

Engineers who play with Data

...

Recommendation System

三种人、三种需求

1. Data Analyst: *Presentation*
 - SQL, Excel, Data Visualization
2. Data Engineer: *System Development*
 - SQL, Python (R), Bash, Java
3. Data Scientist: *Improve Algorithm*
 - SQL, Python (R), Math, Math, Math



一个服务、两套系统 → 一套中间件

EC Sites

1. Web System

- a. Database (MySQL, PostgreSQL, NoSQL, etc.)



ETL^{*1} (Extraction, Transformation, and Loading)

2. Data System

- a. Data Warehouse (Redshift, etc.)

*1. ETL: The process of extracting data from source systems and bringing it into the data warehouse is commonly called ETL, which stands for extraction, transformation, and loading.

Recommendation System

Amazon: 35% sales from recommendations

Netflix: $\frac{2}{3}$ of the movies watched are recommended

1. Collaborative Filtering

- User-based: Find similar users to me and recommend what they liked
- Item-based: Find similar items to those that I have previously liked

	Item1	Item2	Item3	Item4	Item5
Alice	5	3	4	4	?
User1	3	1	2	3	3
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1

2. Content Filtering

Recommendation System: CF

Limitations:

1. Cold Start:

- There needs to be enough other users already in the system to find a match. New items need to get enough ratings.

2. Popularity Bias:

- Hard to recommend items to someone with unique tastes.
- Tends to recommend popular items (items from the tail do not get so much data)!!!

Recommendation System: Evaluation

1. 番外：

- <https://zhuanlan.zhihu.com/p/21622055>
- 一个推荐系统的好坏不仅仅取决于有一个更好或更坏的CTR, CVR, CXR !

↓

2. 卖推荐系统是没有前途的！好的推荐系统只存在于大公司，有钱任性的公司。
- 日本众多(越来越多)只卖「レコメンドエンジン」的公司。

Recommendation System: A/B Testing

一个不会让 Web Engineer 骂娘的 A/B Testing 系统

+

大量A/B Test

+

经典推荐算法

=

超出世界平均线，甚至可以说不错的推荐效果。

Recommendation System: 好事？坏事？

The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think - Eli Pariser, 2011

A world constructed from the familiar is a world in which there's nothing to learn ... (since there is) invisible auto-propaganda, indoctrinating us with our own ideas.

- Eli Pariser in *The Economist*, 2011

