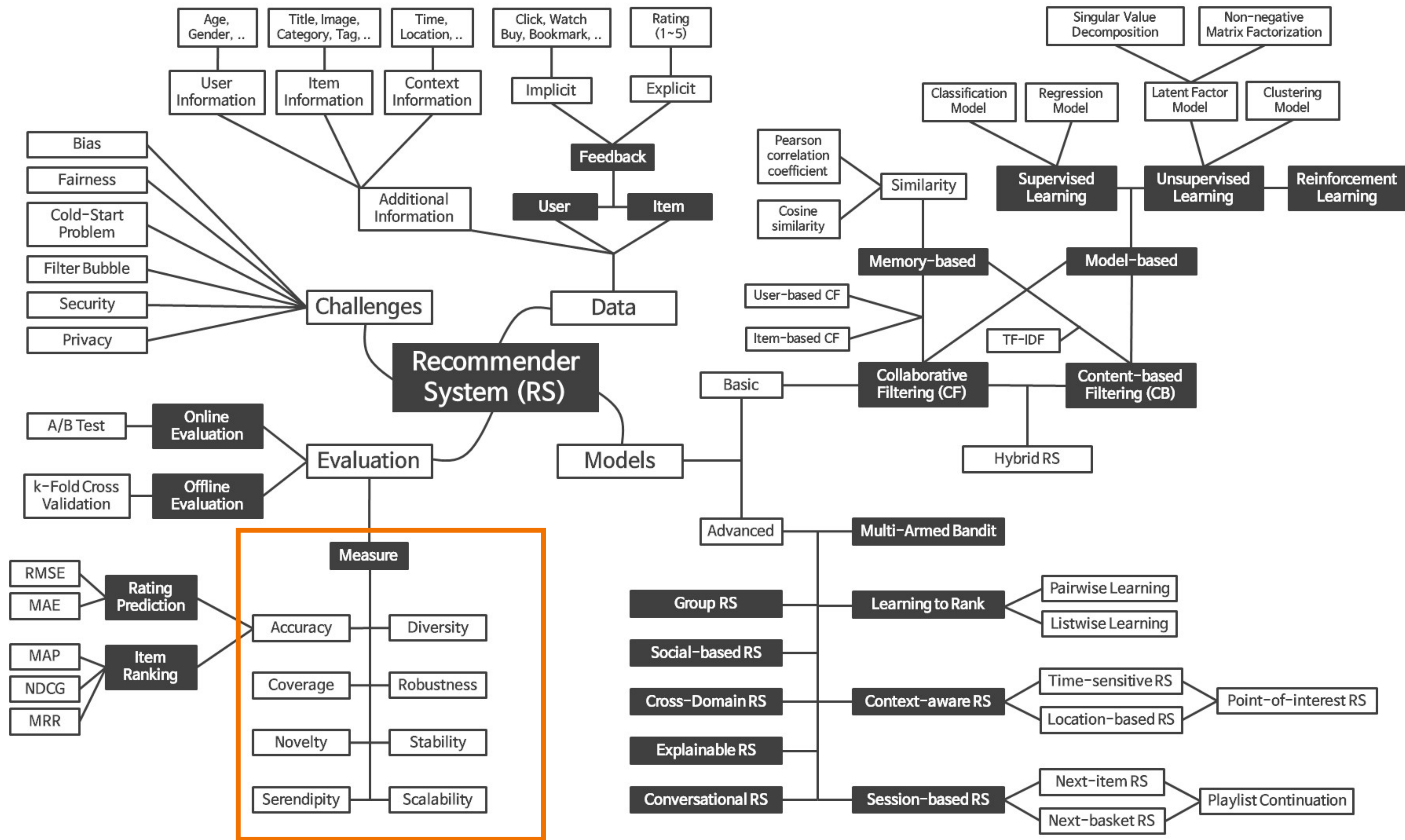


Разнообразие в рекомендательных системах.

