

第一章、推荐系统简介

讲师：武永亮



教学目标

- 了解推荐系统
- 理解推荐系统的应用的的设计理念
- 了解推荐系统评测方法

目录

1 什么是推荐系统

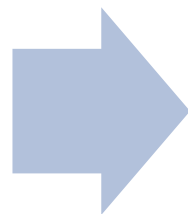
2 个性化推荐系统的应用

3 推荐系统评测

1.1 什么是推荐系统

- 如何找到自己想要的商品？
 - 如何挑选一款电脑？
 - 如何挑选一部电影？

信息匮乏

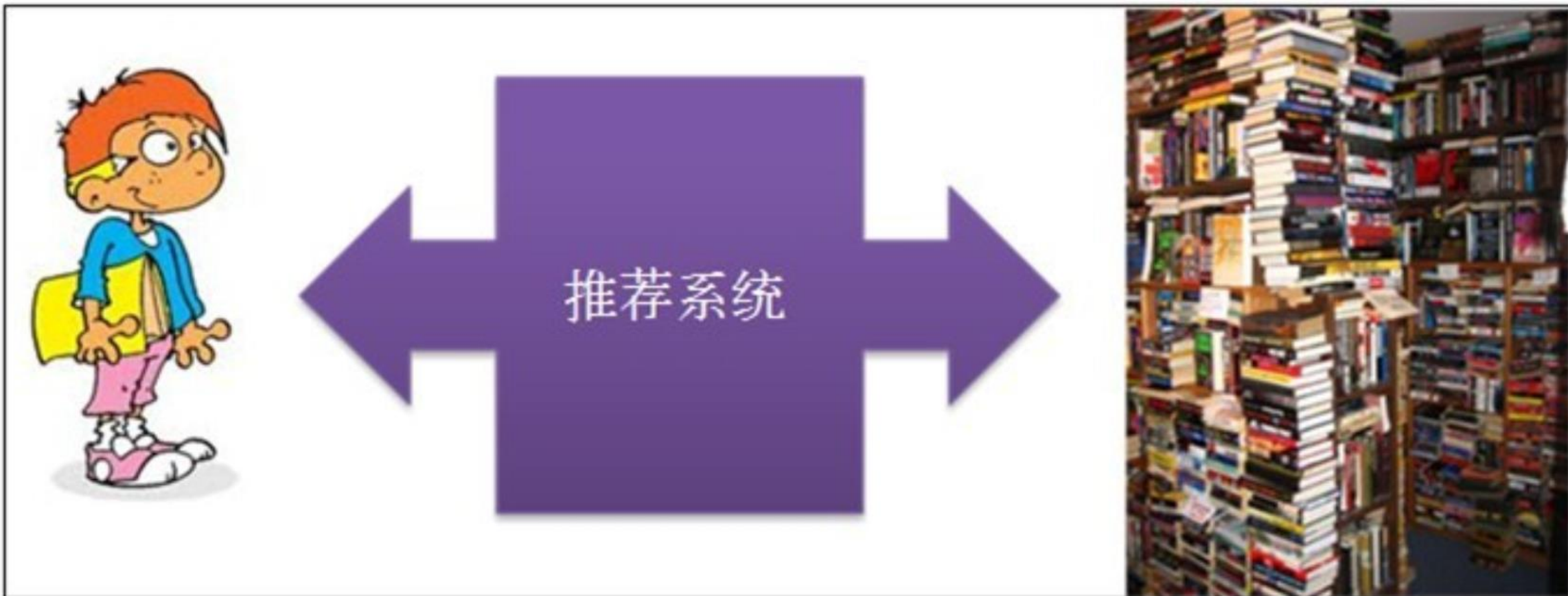


信息过载

1.1 什么是推荐系统

- 为了解决信息过载问题
 - 分类目录：雅虎、DMOZ、Hao123
 - 搜索引擎：谷歌、百度
- 推荐系统也是一种帮助用户快速发现有用信息的工具。和搜索引擎不同的是，推荐系统不需要用户提供明确的需求，而是通过分析用户的历史行为给用户的兴趣建模，从而主动给用户推荐能够满足他们兴趣和需求的信息

