

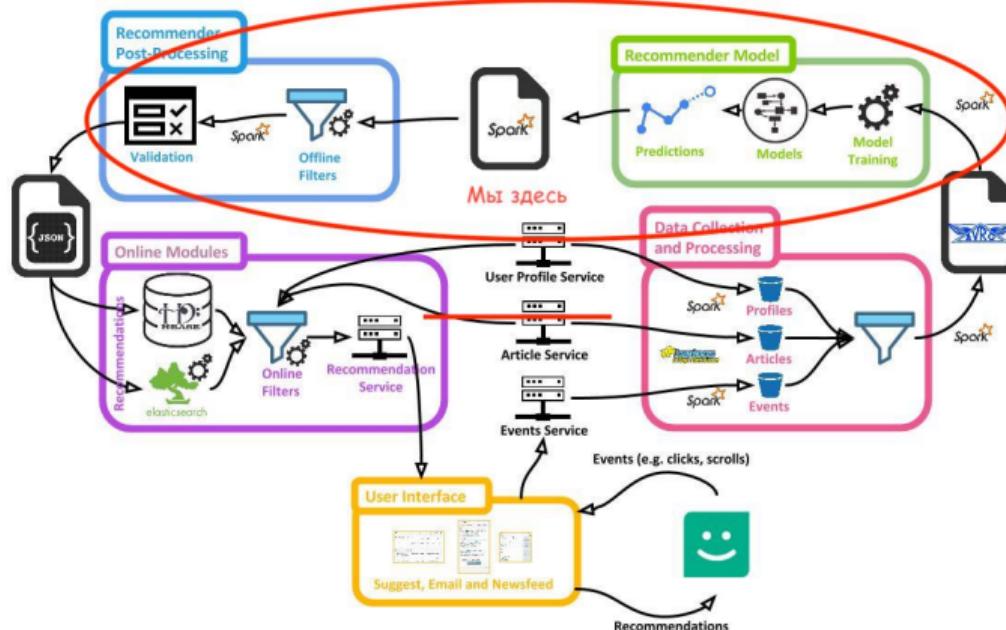
Нерешенные проблемы и новые направления

Николай Анохин

10 ноября 2023 г.



Контекст



Что мы уже умеем

$$\hat{r}_{ui} = f_{\theta}(x_u, x_i, x_c)$$



Проблемы

1. Оцениваем айтемы по-отдельности, а показываем по несколько (лентой)
2. Модель не объясняет, почему именно эти айтемы подходят пользователю
3. Смещение между распределениями на обучении и применении
4. Не учитывается долгострочный эффект рекомендаций