Осиновсков Илья, 22.04.2025. AI masters.

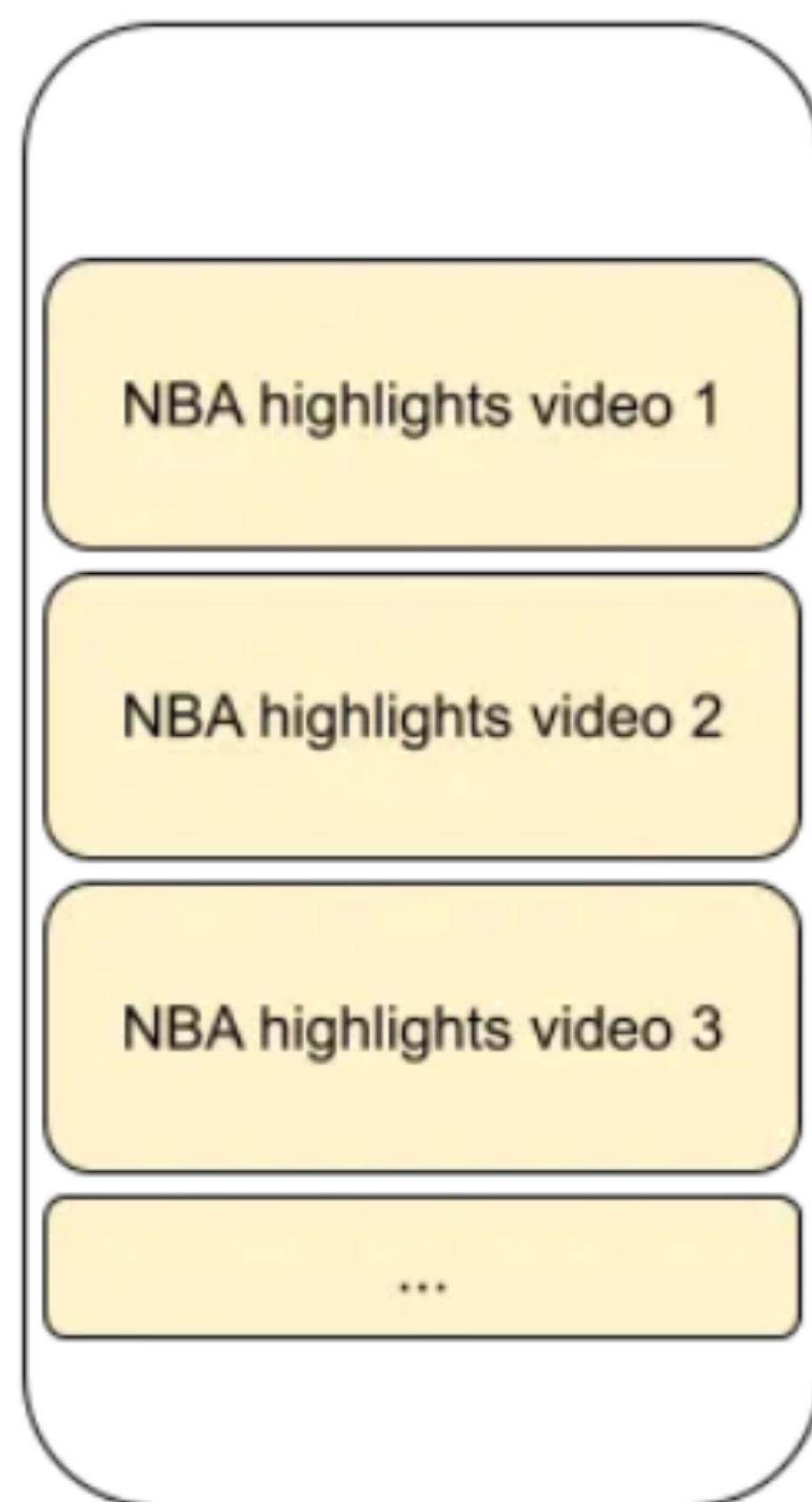
Введение



Разнообразие

- Важный аспект качества рекомендаций, который характеризует степень различия между предлагаемыми пользователю объектами.
- Обычно противопоставляется точности рекомендаций и служит дополнительным критерием оценки эффективности рекомендаций.

