

## 後処理による独立性強化型推薦システム

## Independence Enhanced Recommender System by a Post-Process Approach

神畷 敏弘\*<sup>1</sup> 赤穂 昭太郎\*<sup>1</sup>  
Toshihiro Kamishima Shotaro Akaho

\*<sup>1</sup>産業技術総合研究所

National Institute of Advanced Industrial Science and Technology (AIST)

We develop an independence-enhanced recommender system, which excludes specified information from its recommendation results. Such a system is useful for maintaining neutrality of provided information. Our previous algorithm adopted an *in-process* approach, enhancing recommendation independence while a prediction model is learned. In this paper, we develop a simple algorithm by a *post-process* approach, in which recommendation independence is maintained after a standard recommendation model is learned, and empirically examine this algorithm.

## 1. はじめに

利用者が関心のある情報を予測し、それを提示する推薦システムは、意思決定の支援に広く利用されるようになった。しかし、その情報に偏りがあると、利用者の意思決定が適切なものとならない場合がある。この問題に対して、利用者が与えた情報に関して独立性を保証する独立性強化型推薦システム (independence-enhanced recommender system) を提案した [Kamishima 12, Kamishima 13]。既存手法は、推薦モデルの訓練中に独立性強化の制約を導入する中処理型 (in-process) であった [Ruggieri 10]。これに対し、ここでは通常の推薦モデルを学習した後で、その予測結果を修正する後処理型 (post-process) の手法を開発し、実験的に既存手法と比較する。

以降、2. 節では推薦独立性について述べ、3. 節では独立性強化型推薦アルゴリズムを示し、4. 節で実験により各手法の特徴を分析し、5. 節でまとめを述べる。

## 2. 推薦独立性

ここでは、推薦独立性 (recommendation independence) の形式的定義を与え、この独立性の応用問題を挙げる。

形式的定義を与える前に、推薦独立性について、4. 節の ML1M-Year の映画推薦を例として直感的に説明する。推薦に影響する特徴や因子であるセンシティブ特徴を利用者が指定すると、その特徴に対して推薦独立性は定義される。例えば、映画の公開年は映画への嗜好に影響すると考えられる特徴で、利用者はこの特徴に注目しているでしょう。実際、名作のみが長い年月を経ても鑑賞されるため、古い映画ほど評価値が高くなる傾向があることが知られている。この情報が、推薦結果の生成に全く影響しないようになっている性質を推薦独立という。この映画の例では、映画の公開年が評価値の予測に全く影響しないとき推薦は独立である。そのため、映画の公開年以外の全ての特徴が同一の映画が仮にあったとすると、それらの予測評価値は全く同じになる。

それでは、この推薦独立性の直感的な定義を形式的なものにする。利用者やアイテムの記述や特徴など推薦に必要な全ての情報が与えられている事象を想定し、この情報から推薦結果を予測する。この事象は、三つの確率変数  $R$ ,  $S$ , および  $F$  の実現値として表される。 $R$  は推薦結果で、本稿では予測評価値

にあたり、 $S$  はセンシティブ特徴で、本稿では特に二値である場合のみを考える。 $F$  は通常の特徴で、 $R$  と  $S$  以外の推薦に関連した全ての情報を含むとする。そして、 $S$  と  $F$  が与えられたとき、確率的推薦モデル  $\Pr[R|S, F]$  に基づいて  $R$  を予測することが推薦である。

推薦結果からセンシティブ特徴の情報を除外するというところを、形式的に  $R$  と  $S$  が統計的に独立、すなわち、 $\Pr[R] = \Pr[R|S]$  ( $R \perp S$  と表記) であることと定める。この  $R$  と  $S$  の独立性は  $\Pr[R|S, F] = \Pr[R|F]$  の条件を含意し、推薦結果が  $S$  以外の情報  $F$  のみから予測されることとなる。この独立性からは、 $R$  と  $S$  の間の相互情報量  $I(R; S)$  が 0 となることも導くことができる。これは、情報論理論の観点からは、センシティブ特徴  $S$  の情報が結果  $R$  から完全に排除されていることに他ならない。このように、この推薦独立性形式的定義は、上記の直感的な独立性と良く合致している。