Разнообразие в рекомендательных системах

●ooooooooo

Объяснение рекомендаций

ooooooooooooo

Смещения

oooooooo

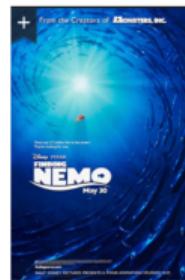
Итоги

oooo

Разнообразие в рекомендательных системах



Разнообразие / Diversity



Два подхода к разнообразию

- **Разнообразие для exploration**

помогает пользователям находить новые айтемы, а рекомендеру собирать данные

- **Разнообразие для utility**

помогает убирать избыточные айтемы



Набираем айтемы с разными аспектами

f - аспект (признак) айтема, $p(f|i)$ – вероятность найти аспект у айтема i

Распределение аспекта у пользователя

$$p(f|u) = \frac{\sum_{i \in I_u} p(f|i)}{|I_u|}$$

Распределение аспекта в рекомендациях

$$q(f|u) = \frac{\sum_{i \in RL} p(f|i)}{|RL|}$$

Формируем список так, чтобы $q(f|u)$ совпало с $p(f|u)$



Жадное переранжирование

Добавляем в список рекомендаций айтем с максимальным значением

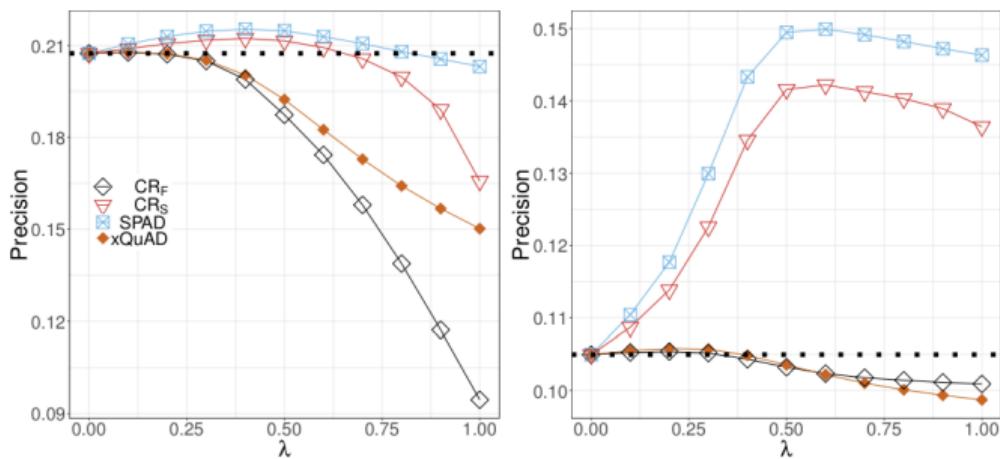
$$(1 - \lambda) \cdot s(u, i) + \lambda \cdot gain(i, RL),$$

пока не получим список нужной длины.

- $s(u, i)$ – релевантность айтема i для пользователя u
- $gain(i, RL) = div(RL \cup \{i\}) - div(RL)$ – улучшение разнообразия при добавлении айтема
- λ – гиперпараметр

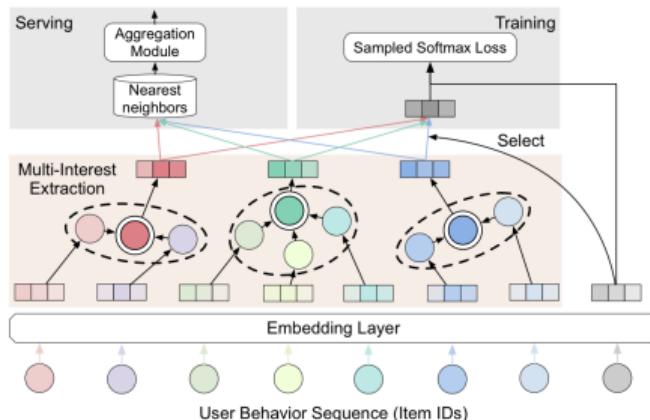


A Comparison of Calibrated and Intent-Aware Recommendations [KB19]