Почему разнообразие важно?



A single slate of recs but not boosted for diversity



A single slate of recs with more diversity of topics

Fig 1: Example of diversity

Почему разнообразие важно?

- Современные рекомендательные системы, стремясь максимизировать точность предсказаний, часто создают "пузыри фильтрации", ограничивая спектр контента, с которым взаимодействует пользователь.

Почему разнообразие важно?

- Способствует открытию нового контента пользователями, расширяя их кругозор и предотвращая когнитивные искажения
- Разнообразие помогает избежать эффекта "популярности", когда популярные элементы становятся еще популярнее, а менее известные - остаются незамеченными.
- Способствует справедливому представлению различных категорий контента и создателей, что особенно актуально в музыкальных сервисах, новостных платформах и социальных сетях.

Типы разнообразия в рекомендательных системах

- Индивидуальное
- Агрегированное

Типы разнообразия в рекомендательных системах

- **Индивидуальное**
Мера разнообразия элементов, рекомендуемых конкретному пользователю. Отражает, насколько различаются между собой предложения в персональном списке рекомендаций по характеристикам, тематикам и атрибутам.
- **Агрегированное**

Типы разнообразия в рекомендательных системах

- Индивидуальное
- Агрегированное
Мера распределения рекомендаций по всей базе элементов для всех пользователей системы.
Показывает, насколько широко используется каталог объектов в рекомендациях, предотвращая концентрацию только на популярных элементах и обеспечивая видимость для элементов "длинного хвоста", что способствует более справедливому распределению внимания пользователей

Метрики индивидуального разнообразия

Intra-list diversity

- Относится к разнообразию элементов внутри одного списка рекомендаций для конкретного пользователя. Например, вместо рекомендации десяти боевиков, система может предложить фильмы разных жанров.

$$ILD(L) = \frac{2}{|L| \cdot (|L| - 1)} \sum_{i=1}^{|L|-1} \sum_{j=i+1}^{|L|} d(i, j)$$

где:

- L — список рекомендаций
- $|L|$ — размер списка
- $d(i, j)$ — функция расстояния (или несходства) между элементами i и j
- Суммы $\sum_{i=1}^{|L|-1} \sum_{j=i+1}^{|L|}$ обеспечивают учет всех уникальных пар элементов в списке
- Множитель $\frac{2}{|L| \cdot (|L| - 1)}$ нормализует сумму, чтобы получить среднее значение расстояния для всех возможных пар в списке. Общее количество таких пар равно $C_{|L|}^2 = \frac{|L| \cdot (|L| - 1)}{2}$

Временное разнообразие

- Рассматривает изменение рекомендаций со временем, предотвращая повторение одних и тех же типов контента день за днем.

Временное разнообразие

- Рассматривает изменение рекомендаций со временем, предотвращая повторение одних и тех же типов контента день за днем.
- Можно использовать меру Жаккара между списками рекомендаций за разные дни

Новизна

- На основе популярности

1. Будем рассматривать собственную информацию item-а как меру его новизны для пользователей.

$$selfInformation(i) = -\log_2 \frac{|u_i|}{N},$$

где i - объект (item), N - число пользователей в обучающей выборке, $|u_i|$ - число пользователей, взаимодействовавших с объектом i в обучающей выборке.

2. Собственная информация определяется для item-а, а значение метрики мы хотим считать для пользователя и k - количества рекомендаций. Для подсчета Surprisal пользователя усредним значения собственной информации в топ- k рекомендациях этого пользователя.