1.1 什么是推荐系统



1.1 什么是推荐系统

- 现实社会中用户面对很多选择时做决定的过程：
 - 向朋友咨询
 - 个人喜好：喜欢的演员和导演
 - 综合排名

1.1 什么是推荐系统

- 推荐算法的本质是通过一定的方式将用户和物品联系起来，而不同的推荐系统利用了不同的方式。



目录

1 什么是推荐系统

2 个性化推荐系统的应用

3 推荐系统评测

1.2 个性化推荐系统的应用

- 和搜索引擎不同，个性化推荐系统需要依赖用户的**行为数据**，因此一般都是作为一个应用存在于不同网站之中。在互联网的各类网站中都可以看到推荐系统的应用，而个性化推荐系统在这些网站中的主要作用是通过**分析大量用户行为日志**，给不同用户提供不同的个性化页面展示，来提高网站的点击率和转化率。
- 广泛利用推荐系统的领域包括：电子商务、电影和视频、音乐、社交网络、阅读、基于位置的服务、个性化邮件和广告等。
- 推荐系统的组成：
 - 前台的展示页面
 - 后台的日志系统
 - 推荐算法系统

1.2 个性化推荐系统的应用

- 电子商务网站是个性化推荐系统的一大应用领域。著名的电子商务网站亚马逊是个性化推荐系统的积极应用者和推广者。

Today's Recommendations For You

Here's a daily sample of items recommended for you. Click here to [see all recommendations](#). Page 6 of 44 (Start over)



[Networks, Crowds, and Markets: R...](#) (Hardcover) by David Easley
★★★★☆ (5) \$39.50
[Fix this recommendation](#)



[R Cookbook \(O'Reilly Cookbooks\)](#) (Paperback) by Paul Teetor
★★★★☆ (11) \$31.49
[Fix this recommendation](#)



[Introduction to Machine Learning](#) (Hardcover) by Ethem Alpaydin
★★★★☆ (6) \$38.00
[Fix this recommendation](#)



[Programming Collective Intelligence](#) (Paperback) by Toby Segaran
★★★★☆ (72) \$26.39
[Fix this recommendation](#)

1.2 个性化推荐系统的应用

- 个性化推荐列表采用了一种基于物品的推荐算法（item-based method），该算法给用户推荐那些和他们之前喜欢的物品相似的物品。

The screenshot shows the Amazon.com interface with a 'Recommended for You' section. The first recommendation is the book 'Introduction to Machine Learning (Adaptive Computation and Machine Learning series)' by Ethem Alpaydin. It includes a 'LOOK INSIDE!' button, a book cover, the title, author, price (\$46.33), and a note that used and new copies start from \$31.97. There are 'Add to Cart' and 'Add to Wish List' buttons. To the right, there is a 'Rate this item' section with a star rating (5 stars selected) and two checkboxes: 'I own it' and 'Not interested'. A red arrow points from the text '对推荐结果的反馈方式' (Feedback method for recommendation results) to the 'Not interested' checkbox. Below this, the 'Because you rated...' section shows two more book recommendations: 'Probabilistic Graphical Models: Principles and Techniques (Adaptive Computation and Machine Learning series)' by Daphne Koller and Nir Friedman, and 'Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)' by Ian H. Witten and Eibe Frank. Each recommendation includes a book cover, title, author, and a star rating (5 stars selected). There are also checkboxes for 'Don't use for recommendations'.

amazon.com[®] [Help](#) | [Close window](#)

Recommended for You

Introduction to Machine Learning (Adaptive Computation and Machine Learning series)
by Ethem Alpaydin (Author)
Our Price: \$46.33
Used & new from \$31.97
[Add to Cart](#) [Add to Wish List](#)

Rate this item
☒ ★★★★★
☐ I own it
☐ Not interested

Because you rated...

Probabilistic Graphical Models: Principles and Techniques (Adaptive Computation and Machine Learning series)
(Hardcover)
by Daphne Koller (Author), Nir Friedman (Author)
☒ ★★★★★
☐ Don't use for recommendations

Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems) (Paperback)
by Ian H. Witten (Author), Eibe Frank (Author)
☒ ★★★★★
☐ Don't use for recommendations

[Help](#) | [Close window](#)

1.2 个性化推荐系统的应用

- 亚马逊还有另外一种个性化推荐列表，就是按照用户在Facebook的好友关系，给用户推荐他们的好友在亚马逊上喜欢的物品。

Popular Among Your Friends on Facebook

Page 1 of 12

The screenshot shows a carousel of product recommendations. The first item is 'The Beatles 1' with a red cover and a large yellow number '1'. The second is 'The Matrix DVD' featuring Keanu Reeves. The third is 'NINETEEN EIGHTY-FOUR' by George Orwell, which has a 'No image available' placeholder. The fourth is 'Michael' by Michael Jackson. The fifth is 'Harry Potter Hardcover Boxed Set' by J.K. Rowling. Each item card displays a star rating, the number of friends who liked it, and the price. Navigation arrows are visible on the left and right sides of the carousel.

