

# 推定購買確率と予測評価値をバランスする意外性指標に基づく推薦システム

1X13C066-0 関口 あゆみ  
指導教員 後藤 正幸

## 1 研究背景と目的

情報技術の発展に伴い、EC サイト等では膨大な量の購買履歴データや購買された商品（以下、アイテム）に対する評価履歴データが蓄積されるようになった。このようなデータを活用することで、それぞれのユーザの嗜好に合うアイテムを推薦するシステムが広く利用されている。推薦システムは購買促進に有効なツールであり、優れた推薦システムの導入はプロモーション戦略において重要である。

一般的な推薦システムは、ユーザ間の嗜好の類似性に着目し、ユーザが未購買かつ購買する可能性の高いアイテムを推薦する仕組みとなっている。しかしながら、このようなアイテムは推薦を行わずとも購買される可能性が高く、推薦を行ったからこそ購買につながるアイテム、すなわち発見性の高いアイテムを推薦する必要がある。これに対して、推薦が人気アイテムに集中しやすいという問題点に着目した、推薦アイテムの多様性を高める手法 [1] などが提案されている。この手法では、人気アイテムに集中せず様々な種類のアイテムを推薦することで、ユーザにとって目新しい、発見性の高いアイテムを推薦することができる。

しかしこのような手法は、多様性を持たせることで間接的に発見性を向上させるものであり、発見性そのものを目的関数とした推薦とはなっていない。この問題を解決するため、本研究では推定購買確率と予測評価値の双方を算出可能な確率的潜在クラスモデルとして Aspect Model [2]（以下、AM）に着目する。購買確率が低いアイテムはユーザが自分では発見し難いアイテムであり、評価値はユーザのアイテムに対する満足度を表すと解釈できるため、このようなモデルを用いることにより発見性を直接的に評価できると考えられる。

以上の議論より、本研究では、発見性が高く、かつユーザの嗜好に合致したアイテム、すなわち意外性を考慮したアイテムの推薦を実現するため、「推定購買確率が低いが、予測評価値が高いアイテム」を推薦するシステムを提案する。提案手法の有効性について、推薦システムのベンチマークデータを用いた実験により検証する。

## 2 準備

### 2.1 評価履歴を用いた Aspect Model

ユーザの嗜好の異質性やアイテムの類似性を潜在クラスとして仮定し、ユーザのアイテムに対する購買確率や評価値予測を行う手法として AM が知られている。AM では、各ユーザ、各アイテムが複数の潜在クラスに確率的に所属することを許容する。

いま、 $J$  人からなるユーザ集合を  $\mathcal{Y} = \{y_j : 1 \leq j \leq J\}$ 、 $I$  個からなるアイテム集合を  $\mathcal{X} = \{x_i : 1 \leq i \leq I\}$  とし、評価値  $r \in \{1, 2, \dots, R\}$  は  $R$  段階の離散評価値とする。ユーザ  $y_j$  がアイテム  $x_i$  を購買し、評価値  $r$  を付ける事象を、 $(y_j, x_i, r)$  と表現する。潜在クラスの集合を  $\mathcal{Z} = \{z_k : 1 \leq k \leq K\}$  としたとき、AM のグラフィカルモデルは図 1 となり、このモデルによる  $(y_j, x_i, r)$  の出現確率  $P(y_j, x_i, r)$  は、式 (1) のように表せる。

$$P(y_j, x_i, r) = \sum_{k=1}^K P(y_j|z_k)P(x_i|z_k)P(r|z_k)P(z_k) \quad (1)$$

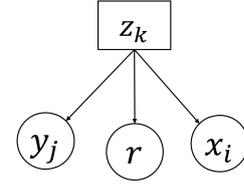


図 1. Aspect Model のグラフィカルモデル

### 2.2 パラメータの推定

AM で仮定している潜在クラスは観測することのできない変数であるため、EM アルゴリズムによってパラメータの推定を行う。ここで、 $N$  個ある購買履歴データのうち  $n$  ( $1 \leq n \leq N$ ) 番目のデータで出現するユーザを  $u_n \in \mathcal{Y}$ 、アイテムを  $v_n \in \mathcal{X}$ 、評価値を  $w_n \in \{1, \dots, R\}$  とする。各パラメータは、以下の式 (3)–(7) で表される E-step と M-step を繰り返し、式 (2) により定義された対数尤度関数  $LL$  が収束するまで更新する。

$$LL = \sum_{n=1}^N \log P(u_n, v_n, w_n) \quad (2)$$

[E-step]

$$P(z_k|u_n, v_n, w_n) = \frac{P(u_n|z_k)P(v_n|z_k)P(w_n|z_k)P(z_k)}{\sum_{k=1}^K P(u_n|z_k)P(v_n|z_k)P(w_n|z_k)P(z_k)} \quad (3)$$

