development at AT&T Wi-Fi AT&T</p>	Dismiss	Connect
 <p>Jo Han Regional Account Manager at AT&T 10 mutual connections</p>	Dismiss	Connect

Ad

Get the latest jobs and industry news



Kuang, explore relevant opportunities with Management Events

[Follow](#)

[About](#)
[Help Center](#)
[Privacy & Terms](#)

[Advertising](#)
[Business Services](#)
[More](#)

[LinkedIn](#) LinkedIn Corporation © 2017

CTRIP: Look for a Hotel - 找一家酒店

The screenshot displays the Ctrip website's hotel search page. The top navigation bar includes the Ctrip logo, links for Flights, Hotels, Trains, Vacation Packages, and Things To Do, along with a 'New' badge. On the right, there are links for Help, SGD, EN, a Sign in button, and an App icon.

The main search area on the left features a sidebar with 'Flights', 'Hotels', and 'Trains' options. The 'Hotels' section is active, showing a search form with fields for 'Destination/Hotel Name' (with a placeholder 'city, airport, region, landmark or hotel name'), 'Check-in' (mm-dd-yyyy), and 'Check-out' (mm-dd-yyyy). Below these are filters for 'Price per night (SGD)' and 'Star rating' (1 to 5 stars).

A large promotional banner for 'Save Big this Spring in Taiwan' is featured, offering 'Up to SGD 30 OFF Your Accommodation' and a 'Get Promo Code' button. The banner includes illustrations of the Taipei skyline and cherry blossoms.

At the bottom, there is a section for 'Get the FREE Ctrip app now' with a QR code and a 'Scan to Download' button. Below this, a 'Travel at your fingertips' section lists benefits: competitive rates, nearby hotel location, real-time flight tracking, train ticket booking in China, and on-the-go travel bookings.

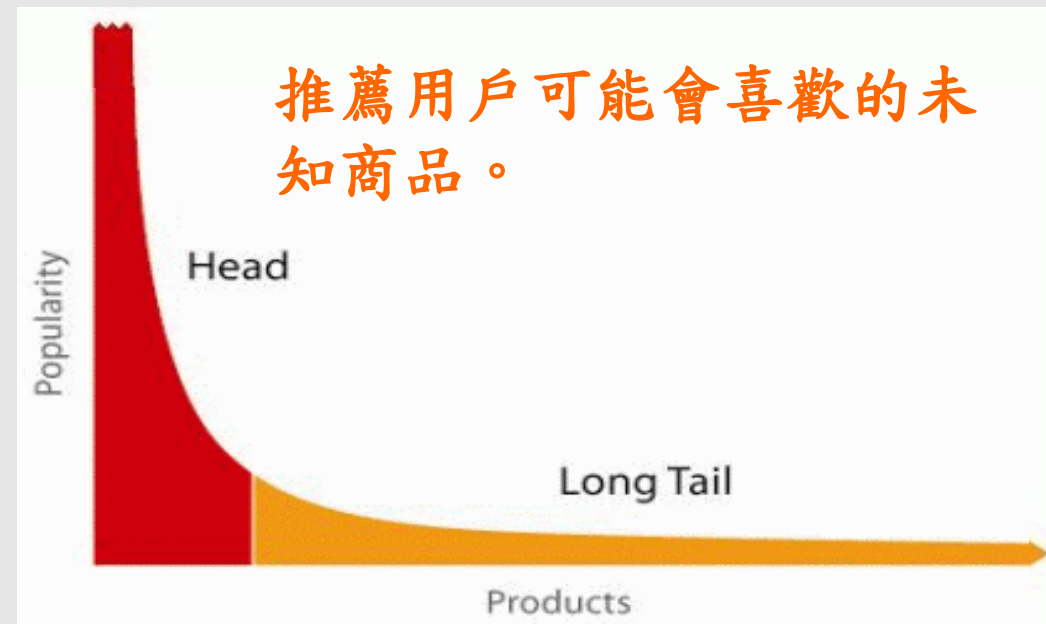
The URL at the bottom of the page is <https://pages.english.ctrip.com/event/taiwan-spring-tour/sg/>.