Item	Price	Rating	Friends Liked
The Beatles 1 ~ The Beatles	\$13.88	★★★★☆ (1,189)	7 friends like this:
The Matrix DVD ~ Keanu Reeves	\$5.49	★★★★☆ (3,019)	6 friends like this:
NINETEEN EIGHTY-FOUR by George Orwell	-	★★★★☆ (1,668)	4 friends like this:
Michael ~ Michael Jackson	\$13.35	★★★★☆ (228)	6 friends like this:
Harry Potter Hardcover Boxed Se... by J. K. Rowling	\$114.07	★★★★☆ (857)	4 friends like this:

1.2 个性化推荐系统的应用

- 除了个性化推荐列表，亚马逊另一个重要的推荐应用就是相关推荐列表。

Customers Who Bought This Item Also Bought

Page 1 of 17

					
CaseCrown Faux Suede Case (Navy) for Pandigital Novel Digital Boo...	HDE 7" Tablet Stand with USB Keyboard - Black Faux Leather Carry...	3 Pack of Universal Touch Screen Stylus Pen (Red +... by Bargaincell	CaseCrown Faux Suede Case (Red) for Pandigital Novel Digital Boo...	3 Pack of Premium Crystal Clear Screen Protectors for Apple iPad	CaseCrown Faux Suede Case (Fuschia) for Pandigital Novel Digital...
★★★★☆ (4)	★★★★☆ (15)	★★★★☆ (636)	★★★★☆ (5)	★★★★☆ (539)	★★★★☆ (9)
\$5.21	\$15.48	\$1.99	\$5.21	\$1.95	\$11.21

1.2 个性化推荐系统的应用

- 除了个性化推荐列表，亚马逊另一个重要的推荐应用就是相关推荐列表。

What Other Items Do Customers Buy After Viewing This Item?



3 Pack of Universal Touch Screen Stylus Pen (Red + Black + Silver) by Bargaincell

★★★★☆ (636)

\$1.99



Summer of Fire (Yellowstone series) by Linda Jacobs Kindle Edition

★★★★★ (5)

\$0.00

SUPER SALE 99



USA SELLER

New 7 inch android 2.2 VIA8650 epad tablet pc 3.0MP Camera, SUPPORTS EXTERNAL 3G MODEM+WIFI. Screen+Flash 10.1 with PRIORITY SHIPPING FROM SUPERSALE.99 USA SELLER by SuperPad

★★★★☆ (28)

\$89.99



SKYTEX Skypad Alpha 7" Touch Screen Cortex-A8 Tablet Android OS 2.3 by SKYTEX

★★★★★ (6)

\$150.92

> [Explore similar items](#)


1.2 个性化推荐系统的应用

- 除了个性化推荐列表，亚马逊另一个重要的推荐应用就是相关推荐列表。

Frequently Bought Together



Price For All Three: **\$108.05**

 Add all three to Cart

Add all three to Wish List

[Show availability and shipping details](#)

- ☒ **This item:** R Cookbook (O'Reilly Cookbooks) by Paul Teetor Paperback **\$25.95**
- ☒ R in a Nutshell: A Desktop Quick Reference by Joseph Adler Paperback **\$35.99**
- ☒ ggplot2: Elegant Graphics for Data Analysis (Use R) by Hadley Wickham Paperback **\$46.11**