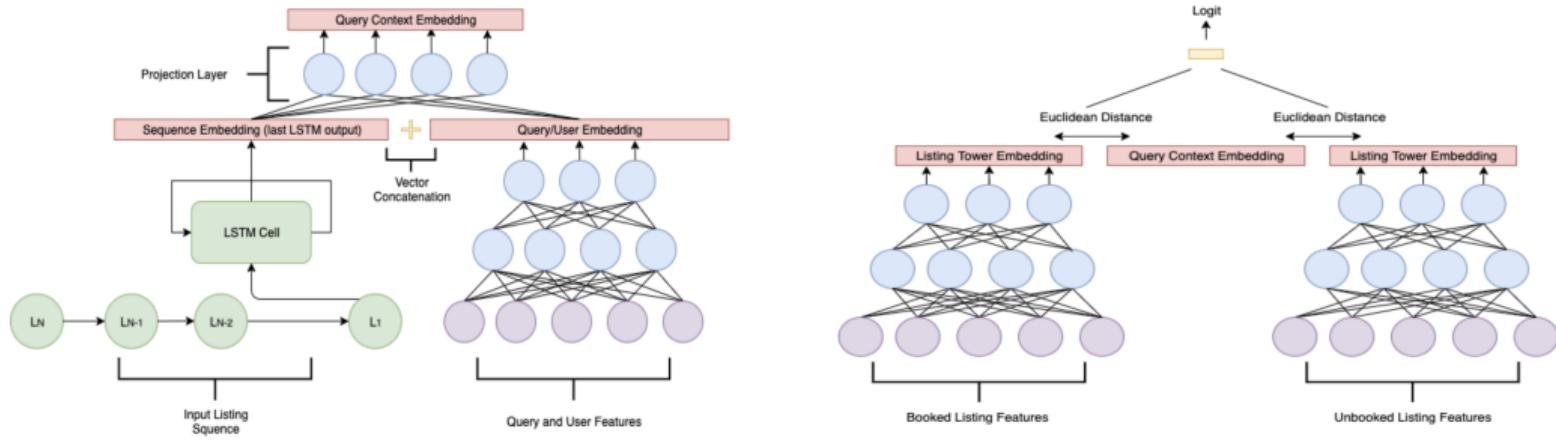
Учим разнообразие вместе с моделью 1

Controllable Multi-Interest Framework for Recommendation [CZZ⁺20]



Учим разнообразие вместе с моделью 2

Managing Diversity in Airbnb Search [AHR⁺20]
Learning To Rank Diversely At Airbnb [HAH⁺23]



Из-за поточечного предсказания релевантности, приходится дополнительно разнообразить списки рекомендаций

Необходимость разнообразия обосновывается A/B экспериментом



Разнообразие в рекомендательных системах
oooooooooo

Объяснение рекомендаций
●oooooooooooo

Смещения
oooooooo

Итоги
oooo

Объяснение рекомендаций



Разнообразие в рекомендательных системах
oooooooooo

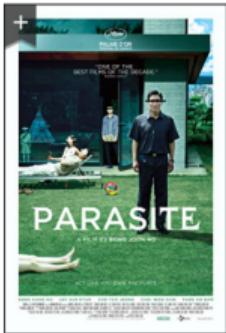
Объяснение рекомендаций
○●oooooooooooo

Смещения
oooooooo

Итоги
oooo

Объяснения

???

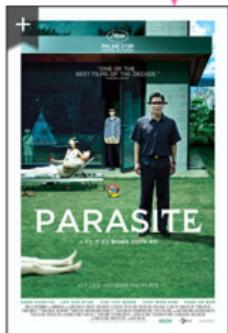


Объяснения

Потому что
вы оценили
Toy Story



Потому что вам
нравятся комедии



Популярное



Зачем объяснять рекомендации?

1. Прозрачность: объяснить пользователю, как работает система
2. Контролируемость: позволить пользователю исправить ошибки
3. Доверие: убедить пользователя, что система работает правильно
4. Убеждение: мотивировать пользователя к покупке



Case-based

Because you have selected or highly rated: Movie A

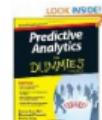
Потому что вы смотрели «ЛЕГО Фильм 2»



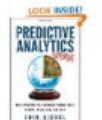
Collaborative

Customers Who Bought This Item Also Bought A

Customers Who Bought This Item Also Bought



Predictive Analytics For Dummies
› Anasse Barri
★★★★★ 29
Paperback
\$17.72



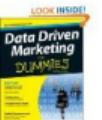
Predictive Analytics: The Power to Predict Who...
› Eric Siegel
★★★★★ 229
#1 Best Seller in Econometrics
Hardcover
\$16.88



Quantifying the User Experience: Practical...
› Jeff Sauro
★★★★★ 8
Paperback
\$40.63



Marketing Analytics: Strategic Models and...
› Stephan Sorger
★★★★★ 29
Paperback
\$50.52

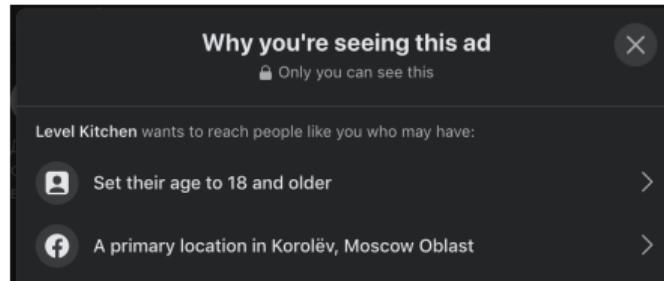


Data Driven Marketing For Dummies
› David Semmelroth
Paperback
\$20.49



Content-based

Recommended because you said liked science fiction



Knowledge-based

Less Memory and Lower Resolution and Cheaper



This item **Lenovo IdeaPad 3 14"** Laptop, Intel Core i3-1005G1 Processor, 4GB DDR4 RAM, 128GB M.2 SSD Storage, 14.0" FHD (1920 x 1080) Display, Integrated Graphics, Windows 10 S, 81WD010QUS, Platinum Grey