推薦的價值和成功的因素

□ 推薦系統可以幫助用戶找到他們感興趣的產品或服務，並增加用戶消費。

- Netflix: 70% 用戶觀看的電影。
- Amazon: 35% 的銷售。
- Google news: 38% 更多的點擊率。

□ 企業可以藉由推薦系統增加更多的用戶及營收，前提是企業要有大量的客戶消費數據，才能有效發揮推薦系統的價值。



《 Into Thin Air 》 and 《 Touching the Void 》

- ❑ 《Touching the Void 》,是一本登山書,在1988年首次出版時並不是很受歡迎。
- ❑ 《Into Thin Air 》,也是一本登山書,是1998年的暢銷書。
- ❑ 在1998年,亞馬遜的推薦系統注意到有幾個人同時購買了這兩本書,因此,開始推薦 《Touching the Void 》 給那些已經購買或正在考慮購買 《 Into Thin Air 》 的人。
- ❑ 《Touching the Void 》 最終變得非常受歡迎,實際上比 《 Into Thin Air 》 更受歡迎。
 - 基於這本書拍的電影 《 Touching the Void 》 也被列入PBS 史上最偉大的百部紀錄片之一。
- ❑ 今天,亞馬遜仍然以\$29.99美元的價格出售 《Touching the Void 》 。

推薦系統的歷史

- 在20世紀90年代早期，協同過濾開始成為處理網路訊息空間過載的解決方案。
- GroupLens 使用這種技術來識別特定用戶感興趣的Usenet文章。
- 在20世紀90年代後期，推薦技術的商業部署開始出現。（例如Amazon.com）
- 推薦的技術也已經超越了協同過濾的範疇，包括基於內容的方法，貝葉斯 (Bayes) 推理, 和基於案例的推理方法等等。
- 隨著各種推薦策略的成熟，混合推薦系統也出現了，它將多種算法結合到一個整合系統中，以利用其中各個算法的優勢。
- 當Netflix推出Netflix挑戰賽以改善電影推薦算法時，2006年對推薦算法的研究引起了極大關注。

Netflix的挑戰賽

- 2006年，Netflix舉行一個挑戰賽。任何一個人或團隊能夠率先提高Netflix內部使用的推薦算法-CineMatch 10%準確度，就可以獲得100萬美元的獎金。經過三年多的比賽，該獎項於2009年9月頒發給由AT & T Labs和GmbH組成的BellKor團隊。。
- Netflix的挑戰賽包括一個訓練(Training)數據集， 裡面包含大約17,000部電影和大約50萬用戶的評分。此外，還有測驗(Quiz)數據集和測試(Test)數據集。
- 已公布數據集中每對（用戶，看過的電影）的數據包括評分（1-5星）和評分的日期。
- 關於這個挑戰賽的一些有趣的，也許是不顯而易見的事實：ka
 - UV分解算法的準確度比CineMatch提高了7%。
 - 獲獎的算法實際上是獨立開發的幾種不同算法的組合。
 - 一些團隊嘗試使用IMDB中包含的數據；但是，發現類型和其他信息是對準確度提高沒有幫助的。
 - 評分的時間變得有用。

推薦系統的目標

- 主要目標是增加商品銷售
- 但是，我們經常無法判斷銷售的增加是否與推薦系統直接相關，因此，我們通常是利用較不明顯的運營和技術目標來當作推薦系統的目標：
 - 是準確的
 - 有關聯的
 - 是新奇的
 - 帶來驚喜的
 - 增加推薦多樣性的



一些真實的推薦系統例子

System	Product Goal
Amazon.com	書籍和其他產品
博客來 books.com.tw	書籍和其他產品
YouTube	線上頻道
Netflix	線上頻道
Pandora	音樂
KKBOX	音樂
Google News	新聞
Facebook	朋友、廣告
momo購物網	線上購買產品
IMDB	電影
Expedia	旅遊產品

推薦系統的類型

- ❑ 非個性化 Non-personalized
- ❑ 人口特點（半個性化） Demographic (semi-personalized)
- ❑ 協同過濾 Collaborative filtering
- ❑ 基於內容的 Content-based
- ❑ 基於知識的 Knowledge-based
- ❑ 混合的 Hybrid
- ❑ 先進的推薦系統 Advanced Recommender Systems

非個性化推薦

□ 每個用戶相同。

□ 通常推薦最受歡迎的商品。
(基於平均評級或銷售數量)

Top10 APRIL PROMO ITEMS

2018
Cathay Big Data Competition



人口特點（半個性化）推薦

- 與同一個人口統計組中的每個用戶相同。
- 這些推薦通常基於同一人口組中最受歡迎的商品。
(基於平均評級或銷售數量)。

人口特點：

人口、年齡、性別、教育、收入、家庭規模、種族 ...等

協同過濾推薦

(COLLABORATIVE FILTERING RECOMMENDATION)