次に、推薦独立性に配慮した推薦システムが活用される潜在的な状況を挙げておく。

**法や規定の遵守** 推薦システムは法や規定を遵守して運用する必要がある。Sweeney は人種差別的であることが疑われる広告の配信の例について言及している [Sweeney 13]。この場合、人種情報とは独立に個人化広告を選択すべきである。

**情報提供者の公平な扱い** 情報提供者に対する公平性が必要になる場合がある。米連邦取引委員会 (FTC) は、競合各社のサービスよりも自社のものを検索エンジンで上位に提示しているとの懸念により Google 社を調査した [Forden 12]。作為的な操作の有無が示せれば、情報提供者を公平に扱っていることを示すには現状では十分とされている。しかし、情報提供者が競合会社かどうかの情報を明示的にアルゴリズムで排除できれば、競合会社が不公正な扱いをされているという懸念に技術的に対処できるだろう。

**利用者が望まない情報の影響の排除** 独立性の強化は、望ましくない情報を排除するのにも役立つ。よく消費される人気アイテムがより頻繁に推薦されやすいという人気バイアス [Celma 08] は推薦システムをよく知られた問題の一つである。アイテムの人気についての情報を除外できれば、利用者の嗜好とは関係ない人気バイアスの影響を受けていない情報を提供できるようになるだろう。

連絡先: <http://www.kamishima.net/>

### 3. 手法

この節では、独立正強化型推薦問題を定式化した後、この問題を解くアルゴリズムとして、中処理型の文献 [Kamishima 13] の方法と、今回提案する後処理型の方法を示す。

#### 3.1 独立性強化型推薦問題の形式的定義

推薦システムの中でも利用者のアイテムへの評価値を予測する評価値予測 (predicting ratings) 問題を対象に、独立性を強化する。  $X \in \{1, \dots, n\}$  と  $Y \in \{1, \dots, m\}$  はそれぞれ利用者とアイテムを示す確率変数。イベント  $(x, y)$  は、確率変数の対  $(X, Y)$  の実現値。利用者  $X$  によるアイテム  $Y$  への評価値を示す確率変数を  $R$  とし、その実現値を  $r$  とする。評価値の定義域は、  $\{1, \dots, 5\}$  といった離散値であることが多いが、ここでは評価値の実数集合として扱う。以上の変数は、独立性強化型ではない評価値推定問題と共通である。

さらに独立性強化型推薦では、センシティブ特徴を表す変数  $S$  を導入する。この変数が何を表すかは利用者が指定し、その値はイベントに依存して決まる。ここでは、センシティブ特徴の定義域は二値、すなわち  $\{0, 1\}$  とする。訓練事例は、イベント  $(x, y)$  と、このイベントに対するセンシティブ特徴の値  $s$  と利用者が与えた評価値  $r$  とで構成される。この訓練事例を  $N$  個含むのが訓練集合  $\mathcal{D} = \{(x_i, y_i, s_i, r_i)\}$ ,  $i = 1, \dots, N$  である。ある新しいイベント  $(x, y)$  とそれに対応するセンシティブ特徴の値  $s$  に対して、利用者  $x$  がアイテム  $y$  に与えた評価値を予測するのが評価値予測関数  $\hat{r}(x, y, s)$  である。そして、推薦独立性を満たす評価値予測関数を、訓練集合から獲得するのが、独立性強化型推薦問題の目標である。