#1 Best Seller

Add to Cart



Lenovo IdeaPad 3 14 Laptop, AMD Ryzen 5 5500U Processor, 8GB DDR4 RAM, 256GB NVMe SSD Storage, 14.0" FHD (1920 x 1080) Display, AMD Radeon 7 Graphics, Windows 10 Home, 82KT00AMUS, Abyss Blue

Add to Cart



Lenovo IdeaPad 1 14 14.0" Laptop, 14.0" HD (1366 x 768) Display, Intel Celeron N4020 Processor, 4GB DDR4 RAM, 64 GB SSD Storage, Intel UHD Graphics 600, Win 10 in S Mode, 81VU0079US, Ice Blue

Add to Cart



Lenovo IdeaPad Gaming 3 15 15.6" Laptop, 15.6" FHD (1920 x 1080) Display, AMD Ryzen 5 5600H Processor, NVIDIA GeForce GTX 1650, 8GB DDR4 RAM, 256GB SSD Storage, Windows 10H, 82K20015US, Shadow Black

Add to Cart

Customer Rating	★★★★☆ (32)	★★★★★ (118)	★★★★★ (590)	★★★★☆ (134)
Price	\$399 ⁰⁰	\$539 ⁰⁰	\$266 ⁴²	\$769 ²²
Sold By	eSales Plus	Xocean	Mohawk Shop	ETRON INC - ELECTRONICS SUPPLIER
Computer Memory Size	128 GB	8	4	8
CPU Model Manufacturer	Intel	AMD	Intel	AMD
CPU Speed	1.2 GHz	2.1	1.1	3.3

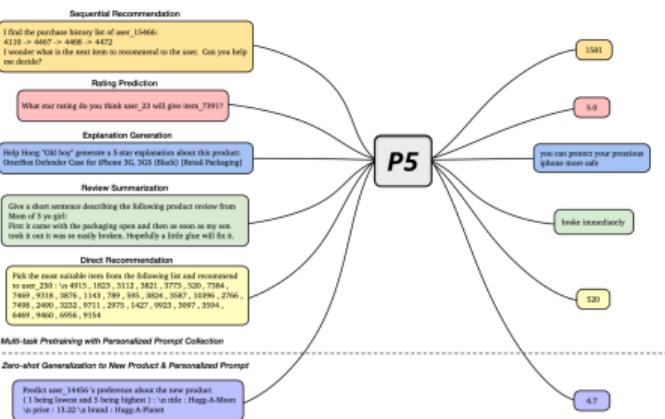
Explore, Exploit, and Explain: Personalizing Explainable Recommendations with Bandits [MLH⁺18]

Explanation	# Impressions
Because it's [day of week]	140.3K
Inspired by [user]'s recent listening	138.4K
Because it's a new release	140.5K
Because [user] likes [genre]	130.7K
Because it's popular	140.5K
Mood	140.7K
Focus	140.5K

$$r(j, e, x) = \sigma(\theta_{global} + \theta_j \times 1_j + \theta_e \times 1_e + \theta_x \times 1_x)$$



Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt and Predict Paradigm (P5) [GLF⁺22]



Rating / Review / Explanation raw data for Beauty

```

user_id: 7641      user_name: stephanie
item_id: 2051
item_title: SHAWY Nail Art Set (24 Fauxsue Colors
Nail Art Polish, Nail Art Decoration)
review: Absolutely great product. I bought this for my fourteen year
old niece for Christmas and of course I had to try it out, then I
tried another one, and another one and another one. So much fun!
I even contemplated keeping a few for myself!
star_rating: 5
summary: Perfect!
explanation: Absolutely great product
feature_word: product
  
```

Which star rating will user_{{user_id}} give item_{{(item_id)}}?
(1 being lowest and 5 being highest) → {{star_rating}}

