

Maximal Marginal Relevance (MMR)

王树森

<http://wangshusen.github.io/>



多样性

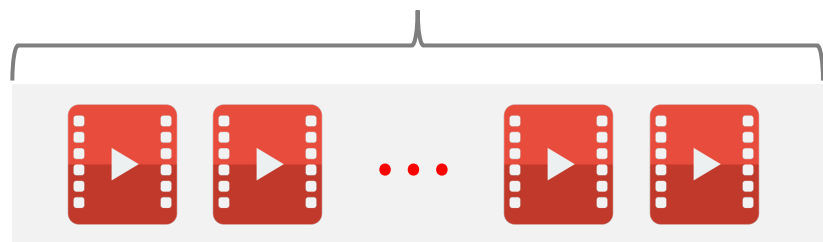
- 精排给 n 个候选物品打分，融合之后的分数为

$$\text{reward}_1, \dots, \text{reward}_n$$

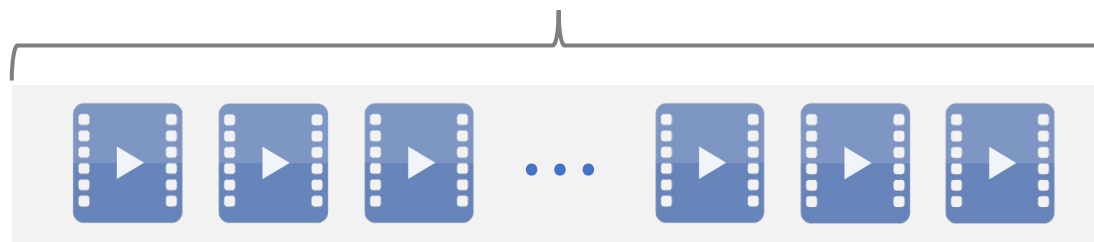
- 把第 i 和 j 个物品的相似度记作 $\text{sim}(i, j)$ 。
- 从 n 个物品中选出 k 个，既要有高精排分数，也要有多样性。

MMR多样性算法

选中的物品（记作 \mathcal{S} ）



未选中的物品（记作 \mathcal{R} ）



MMR多样性算法



- 计算集合 \mathcal{R} 中每个物品 i 的 Marginal Relevance 分数：

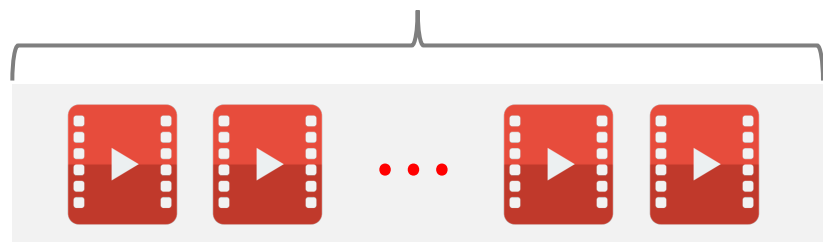
$$\text{MR}_i = \theta \cdot \text{reward}_i - (1 - \theta) \cdot \max_{j \in \mathcal{S}} \text{sim}(i, j).$$

物品 i 的
精排分数

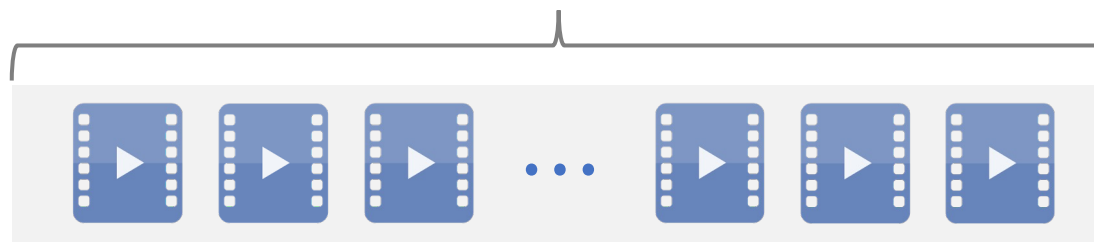
物品 i 的
多样性分数

MMR多样性算法

选中的物品（记作 \mathcal{S} ）



未选中的物品（记作 \mathcal{R} ）



- 计算集合 \mathcal{R} 中每个物品 i 的 Marginal Relevance 分数：

$$\text{MR}_i = \theta \cdot \text{reward}_i - (1 - \theta) \cdot \max_{j \in \mathcal{S}} \text{sim}(i, j).$$