#### 3.2 確率的行列分解

ここで示す独立性強化型推薦アルゴリズムの元となる通常の確率的行列分解モデルを紹介する。このモデルには細部が異なるものがあるが、文献 [Koren 08] の式 (3) の次のものを用いる：

$$\hat{r}(x, y) = \mu + b_x + c_y + \mathbf{p}_x^\top \mathbf{q}_y \quad (1)$$

ただし、 $\mu$ ,  $b_x$ ,  $c_y$  はそれぞれ、大域、利用者ごと、およびアイテムごとのバイアス項である。 $\mathbf{p}_x$  と  $\mathbf{q}_y$  はそれぞれ  $K$  次元ベクトルのパラメータで、利用者とアイテムの交差的な効果を表現する。損失関数  $\text{loss}(r, \hat{r})$  として、正則化項付きの二乗誤差を導入する：

$$\sum_{(x_i, y_i, r_i) \in \mathcal{D}} (r_i - \hat{r}(x_i, y_i))^2 + \text{reg}(\Theta) \quad (2)$$

この目的関数は非凸ではあるが、ほとんどの場合に良い解を単純な勾配降下法によって求められることが実験的に知られている [Koren 08]。

#### 3.3 平均一致法

以前に提案した中処理型の独立性強化型推薦アルゴリズムである平均一致法を示す [Kamishima 13]。この方法では、独立性を強化するための制約項である独立性項  $\text{ind}(R, S)$  を導入する。この項は評価値  $R$  とセンシティブ特徴  $S$  が独立であるほど大きな値をとる。この独立性項を加えた目的関数は次式で表され、これを最小化する：

$$\sum_{\mathcal{D}} \text{loss}(r, \hat{r}(x, y, s)) - \eta \text{ind}(R, S) + \lambda \text{reg}(\Theta) \quad (3)$$

ただし、 $\eta > 0$  は独立性と損失の釣り合いを調整する独立性パラメータ、 $\lambda > 0$  は正則化パラメータ、そして  $\Theta$  はモデルパラメータである。

前述の確率的行列分解を独立性強化型推薦に拡張するために、式 (1) のモデルをセンシティブ特徴の値  $s$  に依存するように修正する。 $S$  のそれぞれの値 (0 と 1) に対してパラメータ集合を用意し、これらのパラメータ集合をセンシティブ特徴の値に応じて選択し、次の関数で予測評価値を得る：

$$\hat{r}(x, y, s) = \mu^{(s)} + b_x^{(s)} + c_y^{(s)} + \mathbf{p}_x^{(s)\top} \mathbf{q}_y^{(s)} \quad (4)$$

式 (3) で、この予測評価値に対する二乗損失と  $L_2$  正則化項を採用すれば独立性強化型の目的関数となる。

最後に残った独立性項  $\text{ind}(R, S)$  について述べる。平均一致法では、二つの分布の 1 次モーメント、すなわち平均を一致させるもので、形式的には次式で定義する：

$$- (\text{mean}^{(0)}(\hat{r}) - \text{mean}^{(1)}(\hat{r}))^2 \quad (5)$$

ただし、 $\text{mean}^{(s)}(\hat{r})$  は、 $s_i = s$  であるような  $\mathcal{D}$  中のイベントに対する予測評価値の平均である。以上で目的関数を定めたので、あとはこれを勾配降下法により最小化することでモデルのパラメータを決めれば、学習したモデルで推薦独立性を保った推薦ができる。

#### 3.4 線形分布変換法

ここでは、新たに提案する後処理型の線形分布変換法について述べる。まず 3.2 節の通常の確率的行列分解を用いて予測モデルを計算する。その後、訓練データ集合  $\mathcal{D}$  中で  $s_i = s$  であるようなイベントに対する予測評価値の平均と標準偏差をそれぞれ  $\text{mean}^{(s)}(\hat{r})$  と  $\text{stdev}^{(s)}(\hat{r})$  とする。また、訓練データ集合  $\mathcal{D}$  中の全てのイベントについて求めた予測評価値の平均と標準偏差を、それぞれ  $\text{mean}(\hat{r})$  と  $\text{stdev}(\hat{r})$  とする。