Based on the feature word {{feature_word}}, generate an explanation for user_{{(user_id)}} about this product: {{(item_title)}} → {{(explanation)}}

Give a short sentence describing the following product review from {{(user_name)}}: {{(review)}} → {{(summary)}}

(a)

Sequential Recommendation raw data for Beauty

```

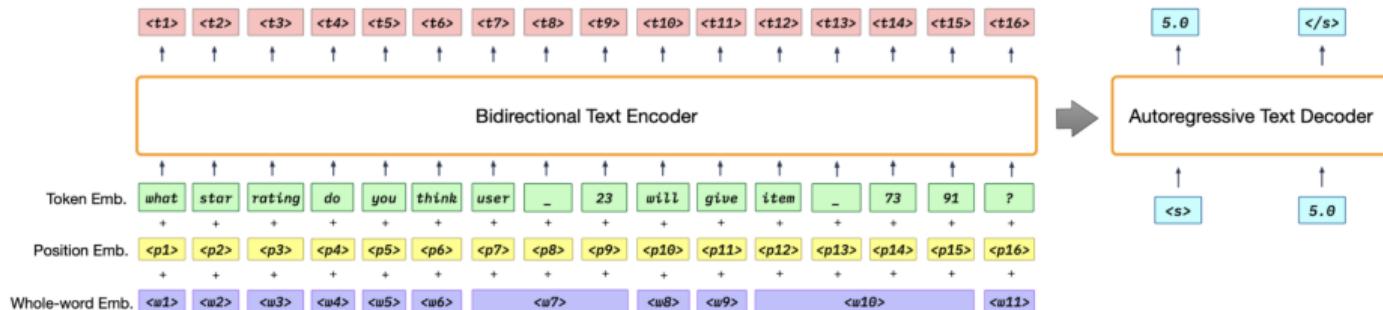
user_id: 7641      user_name: Victor
purchase_history: 652 -> 460 -> 447 -> 653 -> 654 -> 655 -> 656 -> 8
-> 657
next_item: 652
candidate_items: 4885, 4280, 4886, 1997, 878, 4281, 4222,
4887, 2892, 4888, 2879, 3147, 2195, 3148, 3179, 1963,
..., 1982, 552, 2754, 2481, 1916, 2822, 1325
  
```

Here is the purchase history of user_{{(user_id)}}:
({{(purchase_history)}})
What to recommend next for the user? → {{(next_item)}}

(b)



Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt and Predict Paradigm (P5) [GLF⁺22]



Cross-entropy Loss $\mathcal{L}_{\theta}^{P5} = - \sum_{j=1}^{|y|} \log P_{\theta}(y_j | y_{<j}, x)$



Если хотим делать объяснения рекомендаций, нужно ответить на вопросы:

- Какую цель мы достигнем объяснениями?
- Какие объяснения можно получить из модели?
- Как правильно представить объяснения пользователю?



Разнообразие в рекомендательных системах
oooooooooo

Объяснение рекомендаций
oooooooooooo

Смещения
●oooooooo

Итоги
oooo

Смещения



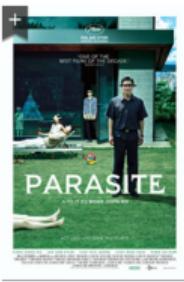
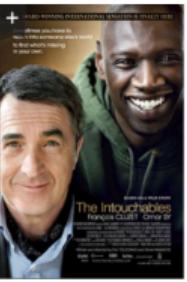
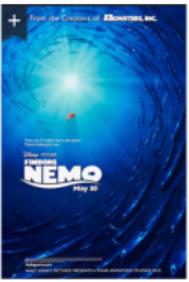
Разнообразие в рекомендательных системах
oooooooooo