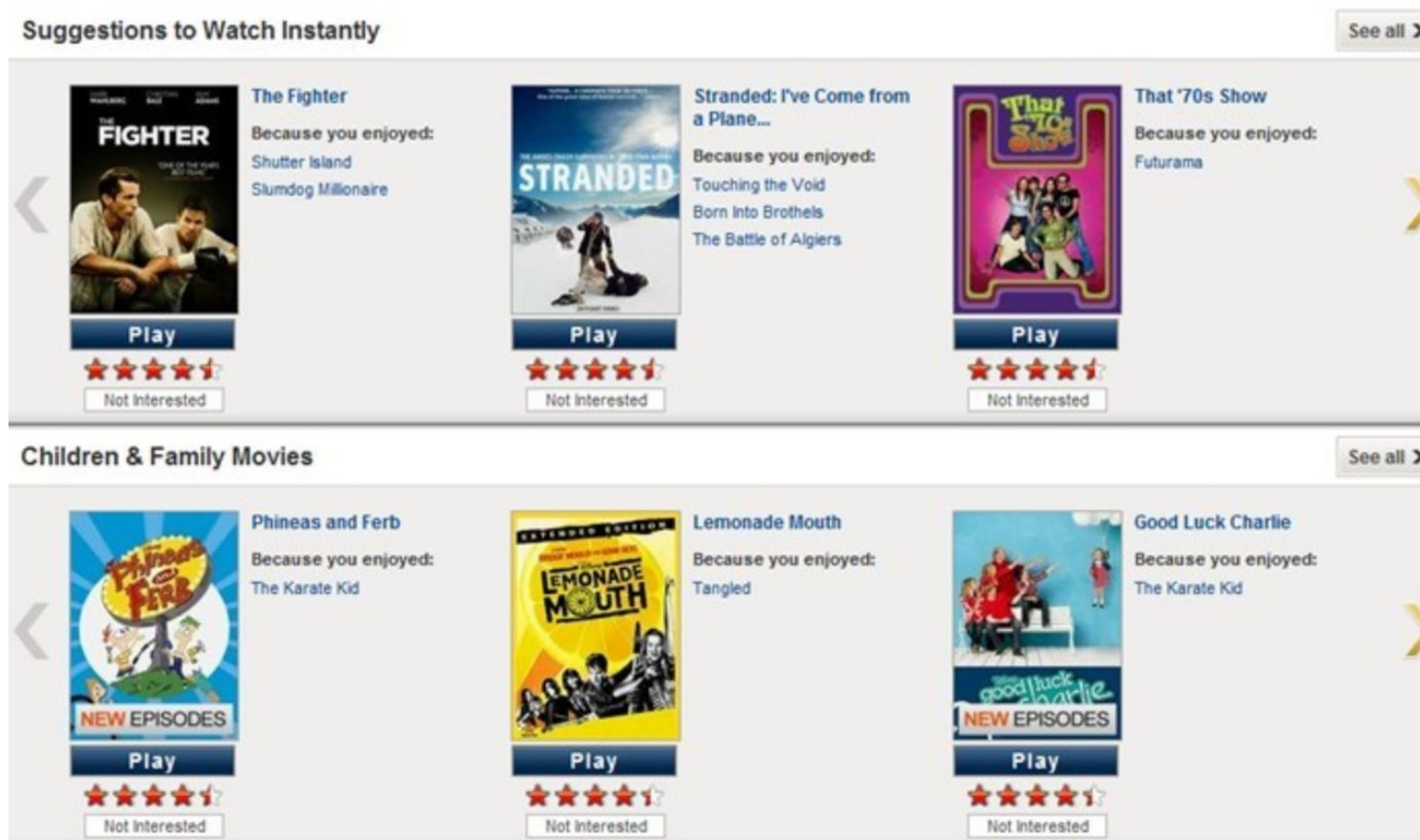
1.2 个性化推荐系统的应用

- 这些推荐的应用，究竟给亚马逊带来了多少商业利益？

亚马逊有20% ~ 30%的销售来自于推荐系统。

1.2 个性化推荐系统的应用

- 在电影和视频网站中，个性化推荐系统也是一种重要的应用。在该领域成功使用推荐系统的一家公司就是Netflix，它和亚马逊是推荐系统领域最具代表性的两家公司。



1.2 个性化推荐系统的应用

- 个性化推荐的成功应用需要两个条件。第一是存在信息过载，因为如果用户可以很容易地从所有物品中找到喜欢的物品，就不需要个性化推荐了。第二是用户大部分时候没有特别明确的需求，因为用户如果有明确的需求，可以直接通过搜索引擎找到感兴趣的物品。

1.2 个性化推荐系统的应用

- 其他推荐系统的应用场景：
 - 个性化音乐网络电台：Pandora个性化网络电台、豆瓣个性化网络电台
 - 社交网络：Facebook、Twitter、QQ
 - 个性化阅读：国际知名的有Google Reader、国内有鲜果网等
 - 基于位置的服务：微信附近的人、美团外卖附近的店
 - 个性化邮件：Gmail的优先级邮箱
 - 个性化广告：广告的精准投放