[M-step]

$$P(z_k) = \frac{\sum_{n=1}^N P(z_k|u_n, v_n, w_n)}{N} \quad (4)$$

$$P(y_j|z_k) = \frac{\sum_{n=1}^N P(z_k|u_n, v_n, w_n)\delta(u_n = y_j)}{P(z_k)N} \quad (5)$$

$$P(x_i|z_k) = \frac{\sum_{n=1}^N P(z_k|u_n, v_n, w_n)\delta(v_n = x_i)}{P(z_k)N} \quad (6)$$

$$P(r|z_k) = \frac{\sum_{n=1}^N P(z_k|u_n, v_n, w_n)\delta(w_n = r)}{P(z_k)N} \quad (7)$$

ただし、 $\delta(x = y)$  は、 $x = y$  ならば 1、それ以外は 0 をとるインジケータ関数である。

### 2.3 推定購買確率と予測評価値の算出

EM アルゴリズムで推定した各パラメータを用いて、ユーザ  $y_j$  の未購買アイテム  $x_i$  に対する推定購買確率  $\hat{P}(x_i|y_j)$ 、及び、ユーザ  $y_j$  の未購買アイテム  $x_i$  に対する予測評価値  $\hat{r}(y_j, x_i)$  を以下の式 (8)、(9) を用いて求める。

$$\hat{P}(x_i|y_j) = \frac{\sum_{r=1}^R \hat{P}(y_j, x_i, r)}{\sum_{j=1}^J \sum_{r=1}^R \hat{P}(y_j, x_i, r)} \quad (8)$$

$$\hat{r}(y_j, x_i) = \frac{\sum_{r=1}^R r \hat{P}(y_j, x_i, r)}{\sum_{r=1}^R \hat{P}(y_j, x_i, r)} \quad (9)$$

## 3 提案手法

### 3.1 着眼点

購買履歴のみを用いた推薦システムでは、未購買アイテムに対して購買確率を推定し、推定購買確率の高いアイテムを推薦する。しかし、購買確率が高いアイテムは、推薦を行わなくとも購買される可能性が高く、推薦を行ったからこそ購

買されるアイテムを推薦すべきである。一方で、例えば購買確率が低くとも、購買した際にはユーザの評価値が高くなるアイテムも存在すると考えられる。評価値はユーザの満足度を表すと考えられるため、評価値が高くなるアイテムを推薦することがユーザの満足度向上につながるという。そこで本研究では、購買確率と評価値の双方を考慮し、「推定購買確率は低いが、予測評価値が高いアイテム」を意外性の高いアイテムと定義し、このアイテムを推薦するための方法を考えることで、これらの解決を図る。

### 3.2 推薦アイテムの選出方法

提案手法では、購買確率と評価値の双方を考慮した意外性のあるアイテムの推薦を目的とする。ユーザ  $y_j$  の未購買アイテム  $x_i$  に対する推定購買確率  $\hat{P}(x_i|y_j)$  に関する単調増加関数と、予測評価値  $\hat{r}(y_j, x_i)$  に関する単調増加関数をそれぞれ  $f(y_j, x_i)$ ,  $g(y_j, x_i)$  とした上で、これらの値のバランスを考慮した指標を基に推薦を行うことを考える。そこで、これらの差、商による手法をそれぞれ DIF (Difference Model), QUO (Quotient Model) と呼び、以下のように定義する。

$$DIF(x_i|y_j) = g(y_j, x_i) - f(y_j, x_i) \quad (10)$$

$$QUO(x_i|y_j) = \frac{g(y_j, x_i)}{f(y_j, x_i)} \quad (11)$$

これらの指標が高いアイテムを推薦することで、購買確率が低く、評価値が高いアイテム、つまり、意外性のあるアイテムの推薦が可能になる。

## 4 実験

提案手法の有効性を検証するため、推定購買確率  $\hat{P}(x_i|y_j)$  のみを用いる手法 (PUR), 予測評価値  $\hat{r}(y_j, x_i)$  のみを用いる手法 (EVA) と、2つの提案指標 (DIF, QUO) を用いる手法の比較実験を行った。

### 4.1 実験条件

実験には、Movielens の映画評価データ 100 万件を用いた。ユーザ数  $J$  は 6,040, アイテム数  $I$  は 3,952 であり、評価値は 1 から 5 までの 5 段階 ( $R=5$ ) である。このデータをランダムに、学習データ 90 万件とテストデータ 10 万件に分割した。本実験では、提案手法における  $f(y_j, x_i)$ ,  $g(y_j, x_i)$  は、以下のように定義した。

$$g(y_j, x_i) = \frac{1}{-\log \frac{\hat{r}(y_j, x_i)}{R}} \quad (12)$$

$$f(y_j, x_i) = \frac{1}{-\log \hat{P}(x_i|y_j)} \quad (13)$$

また、予備実験より、潜在クラス数は  $K=90$ , 推薦アイテム数は  $M=100$  とした。意外性のあるアイテムを「被購買数が少なく、テストデータで評価値が高いアイテム」とし、式 (14) を満たすアイテムと仮定する。

$$\begin{cases} \text{被購買数} \leq \alpha \\ r(y_j, x_i) \geq \beta \end{cases} \quad (14)$$