Чтобы получить значения от 0 до 1, разделим значение метрики для пользователя на максимально возможно значение метрики равное $\log_2 N$.

$$Novelty(u, k) = \frac{1}{k \log_2 N} \sum_i^k selfInformation(i)$$

3. Результат усредним по пользователям.

Новизна

- На основе истории пользователя

$$Novelty(L, H) = 1 - \frac{|L \cap H|}{|L|}$$

где:

L — список рекомендаций

H — история предыдущих взаимодействий пользователя

Неожиданность

- Неожиданность (Unexpectedness) измеряет степень отклонения рекомендаций от ожидаемых пользователем или известных ему элементов.

- Базовая формула неожиданности

$$Serendipity(L) = \frac{1}{|L|} \sum_{i \in L} rel(i) \cdot unexp(i)$$

где:

L — список рекомендаций

$rel(i)$ — релевантность элемента i для пользователя

$unexp(i)$ — неожиданность элемента i для пользователя

Метрики агрегированного разнообразия

Покрытие каталога

- Покрытие каталога (Catalog Coverage) измеряет долю всех доступных элементов, которые когда-либо рекомендуются пользователям, что особенно важно для борьбы с предвзятостью в сторону популярного контента.

Покрытие каталога

- Покрытие каталога (Catalog Coverage) измеряет долю всех доступных элементов, которые когда-либо рекомендуются пользователям, что особенно важно для борьбы с предвзятостью в сторону популярного контента.
- Индекс Джини и коэффициент энтропии Шеннона используются для оценки равномерности распределения рекомендаций между элементами каталога.

$$H = - \sum_{i \in \mathcal{I}} p(i) \log_2 p(i), \quad p(i) = \frac{|\{u \in \mathcal{U} | i \in \mathcal{R}_u\}|}{\sum_{j \in \mathcal{I}} |\{u \in \mathcal{U} | j \in \mathcal{R}_u\}|}$$

**Высокое индивидуальное разнообразие !=
высокое агрегированное разнообразие**

Компромисс между точностью и разнообразием

- Фундаментальный вызов в разработке рекомендательных систем — поиск баланса между точностью предсказаний и разнообразием рекомендаций. Точность, традиционно измеряемая такими метриками как RMSE, Precision@k и Recall@k, отражает способность системы предсказывать предпочтения пользователя и рекомендовать релевантные элементы.

Компромисс между точностью и разнообразием

- Однако стремление к максимальной точности часто приводит к однородным рекомендациям, сконцентрированным вокруг явных интересов пользователя.

Компромисс между точностью и разнообразием

- Бюджет разнообразия — количество разнообразия, которое можно ввести без значительной потери точности.
- При этом оптимальная точка этого компромисса индивидуальна для каждого пользователя и зависит от контекста использования системы. Например, в новостных рекомендациях пользователи обычно ценят разнообразие выше, чем в специализированных технических рекомендациях.
- Современные подходы направлены на персонализацию этого компромисса, где система адаптивно определяет оптимальный баланс для каждого пользователя на основе его поведения и явных предпочтений.

Алгоритмические подходы к увеличению разнообразия

Ручные правила

- Не ставить объекты одного кластера ближе чем k позиций

Квотирование

- Распределить N мест среди всех кластеров объектов
- Кластеры
 - Обучаемые
 - Категории/жанры/etc
- Выбор размера квоты для кластеры
 - Одна квота на все кластеры
 - Распределение интересов пользователя по кластерам

Maximum marginal Relevance

Algorithm 1: Greedy Algorithm for Diversity Enhancement

Input: Recommendation number top K , coefficient λ

```
1 return Recommendation list  $\mathcal{R}_u$ 
2 Init recommendation list  $\mathcal{R}_u = \{\}$ 
3 for  $|\mathcal{R}_u| < K$  do
4    $i^* = \operatorname{argmax}_{i \in I \setminus \mathcal{R}_u} s(\mathcal{R}_u \cup \{i\}, \lambda)$ 
5    $\mathcal{R}_u = \mathcal{R}_u \cup \{i^*\}$ 
6 return  $\mathcal{R}_u$ ;
```

$$s(\mathcal{R}_u, \lambda) = \frac{1 - \lambda}{|\mathcal{R}_u|} \sum_{i \in \mathcal{R}_u} f_{\text{rec}}(i) + \lambda f_{\text{div}}(\mathcal{R}_u)$$