目录

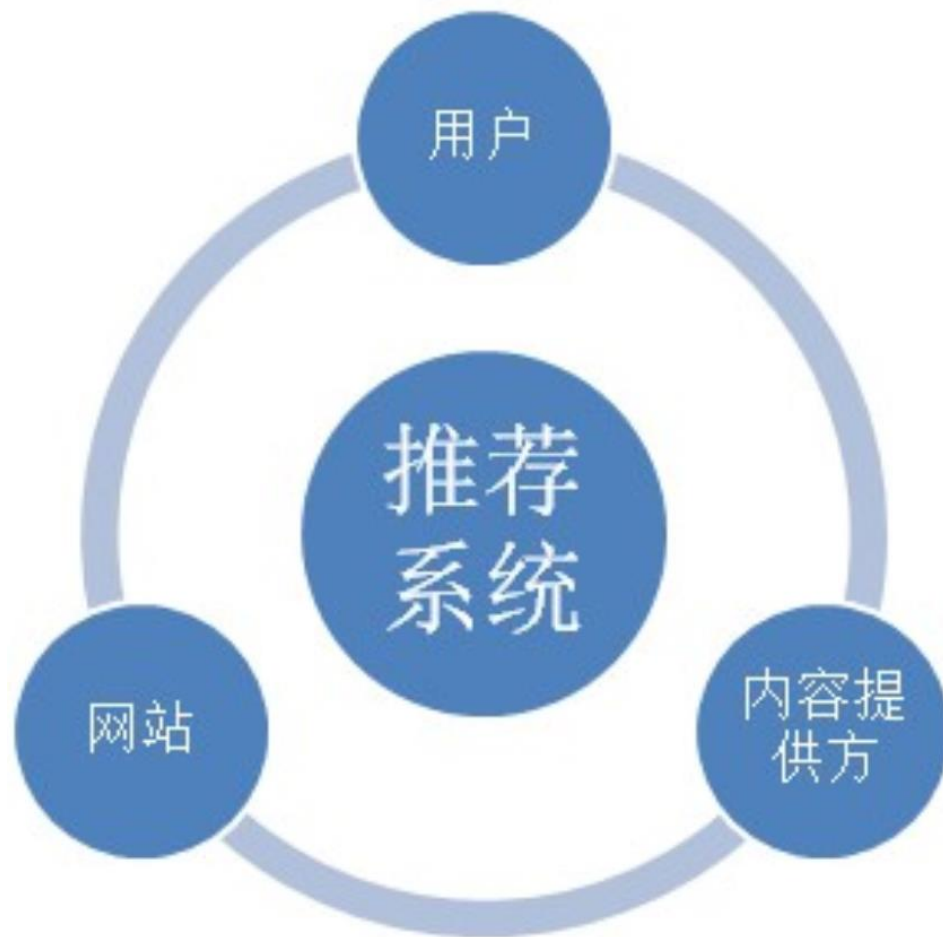
1 什么是推荐系统

2 个性化推荐系统的应用

3 推荐系统评测

1.3 推荐系统评测

- 如何评价推荐系统？这是推荐系统评测需要解决的首要问题。一个完整的推荐系统一般存在3个参与方：用户、物品提供者和提供推荐系统的网站。



1.3 推荐系统评测

- 以图书推荐为例，好的推荐系统需要满足：
 - 首先，推荐系统需要满足**用户**的需求，给用户推荐那些令他们感兴趣的图书。
 - 其次，推荐系统要让各**出版社**的书都能够被推荐给对其感兴趣的用户，而不是只推荐几个大型出版社的书。
 - 最后，好的推荐系统设计，能够让推荐系统本身收集到高质量的用户反馈，不断完善推荐的质量，增加用户和网站的交互，**提高网站的收入**。
- 好的推荐系统不仅仅能够**准确预测用户的行为**，而且能够扩展用户的视野，帮助用户发现那些他们可能会感兴趣，但**却不那么容易发现**的东西。

1.3 推荐系统评测

- 在推荐系统中，主要有3种评测推荐效果的实验方法：
 - 即离线实验 (offline experiment)
 - 用户调查 (user study)
 - 在线实验 (online experiment)

1.3 推荐系统评测

- **1.离线实验**的方法一般由如下几个步骤构成：
 - 通过日志系统获得用户行为数据，并按照一定格式生成一个标准的数据集；
 - 将数据集按照一定的规则分成训练集和测试集；
 - 在训练集上训练用户兴趣模型，在测试集上进行预测；
 - 通过事先定义的离线指标评测算法在测试集上的预测结果。
- 优缺点