ただし、被購買数は対象アイテムを購買したユーザ数である。閾値は、意外性のあるアイテムが全データの約 2 割となるように、 $\alpha=540$ ,  $\beta=4$  と決定した。このようにして求めたユーザ  $y_j$  に対する意外性のあるアイテム集合を  $\mathcal{L}_j$  とする。

### 4.2 評価指標

推薦手法の評価指標には、推薦アイテムのテストデータにおける平均評価値、意外性に基づく推薦の適合率、および意外性指標の精度を用いる。

平均評価値は、推薦したアイテムが購買後に評価が高くなっているかを確認するために用いる。そして、ユーザ  $y_j$  へ

の推薦アイテム集合を  $\mathcal{S}_j$ , テストデータ中でユーザ  $y_j$  が購買しているアイテム集合を  $\mathcal{T}_j$  とすると、意外性の観点における推薦の適合率  $Pre$  は、式 (15) で定義できる。推薦の適合率は、推薦したアイテムをテストデータで評価した場合、そのアイテムが人気アイテムではなく、かつユーザが高評価を付けている割合である。

$$Pre = \frac{1}{J} \sum_{j=1}^J \frac{|\mathcal{S}_j \cap \mathcal{L}_j \cap \mathcal{T}_j|}{|\mathcal{S}_j \cap \mathcal{T}_j|} \quad (15)$$

意外性指標の精度は、各ユーザがテストデータで購買しているアイテムに限定して、各推薦指標によりランキングすることで、意外性のあるアイテムを推薦できているかを評価する。具体的には、 $\mathcal{T}_j$  内に限定した各指標による上位  $Q$  件のアイテムの集合を  $\mathcal{Q}_j$  とすると、意外性指標の精度  $Gap$  は式 (16) で表わされる。

$$Gap = \frac{1}{J} \sum_{j=1}^J \frac{|\mathcal{Q}_j \cap \mathcal{L}_j|}{|\mathcal{Q}_j|} \quad (16)$$

### 4.3 実験結果と考察

表 1 に実験結果を示す。ただし、 $Gap$  算出には  $Q=5$  とした。

表 1. 各指標による実験結果の比較

	従来法		提案法	
	PUR	EVA	DIF	QUO
平均評価値	3.80	4.37	<b>4.40</b>	4.39
$Pre$	8.58%	18.8%	18.0%	<b>30.1%</b>
$Gap$	9.85%	17.1%	17.2%	<b>20.1%</b>

従来手法 EVA は予測評価値の高いアイテムを推薦する手法であるため、平均評価値が高くなることを目指した手法である。一方、提案手法 QUO と DIF では、推定購買確率とのバランスを考慮するが、表 1 より、QUO と DIF の平均評価値は、EVA と比較し低下していないことがわかる。これは、推定購買確率と予測評価値をバランスする際に、評価値の高さが重視されたことが要因として考えられる。この結果より、QUO と DIF は評価値の高いアイテムが推薦可能な、ユーザ満足度の高い推薦であるといえる。

また、 $Pre$  と  $Gap$  の値に着目すると、DIF は EVA と比較して、大きな差異が見られない。これは、 $g(y_j, x_i)$  と比較して  $f(y_j, x_i)$  の値が小さかったことが原因として考えられる。一方 QUO は、EVA と比較して、平均評価値を低下させることなく、 $Pre$  と  $Gap$  の双方で大きく改善することができている。すなわち、提案指標 QUO による方法は、意外性のある推薦になっていると考えられる。

## 5 まとめと今後の課題

本研究では、意外性のあるアイテムを、推定購買確率は低いが予測評価値が高いアイテムと定義し、推定購買確率と予測評価値をバランスすることで、意外性のあるアイテムの推薦を可能とする推薦システムの提案を行った。実験により、提案手法による推薦は意外性のあるアイテムの推薦に有効であることを示した。今後の課題として、意外性のあるアイテムをユーザごとに設定することなどが挙げられる。

### 参考文献

- [1] Takeshi Suzuki, Kenta Mikawa, Masayuki Goto, "A Study of Recommender System to Improve Aggregate Diversity based on Latent Class Model," *The 13th Asia Pacific Industrial Engineering and Management Systems Conference (APIEMS 2012)*, No.122-1, 2012.
- [2] T. Hofmann "Probabilistic latent semantic indexing," *Proc. 15th conference on Uncertainty in artificial intelligence*, pp. 289-296, 1999.