Determinantal Point Process

- Математическая модель, которая изначально использовалась в квантовой физике для описания отталкивания заряженных частиц. В контексте рекомендательных систем она помогает моделировать "отталкивание" между похожими элементами.

Determinantal Point Process

- Ключевая особенность DPP в том, что вероятность выбора подмножества элементов пропорциональна определителю матрицы, элементы которой отражают как качество отдельных элементов, так и их сходство между собой.

Determinantal Point Process

- Матрица ядра (L-matrix) - центральный элемент DPP. Для каждой пары элементов i и j матрица содержит значение $L(i,j)$, которое отражает их сходство. Диагональные элементы $L(i,i)$ представляют релевантность элемента i .
- Хотим набрать такую матрицу из N элементов, определитель которой будет максимален
- Точное решение NP-сложное, поэтому будем пользоваться жадным

Determinantal Point Process

- Точное решение NP-сложное, поэтому будем пользоваться жадным
- Даже в жадном нам надо считать определитель матрицы от N элементов, что имеет кубическую сложность относительно N
- В статье предлагается использовать разложение Холецкого для ускорения вычислений

Статья https://lsrs2017.wordpress.com/wp-content/uploads/2017/08/lsrs_2017_lamingchen.pdf

MULTISLOT RERANKER: A GENERIC MODEL-BASED RE-RANKING FRAMEWORK IN RECOMMENDATION SYSTEMS

Algorithm 1 Sequential Greedy Algorithm

Require: $i \geq 0$ and $i < N$ ▷ i is slot or position index, N is total number items.
Select item with top SPR score and place it to slot 0
 $i \leftarrow 1$ ▷ The slot 0 does not have previous slot
 $C \leftarrow \{0 \dots K - 1\}$ ▷ Initialize top K candidate set by second pass ranking scores
while $i < N$ **do**
 $reranking_score_max \leftarrow -\infty$
 $k \leftarrow 0$ ▷ Select item k from Candidates
 for $a \in C$ **do**
 Extract features from item a
 Extract interaction features between item a and the previous slot $i - 1$
 Calculate score by equation 1
 if $score > reranking_score_max$ **then**
 $reranking_score_max \leftarrow score$
 $k \leftarrow a$ ▷ Pick the item with largest re-ranking score, i.e., greedy
 end if
 end for
 Remove item k from Candidate set C and place it to slot i
 Pick the next top item in remaining list and add it to top K Candidate set C
 $i \leftarrow i + 1$
end while

- Стаття <https://arxiv.org/pdf/2401.06293>

MULTISLOT RERANKER: A GENERIC MODEL-BASED RE-RANKING FRAMEWORK IN RECOMMENDATION SYSTEMS

Table 1: A List of Features for Sequential Greedy Algorithm

| Feature Category | Features |
|----------------------------|--|
| Second Pass Ranking Scores | p(Click) Contribution responses such as p(Like), p(Comment), p(Share), p(Skip) p(Contributions) whose label is positive if any of Like, Comment, Share, Skip is positive. |
| Current Slot's Features | Slot index i Embeddings of item at slot i Type of item at slot i such as Video, Image, Job, Company, Article, etc. |
| Interaction Features | Type of items in previous slots Count of each type of items in previous slots Cross feature with item type among slot i and previous slots Embeddings dot product of items among current slot i and previous slots Whether items at current slot i and previous slots are created by the same user |

