

発見性を考慮した協調フィルタリング アルゴリズムに関する基礎検討

清水 拓也 土方 嘉徳 西田 正吾

島部 達哉

富山県立大学 電子・情報工学科

2021年10月29日

推薦システム

インターネット上に存在するアイテムの数は膨大であり、興味があるものを効率よく探すことは困難である。その対処として、近年、そのユーザが興味があるであろうアイテムを予測し、推薦するシステムが注目されている。これを、一般に「推薦システム」という。

推薦システムの問題

推薦システムにおいて、予測には、協調フィルタリングが使用されることが多いが、精度が高い一方、既知なアイテムや、同様なアイテムが推薦されやすい問題がある。

本研究の目的

予測の精度をある程度維持しながら、推薦されるアイテムの発見性や多様性を向上させる協調フィルタリングのアルゴリズムを提案し、評価する。

協調フィルタリング

複数のユーザが、複数のアイテムを評価した統計量から、対象のユーザが評価していないアイテムの評価値を予測する手法である。そのアルゴリズムには、大きく分けて「ユーザベース」と「アイテムベース」の2種類がある。

表 1: 統計量の例

	item 1	item 2	item 3	item 4
user A	5	2	3	4
user B	4		2	5
user C	1	3	5	2
user D	2	4		1

ユーザベースの協調フィルタリング

対象のユーザの評価値と、そのほかの各ユーザの評価値との類似度を計算し、類似度が一定以上大きいユーザの評価値を用いて、対象のユーザが評価していないアイテムの評価の予測値を計算する。

類似度

類似度の算出には、コサイン類似度やピアソン相関関係などが用いられる。

予測値

はじめに
前提知識
提案手法
実験
おわりに

ユーザ a_i が評価していないアイテム b_k の予測値 $p_i(b_k)$ は、以下の式で算出される。

$$p_i(b_k) = \bar{r}_i + \frac{\sum_{a_o \in A'_i} s(a_i, a_o) \cdot (r_o(b_k) - \bar{r}_o)}{\sum_{a_o \in A'_i} |s(a_i, a_o)|} \quad (1)$$

$$A'_i := \{a_o | a_o \in \text{neighbor}(a_i)\}$$

$$r'_i = \sum_{i=1}^n r_i(b_k) / n$$

$$a_o \in A \setminus \{a_i\}$$

A ：全ユーザの集合

$r_i(b_k)$ ：ユーザ a_i のアイテム b_k に対する評価値

$s(a_i, a_o)$ ：ユーザ a_i とユーザ a_o の類似度

$\text{neighbor}(a_i)$ ：ユーザ a_i と類似度が一定以上大きいユーザの集合

アイテムベースの協調フィルタリング

ユーザベースとほとんど同様だが、ユーザ間ではなく、アイテム間の類似度を計算する。

予測値

ユーザ a_i が評価していないアイテム b_k の予測値 $p_i(b_k)$ は、以下の式で算出される。

$$p_i(b_k) = \frac{\sum_{b_e \in B'_k} s(b_k, b_e) \cdot (r_i(b_e))}{\sum_{b_e \in B'_k} |s(b_k, b_e)|} \quad (2)$$

$$B'_k := \{b_e | b_e \in \text{neighbor}(b_k)\}$$

$\text{neighbor}(b_k)$ ：アイテム b_k と類似度が一定以上大きいアイテムの集合

基本的な考え方

対象ユーザが知らないアイテムをより多く推薦するためには、対象ユーザが知らないであろうアイテムを予測しなければならない。

そのため、「知っているアイテムが類似しているほかのユーザが知っているアイテムは、対象ユーザも知っているだろう」という考えに基づき、そのアイテムを知っているかどうかの評価値について、協調フィルタリングを行い、対象ユーザの評価値を予測する。

発見性を考慮したアルゴリズム

- ユーザは、そのアイテムが「どれぐらい好きか（嗜好）」という評価に加え、そのアイテムを「知っているかどうか（既知・不既知）」という評価を行う。
- 嗜好の評価について、協調フィルタリングを行い、予測値を算出する。その予測値を降順にソートし、上位 N 個のアイテムからなるリスト L_{pi} を作成する。
- リスト L_{pi} のすべてのアイテム $b \in B_i$ について、協調フィルタリングを行い、既知・不既知の評価の予測値を算出する。その予測値を昇順にソートしたリスト $L_{pi}^{unknown}$ を作成する。
すなわち、リスト $L_{pi}^{unknown}$ は、リスト L_{pi} を、知らないと予測される順にソートしている。

- 各アイテム b について、リスト $L_{pi}^{unknown}$ における順位を $L_{pi}^{unknown}(b)$ 、リスト L_{pi} における順位を $L_{pi}(b)$ としたとき、

$$p_i^*(b) = L_{pi}^{unknown}(b) \cdot \alpha + L_{pi}(b) \cdot (1 - \alpha) \quad (3)$$

を計算し、 $p_i^*(b)$ を昇順でソートしたリスト L_{pi}^* を作成する。

ここで、 $\alpha \in [0, 1]$ は重みであり、式 (3) より、 α が大きいほど、知らないであろうアイテムがより多く推薦され、 α が小さいほど、嗜好に基づいたアイテムがより多く推薦される。