□ 最突出的推薦方法：

- 大型商業電子商務網站經常使用。
- 很好理解，存在各種算法和變化。
- 適用於許多領域（書籍，電影，DVD，..）。

買了這個商品的其他顧客
也買了這個商品

□ 方法：

- 使用“人群的智慧”來推薦商品

□ 基本假設和想法：

- 用戶給商品的評分（隱式或顯式）。
- 過去曾經有類似喜好的顧客將來也會有類似的喜好。

協同過濾推薦 - 亞馬遜釋例

Customers who bought this item also bought 買了這個的顧客還買了

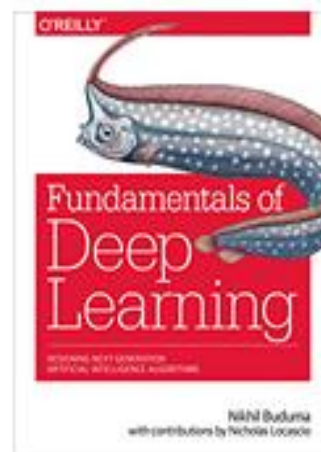


Hands-On Machine Learning with Scikit-Learn and TensorFlow:...

› Aurélien Géron

★★★★★ 116

#1 Best Seller in Business Software

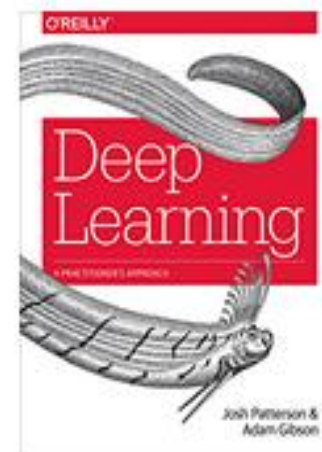


Fundamentals of Deep Learning: Designing Next-Generation Machine...

› Nikhil Buduma

★★★★★ 24

Kindle Edition
\$22.49

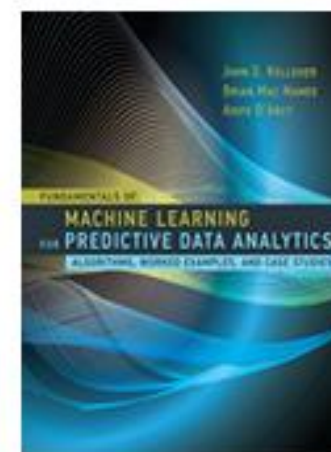


Deep Learning: A Practitioner's Approach

› Josh Patterson

★★★★★ 9

Kindle Edition
\$24.99



Fundamentals of Machine Learning for Predictive Data Analytics:...

› John D. Kelleher

★★★★★ 25

Kindle Edition
\$63.99

基於內容的推薦 (Content-Based Recommendation)

□ 雖然CF方法不需要關於這些商品的任何訊息：

- 利用這些訊息可能是合理的。
- 例如：向過去喜歡科幻電影的人推薦科幻電影。

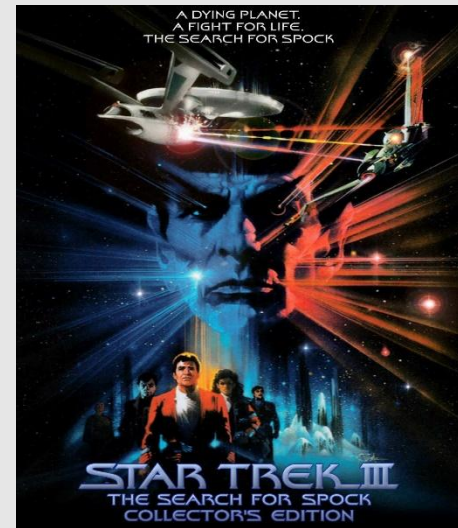


□ 我們需要什麼：

- 有關商品的一些訊息，例如類型（內容）。
- 某種描述用戶喜歡商品的文件（偏好）。

□ 解決方案：

- 學習用戶偏好。
- 找到/推薦與用戶自己的偏好“相似”的商品。
- 使用許多文本挖掘技術。



基於知識的推薦 (Knowledge-Based Recommender Systems)