MMR多样性算法



- 计算集合 \mathcal{R} 中每个物品 i 的 Marginal Relevance 分数：

$$\text{MR}_i = \theta \cdot \text{reward}_i - (1 - \theta) \cdot \max_{j \in \mathcal{S}} \text{sim}(i, j).$$

- Maximal Marginal Relevance (MMR) :

$$\text{argmax}_{i \in \mathcal{R}} \text{MR}_i.$$

MMR多样性算法

1. 已选中的物品 \mathcal{S} 初始化为空集，未选中的物品 \mathcal{R} 初始化为全集 $\{1, \dots, n\}$ 。
2. 选择精排分数 reward_i 最高的物品，从集合 \mathcal{R} 移到 \mathcal{S} 。
3. 做 $k - 1$ 轮循环：
 - a. 计算集合 \mathcal{R} 中所有物品的分数 $\{\text{MR}_i\}_{i \in \mathcal{R}}$ 。
 - b. 选出分数最高的物品，将其从 \mathcal{R} 移到 \mathcal{S} 。

滑动窗口

- MMR : $\operatorname{argmax}_{\substack{i \in \mathcal{R}}} \left\{ \theta \cdot \text{reward}_i - (1 - \theta) \cdot \max_{\substack{j \in \mathcal{S}}} \text{sim}(i, j) \right\}.$

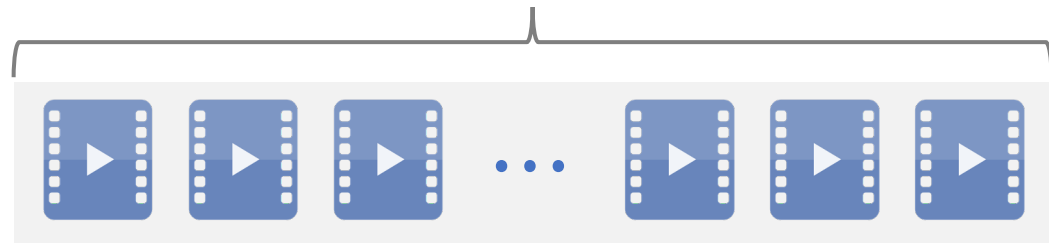
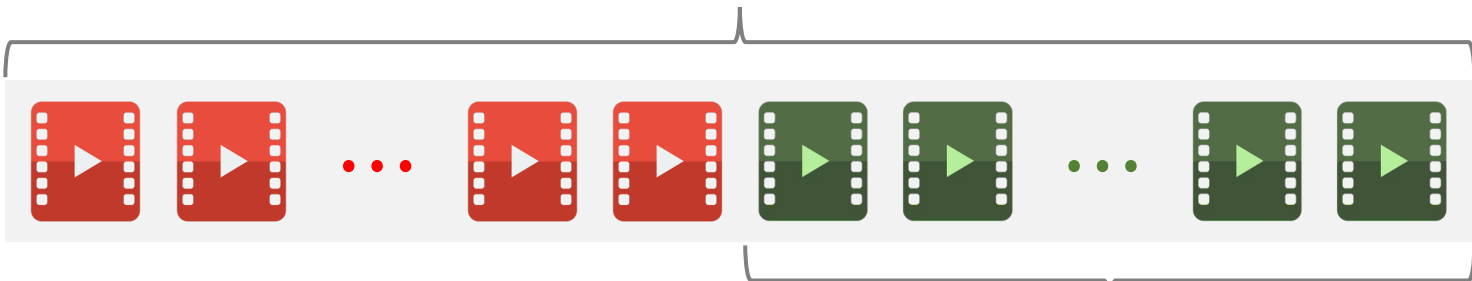
滑动窗口