テストイベント  $(x, y)$  とそのセンシティブ値  $s$  について予測評価値  $\hat{r}(x, y, s)$  を求めたあと、次式によってセンシティブ特徴の違いによる分布の平均と標準偏差を一致させる：

$$\hat{r}_{\text{LDT}}(x, y, s) = \frac{\text{stdev}(\hat{r})}{\text{stdev}^{(s)}(\hat{r})} (\hat{r}(x, y, s) - \text{mean}^{(s)}(\hat{r})) + \text{mean}(\hat{r}) \quad (6)$$

この線形変換により、センシティブ特徴の値ごとの予測評価値の分布の 2 次までのモーメントは、全体のそれに一致させることができ、センシティブ特徴の値の異なる予測評価値の二つの分布を近づけることができる。ただし、3 次以降のモーメントは無視しているので分布が完全に一致するわけではないことに注意されたい。このようにして、評価値  $R$  とセンシティブ特徴  $S$  の独立性を向上させる。

### 4. 実験

ここでは通常の推薦手法である 3.2 節の確率的行列分解をベースラインに、3.3 節の平均一致法と 3.4 節の線形分布変換

表 1: 実験データとアルゴリズムのパラメータの概要

データ	ML1M	Flixster	Sushi
利用者数	6,040	147,612	5,000
アイテム数	3,706	48,794	100
評価値数	1,000,209	8,196,077	50,000
平均評価値	3.58	3.61	1.27
潜在因子数 $K$	7	20	5
正則化パラメータ $\lambda$	1	30	10
独立性パラメータ $\eta$	100	100	100

表 2: 各データ集合に対して 3 手法を適用したときの MAE と KS の変化

手法	ML1M		Flixster		Sushi							
	Year		Gender		MAE	KS	Age		Gender		Seafood	
	MAE	KS	MAE	KS			MAE	KS	MAE	KS	MAE	KS
標準	0.687	0.1662	0.691	0.0319	0.645	0.1416	0.918	0.2445	0.922	0.0739	0.908	0.3306
平均一致法	0.697	0.0271	0.694	0.0050	0.653	0.0165	0.926	0.0256	0.925	0.0321	0.925	0.2002
線形分布変換法	0.696	0.0174	0.692	0.0047	0.650	0.0280	0.915	0.0462	0.925	0.0235	0.912	0.2291

法とを比較する。最初に実験データと評価指標についてまとめたあとで、実験結果とその考察を述べる。

データ集合は、表 1 のような、Movielens 1M, Flixster, および寿司の 3 種類を用いた。それぞれ、ML1M, Flixster, および Sushi と記す。ML1M データは、Grouplens プロジェクト [Gro] が公開しているもので、映画に対する嗜好を 5 段階評価したものである。このデータは、映画の公開年が 1990 年より古いかどうかを表す Year と、利用者の性別を表す Gender の二つのセンシティブ特徴で実験した。Flixster データ (<http://www.sfu.ca/~sja25/datasets/>) [Jamali 10] は映画について 10 段階で評価したものである。被評価値数が上位 1% 以上の人気アイテム (評価値全体の 47.2% に相当) かどうかをセンシティブ特徴として、人気バイアスの補正を試みた。Sushi データ (<http://www.kamishima.net/sushi/>) [Kamishima 03] は、5000 人の被験者にそれぞれ 100 種のうち 10 種の寿司について嗜好を尋ねたものである。被験者が 10 代かどうかの Age, 被験者の性別の Gender, および寿司が魚介類かどうかの Seafood の 3 種類のセンシティブ特徴で実験した。