□ 產品可用的評分數量少



□ 時間是重要的影響因子、時間的影響極為重要：

- 已經超過了五年的電腦評分。
- 用戶生活方式或家庭狀況的變化。

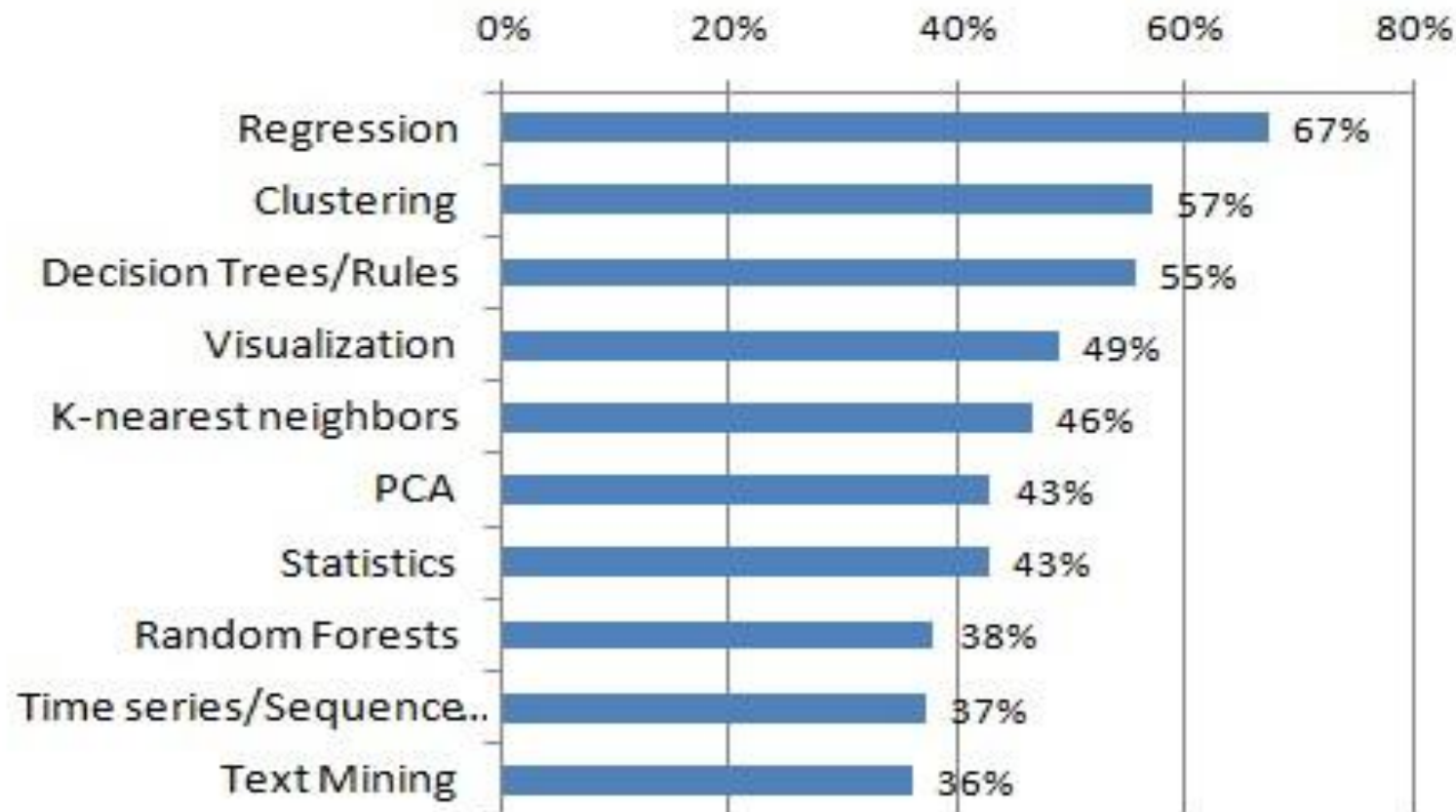
□ 客戶希望明確定義他們的要求：

- 商品通常比較複雜，例如：汽車的顏色應該是黑色的。

推薦系統的比較

推薦方法	概念目標	輸入
協同過濾	一個基於我和其他人對於某些商品的評價/行動的合作格式	用戶評分和社群評分
基於內容	基於我過去的評價/行動和產品內容的格式	用戶評分和商品屬性
基於知識	一個基於我對這些商品的具體要求的格式	用戶特定的要求， 商品屬性和領域知識

Top 10 Algorithms & Methods used by Data Scientists



大多數機器學習算法可用於解決推薦問題。
根據問題的類型，有些算法會比其他更好。

補

充



時

間

先進的推薦系統

□ 情境敏感的推薦系統

(Context-Sensitive Recommender Systems)

□ 社群和以信任為中心的推薦系統

(Social and Trust-Centric Recommender Systems)

□ 防攻擊推薦系統

(Attack-Resistant Recommender Systems)

□ 群組推薦系統

(Group Recommender Systems)

□ 計算廣告

(Computational Advertising)