优点	缺点
不需要有对实际系统的控制权	无法计算商业上关心的指标
不需要用户参与实验	离线实验的指标和商业指标存在差距
速度快，可以测试大量算法	

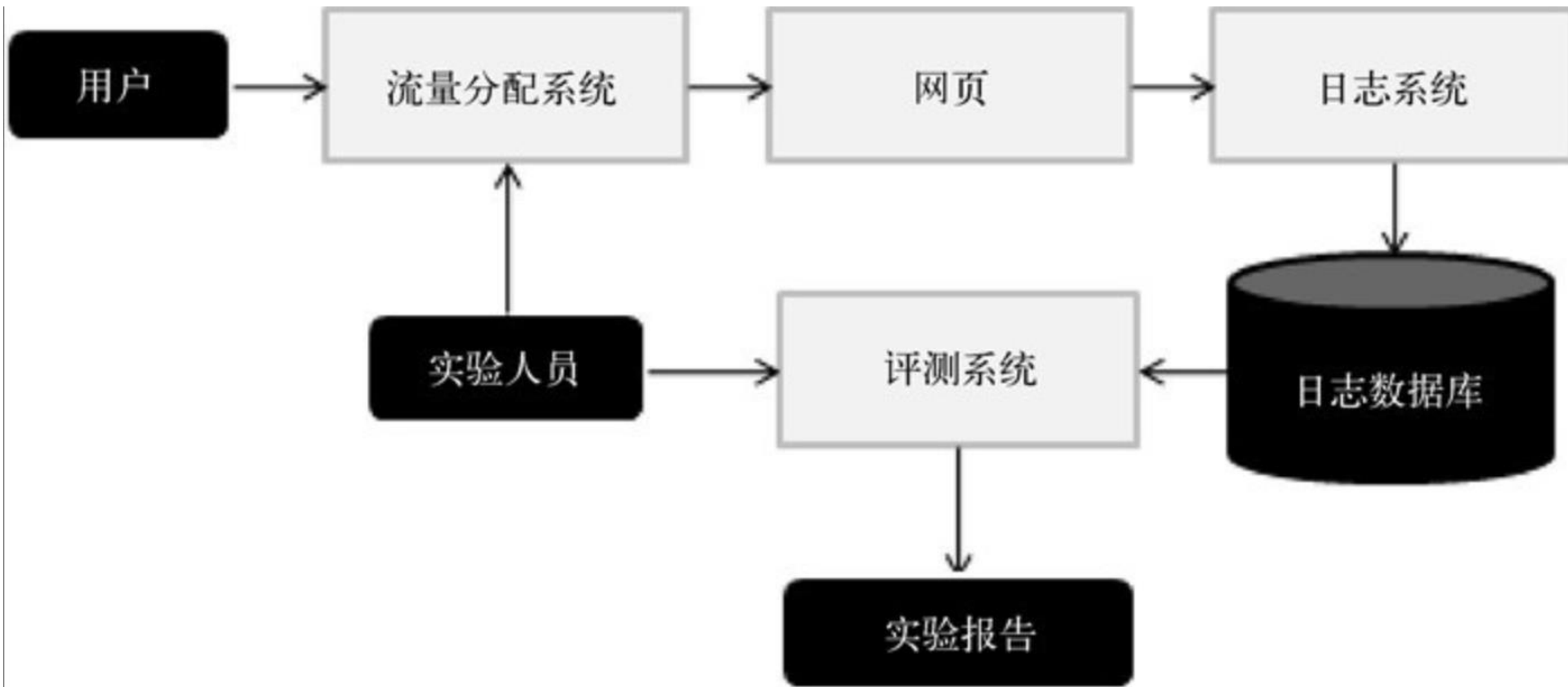
1.3 推荐系统评测

- **2.用户调查**需要有一些真实用户，让他们在需要测试的推荐系统上完成一些任务。在他们完成任务时，我们需要观察和记录他们的行为，并让他们回答一些问题。最后，我们需要通过分析他们的行为和答案了解测试系统的性能。
- 优缺点
 - 优点是可以获得很多体现用户主观感受的指标，相对在线实验风险很低，出现错误后很容易弥补。
 - 缺点是招募测试用户代价较大，很难组织大规模的测试用户，因此会使测试结果的统计意义不足。此外，在很多时候设计双盲实验非常困难，而且用户在测试环境下的行为和真实环境下的行为可能有所不同，因而在测试环境下收集的测试指标可能在真实环境下无法重现。