- Стаття <https://arxiv.org/pdf/2401.06293>

Многоуровневая система

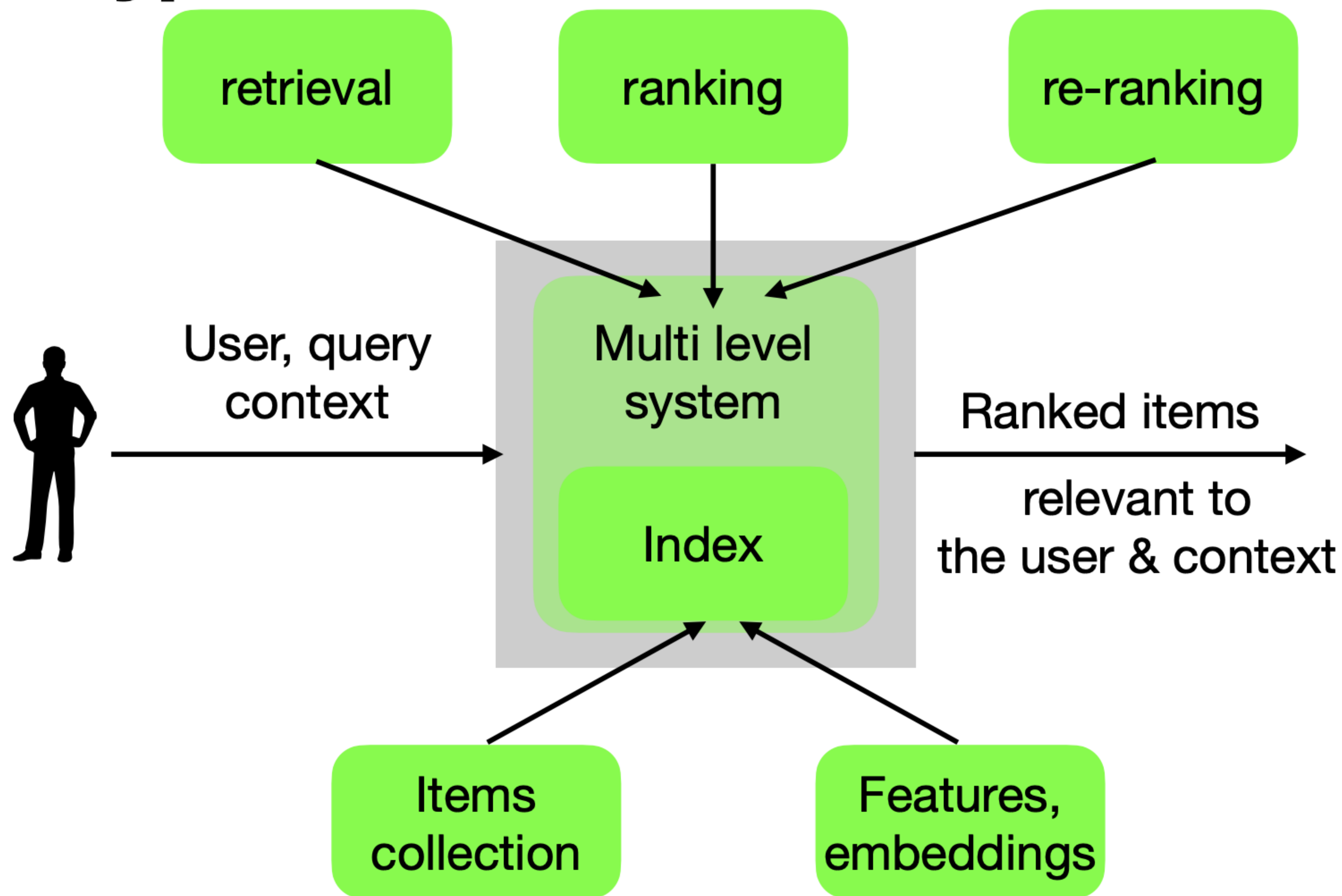
retrieval

filtering

ranking

re-ranking

Многоуровневая система



Вопросы