評価指標は 2 種類あり、一つは予測精度を測るための MAE (mean absolute error) である。予測評価値と真の評価値の絶対誤差の平均値であり、小さいほど推薦の予測精度は高い。もう一つは、分布の一致性を調べるノンパラメトリック検定である 2 標本コルモゴロフスミルノフ検定統計量 (KS と略記) で、独立性を測る。二つのセンシティブ特徴の値それぞれについて予測評価値の経験累積密度関数を求めたとき、この二つの関数の間の面積が KS である。分布が一致するほど、評価値  $R$  とセンシティブ特徴  $S$  は独立になるので、KS が小さいほど独立性は高い。統計量は 5 分割の交差確認を行って求めた。各データに対するアルゴリズムのパラメータも表 1 に示しておく。

実験結果を表 2 に示す。標準の確率的行列分解と、独立性強化型の平均一致法や線形分布変換法とを比較すると、独立性強化型では確かに独立性は強化されており (KS が低下)、その代わりに利用可能な情報量が減少しているので誤差 MAE は若干悪化する傾向がある。平均一致法と線形分布変換法とを比較すると、ML1M-Year, ML1M-Gender, および Sushi-Gender では KS が線形分布変換法が、残りのデータ集合については平均一致法がより独立な分布を得ている。大まかにいって、独立性の強化という点においてこの二つは同等といえるだろう。線形分布変換法が 2 次のモーメントまで考慮しているのに対し、平均一致法では線形以外の評価値変換が可能であるという利点がそれぞれあるためであろう。

## 5. まとめ

本論文では、利用者が指定したセンシティブ特徴に対する独立性を強化させる独立性強化型システムについて論じた。非常に簡潔ではあるが既存手法と同等の精度を後処理型の線形分布変換法で得ることができた。平均一致法と同等の非線形変換が可能で、2 次以上の分布を考慮できる手法を開発でき

ば、予測精度を悪化させずにさらに独立性を強化できる方法が開発できるであろう。

謝辞: 本研究は JSPS 科研費 16700157, 21500154, 24500194, および 15K00327 の助成を受けた。Grouplens research lab. と Dr. Mohsen Jamali による実験データの提供に感謝する。

## 参考文献

- [Celma 08] Celma, Ò. and Cano, P.: From Hits to Niches?: or How Popular Artists Can Bias Music Recommendation and Discovery, in *Proc. of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition* (2008)
- [Forden 12] Forden, S.: Google Said to Face Ultimatum From FTC in Antitrust Talks, Bloomberg (2012), (<http://bloom.bg/PPNEaS>)
- [Gro] GroupLens research lab, University of Minnesota (<http://www.grouplens.org/>)
- [Jamali 10] Jamali, M. and Ester, M.: A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks, in *Proc. of the 4th ACM Conf. on Recommender Systems*, pp. 135–142 (2010)
- [Kamishima 03] Kamishima, T.: Nantonac Collaborative Filtering: Recommendation Based on Order Responses, in *Proc. of The 9th Int'l Conf. on Knowledge Discovery and Data Mining*, pp. 583–588 (2003)
- [Kamishima 12] Kamishima, T., Akaho, S., Asoh, H., and Sakuma, J.: Enhancement of the Neutrality in Recommendation, in *Proc. of the 2nd Workshop on Human Decision Making in Recommender Systems*, pp. 8–14 (2012)
- [Kamishima 13] Kamishima, T., Akaho, S., Asoh, H., and Sakuma, J.: Efficiency Improvement of Neutrality-enhanced Recommendation, in *Proc. of the 3rd Workshop on Human Decision Making in Recommender Systems*, pp. 1–8 (2013)
- [Koren 08] Koren, Y.: Factorization Meets the Neighborhood: A Multifaceted Collaborative Filtering Model, in *Proc. of the 14th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*, pp. 426–434 (2008)
- [Ruggieri 10] Ruggieri, S., Pedreschi, D., and Turini, F.: Data Mining for Discrimination Discovery, *ACM Transactions on Knowledge Discovery from Data*, Vol. 4, No. 2 (2010)
- [Sweeney 13] Sweeney, L.: Discrimination in Online Ad Delivery, *Communications of the ACM*, Vol. 56, No. 5, pp. 44–54 (2013)