1.3 推荐系统评测

- **3.在线实验顾名思义就是在完成离线实验和必要的用户调查后，可以将推荐系统上线做AB测试，将它和旧的算法进行比较。**
- AB测试是一种很常用的在线评测算法的实验方法。它通过一定的规则将用户随机分成几组，并对不同组的用户采用不同的算法，然后通过统计不同组用户的各种不同的评测指标比较不同算法，在完成离线实验和必要的用户调查后，可以将推荐系统上线做AB测试，将它和旧的算法进行比较。
- 优缺点
 - 优点是可以公平获得不同算法实际在线时的性能指标，包括商业上关注的指标。AB测试的缺点主要是周期比较长，必须进行长期的实验才能得到可靠的结果。
 - 缺点是，一个大型网站的AB测试系统的设计也是一项复杂的工程。一个大型网站的架构分前端和后端，从前端展示给用户的界面到最后端的算法，中间往往经过了很多层，这些层往往由不同的团队控制，而且都有可能做AB测试。如果为不同的层分别设计AB测试系统，那么不同的AB测试之间往往会互相干扰。

1.3 推荐系统评测



1.3 推荐系统评测

- 一般来说，一个新的推荐算法最终上线，需要完成上面三个步骤：
 - 首先，需要通过离线实验证明它在很多离线指标上优于现有的算法。
 - 然后，需要通过用户调查确定它的用户满意度不低于现有的算法。
 - 最后，通过在线的AB测试确定它在我们关心的指标上优于现有的算法。

1.3 推荐系统评测

- 这些评测指标可用于评价推荐系统各方面的性能。这些指标有些可以定量计算，有些只能定性描述，有些可以通过离线实验计算，有些需要通过用户调查获得，还有些只能在线评测。接下来出不同的评测指标：
 - 用户满意度
 - 预测准确度
 - 覆盖率
 - 多样性
 - 新颖性
 - 惊喜度
 - 信任度
 - 实时性
 - 健壮性
 - ...

1.3 推荐系统评测

- 用户满意度:用户作为推荐系统的重要参与者，其满意度是评测推荐系统的最重要指标。但是，用户满意度没有办法离线计算，只能通过**用户调查**或者**在线实验**获得。



Starhunter 2300

Recommended because you
have shown interest in Total
Recall 2070

Are you interested?

yes

no

I've seen it



Brooklyn Rules

Recommended because you
have shown interest in One
Eyed King

Are you interested?

yes

no

I've seen it

1.3 推荐系统评测

- 用户满意度:用户作为推荐系统的重要参与者，其满意度是评测推荐系统的最重要指标。但是，用户满意度没有办法离线计算，只能通过**用户调查**或者**在线实验**获得。



1.3 推荐系统评测

- 预测准确度度量一个推荐系统或者推荐算法预测用户行为的能力。这个指标是最重要的推荐系统离线评测指标，从推荐系统诞生的那一天起，几乎99%与推荐相关的论文都在讨论这个指标。这主要是因为该指标可以通过离线实验计算，方便了很多学术界的研究人员研究推荐算法。
- 计算步骤：
 - 获取离线的数据集，该数据集包含用户的历史行为记录。
 - 将该数据集通过时间分成训练集和测试集。
 - 通过在训练集上建立用户的行为和兴趣模型预测用户在测试集上的行为，并计算预测行为和测试集上实际行为的重合度作为预测准确度。

1.3 推荐系统评测

- 由于离线的推荐算法有不同的研究方向，常见的预测准确度指标。
- 评分预测



1.3 推荐系统评测

- 由于离线的推荐算法有不同的研究方向，常见的预测准确度指标。
- 评分预测的预测准确度一般通过均方根误差（RMSE）和平均绝对误差（MAE）计算。

对于测试集中的一个用户 u 和物品 i ，令 r_{ui} 是用户 u 对物品 i 的实际评分，而 \hat{r}_{ui} 是推荐算法给出的预测评分，那么RMSE的定义为：