Объяснение рекомендаций
oooooooooooo

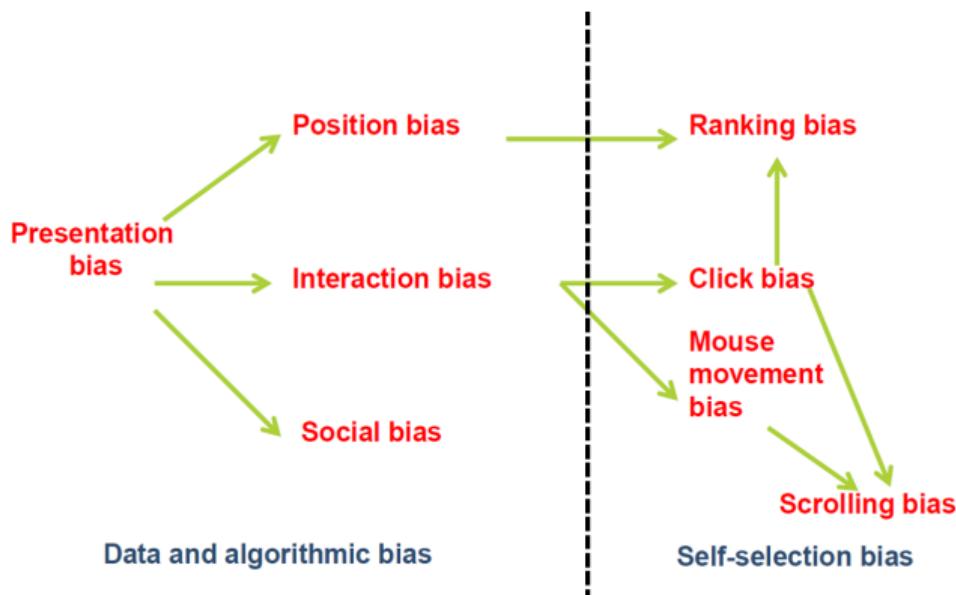
Смещения
o●oooooooo

Итоги
oooo

Удачные рекомендации



Смещения в рекомендациях [Jar19]



Пример self-selection bias

γ	Horror	Romance	Drama	p	Horror	Romance	Drama	O	Horror	Romance	Drama
Horror Lovers	5	1	3		p	$p/10$	$p/2$		1	1	1
	1	5	3		$p/10$	p	$p/2$		1	1	1
$\hat{\gamma}_1$	5	1	5		5	5	3		1	1	1
	1	5	5		5	5	3		1	1	1
$\hat{\gamma}_2$									1	1	1
									1	1	1
$\hat{\gamma}_3$									1	1	1
									1	1	1

Figure 1. Movie-Lovers toy example. Top row: true rating matrix Y , propensity matrix P , observation indicator matrix O . Bottom row: two rating prediction matrices \hat{Y}_1 and \hat{Y}_2 , and intervention indicator matrix \hat{Y}_3 .

$$R(\hat{Y}) = \frac{1}{UI} \sum_u \sum_i \delta_{ui}(Y, \hat{Y}), \quad R_{naive}(\hat{Y}) = \frac{1}{N} \sum_{(u,i) \in D} \delta_{ui}(Y, \hat{Y})$$



Inverse Propensity Scored Estimator [SSS⁺16]

$P_{ui} = P((u, i) \in D)$ – вероятность, что пользователь u поставит оценку айтему i

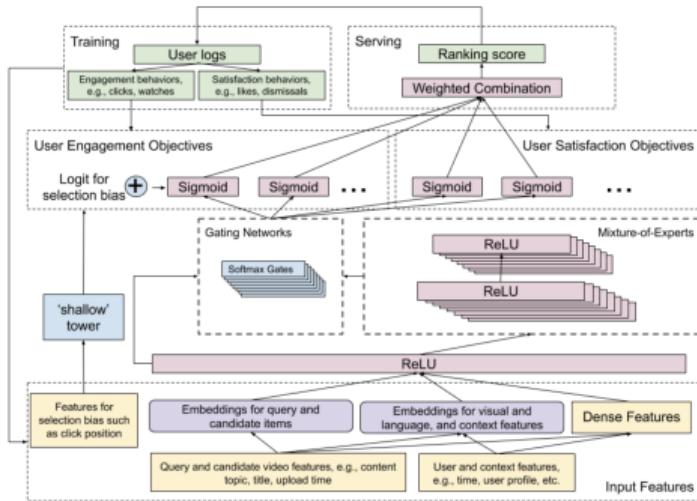
$$R_{IPS}(\hat{Y}|P) = \frac{1}{UI} \sum_{(u,i) \in D} \frac{\delta_{ui}(Y, \hat{Y})}{P_{ui}}$$