- MMR : $\operatorname{argmax}_{i \in \mathcal{R}} \left\{ \theta \cdot \text{reward}_i - (1 - \theta) \cdot \max_{j \in \mathcal{S}} \text{sim}(i, j) \right\}$.
- 已选中的物品越多（即集合 \mathcal{S} 越大），越难找出物品 $i \in \mathcal{R}$ ，使得 i 与 \mathcal{S} 中的物品都不相似。
- 设 sim 的取值范围是 $[0, 1]$ 。当 \mathcal{S} 很大时，多样性分数 $\max_{j \in \mathcal{S}} \text{sim}(i, j)$ 总是约等于 1，导致 MMR 算法失效。
- 解决方案：设置一个滑动窗口 \mathcal{W} ，比如最近选中的 10 个物品，用 \mathcal{W} 代替 MMR 公式中的 \mathcal{S} 。

滑动窗口

选中的物品（记作 \mathcal{S} ）

未选中的物品（记作 \mathcal{R} ）



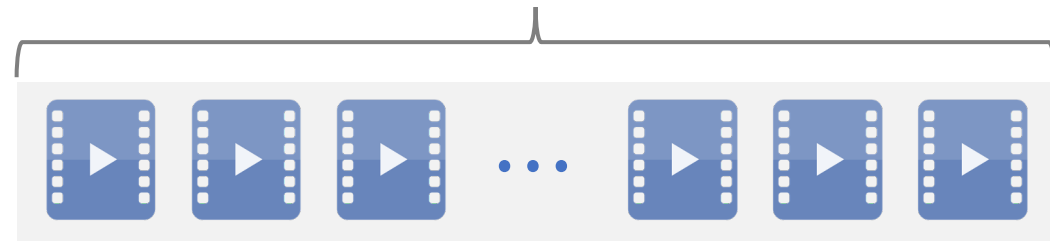
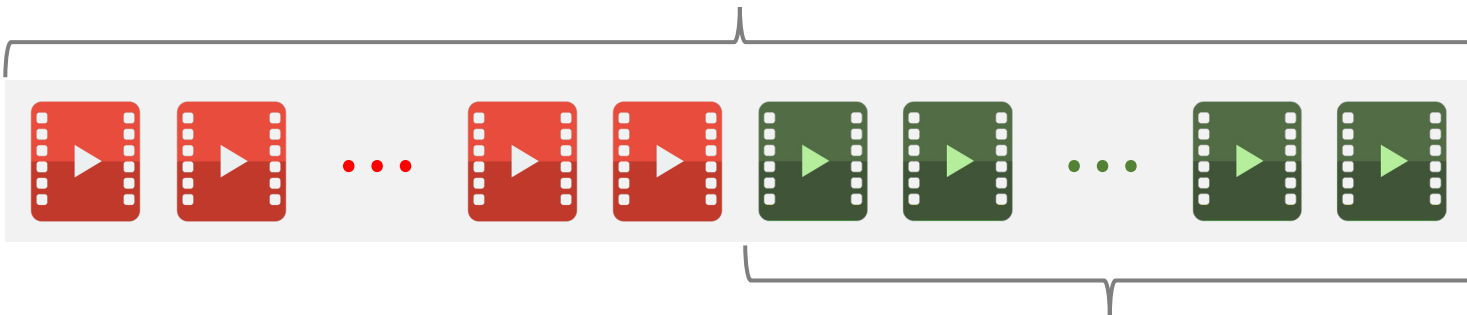
滑动窗口（记作 \mathcal{W} ）

滑动窗口

- 标准 MMR : $\operatorname{argmax}_{i \in \mathcal{R}} \left\{ \theta \cdot \text{reward}_i - (1 - \theta) \max_{j \in \mathcal{S}} \text{sim}(i, j) \right\}.$
- 用滑动窗口 : $\operatorname{argmax}_{i \in \mathcal{R}} \left\{ \theta \cdot \text{reward}_i - (1 - \theta) \cdot \max_{j \in \mathcal{W}} \text{sim}(i, j) \right\}.$

选中的物品 (记作 \mathcal{S})

未选中的物品 (记作 \mathcal{R})



滑动窗口 (记作 \mathcal{W})

Thank You!

<http://wangshusen.github.io/>