$$\text{RMSE} = \frac{\sqrt{\sum_{u,i \in T} (r_{ui} - \hat{r}_{ui})^2}}{|T|}$$

MAE采用绝对值计算预测误差，它的定义为：

$$\text{MAE} = \frac{\sum_{u,i \in T} |r_{ui} - \hat{r}_{ui}|}{|T|}$$

1.3 推荐系统评测

- 覆盖率（coverage）描述一个推荐系统对物品长尾的发掘能力。覆盖率有不同的定义方法，最简单的定义为推荐系统能够推荐出来的物品占总物品集合的比例。假设系统的用户集合为 U ，推荐系统给每个用户推荐一个长度为 N 的物品列表 $R(u)$ 。那么推荐系统的覆盖率可以通过下面的公式计算：

$$\text{Coverage} = \frac{|\bigcup_{u \in U} R(u)|}{|I|}$$

1.3 推荐系统评测

- 多样性：为了满足用户广泛的兴趣，推荐列表需要能够覆盖用户不同的兴趣领域，即推荐结果需要具有多样性。
- 多样性描述了推荐列表中物品两两之间的不相似性。因此，多样性和相似性是对应的。

假设 $s(i, j) \in [0, 1]$ 定义了物品 i 和 j 之间的相似度，那么用户 u 的推荐列表 $R(u)$ 的多样性定义如下：

$$\text{Diversity} = 1 - \frac{\sum_{i, j \in R(u), i \neq j} s(i, j)}{\frac{1}{2} |R(u)| (|R(u)| - 1)}$$

而推荐系统的整体多样性可以定义为所有用户推荐列表多样性的平均值：

$$\text{Diversity} = \frac{1}{|U|} \sum_{u \in U} \text{Diversity}(R(u))$$

1.3 推荐系统评测

- 新颖的推荐是指给用户推荐那些他们以前没有听说过的物品。在一个网站中实现新颖性的最简单办法是，把那些用户之前在网站中对其有过行为的物品从推荐列表中过滤掉。

1.3 推荐系统评测

- 惊喜度（ serendipity ）是最近这几年推荐系统领域最热门的话题。
- 惊喜度与新颖性有什么区别是首先需要弄清楚的问题。注意，这里讨论的是惊喜度和新颖度作为推荐指标在意义上的区别，而不是这两个词在中文里的含义区别，所以我们首先要摒弃大脑中关于这两个词在中文中的基本含义。

1.3 推荐系统评测

- 实时性：在很多网站中，因为物品（新闻、微博等）具有很强的时效性，所以需要在物品还具有时效性时就将它们推荐给用户。比如，给用户推荐昨天的新闻显然不如给用户推荐今天的新闻。因此，在这些网站中，推荐系统的实时性就显得至关重要。

1.3 推荐系统评测

- 健壮性：任何一个能带来利益的算法系统都会被人攻击，这方面最典型的例子就是搜索引擎。搜索引擎的作弊和反作弊斗争异常激烈，这是因为如果能让自己的商品成为热门搜索词的第一个搜索果，会带来极大的商业利益。推荐系统目前也遇到了同样的作弊问题，而健壮性（即 robust, 鲁棒性）指标衡量了一个推荐系统抗击作弊的能力。

1.3 推荐系统评测

- 商业目标：很多时候，网站评测推荐系统更加注重网站的商业目标是否达成，而商业目标和网站的盈利模式是息息相关的。一般来说，最本质的商业目标就是平均一个用户给公司带来的盈利。不过这种指标不是很难计算，只是计算一次需要比较大的代价。因此，很多公司会根据自己的盈利模式设计不同的商业目标。

1.3 推荐系统评测

- 本节提到了很多指标，其中有些指标可以离线计算，有些只能在线获得。但是，离线指标很多，在线指标也很多，那么如何优化离线指标来提高在线指标是推荐系统研究的重要问题。关于这个问题，目前仍然没有什么定论，只是不同系统的研究人员有不同的感性认识。

	离线实验	问卷调查	在线实验
用户满意度	×	√	○
预测准确度	√	√	×
覆盖率	√	√	√
多样性	○	√	○
新颖性	○	√	○
惊喜度	×	√	×

1.3 推荐系统评测

- 推荐系统还需要通过不同维度来进行评测：
 - 用户维度 主要包括用户的人口统计学信息、活跃度以及是不是新用户等。
 - 物品维度 包括物品的属性信息、流行度、平均分以及是不是新加入的物品等。
 - 时间维度 包括季节，是工作日还是周末，是白天还是晚上等。

内容回顾

1 什么是推荐系统

2 个性化推荐系统的应用

3 推荐系统评测

Thank you !