$$\begin{aligned} E_D[R_{IPS}(\hat{Y}|P)] &= \frac{1}{UI} \sum_u \sum_i E_D \left[\frac{\delta_{ui}(Y, \hat{Y})}{P_{ui}} \mathbb{I}\{(u, i) \in D\} \right] = \\ &= \frac{1}{UI} \sum_u \sum_i \delta_{ui}(Y, \hat{Y}) = R(\hat{Y}) \end{aligned}$$

IPS Estimator: проблемы

1. Когда P_{ui} неизвестно, его приходится оценивать
2. Большая дисперсия при оценке P_{ui}
3. Непонятно, как быть с рекомендациями списков

Recommending What Video to Watch Next: A Multitask Ranking System [ZHW⁺19]



Идея. При обучении модель “видит” признак-позицию айтема, а при инференсе признак зануляется.

Разнообразие в рекомендательных системах
oooooooooo

Объяснение рекомендаций
oooooooooooo

Смещения
oooooooo●

Итоги
oooo

Из-за специфики сбора данных рекомендации подвержены смещениям.

Существуют техники для корректировки, но они несовершенны.



Разнообразие в рекомендательных системах
oooooooooo

Объяснение рекомендаций
oooooooooooo

Смещения
oooooooo

Итоги
●ooo

Итоги



Итоги

При построении моделей мы делаем упрощающие предположения. Из-за этих предположений в продакшен системах могут возникать негативные эффекты. Эти эффекты нужно учитывать и пытаться исправить.



Разнообразие в рекомендательных системах
oooooooooo

Объяснение рекомендаций
oooooooooooo

Смещения
oooooooo

Итоги
ooo•o



<https://t.me/mlvok>



Литература I

-  Mustafa Abdool, Malay Haldar, Prashant Ramanathan, Tyler Sax, Lanbo Zhang, Aamir Manaswala, Lynn Yang, Bradley Turnbull, Qing Zhang, and Thomas Legrand, *Managing diversity in airbnb search*, Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (New York, NY, USA), KDD '20, Association for Computing Machinery, 2020, p. 2952–2960.
-  Yukuo Cen, Jianwei Zhang, Xu Zou, Chang Zhou, Hongxia Yang, and Jie Tang, *Controllable multi-interest framework for recommendation*, Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2020).



Литература II

-  Shijie Geng, Shuchang Liu, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang, *Recommendation as language processing (rlp): A unified pretrain, personalized prompt and predict paradigm (p5)*, Proceedings of the 16th ACM Conference on Recommender Systems (New York, NY, USA), RecSys '22, Association for Computing Machinery, 2022, p. 299–315.
-  Malay Haldar, Mustafa Abdoor, Liwei He, Dillon Davis, Huiji Gao, and Sanjeev Katariya, *Learning to rank diversely at airbnb*, 2023.
-  Greg Jarboe, *Understanding biases in search and recommender systems*, 2019.
-  Mesut Kaya and Derek Bridge, *A comparison of calibrated and intent-aware recommendations*, Proceedings of the 13th ACM Conference on Recommender Systems (New York, NY, USA), RecSys '19, Association for Computing Machinery, 2019, p. 151–159.



Литература III

-  James McInerney, Benjamin Lacker, Samantha Hansen, Karl Higley, Hugues Bouchard, Alois Gruson, and Rishabh Mehrotra, *Explore, exploit, and explain: Personalizing explainable recommendations with bandits*, Proceedings of the 12th ACM Conference on Recommender Systems (New York, NY, USA), RecSys '18, Association for Computing Machinery, 2018, p. 31–39.
-  Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims, *Recommendations as treatments: Debiasing learning and evaluation*, ICML'16, JMLR.org, 2016, p. 1670–1679.
-  Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, and Ed Chi, *Recommending what video to watch next: A multitask ranking system*, Proceedings of the 13th ACM Conference on Recommender Systems (New York, NY, USA), RecSys '19, Association for Computing Machinery, 2019, p. 43–51.