- このアルゴリズムでは、このリスト L_{pi}^* を、ユーザに推薦するアイテムのリストとする。

実験の概要

さまざまな年代からなるユーザ 100 人, 一人ひとりに対して,
さまざまなジャンルからなる合計 1000 個の楽曲から,
ランダムに 200 曲を選出し,

- 嗜好の評価（「1」,「2」,「3」,「4」,「5」の 5 段階）
- 既知・不既知の評価（「0」,「1」の 2 段階）

を行ってもらった。

各ユーザが評価した 200 曲を,

- 訓練事例（150 曲）
- データセット（50 曲）

の 2 つにランダムで分割し, 「訓練事例」となった楽曲の評価のみを用いて, 前述のアルゴリズムを適用した。

このとき, 評価値を予測するアイテムは「データセット」の楽曲とし, 上位 10 個の推薦リストを, 重み α が $\alpha = 0.0, 0.1, 0.2, \dots, 1.0$ であるときのそれぞれについて, 出力した。

評価指標

あるシステムの質や性能などを判断するときは、そのシステムの出力結果を、評価する内容に沿って数値化する。これを、一般に「評価指標」という。

この実験では、以下の5つの評価指標を使用、または定義した。

■ 嗜好の精度

$$\text{嗜好 (精度)} = \frac{\text{推薦リストにある好みなアイテムの数}}{\text{推薦リストのアイテムの数}}$$

■ 嗜好の再現率

$$\text{嗜好 (再現率)} = \frac{\text{推薦リストにある好みなアイテムの数}}{\text{テストセットにある好みなアイテムの数}}$$

■ 発見性

$$\text{発見性} = \frac{\text{推薦リストにある知らないアイテムの数}}{\text{テストセットにあるアイテムの数}}$$

■ *Novelty*¹ の精度

$$\text{Novelty (精度)} = \frac{\text{推薦リストにある Novelty なアイテムの数}}{\text{推薦リストのアイテムの数}}$$

■ *Novelty* の再現率

$$\text{Novelty (再現率)} = \frac{\text{推薦リストにある Novelty なアイテムの数}}{\text{テストセットにある Novelty なアイテムの数}}$$

¹対象ユーザが、推薦されたアイテムを「知らず、かつ好みである」とき、そのアイテムは「Novelty である」という。

実験の結果

はじめに
前提知識
提案手法
実験
おわりに

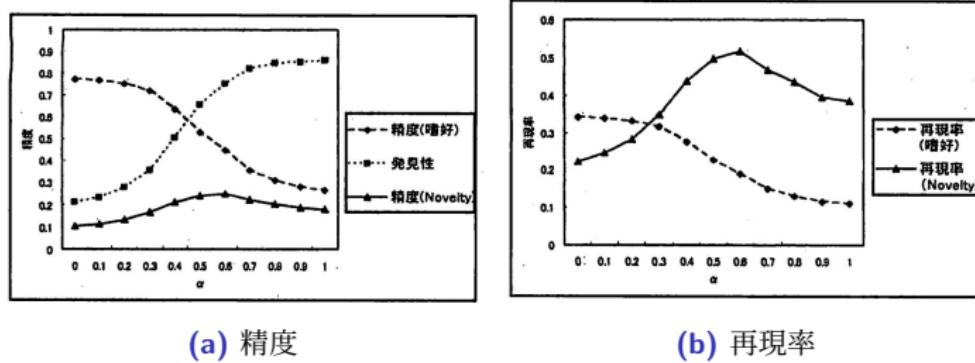


図 1: ユーザベースの協調フィルタリングを適用したときの推移

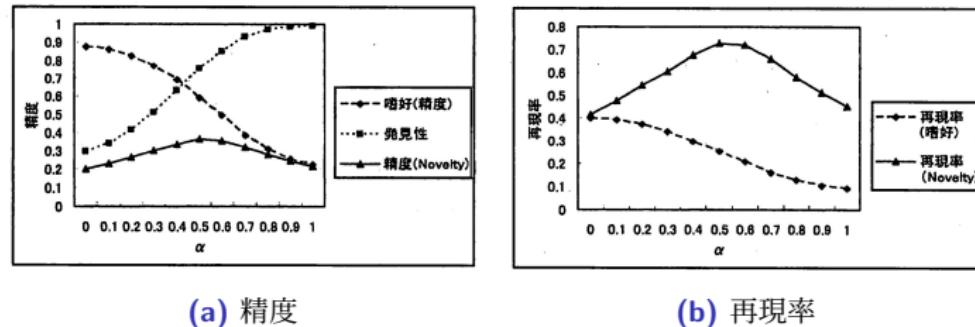


図 2: アイテムベースの協調フィルタリングを適用したときの推移

実験の結果

■ 本手法と従来手法との比較

嗜好のみで推薦リストを作成する従来手法 ($\alpha = 0$) と比較すると、本手法の目的である「発見性」や「Novelty」の向上が、 $0.1 \leq \alpha \leq 1.0$ のすべての範囲で確認できた。

■ *Novelty*

Novelty は、「嗜好の精度」と「発見性」が交差するときの α 付近で、最大値を示した。これは、嗜好の予測と、既知・不既知の予測を、バランスよく組み合わせることで、*Novelty* が向上することを意味する。

まとめ

本研究で提案した協調フィルタリングのアルゴリズムは、推薦されるアイテムの発見性や *Novelty* を向上させることができた。

今後

- 実際にユーザにアイテムを推薦し、満足度を検証する

本研究の実験では、効果を検証するために、すでに評価されているアイテムに対して予測したが、今後は、実際にユーザに推薦するために、評価されていないアイテムに対して予測し、満足度を検証する。