

Laptop Recommendation Algorithm Documentation

English Version

Overview

The recommendation system employs weighted cosine similarity to match user preferences with laptop characteristics. The algorithm combines hard constraints and soft preferences to generate personalized recommendations.

Hard Constraints

Before calculating similarity scores, laptops must satisfy:

- Operating System match
- Touchscreen requirement match

Laptops failing these constraints receive a similarity score of 0.

Feature Weights

Features contribute differently to the final similarity score:

- Price: 20%
- Type, RAM, Storage, Rating: 10% each
- Brand, CPU Brand, GPU Brand, Speed Rating, Ports, Display, Weight: 5% each

Mathematical Foundation

1. Feature Normalization

For numerical features, values are normalized to a [0,1] range:

$$NormalizedValue = \frac{value - min}{max - min}$$

Example: For a 2000laptopinrange[500, \$5000]: $NormalizedValue = \frac{2000 - 500}{5000 - 500} = 0.33$

2. Weighted Cosine Similarity

The similarity between user preferences and laptop features is calculated as:

$$Similarity = \frac{\sum_{i=1}^n w_i u_i l_i}{\sqrt{\sum_{i=1}^n w_i u_i^2} \sqrt{\sum_{i=1}^n w_i l_i^2}}$$

Where:

- w_i is the weight of feature i
- u_i is the normalized user preference for feature i
- l_i is the normalized laptop value for feature i

Detailed Example

User Preferences:

- Price Range: 1000–2000 (midpoint \$1500)
- RAM: 16GB
- Storage: 512GB

Laptop Specifications:

- Price: \$1800
- RAM: 16GB
- Storage: 1TB

Step 1 - Normalize Values:

Price:
User: $(1500 - 500)/(5000 - 500) = 0.22$
Laptop: $(1800 - 500)/(5000 - 500) = 0.29$

RAM:
User: $(16 - 4)/(64 - 4) = 0.20$
Laptop: $(16 - 4)/(64 - 4) = 0.20$

Storage:
User: $(512 - 128)/(2048 - 128) = 0.20$
Laptop: $(1024 - 128)/(2048 - 128) = 0.47$

Step 2 - Weighted Contributions: $Price_{contribution} = 0.20 \times 0.22 \times 0.29 = 0.013$
 $RAM_{contribution} = 0.10 \times 0.20 \times 0.20 = 0.004$ $Storage_{contribution} = 0.10 \times 0.20 \times 0.47 = 0.009$

Implementation Notes

1. Categorical Features:
$$Score_{categorical} = \begin{cases} 1 & \text{if match} \\ 0 & \text{if no match} \end{cases}$$
2. Multi-value Features:
$$Score_{multivalue} = \frac{\text{matching items}}{\text{total requested items}}$$
3. Missing values are excluded from calculation
4. Results are cached for performance

概述

推荐系统使用加权余弦相似度来匹配用户偏好与笔记本电脑特征。该算法结合硬性约束和软性偏好来生成个性化推荐。

硬性约束

在计算相似度之前，笔记本电脑必须满足：

- 操作系统匹配

- 触摸屏要求匹配

不满足这些约束的笔记本电脑相似度分数为0。

特征权重

不同特征对最终相似度分数的贡献：

- 价格：20%
- 类型、内存、存储、评分：各10%
- 品牌、CPU品牌、GPU品牌、速度评分、接口、显示屏、重量：各5%

数学基础

1. 特征归一化

对于数值型特征，值被归一化到[0,1]范围：

$$\text{归一化值} = \frac{\text{值} - \text{最小值}}{\text{最大值} - \text{最小值}}$$

示例：对于价格2000美元，范围[500美元, 5000美元]：归一化值 = $\frac{2000-500}{5000-500} = 0.33$

2. 加权余弦相似度

用户偏好和笔记本特征之间的相似度计算为：

$$\text{相似度} = \frac{\sum_{i=1}^n w_i u_i l_i}{\sqrt{\sum_{i=1}^n w_i u_i^2} \sqrt{\sum_{i=1}^n w_i l_i^2}}$$

其中：

- w_i 是特征i的权重
- u_i 是特征i的归一化用户偏好值
- l_i 是特征i的归一化笔记本电脑值

详细示例

用户偏好：

- 价格范围：1000-2000美元（中点1500美元）
- 内存：16GB
- 存储：512GB

笔记本规格：

- 价格：1800美元
- 内存：16GB
- 存储：1TB

步骤1 - 归一化值：

价格:

用户: $(1500 - 500)/(5000 - 500) = 0.22$

笔记本: $(1800 - 500)/(5000 - 500) = 0.29$

内存:

用户: $(16 - 4)/(64 - 4) = 0.20$

笔记本: $(16 - 4)/(64 - 4) = 0.20$

存储:

用户: $(512 - 128)/(2048 - 128) = 0.20$

笔记本: $(1024 - 128)/(2048 - 128) = 0.47$

步骤2 - 加权贡献: 价格贡献 $= 0.20 \times 0.22 \times 0.29 = 0.013$ 内存贡献 $= 0.10 \times 0.20 \times 0.20 = 0.004$

存储贡献 $= 0.10 \times 0.20 \times 0.47 = 0.009$

实现注意事项

1. 类别特征:

$$\text{分数}_{\text{类别}} = \begin{cases} 1 & \text{如果 匹配} \\ 0 & \text{如果 不匹配} \end{cases}$$

2. 多值特征:

$$\text{分数}_{\text{多值}} = \frac{\text{匹配项数}}{\text{请求项总数}}$$

3. 计算时排除缺失值

4. 结果被缓存以提高性能